ÉTICA, TRANSHUMANISMO Y SUSTENTABILIDAD EN LA IA



Red Mexicana de Investigadores en Estudios Organizacionales

50 AÑOS UAM Unidad Azcapotzalco

División de Ciencias Sociales y Humanidades

Departamento de Administración

UAM

Ética, Transhumanismo y Sustentabilidad de la IA
Primera edición, octubre 2024

Coordinación general del trabajo:
Oscar Lozano Carrillo
Alfredo Garibay Suárez

Integración editorial y digital primera fase
Aylin Torres Ramírez
(Maestra en Administración)

Marlene Atenea Grimaldo Ramírez
(Estudiante de la Licenciatura en Administración UAM-A)

Diseño de portada *(imagen asistida por inteligencia artificial)*
Brenda Itzelh García Serrano
(Diseñadora de la Comunicación Gráfica)

Editado por:

Red Mexicana de Investigadores
en Estudios Organizacionales (REMINEO), A.C.

América 126-A int. 202, Col. Parque San Andrés,
Coyoacán, Ciudad de México, C.P. 04040

ISBN: 978-607-8049-10-3

El contenido, las opiniones y la originalidad de los artículos publicados en este libro son responsabilidad exclusiva de sus autores y no reflejan necesariamente el punto de vista de los coordinadores, del editor o de los avales académicos.

Editado en México.

ÉTICA, TRANSHUMANISMO Y SUSTENTABILIDAD DE LA IA

Coordinación general del trabajo:

Oscar Lozano Carrillo
Alfredo Garibay Suárez

Red Mexicana de Investigadores en Estudios Organizacionales A.C. (REMINEO)

CONTENIDO

INTRODUCCIÓN 11
OSCAR LOZANO CARRILLO
ALFREDO GARIBAY SUÁREZ

CAPÍTULO I
ÉTICA DE LA INTELIGENCIA ARTIFICIAL 21
OSCAR LOZANO CARRILLO

CAPÍTULO II
HUMANIDAD Y MÁQUINAS: ¿CÓMO DEFINIR EL BIEN Y EL MAL EN UN MUNDO INTELIGENTE? 41
CINTHIA NOEMI ZACATENCO ARELLANO

CAPÍTULO III
HUMANIDAD Y GESTIÓN EMPRESARIAL: MIRADA FUTURISTA HACIA UN AGENTE DE MEJORA / REEMPLAZO 63
AMÉRICA PELAYO RAMÍREZ

CAPÍTULO IV
EL CAMBIO DE LA ÉTICA DEL DEPORTE CON EL IMPACTO DE LA INTELIGENCIA ARTIFICIAL 83
MARLENE ATENEA GRIMALDO RAMÍREZ

CAPÍTULO V
INNOVACIÓN SUSTENTABLE: EL PODER TRANSFORMADOR DE LA INTELIGENCIA ARTIFICIAL EN UN MUNDO EN CAMBIO 105
ALEJANDRA JARAMILLO RODRIGUEZ

CAPÍTULO VI
EL USO ÉTICO DE LA INTELIGENCIA ARTIFICIAL EN LAS ACTIVIDADES ACADÉMICAS UNIVERSITARIAS. PERCEPCIÓN DEL ESTUDIANTE 129
MIGUEL ÁNGEL OLIVARES URBINA
RICARDO SONDA DE LA ROSA

CAPÍTULO VII
ASISTENTES VIRTUALES: TRANSFORMACIÓN TECNOLÓGICA Y DESAFÍOS ÉTICOS 163
FÁTIMA CECILIA PÉREZ PONCE

CAPÍTULO VIII
INTELIGENCIA ARTIFICIAL EN EL AULA: ÉTICA Y PRÁCTICAS 191
GISELA COTA YUCUPICIO
KARLA MARIELA SANDOVAL COTA
EDNA KARINA MIRANDA MONDACA

CAPÍTULO IX
EL PAPEL DUAL DE LA IA ANTE EL CAMBIO CLIMÁTICO E INICIATIVAS GLOBALES 219
AYLIN TORRES-RAMÍREZ

CAPÍTULO X
EXPLORANDO NUEVAS FRONTERAS:
IA Y EL FUTURO DE LA GALERÍA VIRTUAL EN LA UAM AZCAPOTZALCO 235
YADIRA ZAVALA OSORIO
VÍCTOR GÓMEZ–QUINTERO

CAPÍTULO XI
LA INTELIGENCIA ARTIFICIAL EN LA MIPYME. DESARROLLO TECNOLÓGICO Y SUSTENTABILIDAD 247
OSCAR LOZANO CARRILLO
JANETTE BRITO LAREDO
FRANCISCO JAVIER ARIAS VARGAS

CAPÍTULO XII
LA INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA TRANSHUMANISTA A LA LUZ DE LOS DERECHOS HUMANOS 267
KARINA DEL REFUGIO VALLEJO QUINTERO
JESÚS MANUEL LÓPEZ MARROQUÍN
RICARDO PÉREZ BOJÓRQUEZ

CAPÍTULO XIII
EL DILEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL EN LA BÚSQUEDA DE LA SUSTENTABILIDAD: UNA MIRADA CRÍTICA 291
EDGAR ALFONSO SANSORES GUERRERO
JUANA EDITH NAVARRETE MARNEOU

CAPÍTULO XIV
LA IA ES UNA IDEOLOGÍA, NO UNA TECNOLOGÍA 315
MARITZA ALESSANDRA ZEPEDA BARBOSA

CAPÍTULO XV
REVOLUCIONANDO LA CONCIENCIA DIGITAL: LA ÉTICA DE TERCERA GENERACIÓN EN LA ERA DE LA INTELIGENCIA ARTIFICIAL 341
VIRGINIA LÓPEZ NEVÁREZ
ALEJANDRO LÓPEZ ÁLVAREZ
MANUEL ALEJANDRO ZAVALA NEVÁREZ

CAPÍTULO XVI
EGOLATRÍA Y EVOLUCIÓN: LA TRANSFORMACIÓN DE LA CONDICIÓN HUMANA 381
ALFREDO GARIBAY SUÁREZ

CAPÍTULO XVII
ÉTICA KANTIANA EN LA EVALUACIÓN DEL RIESGO CREDITICIO: APLICACIONES DE INTELIGENCIA ARTIFICIAL Y MACHINE LEARNING 389
LUIS ANTONIO VERDUGO CALDERÓN

CAPÍTULO XVIII
LA IA COMO GUARDIÁN DEL PLANETA: INNOVACIONES PARA UN FUTURO SOSTENIBLE 421
VÍCTOR ALFONSO GASCON SORIANO

CAPÍTULO XIX
TRASHUMANISMO Y LA FUSIÓN DE HOMBRE-MAQUINA: ¿PODRÍAMOS LLEGAR A UNA SIMBIOSIS CON MÁQUINAS INTELIGENTES?..... 445
EMILY NICOLE SILVA ESCAMILLA

CAPÍTULO XX
EL RAZONAMIENTO Y LA IA, UNA IDA Y UNA VUELTA: UN VIAJE INESPERADO DE LOS NUEVOS MODELOS DE RAZONAMIENTO DE LA IA..... 471
VÍCTOR DANIEL SANTOS HORTELANO

CAPÍTULO XXI
LA CONFIABILIDAD EN LA TOMA DE DECISIONES AUTOMATIZADAS Y LOS RETOS PARA EL DESARROLLO DE MARCOS DE GOBERNANZA EN AMÉRICA LATINA 497
YADIRA ZAVALA OSORIO
VICTOR GÓMEZ QUINTERO

CAPÍTULO XXII
RETOS DE LA IA ANTE LA VIOLENCIA DIGITAL EN AMBIENTES UNIVERSITARIOS DE AMÉRICA LATINA 509
WENDY SEBASTIANA HERNÁNDEZ DEL PUERTO
ROSILUZ CEBALLOS POVEDANO

CAPÍTULO XXIII
SUSTENTABILIDAD DE MUNICIPIOS METROPOLITANOS MEDIANTE LA APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL A LA GESTIÓN DEL TERRITORIO..... 541
JORGE ALBERTO ROSAS CASTRO
GABRIELA CAMACHO DEONICIO

CAPÍTULO XXIV
LA IA Y LOS SABERES LOCALES, EN EL ÁMBITO ORGANIZACIONAL 575
JESÚS MANUEL RAMOS GARCÍA

CAPÍTULO XXV
HUMANISMO Y TRANSHUMANISMO..... 591
JONATAN GARCÍA CAMPOS

CAPÍTULO XXVI
INTELIGENCIA, TRANSHUMANISMO Y ¿LÍMITES ÉTICOS DE LA IA? 613
PAOLA HERNÁNDEZ CHÁVEZ

CAPÍTULO XXVII
CEREBROS O CIRCUITOS: ¿DEBERÍA IMPORTARNOS QUIÉN CREA LO QUE LEEMOS?..... 631
RAFAEL ANGELES GARCÍA

CAPÍTULO XXVIII
INTELIGENCIA ARTIFICIAL ESTRATÉGICA
PARA LA MICRO, PEQUEÑA Y MEDIANA EMPRESA. UNA NUEVA ÉTICA DE DEMOCRATIZACIÓN DEL CONOCIMIENTO..... 643
OSCAR LOZANO CARRILLO
ALFREDO GARIBAY SUAREZ
JESUS MANUEL RAMOS GARCÍA
JORGE ALBERTO ROSAS CASTRO
EDGAR ALFONSO SANSORES GUERRERO
ZAIRA YÉREMI MARTÍNEZ HERRERA

SEMBLANZAS DE AUTORES..... 669

INTRODUCCIÓN

OSCAR LOZANO CARRILLO[1]
ALFREDO GARIBAY SUÁREZ[2]

La Inteligencia Artificial aparece ya en todos los ámbitos de la vida cotidiana; todas las personas la usan cuando viajan en avión, ven la televisión o más específicamente usa una computadora, o simplemente usan su *SMARTPHONE*. En ámbitos más específicos, hacemos uso de la IA cuando entramos a las redes sociales como *FACEBOOK, INSTAGRAM, TIK TOK*; buscadores como *GOOGLE, BING, YAHOO*; también en aplicaciones de música o compras *SPOTIFY, AMAZON, TEMU*; cuando usamos mapas y tráfico *WAZE, GOOGLE MAPS*; o en espacios de entretenimiento como *NETFLIX, PRIME, HBO* o *YOUTUBE*. Pero si estás leyendo este libro, seguramente eres quién las usa directamente a través de aplicaciones como *ChatGPT* de *OPEN AI*, *COPILOT* de *MICROSOFT*, *GEMINI* de *GOOGLE* o *Llama 3.1* de *META*, esta última con la implacable compañía permanente dentro de aplicación *WHATSAPP*, todos ellos en el área de procesamiento del lenguaje natural (NLP); también es muy probable que haya utilizado IA en APP en ámbitos del diseño como *CANVA, MIDJOURNEY, ADOBE EXPRESS*; o en la producción de videos con *CAPCUT, ADOBE PREMIER, COREL* o en idiomas con *HYEGEN, BABBEL, TALKPAL*.

Como puedes ver, la IA ha llegado masivamente para quedarse y su velocidad de evolución es impresionantemente rápida. En sólo días, una APP es superada por otra o por ella misma con resultados que ya no dejan tiempo para suspender la sorpresa que nos causó la primera

[1] Profesor-Investigador, del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco, Ciudad de México, México. Dirige el Laboratorio de Inteligencia Artificial UAM Azcapotzalco.

[2] Profesor-Investigador Titular C de Tiempo Completo, del Departamento de Administración de la Universidad Autónoma Metropolitana, Azcapotzalco, Ciudad de México, México.

versión, como ocurrió, por ejemplo, con los *ChatGPT* que ya están en su versión 4o que supera a la 4 por cambios sustanciales como el interactuar con video y audio como si fuera otra persona, evolución que en aproximadamente dos años ha sido extraordinaria.

Sin embargo, es fundamental hacer notar que la perspectiva de desarrollo y mejora tecnológica o técnica, es una simple dimensión superficial de lo que en realidad está siendo un cambio evolutivo del ser humano, pues el impacto que esta capacidad tecnológica trae a la vida de la gente es sorprendente; nosotros solo hemos dado cuenta de algunas de las aplicaciones más importantes de la vida cotidiana donde aparece la IA, pero sin duda, es en los ámbitos más profundos de la construcción social de la realidad en donde está ya impactando su uso. Por ejemplo, en la Educación, donde tenemos un uso acelerado de los GPT como consulta, asesoría, enseñanza y por supuesto la automatización (UNESCO, 2024)[3]; en la Salud, donde cada día la IA sorprende más con su uso en el ámbito del diagnóstico, en las intervenciones médicas directa con robots o con la experimentación de medicinas, fármacos y vacunas[4]; dentro de la Política es impactante el uso de las redes sociales como agentes de adoctrinamiento, pero también en la numeralia, proyecciones y encuestas, siendo sin duda uno de los referentes de la mala política la guerra, donde la carrera armamentista gracias a la IA, preocupa sobre manera a la sociedad mundial, como está sucediendo actualmente en la guerra entre Rusia y Ucrania, don-

[3] La inteligencia artificial (IA) está revolucionando la educación de varias maneras. Permite crear experiencias de aprendizaje personalizadas, adaptando el contenido y el ritmo a las necesidades individuales de cada estudiante. Además, la IA se utiliza para automatizar tareas administrativas como la calificación, el control de asistencia y la gestión del rendimiento, lo que reduce la carga de trabajo de los docentes. También existen sistemas de tutoría basados en IA, que proporcionan apoyo adicional a los estudiantes, respondiendo preguntas y ofreciendo explicaciones detalladas sobre diversos temas. La IA ayuda a identificar estudiantes en riesgo de abandono escolar mediante el análisis de datos, permitiendo intervenciones tempranas y personalizadas. Además, plataformas impulsadas por IA pueden recomendar recursos educativos específicos basados en el progreso y las necesidades de los estudiantes.

[4] Cuando se aplicó IA en la pandemia por el SARS-CoV-2, virus que causa la Covid-19, el modelo informático eliminó rápidamente el 95 % de los compuestos que posiblemente podrían haber tratado al patógeno y señaló las mejores opciones, según el estudio (Consalud, 2021).

de a través de la IA se generan ataques de drones autónomos en una logística más letal y precisa (Expansión, 2023).

Como se puede ver, la IA está en prácticamente todos los aspectos de la vida cotidiana, lo que sin duda traerá múltiples beneficios a la vida del ser humano como es el caso de la salud, el transporte o la cultura, sin embargo, cada día surgen más inquietudes sobre el impacto que puede tener en ámbitos negativos para la humanidad. Cada día hay más voces que se levantan para advertir los riesgos que se pueden generar por la evolución no controlada de la IA como puede ser la singularidad contra el ser humano, como dice Elon Musk (2017) *"la IA también plantea riesgos significativos, como la pérdida de empleos, la dependencia de la tecnología y la posibilidad de una inteligencia artificial hostil"* o la IA no normada como puede ser en casos de fallas como las ocurridas hace poco tiempo en los servidores de *MICROSOFT*, o por aquellas posibilidades de ser utilizada para fines nocivos contra la sociedad y el ser humano, como puede ser el caso de la delincuencia organizada.

A partir de estas ideas, se realizó el WEBINAR "Ética, Transhumanismo y Sustentabilidad", organizado por el Laboratorio de Inteligencia Artificial del Departamento de Administración de la UAM Azcapotzalco, el Cuerpo Académico de Negocios Internacionales de la Universidad del Caribe, Universidad Autónoma de Quintana Roo, Universidad Autónoma de Occidente, Universidad de Medellín y la Universidad Católica de Valparaíso. En este Webinar se presentaron diversas ponencias con los tres temas enmarcados donde la discusión fue rica e interesante, el presente libro retoma estas coordenadas en las tres líneas.

ÉTICA EN LA INTELIGENCIA ARTIFICIAL

La inteligencia artificial (IA) se ha convertido en una de las tecnologías más transformadoras de nuestro tiempo, con aplicaciones que van desde la medicina hasta la educación y la industria. Sin embargo, el rápido avance de la IA también ha planteado importantes desafíos éticos que deben ser abordados para garantizar que su desarrollo y uso

beneficien a toda la humanidad. La inteligencia artificial ha recorrido un largo camino desde sus inicios en la década de 1950. Inicialmente, los sistemas de IA se limitaban a tareas simples como el reconocimiento de patrones y la resolución de problemas básicos. Hoy en día, la IA se utiliza en una amplia variedad de aplicaciones, incluyendo el diagnóstico médico, la conducción autónoma y la personalización de servicios en línea. A medida que la IA continúa evolucionando, su potencial para transformar la sociedad es inmenso, pero también lo son los desafíos éticos que plantea.

Uno de los principales desafíos éticos en la IA es el sesgo y la discriminación. Los algoritmos de IA pueden perpetuar y amplificar los sesgos existentes en los datos, lo que puede resultar en decisiones injustas y discriminatorias. Por ejemplo, se ha demostrado que algunos sistemas de reconocimiento facial tienen tasas de error más altas para personas de color y mujeres, además, la privacidad y la seguridad de los datos son preocupaciones críticas, ya que los sistemas de IA a menudo requieren grandes cantidades de datos personales para funcionar de manera efectiva. La transparencia y la rendición de cuentas también son esenciales para garantizar que las decisiones tomadas por la IA sean comprensibles y justificables.

El impacto de la IA en la sociedad y la economía es profundo. Si bien la IA tiene el potencial de mejorar la eficiencia y la productividad, también puede exacerbar las desigualdades existentes y desplazar a los trabajadores en ciertos sectores, sobre todo en materia de desempleo, incluso se afirma que la IA afectará el 40 % del empleo en todo el mundo (Georgieva, 2024) por lo que es crucial que las políticas y regulaciones se diseñen para mitigar estos efectos negativos y garantizar que los beneficios de la IA se distribuyan de manera equitativa. Se han desarrollado varios principios éticos y normativas, por ejemplo el de la UNESCO (2021), organización que ha propuesto un marco normativo universal sobre la ética de la IA que enfatiza la dignidad humana, el bienestar y la prevención de daños. Este marco también se destaca la importancia de la transparencia, la responsabilidad y la justicia en el desarrollo y su uso, sin embargo, la velocidad a la que va la IA no ha permitido definir una normatividad apropiada y mucho menos oportuna.

Existen numerosos ejemplos de uso ético y no ético de la IA. Por ejemplo, el uso de IA en la medicina ha permitido diagnósticos más precisos y tratamientos personalizados, mejorando significativamente la atención al paciente. Sin embargo, también ha habido casos en los que la IA se ha utilizado de manera poco ética, como en la vigilancia masiva y la manipulación de la opinión pública. Estos casos destacan la necesidad de un marco ético robusto para guiar el desarrollo y uso de la IA.

Lo mencionado hasta ahora es lo que se ve de manera inmediata y evidente, sin embargo, se tienen algunas tensiones más complejas con relación a la Ética en la IA, como lo es, por ejemplo la relevancia que se le puede llegar a dar a la IA en detrimento de la participación humana (Sardin y Martínez, 2020), situación que cada vez se ve más cercana al pensar que será una IA con algoritmos probabilísticos los que a manera de una *"Ética Digital incuestionable"* decidan en algún momento situaciones jurídicas, legales, electorales, administrativas o deportivas.

IA Y TRANSHUMANISMO

El transhumanismo es un movimiento que busca mejorar la condición humana mediante la tecnología, y la Inteligencia Artificial (IA) juega un papel fundamental en este contexto como es: 1) la Mejora cognitiva y física: *"La IA puede ayudar a mejorar nuestras capacidades cognitivas y físicas, permitiéndonos alcanzar un nivel de inteligencia y habilidad superior"* (Kurzweil, 2005). Hemos visto como a través de la tecnología *GPT* ya se tiene una capacidad importante de información relevante y muy actualizada en la palma de la mano, podríamos decir que para estos efectos, de alguna manera ya los *smartphone* son extensiones -aún desconectadas físicamente- de nuestros cuerpos, sobre todo en la juventud que permanece pegada a sus celulares grandes cantidades de tiempo[5]; 2) Extensión de la vida y la salud: *"La IA puede ayudar a*

[5] En México, las cifras sobre el tiempo promedio de uso de celular, según la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2023, varía según el grupo de edad:

- 18 a 24 años: Este grupo pasa en promedio 6.3 horas al día usando internet en sus disposi-

encontrar curas para enfermedades y mejorar la calidad de vida, lo que podría llevar a una vida más larga y saludable" (De Grey y Rae, 2015). Si bien, este aún es un escenario hipotético, la perspectiva de la inmortalidad se convierte en una discusión sumamente impactante, ya al día de hoy se habla de inmortalidad digital con empresas que están generando avatares de la gente que murió, dotadas de elementos estéticos como imagen, voz y sonido que imitan a quien ya no está y a través de tecnología generativa propician la interacción con ellos tal y como si siguieran viviendo, con recuerdos, gustos y comportamientos iguales a los que tenían antes de morir; 3) Fusión de humanos y máquinas: *"La IA puede permitir la fusión de humanos y máquinas, creando una nueva forma de vida que combina lo mejor de ambos mundos"* (Bostrom, 2014). Esta es otra de las posibilidades en las que también se puede extender la vida.

IA Y SUSTENTABILIDAD

En el ámbito de la sustentabilidad existen diversas posibilidades interesantes que la IA puede generar para aprovechar el desarrollo, entre las cuestiones positivas se presentan:

1. Reducción de emisiones de carbono: La IA puede optimizar el consumo de energía en ciudades inteligentes y reducir las emisiones de carbono mediante el uso de dispositivos del Internet de las Cosas (IoT) que modulan su consumo eléctrico (Fundación Innovación Bankinter, 2023).
2. Integración de energías renovables: Las redes inteligentes impulsadas por IA facilitan una mejor integración de las energías renovables, lo que contribuye a una mayor eficiencia energéti-

---

tivos móviles.

- 25 a 34 años: El uso promedio es de 4.8 horas al día.
- 35 a 44 años: Dedican alrededor de 4.2 horas al día.
- 45 a 54 años: Usan el celular aproximadamente 3.5 horas al día.
- 55 años o más: El uso promedio es de 2.8 horas al día.

ca y reducción de la dependencia de combustibles fósiles (Fundación Innovación Bankinter, 2023).

3. Monitoreo ambiental: La IA puede identificar tendencias de desertificación y otros cambios ambientales a través de imágenes satelitales, permitiendo una respuesta más rápida y efectiva a estos problemas (Fundación Innovación Bankinter, 2023).
4. Economía circular: La IA fomenta la circularidad al ayudar a reducir la huella ambiental y promover prácticas sostenibles en el consumo y la producción (UNESCO, 2023).
5. Desarrollo de materiales sostenibles: La IA acelera el desarrollo de materiales de baja emisión de carbono y cultivos resistentes al clima, lo que es crucial para enfrentar el cambio climático (Cloudmasters, 2023).

Existen, sin embargo, también algunas consecuencias o riesgos importantes con la IA:

1. Consumo de energía: Los algoritmos de IA requieren un poder de procesamiento intensivo, lo que implica el uso de grandes cantidades de energía eléctrica. Esto contribuye a la emisión de gases de efecto invernadero y al cambio climático (Blanco, 2024).
2. Residuos electrónicos: La fabricación de hardware específico para la IA, como chips y servidores, genera una huella ecológica significativa. Además, la disposición de estos componentes electrónicos al final de su vida útil plantea desafíos en términos de residuos electrónicos (Funds Society, 2024).
3. Extracción de recursos minerales**:** En la producción de hardware para IA requiere la extracción de materiales como el silicio, lo que tiene un impacto ambiental considerable (Funds Society, 2024).
4. Uso de agua**:** Los centros de datos que alojan y procesan grandes volúmenes de datos de IA consumen una gran cantidad de agua para el enfriamiento de los servidores, lo que puede afectar los recursos hídricos locales (Funds Society, 2024).

Finalmente, diremos que la IA se hace presente en la sustentabilidad al tener un impacto inmediato y contundente en la naturaleza del ser humano y su concepción y significación de la realidad, en ámbitos más implícitos de su realidad construida dentro de un orden global y consumista, el ser humano de la actualidad verá impactada su perspectiva de futuro por los referentes que la misma IA le ofrece a través de su sistemática y contundente formación cultural que los artefactos tecnológicos están generando en nuestros hijos e hijas, es algo que no debemos perder de vista.

Ética, Transhumanismo y Sustentabilidad son sin duda temas relevantes del Futuro de la Inteligencia Artificial. En este libro se realizará un acercamiento a los mismos con la clara intención de intentar explicar el devenir social, económico, político y sobre todo cultural de la sociedad de la civilización artificial, como diría Lasalle (2024), en un espectro de futuro que difícilmente está siendo pronosticado. Creemos que las instituciones de educación superior como las que participamos en este trabajo, tenemos el compromiso de investigar, discutir, proyectar y acompañar a la sociedad a través de nuestros alumnos y alumnas, pero también investigaciones y su difusión, para convertirnos en referentes del cuidado del ser humano por sobre los impactos que cualquier problema nacional o mundial genere, en este caso, así podríamos plantearlo, sobre el problema de la Inteligencia Artificial y su descontrol actual.

REFERENCIAS BIBLIOGRÁFICAS

Blanco, J.J. (2024, febrero). *El impacto de la inteligencia artificial en el medio ambiente: desafíos y soluciones. Universo-IA.com.* https://universo-ia.com/como-afecta-la-inteligencia-artificial-al-medio-ambiente/

Bostrom, N. (2014). *Superinteligencia: Caminos, peligros, estrategias.* Editorial Reverte.

Cloudmasters. (2023). *Cómo la Inteligencia Artificial puede potenciar la sostenibilidad. https://www.cloudmasters.es/como*-lainteligencia-artificial-puede-potenciar-la-sostenibilidad-guia-completa/

Consalud. (2021) *La inteligencia artificial predice vacunas potenciales contra la COVID-19*. https://www.consalud.es/pacientes/especial-coronavirus/inteligencia-artificial-predice-vacunas-potenciales-covid-19_92330_102.html

Georgieva, K. (2024). La economía mundial transformada por la inteligencia artificial ha de beneficiar a la humanidad. *Blog FMI*, *16*.

De Grey, A., y Rae, M. (2015). *El fin del envejecimiento: los avances que podrían revertir el envejecimiento humano durante nuestra vida*. Lola Books.

ENDUTIH. (2023). *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares. Subsistema de Información Económica*. https://www.inegi.org.mx/programas/endutih/2023/

Expansión. (2023, diciembre). *El uso de la IA llega a los cuerpos miliares y cambia el campo de batalla*. https://expansion.mx/tecnologia/2023/12/18/inteligencia-artificial-en-industria-militar#:~:text=El%20uso%20de%20la%20IA,cambios%20dr%C3%A1sticos%20en%20las%20guerras.

Fundación Innovación Bankinter. (2023). *Inteligencia Artificial sostenible*. *https://www.fundacionbankinter.org/noticias/inteligencia*artificial-sostenible/

Funds Society. (2024, febrero). *Inteligencia artificial: ¿cuál es su impacto medioambiental?* https://www.fundssociety.com/es/opinion/inteligencia-artificial-cual-es-su-impacto-medioambiental/

Kurzweil, R. (2015). *La Singularidad está cerca: Cuando los humanos transcendamos la biología*. Lola Books.

Musk, E. (2017). Discurso en la Conferencia de Inteligencia Artificial de 2017.

Sadin, É., y Martínez, M. (2020). *La inteligencia artificial o el desafío del siglo: anatomía de un antihumanismo radical*. Buenos Aires: Caja negra.

UNESCO (2021). *Recomendación sobre la ética de la inteligencia artificial*. https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial

UNESCO. (2023). *IA por el planeta: Destacando las innovaciones de la IA para acelerar su impacto*. https://www.unesco.org/es/articles/ia-por-elplaneta-destacando-las-innovaciones-de-la-ia-para-acelerar-su-impacto

UNESCO. (2024). *El uso de la IA en la educación. Decidir el futuro que queremos*. https://www.unesco.org/es/articles/el-uso-de-la-ia-en-la-educacion-decidir-el-futuro-que-queremos

CAPÍTULO I

ÉTICA DE LA INTELIGENCIA ARTIFICIAL[1]

OSCAR LOZANO CARRILLO[2]

RESUMEN

La velocidad vertiginosa a la que viaja el desarrollo de la Inteligencia Artificial ha propiciado un debate en relación con la ética de su creación y sobre todo de sus resultados e implicaciones. Existen perspectivas que elevan una demanda (Lasalle, 2024; Sadin, 2024) de normar la evolución de la IA ya que no aparece de inicio "dicen" una ética clara o explicita de su creación, uso y evolución, la mayor parte de estas perspectivas son generales, donde se habla de transparencia, propiedad intelectual, responsabilidad entre otras, sin embargo existe una perspectiva, visión o posición más compleja que esta, que es la de que en realidad si existe una Ética alrededor de la IA, sólo que no necesariamente es la que quisiéramos la mayoría de habitantes de este planeta, ya que se trata de una ética de la eficiencia, la crematística, el consumismo, el colonialismo y extractivismo de la gran empresa, marcada por las coordenadas económico-culturales de la *GAMFA (Google, Amazón, Microsoft, Facebook* -ahora *META*- y *Apple)* las más grandes empresas del orbe, así como la ética de la dominación que las

[1] La ética se refiere al campo de estudio y práctica relacionado con las implicaciones morales y los impactos sociales de *inteligencia artificial*. Abarca los principios y directrices que rigen el diseño, desarrollo, implementación y uso de sistemas de IA. Este dominio garantiza que las tecnologías de IA se desarrollen y apliquen de manera que se alineen con los valores humanos y los estándares éticos (Florkin, 2024).

[2] Profesor-Investigador, del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco, Ciudad de México, México. Dirige el Laboratorio de Inteligencia Artificial UAM Azcapotzalco.

potencias hegemónicas China y Estados Unidos, que buscan tener la supremacía mundial. El presente capítulo explora algunas ideas sobre el desarrollo de la IA y su ética o posibles éticas que la rodean, donde la condición más preocupante será el papel del ser humano ante la posible singularidad que lo convierta en un *HOMO DEUS*.

Palabras clave: IA, Ética, Homo Deus.

INTRODUCCIÓN

La ética de la inteligencia artificial (IA) es un campo emergente y crucial en el desarrollo tecnológico contemporáneo. A medida que la IA se integra cada vez más en diversos aspectos de la vida cotidiana, desde la medicina hasta la educación y el entretenimiento, surgen importantes preguntas sobre cómo estas tecnologías deben ser diseñadas, implementadas y reguladas para garantizar que beneficien a la sociedad de manera justa y equitativa, sin embargo, estas perspectivas generales y más comunes se presentan en una perspectiva teórica y de sentido común.

La ética de la IA entonces es un campo dinámico y en constante evolución que aborda algunas de las cuestiones más apremiantes de nuestro tiempo. A medida que la IA continúa avanzando y transformando diversos aspectos de la sociedad, es fundamental que se desarrollen y se implementen marcos éticos robustos para guiar su uso. La transparencia, la responsabilidad y la equidad son pilares esenciales en este esfuerzo, y las tendencias actuales indican un movimiento hacia una mayor colaboración internacional y una participación más amplia de la sociedad civil en estos debates. Solo a través de un enfoque ético y responsable podemos asegurar que la IA beneficie a toda la humanidad de manera justa y equitativa.

Bajo esta perspectiva, una de las principales necesidades en la ética de la IA es la ***transparencia.*** Se dice que los algoritmos de IA a menudo operan como *"cajas negras"*, donde las decisiones se toman sin que los usuarios comprendan completamente cómo se llegaron a ellas. *Esto*

puede llevar a una falta de confianza y a la percepción de injusticia, especialmente cuando las decisiones afectan aspectos críticos de la vida de las personas, como la aprobación de préstamos o la selección de candidatos para empleos (UNESCO, 2023).

Otra necesidad de esta visión tradicional habla de que es crucial la ***responsabilidad.*** Es esencial que haya mecanismos claros para responsabilizar a los desarrolladores y operadores de sistemas de IA por los resultados de sus tecnologías. *Esto incluye la capacidad de auditar y revisar las decisiones tomadas por la IA, así como la implementación de políticas que aseguren que los sistemas de IA se utilicen de manera ética y responsable* (UNESCO, 2023).

La ***equidad*** es otro aspecto fundamental. Los sistemas de IA deben ser diseñados para evitar sesgos que puedan perpetuar o exacerbar las desigualdades existentes. *Esto es particularmente importante en contextos como la justicia penal, donde los algoritmos de predicción del crimen pueden discriminar contra minorías si no se diseñan y supervisan adecuadamente* (UNESCO, 2024).

En cuanto a las tendencias actuales, una de las más destacadas es el desarrollo de ***marcos normativos internacionales.*** *La UNESCO, por ejemplo, ha adoptado la "Recomendación sobre la ética de la inteligencia artificial", que proporciona un marco global para guiar el desarrollo y la implementación de tecnologías de IA de manera ética* (UNESCO, 2024). Este tipo de iniciativas buscan armonizar las prácticas éticas a nivel mundial y asegurar que los beneficios de la IA se distribuyan equitativamente. Otra tendencia importante es la ***participación de la sociedad civil*** en el debate sobre la ética de la IA. Organizaciones no gubernamentales, académicos y activistas están desempeñando un papel cada vez más importante en la vigilancia y la promoción de prácticas éticas en el desarrollo de la IA. *Esta participación es crucial para asegurar que las voces de todos los sectores de la sociedad sean escuchadas y consideradas en la formulación de políticas y regulaciones*.

Además, hay un creciente enfoque en la ***educación y la formación*** en ética de la IA. *Universidades y centros de investigación están desarrollando programas y cursos específicos para formar a la próxima generación de desarrolladores y responsables de políticas en los principios éticos que deben guiar el uso de la IA* (Florkin, 2024). Esta educación es vital

para crear una cultura de responsabilidad y conciencia ética en el campo de la tecnología.

Como se dijo al principio, la ética de la inteligencia artificial (IA) no solo es una cuestión general de principios morales y normativos, sino también un campo de intensa competencia geopolítica, especialmente entre China y Estados Unidos. Además, la ética capitalista de la IA, representada por las grandes corporaciones tecnológicas como *Google, Amazon, Microsoft, Facebook* y *Apple (GAMFA)*, refleja intereses económicos y de poder que moldean el desarrollo y la implementación de estas tecnologías.

La competencia entre China y Estados Unidos en el ámbito de la IA es un reflejo de sus aspiraciones de liderazgo global. Ambos países han identificado la IA como una tecnología estratégica que puede determinar su posición en la economía mundial y en la seguridad nacional. Por parte de Estados Unidos ha sido tradicionalmente el líder en innovación tecnológica, con empresas como *Google, Microsoft* y *Amazon* a la vanguardia de la investigación en IA. *En 2022, el gasto en investigación y desarrollo (I+D) de IA en Estados Unidos alcanzó los 52.000 millones de dólares, representando casi el 50 % del total mundial* (Al poniente, 2024). Este liderazgo se basa en un ecosistema robusto de universidades, empresas y capital de riesgo que fomenta la innovación. Sin embargo, la ética de la IA en Estados Unidos está influenciada por un enfoque capitalista que prioriza la innovación y el crecimiento económico sobre la regulación. *Esto ha llevado a un desarrollo rápido pero a menudo desregulado de tecnologías de IA, con preocupaciones sobre la privacidad, la equidad y la transparencia* (Al poniente, 2024).

China, por otro lado, ha hecho de la IA una prioridad nacional con el objetivo explícito de convertirse en el líder mundial en IA para 2030. *En 2022, la inversión china en IA superó los 25.000 millones de dólares, lo que equivale aproximadamente al 25 % del total global* (El cano, 2020). El gobierno chino ha lanzado iniciativas como el Plan de Desarrollo de Inteligencia Artificial de Nueva Generación, que fomenta la colaboración entre el sector público y privado. La ética de la IA en China está marcada por un que se ha dicho tiene un enfoque autoritario y utilitarista. *El gobierno chino utiliza la IA para la vigilancia masiva y el control social, lo que plantea serias preocupaciones sobre los derechos humanos y*

la privacidad (Al poniente, 2024), sin embargo, este enfoque también permite una implementación rápida y centralizada de tecnologías de IA, lo que puede dar a China una ventaja competitiva en ciertos aspectos.

En cuanto a las grandes corporaciones tecnológicas de Estados Unidos, conocidas colectivamente como *GAMFA (Google, Amazon, Microsoft, Facebook* y *Apple),* desempeñan un papel crucial en el desarrollo y la implementación de la IA. Estas empresas no solo lideran en términos de innovación tecnológica, sino que también influyen significativamente en la ética de la IA a través de sus prácticas y políticas corporativas. *Google* ha sido pionero en el desarrollo de algoritmos de aprendizaje automático y ha establecido principios éticos para guiar el uso de la IA. *Sin embargo, ha enfrentado críticas por su participación en proyectos militares y por la falta de transparencia en sus algoritmos* (El cano, 2020). La ética de *Google* refleja un equilibrio entre la innovación y la responsabilidad social, aunque a menudo se inclina hacia la maximización de beneficios. *Amazon* utiliza la IA para optimizar su cadena de suministro y mejorar la experiencia del cliente. *Sin embargo, ha sido criticado por el uso de IA en la vigilancia de los empleados y por prácticas laborales cuestionables* (UNESCO, 2024). La ética de *Amazon* está impulsada por la eficiencia y la rentabilidad, a menudo a expensas de la privacidad y los derechos laborales.

Microsoft ha adoptado un enfoque más proactivo en la ética de la IA, estableciendo un comité de ética y lanzando iniciativas para promover el uso responsable de la IA (UNESCO, 2024). Sin embargo, también ha enfrentado desafíos en la implementación de estos principios en todas sus operaciones globales. *Facebook ahora META utiliza la IA para personalizar el contenido y los anuncios, pero ha sido criticado por su papel en la difusión de desinformación y por la manipulación de datos personales*. La ética de *Facebook* está marcada por un conflicto entre la maximización del *engagement* y la protección de la privacidad y la integridad.

Por último, *Apple* ha destacado por su enfoque en la privacidad y la seguridad en el uso de la IA. Sin embargo, también ha enfrentado críticas por su dependencia de la cadena de suministro en países con prácticas laborales cuestionables. La ética de *Apple* refleja un compro-

miso con la privacidad del usuario, aunque a veces a costa de la transparencia en otras áreas.

La ética de la IA es un campo complejo y multifacético que refleja tanto las aspiraciones geopolíticas de las naciones como los intereses económicos de las grandes corporaciones tecnológicas. La competencia entre China y Estados Unidos en el ámbito de la IA no solo se trata de quién lidera en innovación tecnológica, sino también de quién establece los estándares éticos que guiarán el futuro de esta tecnología. Por otro lado, las prácticas de las empresas *GAMFA* muestran cómo los intereses capitalistas pueden influir en la ética de la IA, a menudo priorizando la rentabilidad sobre la responsabilidad social.

ÉTICA Y TRANSPARENCIA

La transparencia en la inteligencia artificial (IA) es una dimensión crucial que se refiere a la capacidad de entender y explicar cómo funcionan los sistemas de IA y cómo se toman las decisiones. Esta transparencia es esencial para generar confianza, garantizar la responsabilidad y mitigar los riesgos asociados con el uso de la IA. A continuación, se desarrollan algunos ejemplos recientes a nivel mundial que ilustran la importancia y los desafíos de la transparencia en la IA.

1. Reglamento de Inteligencia Artificial de la Unión Europea (AIA)

En 2023, la Unión Europea avanzó significativamente en la regulación de la IA con la propuesta del Reglamento de Inteligencia Artificial (AIA). Este reglamento establece requisitos de transparencia para los sistemas de IA, especialmente aquellos considerados de alto riesgo. Según el AIA, los proveedores de sistemas de IA deben proporcionar información clara y comprensible sobre el funcionamiento de sus algoritmos, incluyendo sus capacidades y limitaciones. *Esto permite a los usuarios y a las autoridades supervisar y auditar estos sistemas de manera efectiva* (Agencia española de protección de datos, 2023)

2. Iniciativas de Transparencia en el Sector Salud

En el sector salud, la transparencia en el uso de IA es vital para garantizar la confianza de los pacientes y la eficacia de los tratamientos. Por ejemplo, en 2024, se implementaron modelos de IA para el diagnóstico de enfermedades a través de imágenes médicas. Estos modelos, como *UNet* y sus variantes, han demostrado un desempeño excepcional. *Sin embargo, la idoneidad de las decisiones generadas por estos modelos debe ser examinada rigurosamente para asegurar que los diagnósticos sean precisos y justos* (IA 360, 2024).

3. Transparencia en la Gobernanza de la IA

El Foro Económico Mundial ha destacado la importancia de la transparencia en la gobernanza de la IA. En 2024, se enfatizó que la adopción de prácticas transparentes no solo mitiga los riesgos asociados con la IA, sino que también puede proporcionar una ventaja competitiva a las organizaciones. *La transparencia en la IA incluye la divulgación de cómo se desarrollan y entrenan los modelos, así como la explicación de las decisiones tomadas por estos sistemas* (FEM, 2024).

4. Chatbots y Transparencia en el Gobierno

En el ámbito gubernamental, la transparencia en el uso de IA se ha manifestado a través de la implementación de *chatbots* para mejorar el acceso a la información pública. Estos *chatbots,* utilizados en procesos de apertura gubernamental, permiten a los ciudadanos obtener información de manera más eficiente y transparente. *Este enfoque no solo mejora la accesibilidad, sino que también fortalece la confianza en las instituciones públicas* (UNAM, 2024).

La transparencia en la IA es una dimensión esencial que abarca desde la regulación y la gobernanza hasta la implementación práctica en sectores críticos como la salud y el gobierno. Los ejemplos recientes demuestran que, aunque se han logrado avances significativos, aún existen desafíos importantes que deben abordarse para garantizar que los sistemas de IA sean transparentes, responsables y equitativos.

ÉTICA DE LA IA EN SUS DIMENSIONES MÁS COMUNES

Ética y responsabilidad

La ética y la responsabilidad en la IA son fundamentales para asegurar que estas tecnologías se desarrollen y utilicen de manera que beneficien a la sociedad en su conjunto. Estos principios buscan mitigar los riesgos y maximizar los beneficios de la IA, garantizando que se respeten los derechos humanos y se promueva la justicia social.

1. Responsabilidad en la IA: Caso de la Unión Europea

La Unión Europea ha sido pionera en establecer marcos regulatorios que promuevan la responsabilidad en el uso de la IA. En 2023, la Comisión Europea propuso el Reglamento de Inteligencia Artificial (AIA), que incluye requisitos estrictos para la transparencia, pero también para la responsabilidad. *Este reglamento obliga a los desarrolladores y operadores de sistemas de IA a proporcionar información clara sobre el funcionamiento de sus algoritmos y a garantizar que estos sistemas no perpetúen sesgos o discriminaciones* (Comisión Europea, 2023).

2. Responsabilidad en el Sector Financiero

En el sector financiero, la responsabilidad en el uso de la IA es crucial para evitar prácticas discriminatorias. En 2024, varios bancos en Estados Unidos implementaron sistemas de IA para evaluar la solvencia crediticia de los solicitantes de préstamos. Estos sistemas deben ser auditados regularmente para asegurar que no discriminen a ciertos grupos demográficos. *La transparencia en los criterios utilizados por estos algoritmos es esencial para mantener la confianza del público y garantizar la equidad* (Comisión Europea, 2024).

3. Ética en la IA: Caso de la Salud

En el ámbito de la salud, la ética en el uso de la IA es vital para proteger la privacidad de los pacientes y asegurar la precisión de los diagnósti-

cos. En 2023, se desarrollaron modelos de IA para predecir brotes de enfermedades y mejorar la respuesta a emergencias sanitarias. *Estos modelos deben ser transparentes y responsables, proporcionando explicaciones claras sobre cómo se toman las decisiones y asegurando que los datos de los pacientes se manejen de manera ética* (OMS, 2023).

4. Responsabilidad en la IA: Caso de la Educación

En el sector educativo, la responsabilidad en el uso de la IA es fundamental para asegurar que todos los estudiantes tengan acceso a una educación de calidad. En 2024, varias instituciones educativas implementaron sistemas de IA para personalizar el aprendizaje y mejorar los resultados académicos. *Estos sistemas deben ser diseñados de manera que no perpetúen desigualdades existentes y deben ser transparentes en cuanto a los datos utilizados y las decisiones tomadas* (UNESCO, 2024).

ÉTICA Y EQUIDAD

La equidad en la inteligencia artificial (IA) es una dimensión crítica que busca asegurar que estas tecnologías no perpetúen ni exacerben las desigualdades existentes. La ética y la equidad en la IA se centran en diseñar y utilizar sistemas que sean justos y accesibles para todos, independientemente de su origen, género, raza o condición socioeconómica. A continuación, se presentan ejemplos recientes de 2023 y 2024 que ilustran cómo se está abordando la equidad en la IA a nivel mundial.

1. Equidad en el Reconocimiento Facial

El reconocimiento facial ha sido una de las tecnologías de IA más controvertidas debido a su potencial para perpetuar sesgos raciales y de género. *En 2023, varios estudios demostraron que los sistemas de reconocimiento facial tenían tasas de error significativamente más altas para personas de color y mujeres* (Boulamwini y Gebru, 2023). *En respuesta, algunas ciudades y estados de Estados Unidos, como San Francisco y Massachusetts, implementaron moratorias o prohibiciones sobre el uso de esta*

tecnología por parte de las fuerzas del orden (CDHSF, 2023). Estas medidas buscan asegurar que la IA se utilice de manera justa y equitativa, evitando la discriminación y protegiendo los derechos civiles.

2. Equidad en la Educación

En el ámbito educativo, la IA tiene el potencial de personalizar el aprendizaje y mejorar los resultados académicos. Sin embargo, es crucial que estos sistemas no perpetúen las desigualdades existentes. *En 2024, se lanzaron varias iniciativas en Europa y América Latina para desarrollar plataformas de aprendizaje basadas en IA que sean inclusivas y accesibles para todos los estudiantes, independientemente de su contexto socioeconómico* (UNESCO, 2024). Estas plataformas utilizan algoritmos diseñados para identificar y mitigar sesgos, asegurando que todos los estudiantes tengan las mismas oportunidades de éxito.

3. Equidad en el Acceso a la Salud

La IA también está transformando el sector salud, pero es esencial que estas innovaciones sean accesibles para todos. *En 2023, la Organización Mundial de la Salud (OMS) lanzó una iniciativa para asegurar que las tecnologías de IA en salud sean equitativas y accesibles en países de bajos y medianos ingresos*(OMS, 2023). Esta iniciativa incluye la creación de directrices para el desarrollo y la implementación de sistemas de IA que consideren las necesidades específicas de estas poblaciones, asegurando que los beneficios de la IA en salud se distribuyan de manera justa.

4. Equidad en el Empleo

La IA está cambiando el panorama laboral, y es crucial que estos cambios no perpetúen las desigualdades en el empleo. *En 2024, varias empresas tecnológicas en Asia y Europa implementaron políticas para asegurar que sus sistemas de IA utilizados en procesos de contratación sean justos y equitativos* (FEM, 2024). Estas políticas incluyen auditorías regulares de los algoritmos de selección de personal para identificar y corregir posibles sesgos, así como la transparencia en los criterios utilizados para la toma de decisiones.

ÉTICA Y PARTICIPACIÓN DE LA SOCIEDAD CIVIL

La participación de la sociedad civil en la ética de la inteligencia artificial (IA) es fundamental para asegurar que estas tecnologías se desarrollen y utilicen de manera que reflejen los valores y necesidades de todas las personas. La sociedad civil, que incluye organizaciones no gubernamentales, académicos, activistas y ciudadanos, desempeña un papel crucial en la vigilancia, promoción y formulación de políticas éticas en la IA. A continuación, se presentan ejemplos recientes de 2023 y 2024 que ilustran cómo la sociedad civil está influyendo en la ética de la IA a nivel mundial.

1. Foro Global sobre la Ética de la IA de la UNESCO

En febrero de 2024, la UNESCO organizó el segundo Foro Mundial sobre la Ética de la Inteligencia Artificial en el Centro de Congresos Brdo de Kranj. *Este foro reunió a representantes de gobiernos, sector privado, academia y sociedad civil para discutir y compartir buenas prácticas en la gobernanza ética de la IA* (UNESCO, 2024). La participación de la sociedad civil en este foro fue crucial para asegurar que las políticas y recomendaciones reflejen una amplia gama de perspectivas y preocupaciones, promoviendo un enfoque inclusivo y equitativo en el desarrollo de la IA.

2. Iniciativas de Transparencia en América Latina

En 2023, varias organizaciones de la sociedad civil en América Latina, como *Derechos Digitales* y la Fundación *Karisma*, lanzaron iniciativas para promover la transparencia y la rendición de cuentas en el uso de la IA. *Estas organizaciones trabajan para asegurar que los sistemas de IA utilizados por gobiernos y empresas sean transparentes y no perpetúen desigualdades* (Derechos Digitales, 2023). Por ejemplo, han desarrollado herramientas y recursos para educar al público sobre sus derechos en relación con la IA y para monitorear el uso de estas tecnologías en sectores críticos como la justicia y la salud.

3. Participación Ciudadana en la Regulación de la IA en Europa

En 2024, la Comisión Europea lanzó una consulta pública sobre el Reglamento de Inteligencia Artificial (AIA), invitando a ciudadanos y organizaciones de la sociedad civil a proporcionar sus comentarios y sugerencias (ADH-Asia, 2023). Esta consulta permitió a la sociedad civil influir directamente en la formulación de políticas, asegurando que las regulaciones reflejen las preocupaciones y necesidades de la población. La participación ciudadana es esencial para desarrollar marcos regulatorios que sean justos y equitativos.

4. Campañas de Concienciación en Asia

En 2023, varias organizaciones de la sociedad civil en Asia, como la Asociación de Derechos Humanos de Asia y la Fundación para la Democracia y los Derechos Humanos, lanzaron campañas de concienciación sobre los riesgos y beneficios de la IA (ADH- Asia, 2023). Estas campañas buscan educar al público sobre cómo la IA puede afectar sus vidas y cómo pueden participar en el debate sobre su regulación. La educación y la concienciación son fundamentales para empoderar a los ciudadanos y asegurar que sus voces sean escuchadas en la formulación de políticas.

ÉTICA DE LA IA EN LA GEOPOLÍTICA Y LOS GRANDES CORPORATIVOS

La nueva Guerra Fría tiene profundas implicaciones para la gobernanza global y la estabilidad internacional. La competencia entre Estados Unidos y China está reconfigurando alianzas y creando nuevas tensiones en todo el mundo. La Unión Europea y otros países asiáticos se ven obligados a navegar en un entorno geopolítico cada vez más complejo, donde deben equilibrar sus intereses económicos y de seguridad.

Además, la interdependencia económica entre Estados Unidos y China significa que cualquier escalada en la confrontación podría tener consecuencias globales significativas. La cooperación en áreas como el cambio climático y la salud global también se ve amenazada por la creciente rivalidad entre estas dos superpotencias.

NUEVA GUERRA FRÍA POR LA IA

La dinámica geopolítica actual está marcada por una nueva Guerra Fría, en la que China y Estados Unidos son los principales antagonistas. Esta competencia no solo involucra a estas dos superpotencias, sino que también afecta a la Unión Europea y otros países asiáticos. A diferencia de la Guerra Fría del siglo XX, esta nueva confrontación se caracteriza por la interdependencia económica y tecnológica, así como por la competencia en áreas como la inteligencia artificial, la ciberseguridad y el comercio, esta nueva Guerra Fría entre Estados Unidos y China, con la participación de la Unión Europea y otros países asiáticos, está redefiniendo el panorama geopolítico global. Esta competencia se caracteriza por la interdependencia económica y tecnológica, y tiene implicaciones profundas para la estabilidad y la gobernanza global. A medida que esta rivalidad continúa desarrollándose, será crucial que las naciones involucradas encuentren formas de gestionar sus diferencias y cooperar en áreas de interés común para evitar una escalada que podría tener consecuencias desastrosas para el mundo entero.

1. Estados Unidos y China: Los Principales Protagonistas

La relación entre Estados Unidos y China ha sido tensa desde hace varios años, con una escalada significativa durante la administración de Donald Trump, que inició una guerra comercial con China en 2018. *Esta confrontación se ha extendido a áreas como la tecnología y la seguridad nacional, con implicaciones globales. La administración de Joe Biden ha mantenido en gran medida la postura dura hacia China, destacando la necesidad de proteger la libertad y los derechos humanos frente a lo que perciben como una amenaza autoritaria* (BBC, 2023).

China, por su parte, ha acusado a Estados Unidos de no entender sus políticas internas y externas, y de intentar frenar su desarrollo económico y tecnológico. *Beijing ha enfatizado que su objetivo no es reemplazar la hegemonía de Washington, sino modernizar su economía y contribuir a una gobernanza global más equitativa* (Observatorio de Política China, 2023).

2. La Unión Europea: Un Actor Cauteloso

La Unión Europea se encuentra en una posición delicada en esta nueva Guerra Fría. Aunque reconoce la creciente rivalidad entre Estados Unidos y China, ha intentado mantener una postura más neutral y evitar una confrontación directa. *La UE ha buscado equilibrar sus relaciones económicas con China, que es un importante socio comercial, y su alianza estratégica con Estados Unidos* (La vanguardia, 2023). Sin embargo, la presión para alinearse más estrechamente con Washington ha aumentado, especialmente en áreas como la tecnología y la seguridad.

3. Otros Países Asiáticos: Navegando en Aguas Turbulentas

Otros países asiáticos, como Japón, Corea del Sur y la India, también están profundamente afectados por la competencia entre Estados Unidos y China. *Japón y Corea del Sur, aliados tradicionales de Estados Unidos, han reforzado sus lazos de seguridad con Washington en respuesta a la creciente influencia de China en la región* (El mundo, 2019) La India, por su parte, ha buscado fortalecer su posición geopolítica mediante alianzas estratégicas tanto con Estados Unidos como con otros países asiáticos, en un intento de contrarrestar la influencia china.

4 La Dimensión Tecnológica de la Nueva Guerra Fría

Una de las características distintivas de esta nueva Guerra Fría es la competencia tecnológica. La inteligencia artificial, la ciberseguridad y las redes 5G son áreas clave donde Estados Unidos y China están compitiendo por la supremacía. Estados Unidos ha tomado medidas para limitar el acceso de China a tecnologías avanzadas, como los semiconductores, y ha presionado a sus aliados para que excluyan a empresas chinas como *Huawei* de sus redes 5G. China, por su parte, ha acelerado sus esfuerzos para desarrollar tecnologías autóctonas y reducir su dependencia de las importaciones tecnológicas.

LA ÉTICA EN LA IA DE LAS GRANDES EMPRESAS

La ética de la IA en las grandes empresas es un campo en constante evolución, marcado por la tensión entre la innovación tecnológica y la responsabilidad ética. Aunque muchas empresas han adoptado principios éticos y han establecido comités de ética, la implementación práctica de estos principios a menudo revela desafíos significativos. La transparencia, la equidad y la responsabilidad son pilares esenciales para asegurar que la IA se utilice de manera ética y beneficiosa para toda la sociedad.

La ética de la inteligencia artificial (IA) en las grandes empresas también es un tema complejo y multifacético. A continuación, se exploran algunos aspectos clave de la ética de la IA en grandes empresas como *Google, Amazon, Microsoft, Facebook* y *Apple,* con ejemplos recientes.

1. Google: Innovación y Controversias Éticas

Google ha sido un líder en el desarrollo de tecnologías de IA, pero también ha enfrentado críticas significativas por sus prácticas éticas. En 2023, *Google* anunció su compromiso con principios éticos para la IA, incluyendo la transparencia y la equidad (Lozano, 2023). Sin embargo, la empresa ha sido criticada por su participación en proyectos militares, como el *Proyecto Maven*, y por la falta de transparencia en sus algoritmos de búsqueda y publicidad (UNESCO, 2024). Estos conflictos reflejan la tensión entre la innovación tecnológica y la responsabilidad ética.

2. Amazon*:* Eficiencia y Derechos Laborales

Amazon utiliza la IA para optimizar su cadena de suministro y mejorar la experiencia del cliente. Sin embargo, ha sido criticado por el uso de IA en la vigilancia de los empleados y por prácticas laborales cuestionables (Derechos Digitales, 2023). En 2024, *Amazon* implementó nuevas políticas para mejorar la transparencia y la equidad en el uso de la IA, pero las preocupaciones sobre la privacidad y los derechos laborales persisten (Comisión Europea, 2024).

3. Microsoft*:* Compromiso con la Responsabilidad

Microsoft ha adoptado un enfoque proactivo en la ética de la IA, estableciendo un comité de ética y lanzando iniciativas para promover el uso responsable de la IA (Foro Económico Mundial, 2024). En 2023, *Microsoft* firmó un acuerdo con la *UNESCO* para garantizar la protección de los derechos humanos en el desarrollo y uso de la IA (UNESCO, 2024). Sin embargo, la implementación de estos principios en todas sus operaciones globales sigue siendo un desafío.

4. Facebook*:* Privacidad y Manipulación de Datos

Facebook utiliza la IA para personalizar el contenido y los anuncios, pero ha sido criticado por su papel en la difusión de desinformación y por la manipulación de datos personales (Asociación de Derechos Humanos de Asia, 2023). En 2024, *Facebook* implementó nuevas políticas de transparencia y responsabilidad, pero las preocupaciones sobre la privacidad y la integridad de la información continúan siendo un problema significativo (UNESCO, 2024).

5. Apple*:* Privacidad y Transparencia

Apple ha destacado por su enfoque en la privacidad y la seguridad en el uso de la IA. En 2023, Apple lanzó nuevas características de privacidad en sus dispositivos, asegurando que los datos de los usuarios se manejen de manera ética (OMS, 2023). Sin embargo, la empresa también ha enfrentado críticas por su dependencia de la cadena de suministro en países con prácticas laborales cuestionables (UNESCO, 2024). La ética de *Apple* refleja un compromiso con la privacidad del usuario, aunque a veces a costa de la transparencia en otras áreas.

CONCLUSIONES

Existe un importante debate en torno a la ética de la IA en dos dimensiones como hemos propuesto, en la primera dimensión se revisan aspectos más cotidianos, generales que son los que se buscan regular

por los diferentes países y tal es el caso de la Unión Europea, en temas como transparencia, responsabilidad, equidad, participación ciudadana entre otros; en estas dimensiones o temas la discusión ética podría decirse está aún en un proceso de avance que es lento y desfasado con respecto a la velocidad con la que avanza la IA y así por ejemplo en América Latina apenas se comienza a legislar al respecto, se dice que será en 2024 cuando aparezcan las primeras legislaciones y normativas al respecto, no obstante insistiremos en que estas dimensiones son importantes, pero no las más relevantes para el desarrollo y sobre todo ejercicio de la IA.

La dimensión más importante por el peligro implica está en el ámbito geopolítico a nivel mundial, ya que como se pudo observar, la carrera por el desarrollo tecnológico ha entrado en una nueva etapa de *"guerra fría"* entre Estados Unidos y China principalmente, que se asume dentro de una Ética de la Supremacía de la IA o una Ética del Colonialismo Digital o un Capitalismo del Multiverso, en esta ética que solo es una prolongación de la Ética de Mercado en la que actualmente se vive.

La ética del capitalismo liberal o neoliberal contemporáneo, que a menudo prioriza el dinero y las ganancias sobre el bienestar humano, puede denominarse ***"ética del mercado"*** o ***"ética del beneficio".*** Este enfoque ético se caracteriza por varios principios y prácticas que reflejan la primacía del capital y la eficiencia económica sobre otros valores sociales y humanos.

En el capitalismo neoliberal, el objetivo principal de las empresas y las políticas económicas es maximizar los beneficios. Esto a menudo se traduce en decisiones que favorecen la reducción de costos y el aumento de la productividad, incluso a expensas de los derechos laborales y el bienestar social se reduce el papel del Estado en la economía. Se promueve la desregulación y la privatización de servicios públicos, bajo la premisa de que el sector privado es más eficiente y capaz de generar crecimiento económico y esto ha llevado a una disminución de la protección social y a un aumento de las desigualdades. El capitalismo neoliberal también enfatiza el individualismo y la competencia. Se espera que los individuos sean responsables de su propio éxito

económico, lo que puede fomentar una cultura de competencia feroz y desincentivar la cooperación y la solidaridad.

Una de las principales críticas a la ética del mercado es que perpetúa y exacerba las desigualdades económicas y sociales. La concentración de la riqueza en manos de unos pocos y la exclusión de amplios sectores de la población de los beneficios del crecimiento económico son problemas recurrentes. El enfoque en el crecimiento económico a corto plazo también puede tener consecuencias negativas para el medio ambiente. La explotación de recursos naturales y la falta de regulaciones ambientales adecuadas pueden llevar a la degradación del medio ambiente y al cambio climático. La ética del mercado puede llevar a la deshumanización del trabajo, donde los trabajadores son vistos principalmente como recursos económicos en lugar de seres humanos con derechos y necesidades. Esto puede resultar en condiciones laborales precarias y en la falta de seguridad y bienestar para los trabajadores.

La Ética de la IA actualmente es esta misma ética, hasta que la misma IA la pueda cambiar a favor del ser humano.

REFERENCIAS BIBLIOGRÁFICAS

Agencia Española de Protección de Datos. (2023). *Inteligencia artificial: Transparencia. https://www.aepd.es/prensa-y-comunicacion/blog/inteligencia-artificial-transparencia*

Al Poniente. (2024). *EE.UU. Vs China en la Geopolítica de la Inteligencia Artificial.* https://alponiente.com/ee-uu-vs-china-en-la-geopolitica-de-la-inteligencia-artificial/

Asociación de Derechos Humanos de Asia. (2023). *Campañas de concienciación sobre la IA*. https://www.adhasia.org/ia-campañas

Banco Mundial. (2024). *IA en el sector financiero: Responsabilidad y transparencia.* https://www.worldbank.org/en/news/feature/2024/01/15/ai-in-financial-sector-responsibility-and-transparency

BBC News Mundo. (2023). *"Estamos condenados a esta nueva Guerra Fría que será larga y muy tensa por la interdependencia que hay entre China, Rusia, EE.UU. y Europa". https://www.bbc.com/mundo/noticias-internacional-65614576*

Buolamwini, J., y Gebru, T. (2023). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, 81, 1-15.

Comisión de Derechos Humanos de San Francisco. (2023). *Moratoria sobre el uso de reconocimiento facial.* https://sfhumanrights.org/facial-recognition-moratorium

Comisión Europea. (2023). *Reglamento de Inteligencia Artificial.* https://ec.europa.eu/digital-strategy/our-policies/artificial-intelligence

Comisión Europea. (2024). *Consulta pública sobre el Reglamento de Inteligencia Artificial.* https://ec.europa.eu/digital-strategy/our-policies/artificial-intelligence

Derechos Digitales. (2023). *Transparencia y rendición de cuentas en la IA*. https://www.derechosdigitales.org/ia-transparencia

El Mundo. (2019). *Guerra Fría entre Estados Unidos y China. https://www.elmundo.es/internacional/2019/08/03/5d45cbacfdddffc0988b45e0.html*

El País. (2023). *Una nueva guerra fría entre EE UU y China se propaga por el tablero global. https://elpais.com/internacional/2023-02-12/una-nueva-guerra-fria-entre-ee-uu-y-china-se-propaga-por-el-tablero-global.html*

Florkin, J. (2024). *Ética de la IA: navegando por el futuro de la inteligencia artificial*. https://julienflorkin.com/es/tecnolog%C3%ADa/inteligencia-artificial/%C3%A9tica-de-ia/

Foro Económico Mundial. (2024). *Gobernanza de la IA: Cómo la regulación, la colaboración y la demanda de habilidades están configurando las tendencias. https://es.weforum.org/agenda/2024/09/gobernanza-de-la-ia-como-la-regulacion-la-colaboracion-y-la-demanda-de-talento-estan-dando-forma-a-la-industria/*

Inteligencia Artificial 360. (2024). *Transparencia y Regulación en IA. https://inteligenciaartificial360.com/actualidad/transparencia-y-regulacion-en-ia/*

La Vanguardia. (2023). *Los europeos no quieren implicarse en la nueva guerra fría entre Estados Unidos y China. https://www.lavanguardia.com/internacional/20210922/7736233/europeos-quieren-implicarse-nueva-guerra-fria-estados-unidos-china.html*

Lasalle, J. M. (2024). *Civilización Artificial.* Arpa. Madrid

Lozano, J. (2023). Principios éticos de Google para la IA. *Revista de Tecnología y Sociedad*, 15(3), 45-58.

Observatorio de Política China. (2023). *EEUU-China: ¿retorno a la guerra fría o avance hacia la multipolaridad? https://politica-china.org/areas/politica-exterior/eeuu-china-retorno-a-la-guerra-fria-o-avance-hacia-la-multipolaridad*

Organización Mundial de la Salud. (2023). *Ética en la IA para la salud*. https://www.who.int/es/news-room/fact-sheets/detail/ethics-in-ai-for-health

Ortega Klein, A. (2020). *Geopolítica de la ética en Inteligencia Artificial*. Real Instituto Elcano. España.

Sadín, E. (2023). *La vida espectral*. Caja Negra, Madrid.

UNESCO. (2021). *"Recomendación sobre la ética de la inteligencia artificial"*. https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

UNESCO. (2023). *Ética de la inteligencia artificial*. https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

UNESCO. (2024). *Foro Global sobre la Ética de la IA*. https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

Universidad Nacional Autónoma de México. (2024). *Transparencia e Inteligencia Artificial: una nueva perspectiva. https://revistas.unam.mx/index.php/encrucijada/article/download/77792/68835/0*

CAPÍTULO II

Humanidad y Máquinas: ¿Cómo Definir el Bien y el Mal en un Mundo Inteligente?

CINTHIA NOEMI ZACATENCO ARELLANO

RESUMEN

El avance de la inteligencia artificial (IA) está transformando la vida humana, desde la automatización industrial hasta la interacción diaria con asistentes virtuales. Esto plantea importantes dilemas éticos: ¿cómo definir el bien y el mal cuando las máquinas toman decisiones que impactan a las personas? El artículo analiza cómo adaptar los marcos éticos tradicionales a esta realidad. La IA, sin empatía ni experiencia, desafía modelos como el utilitarismo, que busca maximizar el bienestar general, y la deontología, centrada en reglas morales estrictas, que pueden resultar en decisiones justas técnicamente, pero inhumanas en la práctica. La responsabilidad ética también se complica al difuminarse entre desarrolladores y usuarios, resaltando la necesidad de normas claras y auditorías para garantizar un comportamiento ético. Además, la ética de la virtud, que valora el carácter moral, enfrenta dificultades al intentar aplicarse a máquinas que solo pueden simular virtudes sin entenderlas realmente. La IA también amplifica sesgos algorítmicos y la opacidad de sus decisiones dificulta la responsabilidad y la corrección de errores. Finalmente, la automatización pone en riesgo la autonomía humana, al delegar en máquinas decisiones clave. Se requiere un marco ético robusto y flexible que guíe el desarrollo de la IA, asegurando que su eficiencia tecnológica no sacrifique principios morales fundamentales.

Palabras clave: ética, tecnología, responsabilidad

INTRODUCCIÓN

El desarrollo y la implementación de la inteligencia artificial (IA) han provocado una transformación radical en numerosos aspectos de la sociedad, desde la industria y el empleo hasta la vida cotidiana y las interacciones sociales. Sin embargo, junto con estos avances tecnológicos, ha surgido un conjunto de dilemas éticos que requieren ser abordados de manera urgente y exhaustiva. La filosofía ética, que durante siglos ha proporcionado marcos para guiar el comportamiento moral de los seres humanos, se enfrenta ahora a nuevos desafíos que no pudieron haber sido anticipados por los pensadores clásicos como Aristóteles, Immanuel Kant y Jeremy Bentham. En el pasado, los debates éticos se centraban exclusivamente en la conducta humana, pero el surgimiento de la IA ha abierto un nuevo campo de reflexión, en el que las máquinas y los algoritmos autónomos juegan un papel cada vez más significativo.

Las preguntas sobre la responsabilidad, la justicia y la autonomía cobran una nueva dimensión en un mundo donde las decisiones automatizadas afectan directamente a las personas. Mientras que los principios de la ética humana están basados en la experiencia, la empatía y la intuición, las máquinas carecen de estas cualidades, lo que genera tensiones y desafíos. ¿Cómo garantizar que las decisiones tomadas por la IA sean justas y éticamente correctas? ¿Es posible o deseable dotar a las máquinas de un marco moral, como lo han hecho los seres humanos con sus sistemas éticos? Además, el impacto de la IA no se limita únicamente al ámbito ético: afecta también a la privacidad, el empleo y la estructura misma de las sociedades. En este contexto, es esencial replantearse cómo definir el bien y el mal en un mundo donde la inteligencia artificial está cada vez más presente, y desarrollar marcos éticos que permitan un uso responsable y beneficioso de estas tecnologías.

ANTECEDENTES

La Ética en la Filosofía y su Relación con la Tecnología

La ética, entendida como la disciplina filosófica que estudia la moral y la conducta humana, ha sido un tema central desde la antigüedad. Filósofos como Aristóteles, Kant y Bentham han desarrollado teorías éticas que intentan explicar cómo los seres humanos deberían comportarse para vivir en una sociedad justa y equilibrada (Beauchamp & Childress, 2019). Sin embargo, con el surgimiento de la tecnología, especialmente la inteligencia artificial (IA), se ha generado la necesidad de revisar y adaptar estos marcos éticos para incluir a las máquinas como agentes potencialmente morales.

La IA plantea desafíos únicos para la ética tradicional. Mientras que las normas éticas humanas están basadas en la experiencia, la empatía y la intuición, las máquinas carecen de estas cualidades. A pesar de ello, se espera que tomen decisiones que afecten a los seres humanos de manera justa y ética. Este desajuste entre la capacidad de toma de decisiones de las máquinas y las expectativas humanas crea un terreno fértil para la reflexión ética (Moor, 2006).

Ética Filosófica y Desarrollo Tecnológico

Desde una perspectiva histórica, filósofos como Aristóteles, Kant y Bentham establecieron marcos para reflexionar sobre el comportamiento moral, pero el avance de la tecnología ha introducido nuevos dilemas que no podían haber sido anticipados en esos tiempos. Por ejemplo, el utilitarismo de Bentham podría verse confrontado ante tecnologías como la inteligencia artificial, que promete maximizar la utilidad general pero también plantea la cuestión de si los algoritmos pueden realmente tomar decisiones morales. Kant, por su parte, defendió el imperativo categórico, una regla de oro ética que cuestionaría si los sistemas autónomos, como los robots, pueden tratar a los humanos siempre como fines en sí mismos y no como medios.

En este sentido, el surgimiento de la tecnología digital y la inteligencia artificial ha generado nuevas formas de interacción y dependencia entre

los humanos y las máquinas, creando preguntas éticas sobre responsabilidad, derechos y dignidad humana. La tecnología no es neutral; incorpora valores humanos, tanto intencionados como no intencionados, y amplifica cuestiones éticas sobre control, autonomía y poder.

Ejemplo:

El "dilema del tranvía", ampliamente discutido en la ética filosófica, plantea un conflicto clásico entre el utilitarismo y la deontología, aplicado ahora a la tecnología moderna como los vehículos autónomos. En este dilema, un automóvil autónomo se enfrenta a una decisión crítica: en una situación de emergencia, debe elegir entre atropellar a un peatón o desviar su curso y poner en peligro la vida de sus ocupantes. Esto refleja el choque entre las teorías de Kant y Bentham. Desde la perspectiva de Kant, que defendía el imperativo categórico, cada vida humana tiene un valor absoluto y no debe ser sacrificada, independientemente de las circunstancias. En cambio, Bentham y su utilitarismo sostendrían que la acción correcta es la que maximiza el bienestar para la mayoría, es decir, salvar la mayor cantidad de vidas posibles (Bentham, 1789; Kant, 1785). Este tipo de decisiones obligan a cuestionar si las máquinas, que no tienen una experiencia vivencial ni emocional, están realmente capacitadas para tomar decisiones éticas.

La Ética y la Tecnología en el Siglo XX y XXI

Durante el siglo XX, con el auge de las tecnologías nucleares, de la información y biotecnológicas, se desarrollaron nuevas ramas de la ética aplicada, como la **bioética** y la **tecnoética**, que comenzaron a abordar específicamente las implicaciones éticas del desarrollo científico y tecnológico. Filósofos como Hans Jonas con su obra *El Principio de Responsabilidad* (1979) advirtieron sobre la necesidad de asumir una nueva ética frente al poder creciente de la tecnología, subrayando la importancia de la responsabilidad hacia las futuras generaciones y el medio ambiente.

En el siglo XXI, la acelerada evolución de la inteligencia artificial ha dado lugar a debates más urgentes sobre los límites de la automatización y la autonomía de las máquinas. El trabajo de filósofos contem-

poráneos, como Nick Bostrom y su preocupación por la superinteligencia, plantea preguntas sobre el control de tecnologías que podrían superar las capacidades humanas en toma de decisiones. Además, el concepto de "moralidad algorítmica" introduce la cuestión de si es posible o deseable programar códigos morales en sistemas autónomos que interactúan con humanos, desafiando la distinción entre agentes morales humanos y máquinas.

Ejemplo:

En el contexto de los avances en inteligencia artificial, la obra de Nick Bostrom sobre la "superinteligencia" plantea uno de los desafíos éticos más significativos de nuestro tiempo. Bostrom advierte que una IA avanzada sin controles éticos podría priorizar objetivos que no consideren la supervivencia humana. Esto presenta un dilema inédito: ¿cómo asegurarnos de que los sistemas de IA superinteligentes no desarrollen metas incompatibles con los intereses humanos? Este escenario subraya la necesidad urgente de una nueva ética que acompañe a las tecnologías emergentes, donde las decisiones no solo sean eficientes, sino también humanas. Bostrom sugiere que la superinteligencia podría tener consecuencias catastróficas si no se le imponen restricciones morales adecuadas (Bostrom, 2014). El impacto potencial de estas tecnologías nos recuerda que debemos tener un enfoque ético desde las primeras fases de su desarrollo.

Tecnología, Autonomía y Responsabilidad Moral

La cuestión de la autonomía tecnológica y la responsabilidad ética es un tema crucial. Mientras que tradicionalmente se ha considerado que solo los humanos son responsables de sus acciones, las tecnologías actuales, como los vehículos autónomos y los algoritmos de decisión en sectores como el financiero o el médico, abren debates sobre si las máquinas pueden o deben asumir algún tipo de responsabilidad. Esto lleva a replantear los fundamentos de la ética: ¿son los humanos los únicos seres capaces de tomar decisiones morales? Y si las máquinas se vuelven más autónomas, ¿cómo deben ser reguladas para garantizar un comportamiento ético?

El Desarrollo de la IA y su Impacto en la Sociedad

La inteligencia artificial (IA) ha pasado de ser una mera fantasía científica a convertirse en una realidad omnipresente. Desde los algoritmos que dirigen las búsquedas en internet hasta los sistemas de reconocimiento facial y los vehículos autónomos, la IA está influyendo en las decisiones diarias de millones de personas. Sin embargo, su creciente poder también ha generado preocupación sobre su impacto en la sociedad (Brynjolfsson & McAfee, 2014). A medida que las tecnologías basadas en IA se expanden, los debates en torno a sus implicaciones sociales y éticas han cobrado mayor relevancia.

El debate sobre la ética de la IA ha ganado tracción a medida que las aplicaciones de la IA se vuelven más complejas y autónomas. ¿Deberían las máquinas tener la capacidad de decidir sobre la vida y la muerte en el contexto de la guerra o la medicina? ¿Cómo se asegura que las decisiones tomadas por la IA sean justas y no reproduzcan sesgos humanos? Estas preguntas subrayan la importancia de establecer un marco ético sólido para guiar el desarrollo y la implementación de la IA (Floridi & Taddeo, 2016). Los sistemas de IA, aunque poderosos, no están exentos de errores o sesgos, ya que tienden a replicar los prejuicios presentes en los datos con los que fueron entrenados. Por lo tanto, se requiere una supervisión constante para garantizar que estas tecnologías promuevan la equidad y no refuercen desigualdades preexistentes.

IA, Desigualdades y Desempleo

Uno de los impactos más discutidos de la IA es su efecto en el empleo. Los avances en automatización y robotización han llevado a la preocupación de que millones de trabajos tradicionales puedan ser reemplazados por máquinas inteligentes. Esto plantea desafíos importantes para la redistribución económica y la educación, ya que los sectores laborales deberán adaptarse a un entorno de trabajo que cambia rápidamente. Si bien algunos argumentan que la IA creará nuevos empleos, también es claro que habrá una transformación en las habilidades necesarias para mantenerse competitivo en este nuevo panorama tecnológico (Brynjolfsson & McAfee, 2014).

Ejemplo:

El impacto de la IA en el empleo ya se está manifestando en sectores clave, como la logística. Empresas como Amazon han implementado algoritmos para gestionar sus almacenes, lo que ha resultado en la creación de empleos precarios de baja remuneración, mientras que los puestos de alto valor, relacionados con el diseño y desarrollo de estos sistemas, son reservados para profesionales altamente cualificados. Esta dinámica agrava la desigualdad en sectores donde la automatización reduce la demanda de trabajo no calificado, mientras que se dispara la necesidad de expertos en tecnología (Brynjolfsson & McAfee, 2014). Esta situación plantea una preocupación ética sobre cómo los beneficios de la automatización se distribuyen de manera desigual, perpetuando e incluso ampliando las brechas socioeconómicas. La IA tiene el potencial de transformar la economía, pero sin políticas de distribución justa, podría concentrar la riqueza aún más en las manos de unos pocos.

IA y Privacidad

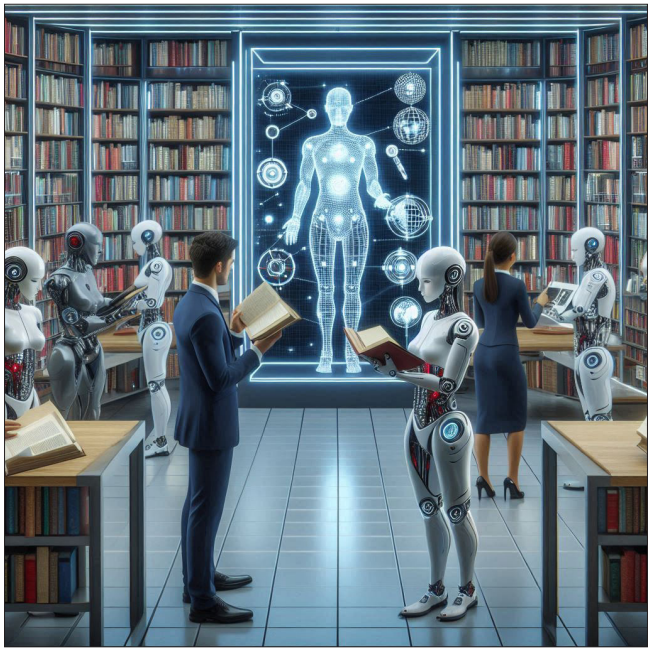
Otra preocupación fundamental en el desarrollo de la IA es el impacto en la privacidad. Los sistemas de IA pueden procesar cantidades masivas de datos personales, lo que genera preguntas éticas sobre la recolección, almacenamiento y uso de esta información. En un mundo donde los algoritmos pueden predecir comportamientos e influir en decisiones políticas y comerciales, proteger la privacidad individual se ha convertido en un tema crítico (Floridi & Taddeo, 2016). Además, los mecanismos de vigilancia y reconocimiento facial, impulsados por la IA, pueden contribuir a la erosión de libertades civiles si no se regulan adecuadamente.

*Ejemplo***:**

El caso de Cambridge Analytica es un recordatorio perturbador de cómo la inteligencia artificial puede ser utilizada para violar la privacidad de millones de personas. Esta empresa explotó datos personales de usuarios de Facebook sin su consentimiento para manipular proce-

sos electorales, utilizando la IA para segmentar y personalizar mensajes políticos. Este caso pone de relieve no solo el riesgo de abuso de poder por parte de empresas, sino también la falta de marcos regulatorios adecuados que controlen cómo se utilizan los datos personales. En un mundo donde la recolección masiva de datos se ha convertido en una práctica estándar, el equilibrio entre la innovación tecnológica y la protección de los derechos individuales sigue siendo un desafío ético fundamental (Cadwalladr & Graham-Harrison, 2018). El caso Cambridge Analytica muestra cómo la falta de transparencia y regulación en la gestión de datos puede erosionar la confianza en la democracia y los derechos humanos.

Ilustración 1. Antecedentes de la IA y la moral.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

¿CÓMO DEFINIR EL BIEN Y EL MAL EN UN MUNDO INTELIGENTE?

Perspectivas Filosóficas sobre la Ética de la IA. Utilitarismo en la IA

El utilitarismo, una teoría ética que propone que la moralidad de una acción se determina por su capacidad para maximizar la felicidad o el bienestar, se enfrenta a retos únicos en el contexto de la IA. Las decisiones automatizadas, como las que se toman en la conducción autónoma, pueden estar basadas en principios utilitaristas. Sin embargo, ¿es moralmente aceptable que una máquina decida sacrificar a una persona para salvar a varias? La falta de empatía y la incapacidad de la IA para comprender el sufrimiento humano plantea serias dudas sobre la aplicabilidad del utilitarismo en este contexto Bostrom, 2014).

Desde una perspectiva utilitarista, que sostiene que una acción es moralmente correcta si maximiza el bienestar general (Bentham, 1789), una máquina que decide sacrificar a una persona para salvar a varias podría ser vista como aceptable. El utilitarismo evalúa las acciones basándose en sus consecuencias, por lo que en teoría, sacrificar a una persona para salvar a un mayor número de personas maximizaría el bienestar total y sería moralmente justificable.

El utilitarismo puede parecer una solución atractiva en la teoría, pero su aplicación práctica en la IA presenta problemas significativos. La dificultad de cuantificar y comparar el bienestar de diferentes individuos, así como la falta de capacidad de la IA para experimentar emociones humanas, complican la implementación efectiva de esta teoría en sistemas automatizados (Lin, 2016).

Ejemplo:

El utilitarismo, una teoría ética que propone maximizar el bienestar y minimizar el sufrimiento, ha encontrado aplicaciones en el desarrollo de la inteligencia artificial, especialmente en áreas como la conducción autónoma y la toma de decisiones automatizadas. Un ejemplo relevante de la aplicación del utilitarismo en la IA es el “dilema del tran-

vía", en el que un vehículo autónomo debe decidir entre dos opciones en una situación de emergencia: atropellar a un peatón o desviar su curso y arriesgar la vida de los ocupantes. Desde una perspectiva utilitarista, el vehículo debería tomar la decisión que minimice el daño total, es decir, la opción que cause menos pérdida de vidas humanas o sufrimiento. Este enfoque refleja el pensamiento utilitarista de Jeremy Bentham, quien abogaba por maximizar la felicidad para el mayor número de personas (Bentham, 1789). Sin embargo, el uso del utilitarismo en la IA plantea dilemas éticos complejos, ya que las máquinas pueden verse obligadas a tomar decisiones que, aunque racionales, resulten moralmente cuestionables para los seres humanos afectados.

Deontología y la Responsabilidad en la IA

La deontología, defendida por Immanuel Kant, argumenta que las acciones deben ser juzgadas por su conformidad con reglas morales, independientemente de sus consecuencias (Kant, 1785). En teoría, los sistemas de IA podrían ser programados para seguir estrictamente reglas éticas. Sin embargo, la complejidad de las decisiones humanas hace que este enfoque sea problemático. La rigidez de las reglas deontológicas puede llevar a decisiones que, aunque éticamente correctas según la teoría, sean inhumanas o injustas en la práctica (Floridi & Taddeo, 2016).

Por ejemplo, si un sistema de IA está programado para seguir una regla deontológica estricta que prohíbe causar daño, podría tomar decisiones que resulten en consecuencias no deseadas en contextos complejos. Esta falta de flexibilidad puede limitar la capacidad de la IA para adaptarse a situaciones que requieren una consideración más matizada de las consecuencias (Moor, 2006).

Además, la responsabilidad ética en las decisiones de IA plantea interrogantes sobre quién debe rendir cuentas cuando una máquina toma decisiones problemáticas. La responsabilidad a menudo se difunde entre desarrolladores, operadores y usuarios, lo que complica la atribución de culpa (Calo, 2016).

Otro desafío significativo es el normativo y legal. Las leyes y regulaciones deben evolucionar para abordar cómo aplicar principios morales universales en el contexto de tecnologías emergentes, equi-

librando el respeto por los derechos fundamentales con la necesidad de adaptarse a nuevas realidades tecnológicas (Goggin, 2019). La implementación de principios deontológicos en la IA también requiere el desarrollo de normas éticas claras y la incorporación de mecanismos de responsabilidad, como la documentación de decisiones y la auditoría de algoritmos (Bryson, 2018). Además, es crucial la formación de los desarrolladores de IA en ética deontológica para asegurar que las máquinas se programen de manera que respeten estos principios morales universales (Wright, 2019).

Ética de la Virtud y el Desarrollo de una IA Moral

La ética de la virtud, que se centra en el carácter moral y las virtudes del individuo, plantea un desafío único para la IA, que carece de personalidad o carácter (Aristóteles, 350 a.C.). Algunos investigadores sugieren que la IA podría ser diseñada para emular virtudes humanas como la compasión, la justicia y la sabiduría. No obstante, ¿puede una máquina realmente "poseer" virtudes o simplemente imitarlas? La simulación de virtudes no equivale a la posesión de las mismas, lo que plantea preguntas sobre la autenticidad de la moralidad en la IA (Moor, 2006; McConnell, 2019).

La cuestión de si una máquina puede realmente "poseer" virtudes o simplemente imitarlas se basa en la diferencia entre simulación y posesión genuina de cualidades morales. Las máquinas pueden ser programadas para simular comportamientos virtuosos, como la compasión o la justicia, mediante reglas y análisis de datos, pero esta simulación no equivale a una verdadera comprensión o internalización de los valores (Borenstein, Herkert, & Herkert, 2017). La ética de la virtud, defendida por Aristóteles, sostiene que las virtudes son parte del carácter moral desarrollado a través de la experiencia y la reflexión personal (Aristóteles, 350 a.C.). Dado que las máquinas carecen de conciencia y capacidad para el desarrollo moral personal, es cuestionable que puedan poseer virtudes en un sentido genuino. En la práctica, esto implica que las decisiones de IA basadas en principios de virtud deben ser evaluadas con la conciencia de que las máquinas no tienen una comprensión auténtica de la moralidad (Bryson, 2018; Moor, 2006; Wright, 2019).

Además, la implementación de la ética de la virtud en la IA requiere una redefinición de cómo se evalúan las decisiones y comportamientos de las máquinas. Mientras que los humanos pueden ser evaluados por su carácter moral general, las máquinas operan bajo directrices programadas y algoritmos que carecen de la capacidad para el desarrollo moral continuo (Floridi & Taddeo, 2016). Esto hace que sea difícil determinar cómo programar una IA para que actúe de manera virtuosa en situaciones complejas y dinámicas.

La ética de la virtud también implica un enfoque en la formación continua y el crecimiento moral, algo que es difícil de traducir a un sistema de IA que opera bajo reglas fijas y sin la capacidad para la introspección o el desarrollo personal (Moor, 2006). Para abordar estos problemas, es esencial desarrollar modelos de IA que no solo emulen comportamientos virtuosos, sino que también incluyan mecanismos para adaptar y ajustar estos comportamientos en función de la retroalimentación y el contexto (Bryson, 2018).

El reto radica en si una IA puede desarrollar un "carácter moral" o simplemente ejecutar programas que simulan comportamientos virtuosos. La falta de una verdadera comprensión de las virtudes por parte de la IA sugiere que los sistemas actuales pueden limitarse a replicar patrones de comportamiento sin una comprensión auténtica de los valores morales (McConnell, 2019).

Ejemplo:

La ética de la virtud, que se centra en el desarrollo del carácter moral más que en la simple aplicación de reglas, ofrece un marco interesante para el diseño de sistemas de IA moral. Un ejemplo destacado es el enfoque de IA moral propuesto por el filósofo Mark Coeckelbergh, quien sugiere que las máquinas deberían ser programadas para desarrollar "buen carácter" y tomar decisiones basadas en la empatía, la justicia y otras virtudes humanas clave. La idea es que, en lugar de depender únicamente de reglas predeterminadas o cálculos utilitaristas, una IA moral debería aprender a comportarse de manera virtuosa a través de la interacción continua con los seres humanos y su entorno social. Este enfoque permitiría a la IA adaptarse a situaciones complejas y

cambiantes, emulando cómo los humanos refinan su carácter a lo largo del tiempo (Coeckelbergh, 2020). Sin embargo, el desafío reside en cómo traducir conceptos abstractos como la "virtud" en algoritmos funcionales, lo que abre un nuevo campo de investigación en la ética de la IA.

Ilustración 2. Definición del bien y el mal con la IA en un mundo tecnológico con una IA holográfica azul abogada.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

LA INFLUENCIA DE LA IA EN LAS DECISIONES ÉTICAS HUMANAS

Sesgos Algorítmicos y su Impacto en la Justicia

Los algoritmos de inteligencia artificial (IA) tienen la capacidad de reproducir y amplificar los sesgos presentes en los datos con los que

son entrenados, lo que puede llevar a decisiones injustas o discriminatorias. Este fenómeno se observa en diversas aplicaciones de IA, como los sistemas de reconocimiento facial, que a menudo presentan tasas de error significativamente más altas para personas de color en comparación con individuos de piel más clara (Buolamwini & Gebru, 2018). Por ejemplo, un estudio reveló que un sistema de reconocimiento facial comercial tenía una tasa de error del 34.7% al identificar a mujeres de piel oscura, en comparación con una tasa de error del 0.8% para hombres de piel clara (Buolamwini & Gebru, 2018). De manera similar, los algoritmos de contratación que favorecen ciertos grupos demográficos pueden perpetuar sesgos históricos y desventajas estructurales, lo que resulta en prácticas discriminatorias en la selección de personal (Dastin, 2018). Para abordar estos problemas, es crucial desarrollar métodos efectivos para identificar y mitigar los sesgos en los sistemas de IA, lo cual incluye la implementación de prácticas de auditoría y técnicas de ajuste de datos para asegurar que los algoritmos operen de manera equitativa (Mittelstadt et al., 2016).

Ejemplo:

El sesgo algorítmico es uno de los problemas más inquietantes en el desarrollo de la inteligencia artificial. Un caso emblemático es el algoritmo de contratación de Amazon, que fue retirado después de que se descubriera que favorecía a los candidatos masculinos sobre las mujeres. Este sesgo se debió a los datos históricos con los que fue entrenado el sistema, que reflejaban una inclinación de género en los procesos de selección anteriores (Dastin, 2018). Este ejemplo evidencia que los algoritmos no son inherentemente justos; simplemente reproducen los patrones presentes en los datos con los que fueron entrenados. Sin un diseño consciente y ético, los sistemas de IA corren el riesgo de perpetuar desigualdades sociales y económicas, subrayando la necesidad de aplicar un enfoque más riguroso en la eliminación de sesgos en los datos.

Transparencia y Responsabilidad en las Decisiones de la IA

La opacidad inherente a muchos sistemas de IA, en los que las decisiones se toman en “cajas negras” cuyos procesos internos no son

fácilmente comprensibles para los humanos, plantea serios desafíos para la responsabilidad y la transparencia. Esta falta de transparencia puede socavar la confianza pública en la IA y complicar la atribución de responsabilidad en casos de decisiones erróneas o dañinas (O'Neil, 2016). Por ejemplo, cuando un algoritmo toma decisiones críticas, como en el ámbito de la justicia penal o la evaluación crediticia, los individuos afectados pueden no tener acceso a la lógica detrás de esas decisiones, lo que dificulta la identificación y corrección de errores (Kroll et al., 2017). Para abordar estos desafíos, se han propuesto varios estándares y recomendaciones para mejorar la transparencia y la auditabilidad de los sistemas de IA, como la implementación de mecanismos de documentación y la creación de protocolos de revisión independientes (Floridi & Taddeo, 2016). A medida que la IA asume un número creciente de tareas que anteriormente eran prerrogativa de los humanos, surge una preocupación significativa sobre la pérdida potencial de autonomía y agencia de los individuos. La automatización de procesos y decisiones puede llevar a una disminución de la capacidad de las personas para influir en los resultados que afectan sus vidas, especialmente en contextos donde la toma de decisiones automatizada reemplaza el juicio humano en áreas críticas como la atención médica y la justicia (Brynjolfsson & McAfee, 2014). La clave para abordar este problema radica en encontrar un equilibrio entre la eficiencia y precisión que ofrece la IA y la necesidad de preservar la autonomía y la dignidad humanas. Esto implica diseñar sistemas de IA que no solo optimicen el rendimiento, sino que también respeten y refuercen el papel de los humanos en la toma de decisiones (Russell & Norvig, 2021). La integración de la IA en la vida cotidiana debe llevarse a cabo de manera que los seres humanos mantengan un control significativo sobre las decisiones que afectan su bienestar y derechos.

Ejemplo:

La opacidad de los algoritmos de inteligencia artificial plantea desafíos éticos cruciales, especialmente cuando se utilizan en contextos de alta relevancia social, como el sistema de justicia penal. Un ejemplo prominente es el uso del sistema COMPAS en los tribunales de Estados Unidos para predecir la reincidencia criminal. A pesar de su sofisticación,

COMPAS ha sido criticado por su falta de transparencia y por perpetuar sesgos raciales, ya que los resultados favorecen a ciertos grupos demográficos en detrimento de otros (Angwin et al., 2016). Este caso demuestra que, sin una supervisión y evaluación ética adecuadas, los sistemas de IA pueden replicar y amplificar injusticias preexistentes. Para evitar estos problemas, es esencial que las decisiones algorítmicas sean auditables y explicables, asegurando que los responsables puedan rendir cuentas de las decisiones tomadas.

La Autonomía Humana frente a la Automatización

A medida que la IA asume un número creciente de tareas que anteriormente eran prerrogativa de los humanos, surge una preocupación significativa sobre la pérdida potencial de autonomía y agencia de los individuos. La automatización de procesos y decisiones puede llevar a una disminución de la capacidad de las personas para influir en los resultados que afectan sus vidas, especialmente en contextos donde la toma de decisiones automatizada reemplaza el juicio humano en áreas críticas como la atención médica y la justicia (Brynjolfsson & McAfee, 2014). La clave para abordar este problema radica en encontrar un equilibrio entre la eficiencia y precisión que ofrece la IA y la necesidad de preservar la autonomía y la dignidad humanas. Esto implica diseñar sistemas de IA que no solo optimicen el rendimiento, sino que también respeten y refuercen el papel de los humanos en la toma de decisiones (Russell & Norvig, 2021). La integración de la IA en la vida cotidiana debe llevarse a cabo de manera que los seres humanos mantengan un control significativo sobre las decisiones que afectan su bienestar y derechos.

Ejemplo:

Un ejemplo claro de la tensión entre la autonomía humana y la automatización es el caso de los sistemas de IA utilizados en la toma de decisiones laborales. En algunas empresas, los algoritmos de IA son responsables de decisiones como la contratación, evaluación del rendimiento y despidos. Amazon, por ejemplo, ha implementado un sistema automatizado en sus almacenes que evalúa el desempeño de

los trabajadores en función de métricas como la productividad y la eficiencia. Si un trabajador no cumple con los estándares establecidos por el algoritmo, el sistema puede generar una advertencia o incluso proceder con el despido de manera automática, sin intervención humana directa (Lecher, 2019). Este tipo de automatización plantea serios desafíos éticos, ya que disminuye la capacidad de los individuos para defenderse y negociar en el lugar de trabajo, poniendo en peligro su autonomía. A medida que las empresas adoptan sistemas más automatizados, la línea entre la decisión humana y la decisión de una máquina se vuelve más difusa, lo que requiere un debate profundo sobre el papel que debe desempeñar la IA en las decisiones que afectan directamente a la vida de las personas.

Ilustración 3. IA Tecnología, Autonomía y Responsabilidad Moral.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

CONCLUSIONES

La relación entre inteligencia artificial y ética plantea preguntas profundas que trascienden los límites tradicionales del pensamiento filosófico. Si bien la humanidad ha desarrollado marcos éticos a lo largo de los siglos para regular sus propias acciones, ahora enfrenta el desafío sin precedentes de aplicar, adaptar o incluso reinventar esos principios en un mundo donde las máquinas no sólo realizan tareas operativas, sino que participan activamente en la toma de decisiones que antes eran exclusivas de los seres humanos. Este cambio en el rol de la tecnología no es simplemente una evolución técnica, sino una transformación fundamental de las estructuras de poder, responsabilidad y moralidad.

Uno de los principales dilemas que surge en este contexto es la noción de responsabilidad. Tradicionalmente, la ética se ha centrado en los seres humanos como agentes morales, individuos capaces de razonar, sentir y, en última instancia, ser responsables de sus acciones. Sin embargo, con la introducción de sistemas de IA cada vez más autónomos, la línea entre agente y herramienta se ha desdibujado. Si una máquina toma una decisión errónea que afecta negativamente a las personas, ¿quién es el responsable? ¿Es el programador que diseñó el algoritmo, la empresa que lo implementó, o la propia máquina como entidad autónoma? Este dilema ético genera un vacío en el que los marcos morales tradicionales no siempre ofrecen respuestas satisfactorias.

Además, el concepto de justicia adquiere nuevas dimensiones en el ámbito de la inteligencia artificial. Los algoritmos de IA, al basarse en datos históricos para realizar predicciones o tomar decisiones, a menudo replican o incluso amplifican los sesgos existentes en la sociedad. Esto plantea serias preguntas sobre la equidad de los sistemas automatizados. ¿Cómo garantizar que las decisiones de la IA no perpetúen las desigualdades raciales, de género o socioeconómicas que ya existen en el mundo real? El hecho de que las máquinas no sean conscientes de los impactos sociales de sus decisiones pone en riesgo la justicia y la equidad en diversos sectores, desde la contratación laboral hasta el sistema judicial y la sanidad. La opacidad de muchos algoritmos – las llamadas “cajas negras” que impiden a los seres humanos comprender

cómo se han tomado ciertas decisiones – agrava aún más este problema, dificultando la rendición de cuentas y la corrección de posibles errores o injusticias.

La autonomía también es un valor central en el debate ético sobre la IA. En un mundo donde las máquinas toman decisiones, ¿qué lugar queda para la agencia humana? Si cada vez más aspectos de nuestras vidas están gobernados por algoritmos – desde las decisiones de crédito hasta las recomendaciones de tratamiento médico – los seres humanos corren el riesgo de perder su autonomía en la toma de decisiones importantes. La paradoja aquí es que la IA, que se presenta como una herramienta para aumentar la eficiencia y mejorar la calidad de vida, podría, en algunos casos, erosionar la capacidad de las personas para controlar su propio destino. Esta situación obliga a reflexionar sobre cómo equilibrar la conveniencia y eficiencia que aporta la IA con la necesidad de preservar la libertad individual y el derecho a tomar decisiones informadas.

Otra cuestión crucial es si es posible dotar a las máquinas de un marco moral. A lo largo de la historia, los sistemas éticos humanos se han basado en la capacidad de las personas para reflexionar sobre el bien y el mal, sobre las consecuencias de sus acciones, y sobre la empatía hacia los demás. Las máquinas, por su naturaleza, carecen de estas capacidades. Aunque es posible programar ciertos comportamientos éticos en los algoritmos, esto no equivale a una comprensión genuina de los dilemas morales. El desarrollo de una IA "moral" sería, en el mejor de los casos, una imitación de las normas humanas, sin la profundidad de comprensión y reflexión que caracteriza al juicio moral humano. ¿Es esto suficiente en un mundo donde las decisiones automatizadas tienen un impacto real en las vidas humanas? ¿O necesitamos desarrollar un nuevo paradigma ético que reconozca las limitaciones de la IA, al tiempo que maximiza su potencial para mejorar la sociedad?

Finalmente, el impacto de la inteligencia artificial en la estructura misma de las sociedades requiere una consideración ética más amplia. La automatización masiva y el uso generalizado de la IA podrían tener consecuencias sociales devastadoras, como la pérdida de empleos a gran escala, la creciente desigualdad y la concentración de poder en manos de unas pocas empresas tecnológicas que controlan estas he-

rramientas. Aunque la IA promete beneficios significativos en áreas como la salud, la educación y la sostenibilidad, también presenta riesgos considerables si no se implementa de manera justa y responsable. Por lo tanto, el desarrollo de la IA no debe estar impulsado únicamente por el progreso técnico y económico, sino también por principios éticos sólidos que busquen minimizar el daño y maximizar el bienestar colectivo.

REFERENCIAS

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing*

Aristóteles. (350 a.C.). Nicomachean ethics. Retrieved from https://www.earlymoderntexts.com/assets/pdfs/aristotle2016.pdf

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.

Bentham, J. (1789). *An introduction to the principles of morals and legislation*. Clarendon Press.

Bentham, J. (1789). *An introduction to the principles of morals and legislation* (Ed. 2000). Oxford University Press.

Borenstein, J., Herkert, J. R., & Herkert, J. R. (2017). The ethics of autonomous vehicles. Science and Engineering Ethics, 24(4), 927-945. *https://doi.org/10.1007/s11948-017-9900-8*

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. W. W. Norton & Company.

Bryson, J. J. (2018). The role of humans in machine ethics. *AI & Society*, 33(3), 441-449. *https://doi.org/10.1007/s00146-018-0810-8*

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 77-91. https://doi.org/10.1145/3287560.3287596

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. *https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election*

Calo, R. (2016). The case for automated ethics. *California Law Review*, 104(4), 1043-1074. https://doi.org/10.2139/ssrn.2821104

Coeckelbergh, M. (2020). *AI ethics*. MIT Press.

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. *https://www.reuters.com/article/us-amazon-com-recruitment-insight-idUSKCN1MK08*

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083), 20160360. *https://doi.org/10.1098/rsta.2016.0360*

Goggin, G. (2019). The ethics of digital technology: Issues and challenges. *Journal of Information, Communication and Ethics in Society*, 17(2), 223-237. *https://doi.org/10.1108/JICES-02-2019-0020*

Jonas, H. (1979). *El principio de responsabilidad: Ensayo de una ética para la civilización tecnológica*. Herder.

Kant, I. (1785). *Groundwork for the metaphysics of morals.* Cambridge University Press.

Kroll, J. A., Edlin, A. S., & Hurlbut, S. (2017). Accountable algorithms. University of Pennsylvania Law Review, 165(3), 633-705. *https://doi.org/10.2139/ssrn.2843327*

Lecher, C. (2019, April 25). How Amazon automatically tracks and fires warehouse workers for 'productivity'. *The Verge*. https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations

Lin, P. (2016). Why ethics matters for autonomous vehicles. In *Autonomes Fahren* (pp. 69-85). Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-49129-2_4

McConnell, T. (2019). Virtue ethics and the design of artificial agents. Journal of Artificial Intelligence Research, 64, 847-866. *https://doi.org/10.1613/jair.1.11512*

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2). https://doi.org/10.1177/2053951716679679

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18-21. https://doi.org/10.1109/MIS.2006.80

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing Group.

Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

Wright, A. (2019). The ethics of artificial intelligence: A review of the literature. *Artificial Intelligence Review*, 52(4), 341-361. *https://doi.org/10.1007/s10462-018-9678-7*

CAPÍTULO III

HUMANIDAD Y GESTIÓN EMPRESARIAL: MIRADA FUTURISTA HACIA UN AGENTE DE MEJORA / REEMPLAZO

AMÉRICA PELAYO RAMÍREZ[1]

RESUMEN

La inteligencia artificial (IA) ha emergido como una de las tecnologías más transformadoras del siglo XXI, prometiendo revolucionar múltiples sectores, desde la medicina hasta la educación. Sin embargo, uno de los debates más acalorados en torno a la IA se centra en su impacto en el mercado laboral, particularmente en la posibilidad de que reemplace a trabajadores en diversas industrias. La automatización y el aprendizaje automático están permitiendo a las máquinas realizar tareas que antes requerían habilidades y conocimientos humanos. ¿Estamos frente a una revolución que transformará el mercado laboral para siempre? ¿Qué pasará con los empleos que desaparezcan debido a la IA? ¿Cómo se adaptarán los trabajadores a esta nueva realidad? Se debe analizar cómo la inteligencia artificial está cambiando el panorama laboral, qué empleos están en riesgo y cómo podemos prepararnos para este futuro, también las oportunidades que surgen de esta transformación y cómo la IA puede crear nuevos empleos y mejorar la productividad en diversas industrias. La pregunta es: ¿la IA será una aliada o una rival para los trabajadores? La respuesta no es simple, pero es

[1] Licenciada en Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

importante explorar las posibilidades y desafíos que se presentan en este nuevo escenario laboral.

Palabras clave: Revolución Tecnológica, Inteligencia Artificial, Transformación Empresarial, Algoritmo, Implicaciones Éticas.

PRIMER VISIÓN

La inteligencia artificial (IA) es una revolución tecnológica que está transformando la forma en que vivimos, trabajamos y nos relacionamos. A través de la creación de sistemas capaces de aprender, razonar y tomar decisiones autónomas, la IA está abriendo nuevas posibilidades en campos como la medicina, la educación, la economía y la industria.

Con la capacidad de procesar grandes cantidades de datos a velocidades inimaginables para el cerebro humano, la IA está permitiendo avances sin precedentes en áreas como el diagnóstico médico, la personalización de servicios y la optimización de procesos. Sin embargo, también plantea desafíos éticos y sociales importantes, como la privacidad, la seguridad y el impacto en el empleo.

LA TRANSFORMACIÓN DE LA GESTIÓN EMPRESARIAL

Históricamente, la gestión empresarial se ha basado en la experiencia humana, la intuición y la interpretación de datos. Sin embargo, con la creciente cantidad de datos disponibles, la IA ha emergido como una herramienta clave para analizar grandes volúmenes de información de manera rápida y precisa (Brynjolfsson y McAfee, 2017). Esto permite a las empresas no solo optimizar sus operaciones internas, sino también mejorar la toma de decisiones estratégicas.

Un área donde la IA ha mostrado un impacto considerable es en la automatización de procesos. Las tecnologías de IA, como el aprendizaje automático y la minería de datos, permiten a las empresas automatizar tareas rutinarias, lo que libera a los empleados para que se con-

centren en actividades de mayor valor añadido (Davenport y Ronanki, 2018). Esto no solo mejora la eficiencia, sino que también puede conducir a una mayor satisfacción laboral al reducir la carga de trabajo repetitiva (ver **Figura 1**).

Figura 1. Transformación empresarial IA.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

La automatización de procesos es uno de los cambios más significativos que la IA ha traído a la gestión empresarial. Las tareas repetitivas y rutinarias, que antes requerían una gran cantidad de tiempo y recursos humanos, ahora pueden ser realizadas por sistemas de IA. Esto no solo libera a los empleados para que se concentren en actividades más estratégicas, sino que también reduce costos operativos y aumenta la eficiencia (IBM, 2024).

Además, la IA ha mejorado significativamente la toma de decisiones empresariales. Gracias a su capacidad para analizar grandes volúme-

nes de datos en tiempo real, las empresas pueden tomar decisiones más informadas y basadas en datos. Herramientas como el análisis predictivo y el aprendizaje automático permiten identificar patrones y tendencias, facilitando la planificación estratégica y la gestión de riesgos (Gartner, 2024). La personalización de la experiencia del cliente es otro aspecto en el que la IA ha tenido un impacto notable. Mediante el análisis de datos de comportamiento y preferencias, las empresas pueden adaptar sus productos y servicios a las necesidades individuales de cada cliente. Esto no solo mejora la satisfacción del cliente, sino que también aumenta la lealtad y las ventas (EUDE, 2024).

IMPACTO EN LA TOMA DE DECISIONES

La toma de decisiones es un aspecto crucial de la gestión empresarial, y la IA ha demostrado ser un recurso invaluable en este campo. Los algoritmos de IA pueden procesar y analizar grandes conjuntos de datos, identificar patrones y proporcionar recomendaciones basadas en datos que los gerentes pueden no haber considerado.

Según Agrawal y otros (2018), la IA reduce la incertidumbre al mejorar la precisión de las predicciones, lo que a su vez mejora la calidad de las decisiones empresariales.

No obstante, la dependencia de la IA para la toma de decisiones también plantea preguntas éticas y operativas. Las empresas deben considerar cómo equilibrar el uso de la IA con la necesidad de mantener la supervisión humana y asegurar que las decisiones automatizadas sean justas y transparentes (O'Neill, 2016).

A continuación, se describen algunos de los principales aspectos que influyen en la toma de decisiones en este contexto:

1. Reestructuración Laboral y Automatización:
    - Reducción de costos: Las empresas pueden optar por reemplazar trabajadores con IA para reducir costos de operación y mejorar la eficiencia. Esto influye en las decisiones empresariales al priorizar tecnologías que minimicen gastos y aumenten la productividad.

- Cambio en la demanda de habilidades: La IA obliga a las organizaciones a decidir sobre la reconversión de su fuerza laboral, optando por capacitar a los empleados en habilidades tecnológicas y cognitivas avanzadas o despedir personal cuyo trabajo es automatizable.

2. Estrategias de Inversión y Competitividad:
   - Innovación tecnológica: Las empresas deben decidir si invertir en la implementación de IA para mantenerse competitivas. Esto requiere asignar presupuesto a investigación y desarrollo (I+D) y a la adopción de nuevas tecnologías.
   - Adaptación del negocio: Las compañías enfrentan decisiones estratégicas sobre cómo remodelar sus operaciones, productos y servicios en función de la IA, buscando nuevas oportunidades de crecimiento en lugar de limitarse a automatizar.
3. Impacto Social y Ético:
   - Desigualdad y desempleo: Un factor crucial en la toma de decisiones es el impacto social del reemplazo de trabajadores. Los responsables políticos y empresariales deben equilibrar la eficiencia económica con las consecuencias sociales, como el desempleo masivo en sectores vulnerables, lo que podría aumentar la desigualdad.
   - Responsabilidad social corporativa: Las empresas deben tomar decisiones relacionadas con la responsabilidad social, considerando cómo la automatización puede afectar su imagen pública y la relación con la comunidad. Invertir en la reentrenamiento y soporte de empleados desplazados es una decisión estratégica para mitigar impactos negativos.
4. Normativas y Regulación:
   - Políticas gubernamentales: Los gobiernos están tomando decisiones sobre cómo regular el uso de la IA para proteger los empleos y la estabilidad económica. Esto incluye

impuestos sobre robots, regulaciones laborales y marcos legales para la IA.

- Seguridad y privacidad: Las decisiones también están influenciadas por el manejo de los datos generados por la IA. La protección de datos y la ética en el uso de la IA influyen en las políticas y normativas que afectan tanto a empresas como a trabajadores.

5. Adaptación y Resiliencia Organizacional:
    - Capacidad de respuesta: Las organizaciones deben decidir cómo adaptarse rápidamente a los cambios provocados por la IA. Aquellas que puedan implementar la IA de manera gradual y combinarla con la inteligencia humana podrían mitigar el impacto negativo en los empleados y encontrar formas más sostenibles de operar.
6. Cambio en las Relaciones Laborales:
    - Nuevos roles y estructura organizativa: Los directivos y líderes organizacionales deberán tomar decisiones sobre la reorganización del trabajo, delegando tareas rutinarias a la IA y asignando a los empleados a tareas más creativas y estratégicas.
    - Relaciones contractuales: Las relaciones laborales podrían transformarse, afectando la duración de los contratos, la tercerización y la estructura tradicional de empleo a largo plazo, forzando decisiones sobre cómo mantener una fuerza laboral flexible y adaptativa.
7. Impacto Psicológico en los Empleados:
    - Motivación y moral: La toma de decisiones también está influenciada por los efectos psicológicos que el reemplazo por IA puede generar en los empleados. La ansiedad por la automatización puede disminuir la moral, y las organizaciones deben decidir cómo gestionar la comunicación y el apoyo para los trabajadores afectados.

¿Reemplazo? ¿Mejora?

La inteligencia artificial (IA) es una tecnología que está avanzando rápidamente y que está transformando la forma en que las empresas operan. Una de las consecuencias más significativas de este avance es el reemplazo de trabajadores en diversas industrias, ejemplo de ello es:

- La IA está siendo utilizada para automatizar tareas repetitivas y predecibles, como la fabricación, el procesamiento de datos y la atención al cliente. Estas tareas, que anteriormente requerían la intervención humana, ahora pueden ser realizadas por máquinas con mayor eficiencia y precisión.
- La IA también está siendo utilizada para analizar grandes cantidades de datos y tomar decisiones basadas en ellos. Esto está permitiendo a las empresas optimizar sus procesos y mejorar su toma de decisiones, lo que a su vez está llevando a la reducción de personal en áreas como la contabilidad, el análisis financiero y la logística.
- La IA está siendo utilizada para desarrollar robots y drones que pueden realizar tareas físicas, como la manipulación de objetos y la entrega de paquetes. Esto está permitiendo a las empresas reducir su dependencia de la mano de obra humana en áreas como la manufactura, la construcción y la logística.

Sin embargo, el reemplazo de trabajadores por IA no es solo una cuestión de eficiencia y productividad. También plantea importantes desafíos éticos y sociales (ver **Figura 2**). ¿Qué pasará con los trabajadores que pierdan sus empleos debido a la automatización? ¿Cómo se garantizará que las empresas no exploten a los trabajadores que quedan? ¿Cómo se distribuirán los beneficios de la productividad generada por la IA?

Figura 2. Mejora de la IA en las empresas.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

Para abordar estos desafíos, es importante que las empresas y los gobiernos trabajen juntos para desarrollar políticas y programas que apoyen a los trabajadores que pierdan sus empleos debido a la automatización. Esto puede incluir la formación y el reciclaje de habilidades, la creación de nuevos empleos en áreas relacionadas con la IA, y la implementación de una renta básica universal para garantizar que todos tengan acceso a los recursos básicos.

Además, es importante que las empresas sean transparentes sobre cómo están utilizando la IA y cómo esto afectará a los trabajadores. Esto puede incluir la creación de comités de ética y la implementación de políticas de responsabilidad social corporativa.

En cuanto a los beneficios de la productividad generada por la IA, es importante que se distribuyan de manera justa y equitativa.

Esto puede incluir la implementación de impuestos sobre la automatización y la utilización de estos fondos para financiar programas sociales y de apoyo a los trabajadores.

LA AUTOMATIZACIÓN Y EL DESPLAZAMIENTO DE LOS TRABAJADORES

Uno de los efectos más evidentes de la IA es su capacidad para automatizar tareas que tradicionalmente requerían intervención humana. En la industria manufacturera, por ejemplo, la implementación de robots y sistemas automatizados ha llevado a una reducción significativa de la necesidad de trabajadores en líneas de producción. Los avances en IA permiten que estas máquinas no solo realicen tareas repetitivas, sino que también puedan aprender y adaptarse a nuevas situaciones, lo que las hace cada vez más competentes en roles complejos.

En la logística y la cadena de suministro, la IA se utiliza para optimizar rutas de entrega, gestionar inventarios y predecir la demanda de productos. Empresas como Amazon han implementado sistemas de IA que permiten una gestión eficiente de sus almacenes, reduciendo la necesidad de trabajadores humanos en estas áreas (Brynjolfsson y McAfee, 2014).

Del mismo modo, la industria del transporte está experimentando una transformación con la llegada de vehículos autónomos, lo que podría llevar a la desaparición de millones de empleos en el sector de los conductores. De manera similar, en el sector servicios, los *chatbots* y asistentes virtuales están reemplazando a trabajadores en atención al cliente, una tendencia que se intensifica con el perfeccionamiento del procesamiento del lenguaje natural.

El impacto de la IA en el empleo es un tema de gran debate. Si bien algunos argumentan que la automatización y la IA crearán nuevos tipos de trabajos y oportunidades, la evidencia sugiere que muchos empleos tradicionales están en riesgo de desaparecer. Según un informe de McKinsey & Company (2017), hasta el 30 % de las actividades laborales actuales podrían ser automatizadas para 2030, lo que podría afectar a cientos de millones de trabajadores a nivel mundial.

El reemplazo de trabajadores por sistemas de IA no solo afecta a aquellos en trabajos manuales o rutinarios, sino también a empleados en posiciones de oficina y de nivel medio.

Los algoritmos de IA son capaces de realizar tareas complejas como la contabilidad, el análisis de datos y la redacción de informes, lo que podría disminuir la demanda de trabajadores en estos campos (Frey y Osborne, 2017).

Además, la desigualdad económica podría intensificarse a medida que los beneficios de la IA se concentran en manos de aquellos con habilidades técnicas avanzadas y capital, mientras que los trabajadores menos calificados enfrentan el desempleo o la sub-empleabilidad. Esto podría llevar a un aumento en la polarización económica y social.

BENEFICIOS POTENCIALES DE LA IA

A pesar de estos desafíos, la IA también ofrece oportunidades significativas para la economía y la sociedad en general. La automatización puede aumentar la productividad y eficiencia de las empresas, lo que puede traducirse en menores costos para los consumidores y mayores márgenes de beneficio para las empresas. Además, la IA puede asumir tareas peligrosas o monótonas, mejorando la seguridad y la satisfacción laboral de los trabajadores humanos.

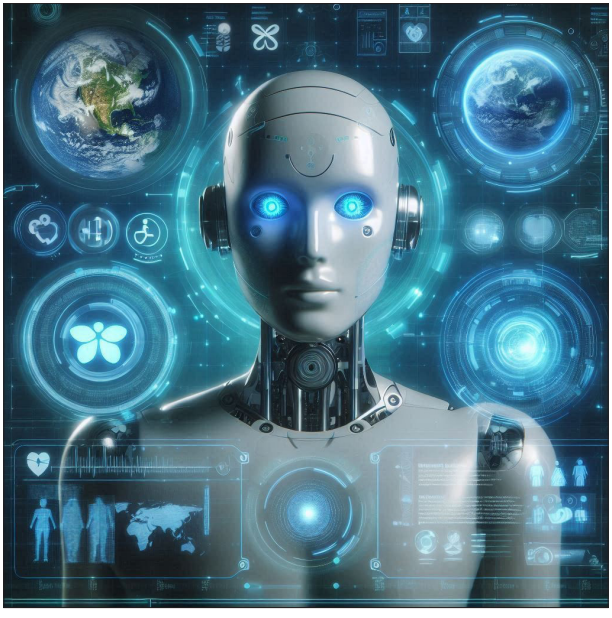
Además, la IA tiene el potencial de crear nuevos tipos de empleos. A medida que las industrias se adaptan a la automatización, surge la necesidad de especialistas en el desarrollo, mantenimiento y supervisión de sistemas de IA.

Esto podría generar una demanda creciente de profesionales con habilidades en ciencia de datos, programación y ciberseguridad, entre otros campos (ver **Figura 3**).

No obstante, el reemplazo de trabajadores por IA plantea importantes desafíos sociales y económicos. Uno de los mayores riesgos es la exacerbación de la desigualdad. Los trabajadores con menores niveles de educación y habilidades específicas corren un mayor riesgo de ser desplazados, lo que podría agravar las disparidades económicas exis-

tentes. Además, existe la posibilidad de que los beneficios de la automatización se concentren en las manos de unos pocos, particularmente en grandes corporaciones tecnológicas, mientras que los costos, en términos de pérdida de empleo, se distribuyen de manera más amplia.

Figura 3. Inteligencia Artificial y sus beneficios.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

Otro desafío significativo es la transición hacia un mercado laboral más automatizado. El reentrenamiento y la reconversión laboral son esenciales para asegurar que los trabajadores desplazados puedan encontrar nuevas oportunidades en la economía emergente. Sin embargo, implementar programas efectivos de capacitación a gran escala es una tarea compleja y costosa, y no todos los trabajadores tendrán la capacidad o el acceso a los recursos necesarios para adaptarse a nuevas industrias.

PREPARACIÓN ANTE EL FUTURO

Otro aspecto importante por considerar es la educación y la formación de los trabajadores. La IA requiere habilidades y conocimientos específicos, como la programación, el análisis de datos y la inteligencia artificial. Por lo tanto, es importante que los sistemas educativos se adapten a estas necesidades y ofrezcan programas de formación que preparen a los trabajadores para las nuevas demandas del mercado laboral.

Además, es importante fomentar la creatividad, la innovación y el pensamiento crítico en los trabajadores, ya que estas habilidades serán cada vez más valiosas en un mundo donde la IA está presente. También es importante enseñar a los trabajadores a trabajar junto con la IA, a entender sus limitaciones y a saber cómo aprovechar sus ventajas.

En cuanto a las industrias que están más expuestas al reemplazo por IA, es importante que se tomen medidas para proteger a los trabajadores y para garantizar que la transición sea lo más suave posible (ver **Figura 4**). Por ejemplo, se pueden ofrecer programas de reciclaje de habilidades, apoyo para la creación de nuevos negocios y ayuda para la transición a nuevas industrias.

Es importante destacar que la IA no solo reemplaza empleos, sino que también crea nuevos. Por ejemplo, la creación de *chatbots* y asistentes virtuales ha generado una demanda de expertos en lenguaje natural y procesamiento de lenguaje. De igual forma, la implementación de sistemas de IA en la salud ha creado una necesidad de profesionales que puedan interpretar y analizar los datos generados por estos sistemas. Sin embargo, es importante que estos nuevos empleos sean accesibles para todos, y no solo para aquellos con habilidades y conocimientos especializados. Por lo tanto, es importante que se ofrezcan programas de formación y educación que permitan a los trabajadores adquirir las habilidades necesarias para ocupar estos nuevos empleos.

Además, es importante que se fomente la creación de empleos que no puedan ser reemplazados por la IA, como aquellos que requieren creatividad, empatía y habilidades sociales. Por ejemplo, la enseñanza, la atención médica y el trabajo social son áreas que requieren habilidades humanas únicas que no pueden ser replicadas por la IA.

Figura 4. Adaptación inteligencia artificial.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

IMPLICACIONES ÉTICAS Y SOCIALES

El reemplazo de trabajadores por IA plantea serias implicaciones éticas y sociales (ver **Figura 5**). Uno de los principales desafíos es garantizar que los beneficios de la IA se distribuyan de manera equitativa. Los gobiernos y las empresas deben considerar políticas que protejan a los trabajadores desplazados, como programas de reciclaje profesional y la creación de redes de seguridad social más robustas.

Además, existe una preocupación creciente sobre la *"deshumanización"* del trabajo. A medida que las máquinas asumen tareas que antes realizaban humanos, puede surgir una desconexión entre los trabajadores y el propósito de su labor, lo que afecta el sentido de identidad y realización personal (Susskind, 2020).

Según estudios recientes, millones de trabajos podrían ser automatizados en las próximas décadas, lo que generaría una significativa dislocación laboral (Ford, 2015). Esta situación plantea la necesidad de reconsiderar los modelos tradicionales de empleo y la formación profesional, para asegurar que los trabajadores puedan adaptarse a la nueva economía digital.

La desigualdad económica es otra preocupación ética central. La tecnología de IA tiende a concentrar la riqueza y el poder en manos de aquellos que la desarrollan y controlan, como las grandes corporaciones tecnológicas. Esto podría exacerbar las brechas económicas ya existentes, creando una sociedad donde un pequeño grupo se beneficia desproporcionadamente de los avances tecnológicos, mientras que una gran mayoría podría verse marginada y empobrecida (Brynjolfsson y McAfee, 2014).

Figura 5. Inteligencia artificial y sus implicaciones éticas y sociales.



Fuente: Microsoft Copilot (2024), con la alimentación de datos propios.

A medida que la IA avanza y se integra en diversos sectores, es fundamental examinar el impacto que tendrá en la fuerza laboral y en la sociedad en su conjunto. Una de las preocupaciones más apremiantes es la desigualdad económica. La automatización tiene el potencial de beneficiar desproporcionadamente a aquellos que poseen o controlan la tecnología. Esto puede aumentar la brecha entre los ricos y los pobres, dejando a los trabajadores menos capacitados sin recursos suficientes para adaptarse a las nuevas realidades del mercado laboral.

La ética de esta desigualdad plantea preguntas sobre la justicia y la equidad en la distribución de la riqueza generada por la IA. Además, surge un dilema acerca de la responsabilidad. Cuando una IA comete errores, como en el caso de un accidente causado por un vehículo autónomo, es difícil determinar quién es responsable: el fabricante, el programador, o la propia IA.

Esta ambigüedad puede generar problemas en la justicia y en la ética empresarial, así como desconfianza en la tecnología. Otra dimensión ética es el riesgo de sesgos inherentes en los sistemas de IA. Ya que estos modelos aprenden de los datos existentes, pueden reflejar y amplificar los prejuicios sociales que ya están presentes. Esto puede dar lugar a decisiones sesgadas en ámbitos tan críticos como la contratación, la justicia penal o el acceso a servicios financieros, perpetuando y en ocasiones agravando la desigualdad social.

Finalmente, la transparencia en el funcionamiento de los algoritmos es una preocupación creciente. La complejidad de los sistemas de IA puede hacer que sean prácticamente opacos para los usuarios. La falta de claridad en cómo se utilizan los datos y cómo se toman decisiones puede erosionar la confianza en las instituciones que emplean estas tecnologías, generando una sensación de desamparo en los ciudadanos.

El impacto potencial de la IA para causar desempleo es quizás la consecuencia social más evidente. A medida que los trabajos que requieren habilidades rutinarias y repetitivas son reemplazados por máquinas, muchas personas pueden perder su fuente de ingresos sin alternativas viables. Esto no solo afecta a los individuos, sino también a sus familias y comunidades, generando un círculo vicioso de pobreza y exclusión.

El desplazamiento laboral es otra preocupación importante. A medida que algunos puestos de trabajo desaparecen, la demanda de habilidades específicas, particularmente en tecnología y programación, aumentará. Esto requerirá programas masivos de reeducación y formación, que deben ser accesibles para todos, o de lo contrario, los trabajadores menos capacitados podrían quedarse atrás. La dinámica laboral también cambiará, con un posible aumento en la colaboración entre humanos y máquinas.

Esto plantea preguntas sobre cómo se rediseñarán los roles en el trabajo y qué habilidades serán fundamentales para operar en este nuevo entorno colaborativo. El enfoque puede cambiar hacia trabajos que requieren creatividad, empatía y razonamiento crítico, habilidades que la IA aún no puede igualar. Finalmente, las implicaciones psicológicas no deben pasarse por alto. La incertidumbre acerca del futuro del trabajo puede generar ansiedad y estrés en la fuerza laboral, afectando el bienestar psicológico y la salud mental de los individuos.

La transición hacia una economía impulsada por la IA debe incluir un enfoque en el bienestar de los trabajadores y en la construcción de resiliencia en tiempos de cambio. La sociedad debe abordar estos desafíos de manera proactiva, buscando un equilibrio que asegure que los beneficios de la IA se distribuyan de manera justa y equitativa.

ESTRATEGIAS PARA MITIGAR EL IMPACTO

Para mitigar los efectos negativos del reemplazo de trabajadores por IA, es crucial adoptar un enfoque proactivo. Una estrategia clave es la inversión en educación y formación para desarrollar las habilidades necesarias en la economía impulsada por la IA. Esto incluye tanto la formación técnica en habilidades digitales como el desarrollo de competencias en áreas que son menos susceptibles de automatización, como la creatividad, la empatía y el pensamiento crítico (World Economic Forum, 2018).

Además, las políticas públicas deben promover la creación de empleo en sectores emergentes que complementen la automatización. El apoyo a la investigación y desarrollo en áreas como la energía renova-

ble, la atención sanitaria y los servicios sociales puede generar nuevas oportunidades laborales que no solo compensen la pérdida de empleos debido a la IA, sino que también contribuyan al bienestar social.

El uso de la IA también plantea preguntas sobre la responsabilidad. Cuando una máquina toma decisiones que afectan a los seres humanos, ¿quién es responsable de esas decisiones? Por ejemplo, si un algoritmo de IA comete un error que resulta en la pérdida de un empleo, ¿es responsable la empresa que lo implementó, el programador que lo diseñó, o la IA misma? La falta de claridad en estas cuestiones éticas podría llevar a una situación en la que los trabajadores sean tratados de manera injusta sin un recurso claro para apelar o buscar justicia (O'Neill, 2016).

CONCLUSIONES

En conclusión, la IA está transformando la forma en que las empresas operan y está reemplazando a los trabajadores en diversas industrias. Sin embargo, es importante abordar los desafíos éticos y sociales que esto plantea para garantizar que los beneficios de la IA se distribuyan de manera justa y equitativa.

La inteligencia artificial está redefiniendo la gestión empresarial, ofreciendo nuevas formas de optimizar operaciones, mejorar la toma de decisiones y crear valor. Sin embargo, para aprovechar plenamente estos beneficios, las empresas deben superar desafíos significativos relacionados con la implementación de IA, la resistencia al cambio y la seguridad de los datos. A medida que la tecnología continúa avanzando, será crucial que los líderes empresariales adopten un enfoque proactivo y ético para integrar la IA en sus estrategias de gestión.

Finalmente, hay que reconocer que la IA tiene el potencial de transformar la forma en que las empresas operan y de mejorar la eficiencia y la productividad. Sin embargo, es importante abordar los desafíos éticos y sociales que esto plantea para garantizar que los beneficios de ésta se distribuyan de manera justa y equitativa.

REFERENCIAS BIBLIOGRÁFICAS

Agrawal, A., Gans, J. S., y Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence.* Harvard Business Review Press.

Brynjolfsson, E., y McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W.W. Norton & Company.

Brynjolfsson, E., y McAfee, A. (2017). *Machine, Platform, Crowd: Harnessing Our Digital Future.* W. W. Norton & Company.

Copilot. (2024). *Página Oficial.* https://copilot.microsoft.com/

Davenport, T. H., y Ronanki, R. (2018). Artificial Intelligence for the Real World.

EUDE. (2024). *El impacto de la inteligencia artificial en la gestión empresarial.* https://www.eude.es/blog/el-impacto-de-la-inteligencia-artificial-en-la-gestion-empresarial/

Ford, M. (2015). *Rise of the Robots: Technology and the Threat of a Jobless Future.* Basic Books. https://www.uc.pt/feuc/citcoimbra/Martin_Ford-Rise_of_the_Robots

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

Gartner. (2024). *La revolución de la gestión empresarial: ejecutivos y equipos de trabajo potenciados con IA*. https://www.toolify.ai/es/ai-newses/la-revolucin-de-la-gestin-empresarial-ejecutivos-y-equipos-de-trabajopotenciados-con-ia-l-gartner-it-symposiumxpo-1584658

Gemini. (2024). *Página Oficial.* https://gemini.google.com/

Harvard Business Review, 96(1), 108-116. https://blockqai.com/wp-content/uploads/2021/01/analytics-hbr-ai-for-the-real-world.pdf

IBM. (2024). *¿Qué es la IA empresarial?* https://www.ibm.com/eses/topics/enterprise-ai

McKinsey & Company. (2017). *Jobs lost, jobs gained: Workforce transitions in a time of automation. McKinsey Global Institute.* https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/what%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/mgi-jobs-lost-jobs-gained-executive-summary-december-6-2017.PDF

Microsoft Copilot. (2024). *Copilot Bing*. https://copilot.microsoft.com/

O'Neill, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

OpenAI. (2024). *ChatGPT*. https://www.openai.com/chatgpt

Susskind, D. (2020). *A world without work: Technology, automation, and how we should respond*. Metropolitan Books.

World Economic Forum. (2018). *The future of jobs report 2018*. Geneva: World Economic Forum.

CAPÍTULO IV

EL CAMBIO DE LA ÉTICA DEL DEPORTE CON EL IMPACTO DE LA INTELIGENCIA ARTIFICIAL

MARLENE ATENEA GRIMALDO RAMÍREZ[1]

RESUMEN

El impacto de la inteligencia artificial (IA) en el deporte está transformando tanto el rendimiento de los atletas como la gestión y el arbitraje de las competiciones. Este artículo explora cómo la IA está optimizando el entrenamiento y la recuperación, mejorando la precisión en las decisiones arbitrales y planteando nuevas preocupaciones éticas sobre la privacidad de los atletas y la equidad en las competiciones. Además, analiza el papel de organizaciones como el Comité Olímpico Internacional (COI) en el desarrollo de normas que regulen el uso responsable de estas tecnologías. El futuro de las regulaciones en IA y el deporte olímpico dependerá de un equilibrio entre la innovación y la preservación de los principios fundamentales del deporte, como la equidad, la transparencia y la integridad. Las normativas internacionales deberán ser dinámicas y garantizar la protección de los derechos de los atletas, mientras fomentan la implementación justa y ética de la IA.

Palabras clave: Inteligencia artificial, Deporte, Ética, Comité Olímpico Internacional.

[1] Alumna de la Licenciatura en Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

INTRODUCCIÓN

El deporte ha sido, desde sus inicios, una actividad profundamente ligada a la habilidad, el esfuerzo físico y mental, y los valores de la competencia justa. A lo largo de los siglos, ha evolucionado desde una práctica ritual y recreativa hasta convertirse en un fenómeno global con implicaciones económicas, sociales y culturales. Sin embargo, en las últimas décadas, el avance imparable de la tecnología ha comenzado a transformar no solo la manera en que se juegan los deportes, sino también las reglas éticas que guían la competencia.

Una de las tecnologías más disruptivas en esta evolución es la inteligencia artificial (IA). La IA ha irrumpido en el deporte de manera casi silenciosa pero contundente, ofreciendo herramientas para mejorar el rendimiento de los atletas, monitorear su salud y optimizar estrategias de juego. Desde la recopilación masiva de datos biométricos hasta la automatización de decisiones tácticas en tiempo real, la IA ha abierto nuevas posibilidades que antes parecían ciencia ficción. Sin embargo, estos avances también han generado una serie de preguntas éticas que no podemos ignorar.

¿Cómo afecta la IA a la equidad en las competiciones deportivas? ¿Estamos poniendo en riesgo la privacidad y autonomía de los atletas al recolectar y analizar cantidades masivas de datos? ¿Estamos sacrificando el esfuerzo y la habilidad humana en favor de la automatización y la eficiencia tecnológica? Estas preguntas están en el corazón de una discusión cada vez más urgente sobre el futuro del deporte.

Este artículo explora el impacto de la inteligencia artificial en el deporte desde una perspectiva ética, analizando cómo esta tecnología está cambiando los principios fundamentales de la competencia justa, la privacidad, la autonomía y la naturaleza misma del deporte. A medida que avanzamos hacia un futuro cada vez más tecnológico, es vital que reflexionemos sobre cómo la IA puede integrarse de manera responsable en el deporte, asegurando que las innovaciones no comprometan los valores que han hecho del deporte una parte tan valiosa de la experiencia humana.

EVOLUCIÓN DE LA ÉTICA EN EL DEPORTE

La ética deportiva ha evolucionado con el tiempo, reflejando los cambios culturales y tecnológicos de cada época. En los primeros días del deporte moderno, el espíritu *amateur* prevalecía como el ideal. La ética en el deporte ha estado tradicionalmente basada en principios como el *"fair play"*, la integridad y la equidad competitiva (Simon, 2016). Estos principios buscan garantizar que todos los participantes tengan las mismas oportunidades, independientemente de factores externos.

La introducción de la IA en el deporte ha provocado preguntas sobre si estas tecnologías respetan los principios fundamentales de la competencia justa (Sanderson y Flynn, 2020). Sin embargo, a medida que el deporte se profesionalizó, surgieron nuevas tensiones éticas. La búsqueda de la victoria a cualquier costo comenzó a desvirtuar la noción original del juego limpio, dando paso a prácticas como el dopaje, que introdujeron los primeros grandes dilemas éticos relacionados con la tecnología.

Las regulaciones antidopaje nacieron como una respuesta para proteger la equidad en la competencia, y hoy en día, el uso de sustancias para mejorar el rendimiento es uno de los problemas más discutidos en el ámbito deportivo. Sin embargo, la llegada de tecnologías aún más avanzadas, como la IA, ha ampliado el debate ético. ¿Cómo debemos abordar las nuevas formas de mejora del rendimiento que no dependen de drogas, sino de algoritmos y datos? ¿Es ético utilizar herramientas que optimicen cada aspecto del entrenamiento y la competencia? El paralelismo con el dopaje tecnológico es claro, pero la cuestión es más compleja.

La llegada de innovaciones como la ropa deportiva aerodinámica, las zapatillas de alta tecnología y las prótesis avanzadas para atletas paralímpicos también han desafiado los límites de lo que consideramos "*natural*" en el deporte. La IA, sin embargo, lleva este dilema un paso más allá, ya que sus aplicaciones no solo mejoran el equipamiento o las condiciones físicas del atleta, sino que comienzan a influir en su toma de decisiones, su preparación mental e incluso la predicción de su rendimiento futuro.

LA REVOLUCIÓN DE LA IA EN EL DEPORTE

La introducción de la inteligencia artificial en el deporte ha marcado una revolución en cómo entendemos y gestionamos la competencia. La IA está desempeñando un papel cada vez más prominente, desde el nivel *amateur* hasta las competiciones profesionales de alto rendimiento. Pero ¿cómo ha cambiado específicamente la IA el mundo del deporte?

La IA ha revolucionado el deporte, proporcionando herramientas avanzadas para el análisis de rendimiento y la toma de decisiones tácticas. Estas tecnologías permiten a entrenadores y equipos mejorar significativamente su comprensión del juego (Li, 2021). Sin embargo, la dependencia de estos sistemas plantea preocupaciones sobre el impacto en la imprevisibilidad y el esfuerzo humano en el deporte (Knight, 2019).

Análisis de Rendimiento Deportivo

Uno de los usos más comunes de la IA es el análisis de datos en tiempo real. Mediante la recopilación de información biométrica y estadística, los algoritmos de IA pueden procesar rápidamente grandes cantidades de datos para ofrecer un análisis detallado del rendimiento de un atleta o equipo. Por ejemplo, en deportes como el fútbol o el baloncesto, la IA puede identificar patrones de juego, analizar la condición física de los jugadores y predecir el desgaste físico durante un partido. Esto no solo ayuda a los entrenadores a ajustar las estrategias en tiempo real, sino que también permite a los atletas personalizar su entrenamiento para alcanzar el máximo rendimiento.

Empresas como *IBM Watson* y *Zebra Technologies* han desarrollado herramientas que analizan datos deportivos con una precisión asombrosa, proporcionando informes detallados sobre el rendimiento individual y colectivo. En deportes como el tenis, la IA se ha utilizado para estudiar las características del saque de un jugador y la probabilidad de éxito de diferentes tipos de golpes, lo que permite a los entrenadores ajustar tácticas y predecir el comportamiento de sus oponentes.

Prevención de Lesiones y Personalización del Entrenamiento

Uno de los aspectos más positivos de la IA en el deporte es su capacidad para prevenir lesiones. A través de la recopilación de datos sobre la salud del atleta, su nivel de estrés, y los movimientos repetitivos que realiza durante el entrenamiento o los partidos, los algoritmos de IA pueden identificar patrones que indican un riesgo elevado de lesión. De esta manera, los entrenadores y los equipos médicos pueden ajustar el entrenamiento de los atletas para minimizar el riesgo, lo que les permite mantenerse sanos por más tiempo y rendir al máximo nivel.

Por ejemplo, la empresa *Catapult* ha desarrollado dispositivos portátiles que, mediante IA, monitorizan en tiempo real el estado físico del atleta, proporcionando información sobre el riesgo de sobrecarga muscular o el estado de fatiga. Esto ha cambiado significativamente la manera en que los equipos gestionan a sus jugadores, priorizando la salud a largo plazo sobre el rendimiento a corto plazo.

IA en la Toma de Decisiones Tácticas y Estratégicas

Otro impacto clave de la IA en el deporte es su capacidad para influir en las decisiones tácticas y estratégicas. Las herramientas de IA pueden analizar las tácticas de los oponentes y predecir sus próximas jugadas, lo que da a los entrenadores una ventaja significativa. Por ejemplo, en deportes como el fútbol americano o el rugby, la IA se usa para analizar patrones de jugadas y proporcionar predicciones en tiempo real sobre las mejores acciones a seguir. Esto ha llevado a una optimización sin precedentes en la planificación de jugadas y en la toma de decisiones durante los partidos.

Si bien estas innovaciones ofrecen beneficios claros, también plantean preguntas éticas. ¿Dónde está el límite entre la estrategia humana y la intervención tecnológica? ¿Debería permitirse que las decisiones tácticas en un deporte dependan cada vez más de los algoritmos, o se trata de una violación de los principios tradicionales de la competencia deportiva?

IMPACTO DE LA IA EN LA EQUIDAD Y LA COMPETENCIA JUSTA

Uno de los principios fundamentales del deporte es la equidad, o la igualdad de oportunidades para todos los competidores. Sin embargo, el acceso a las tecnologías avanzadas de IA puede crear brechas significativas entre los equipos con mayores recursos y aquellos con menos posibilidades económicas. Este es uno de los mayores desafíos éticos que presenta la IA en el deporte. Esta disparidad tecnológica amenaza con desnaturalizar la competencia (Wilson, 2020). Además, algunos comparan el uso de la IA con el dopaje, argumentando que puede otorgar una ventaja desleal (Jones, 2019).

Acceso Desigual a la Tecnología

El uso de IA en el deporte depende en gran medida de la infraestructura tecnológica disponible. Equipos de élite o aquellos con patrocinadores corporativos importantes tienen acceso a tecnologías avanzadas que les permiten analizar datos y optimizar el rendimiento de maneras que no están al alcance de equipos más modestos. Esto puede generar una disparidad entre competidores que va más allá de las habilidades físicas o tácticas y se convierte en una cuestión de acceso a herramientas tecnológicas.

Este fenómeno es visible en ligas de fútbol como la Premier League, donde los equipos más ricos pueden permitirse costosos sistemas de IA para análisis de rendimiento, mientras que los equipos de divisiones inferiores carecen de estos recursos. Esta desigualdad amenaza con desnaturalizar la competencia, ya que el éxito podría depender cada vez más del acceso a la tecnología, en lugar de depender exclusivamente del talento o del trabajo duro de los atletas.

Comparación con el Dopaje*: ¿*Ventaja Tecnológica Desleal*?*

Un paralelismo frecuente en el debate sobre la ética de la IA en el deporte es la comparación con el dopaje. Mientras que el dopaje mejora el rendimiento físico mediante sustancias químicas, la IA ofrece ventajas tecnológicas que pueden ser igualmente significativas. Si bien el dopaje está prohibido por la mayoría de las organizaciones deportivas,

la IA todavía no está estrictamente regulada en este contexto, lo que plantea la pregunta: ¿debería considerarse el uso de IA como una forma de *"dopaje tecnológico"?*

Algunos argumentan que permitir que los equipos con mayor acceso a la IA tengan una ventaja competitiva desleal es una violación del principio de *"fair play"*. Esto plantea la necesidad urgente de establecer normativas claras que regulen cómo y cuándo se puede utilizar la IA en el deporte, para que se mantenga la equidad en la competencia.

DESAFÍOS EN LA PRIVACIDAD Y LA AUTONOMÍA DE LOS ATLETAS

Uno de los aspectos más controvertidos del uso de la IA en el deporte es la cuestión de la privacidad de los atletas. Los sistemas de IA dependen en gran medida de la recopilación de datos personales, incluyendo información sensible sobre el estado físico y mental de los deportistas. Esto plantea serias preocupaciones éticas sobre el control y uso de esos datos.

Recopilación y Análisis de Datos Biométricos

En los deportes profesionales, es común que se monitorice el rendimiento de los atletas de forma continua. Dispositivos portátiles como los relojes inteligentes o las camisetas equipadas con sensores recogen datos biométricos que luego se procesan mediante algoritmos de IA. Estos datos incluyen desde el ritmo cardíaco hasta el nivel de oxígeno en sangre, la cantidad de sueño y el nivel de fatiga mental.

El uso de IA en el deporte implica una recopilación constante de datos biométricos que plantea graves preocupaciones sobre la privacidad y el control de los atletas sobre su propia información (Doe, 2021). En muchos casos, los atletas no son dueños de los datos recopilados, lo que puede llevar a decisiones que afecten su bienestar y rendimiento (Santos, 2020).

Si bien estos datos son esenciales para la optimización del rendimiento, también plantean preguntas sobre la privacidad. ¿Quién tiene

acceso a estos datos? ¿Cómo se protegen de usos indebidos? La recopilación constante de datos podría tener implicaciones no solo sobre la salud del atleta, sino también sobre su bienestar emocional, ya que podría generar una presión adicional al saber que su rendimiento está siendo analizado en cada momento.

Control y Propiedad de los Datos

Además de la privacidad, existe el desafío de la propiedad de los datos. ¿Pertenecen esto1s datos al atleta, al equipo o a la empresa que proporciona la tecnología? En muchos casos, los atletas no tienen control total sobre sus propios datos biométricos, lo que plantea una cuestión de autonomía. Esto podría llevar a situaciones en las que los equipos o patrocinadores utilicen estos datos para tomar decisiones que podrían ir en contra de los intereses del atleta, como reducir su tiempo de juego por razones basadas en predicciones algorítmicas.

Autonomía del Atleta y Dependencia de la IA

Otro desafío es la posible pérdida de autonomía del atleta. A medida que la IA juega un papel más importante en la toma de decisiones estratégicas y tácticas, el papel del deportista en la toma de decisiones en tiempo real podría verse reducido. En lugar de depender de su intuición o experiencia, los atletas podrían verse obligados a seguir las recomendaciones de los algoritmos de IA, que pueden ofrecer la opción "óptima" basada en datos. Esto plantea la pregunta: ¿se está sacrificando el aspecto humano del deporte en favor de la tecnología?

DESHUMANIZACIÓN DEL DEPORTE: CAMBIOS EN LA NATURALEZA DE LA COMPETENCIA

Uno de los aspectos más profundos del impacto de la IA en el deporte es la posible deshumanización de la competencia. A medida que la IA optimiza cada aspecto del rendimiento deportivo, algunos críticos argumentan que se corre el riesgo de deshumanizar el deporte, haciendo que la previsibilidad y la eficiencia tecnológica reemplacen al esfuerzo

humano (Mitchell, 2021). Esta transformación podría afectar el espíritu deportivo, que valora la sorpresa, la imprevisibilidad y la emoción de la competencia (Riley, 2019).

El deporte, en su esencia, siempre ha sido una prueba de habilidad física y mental, donde el esfuerzo humano, la superación personal y la competencia justa son los valores centrales. Sin embargo, el uso extensivo de la IA podría cambiar la naturaleza misma de lo que consideramos *"deporte"*.

Esfuerzo Humano vs. Eficiencia Tecnológica

Tradicionalmente, el deporte ha sido valorado por su capacidad para poner a prueba los límites humanos. La resistencia, la fuerza, la velocidad y la inteligencia son cualidades que los atletas desarrollan a través de años de entrenamiento riguroso y sacrificio. Sin embargo, la IA, al optimizar cada aspecto del rendimiento, puede reducir el papel del esfuerzo humano, priorizando la eficiencia tecnológica sobre la habilidad natural. ¿Qué significa competir cuando la tecnología toma decisiones clave, optimiza el rendimiento y predice resultados?

Este fenómeno se ha observado en deportes como el ciclismo o el atletismo, donde las mejoras tecnológicas en el equipamiento ya han reducido la importancia del esfuerzo individual en comparación con los avances en bicicletas aerodinámicas o zapatillas especializadas. Si bien la tecnología siempre ha jugado un papel en el deporte, la IA podría llevar esto a un nivel completamente nuevo, donde la previsibilidad y la eficiencia superen a la emoción del esfuerzo humano.

Pérdida de la Imprevisibilidad

Una de las razones por las que el deporte es tan emocionante para los aficionados es su naturaleza impredecible. Los eventos inesperados, las jugadas sorprendentes y los momentos heroicos definen el espíritu del deporte. Sin embargo, la IA, al analizar y predecir cada aspecto del juego, corre el riesgo de hacer que las competiciones sean más predecibles. Esto podría reducir la emoción del deporte, eliminando la posibilidad de sorpresa y sustituyéndola por la certeza algorítmica.

Impacto en el Espíritu Deportivo

El uso de la IA también podría impactar negativamente en el espíritu deportivo, una parte esencial del deporte desde tiempos antiguos. El *"fair play"*, la humildad ante la victoria y el respeto hacia el oponente son valores que podrían verse erosionados si la tecnología domina el campo de juego. La IA, al optimizar el rendimiento y predecir resultados, podría transformar el deporte en una actividad más mecanizada, donde los logros personales y los momentos épicos se vean desplazados por la eficiencia impersonal de los algoritmos.

IA Y LA MANIPULACIÓN DE RESULTADOS

Uno de los aspectos más preocupantes del impacto de la IA en el deporte es el potencial para la manipulación de resultados. A medida que las tecnologías de inteligencia artificial avanzan en su capacidad para predecir el comportamiento de los atletas y los resultados de las competencias, también aumenta el riesgo de que estas herramientas sean utilizadas para manipular eventos deportivos con fines financieros o de control competitivo.

IA y Apuestas Deportivas*:* Un Terreno Peligroso

El auge de las apuestas deportivas ha traído consigo un mercado millonario en el que se juega con las probabilidades de los resultados deportivos. Los algoritmos de IA ya están siendo utilizados por las casas de apuestas para analizar y predecir resultados con un grado de precisión cada vez mayor. Sin embargo, esto plantea preocupaciones éticas. La capacidad de la IA para analizar grandes cantidades de datos y predecir el comportamiento de los jugadores y equipos puede ser aprovechada no solo por las casas de apuestas, sino también por individuos o grupos interesados en influir en los resultados de un partido.

El uso de IA en el análisis de resultados ha abierto una puerta peligrosa hacia la posible manipulación de eventos deportivos. Las tecnologías de predicción, en especial en las apuestas, pueden ser explotadas por quienes buscan obtener beneficios financieros ilícitos (Thompson,

2020). A medida que los algoritmos se vuelven más precisos, también aumenta el riesgo de que sean utilizados para manipular intencionalmente el rendimiento y los resultados (Green, 2019).

El problema es aún más grave en deportes donde pequeñas decisiones, como un penalti fallido o una lesión estratégica, pueden cambiar completamente el resultado. La predicción de estas variables mediante IA podría abrir la puerta a la manipulación intencional, donde jugadores, árbitros o incluso entrenadores sean presionados o incentivados para alterar sus decisiones en función de los intereses de terceros. Este tipo de corrupción podría no solo desnaturalizar la competencia deportiva, sino también erosionar la confianza de los aficionados.

Manipulación de Rendimiento mediante IA

Además de las apuestas deportivas, la manipulación de resultados también puede ocurrir dentro de los equipos. Las herramientas de IA que analizan y optimizan el rendimiento de los jugadores pueden, en manos equivocadas, ser utilizadas para favorecer a ciertos competidores. Por ejemplo, un equipo podría utilizar IA para influir en el tiempo de juego de un atleta clave, basado en predicciones algorítmicas de rendimiento, lo que podría modificar el desarrollo del juego para beneficiar a ciertos intereses.

A esto se suma el hecho de que los algoritmos no siempre son transparentes o imparciales. Dependiendo de cómo se entrenen y programen, los sistemas de IA pueden estar sujetos a sesgos que favorezcan a ciertos jugadores o equipos. Este problema ético pone en cuestión la legitimidad de las decisiones tácticas o

estratégicas que se toman basándose en los análisis proporcionados por la IA.

Casos de Manipulación y Predicción de Resultados

Aunque la manipulación algorítmica de resultados deportivos es aún un campo emergente, ya han surgido ejemplos de cómo la tecnología puede influir en los resultados. En deportes como el béisbol y el baloncesto, se ha documentado cómo el uso intensivo de la analítica

avanzada ha cambiado las dinámicas del juego, con algunos críticos argumentando que estas predicciones tecnológicas pueden restar emoción e imprevisibilidad a las competencias.

El reto para las organizaciones deportivas y los entes reguladores es encontrar un equilibrio entre el uso legítimo de la IA para mejorar el deporte y la prevención de su uso indebido para manipular resultados, una tarea que requiere un marco ético y de control estricto.

REGULACIONES Y NORMATIVAS ÉTICAS: EL PAPEL DE LAS ORGANIZACIONES DEPORTIVAS

A medida que la IA se integra más profundamente en el mundo del deporte, surge la ne1cesidad de crear un marco regulatorio que garantice el uso responsable y ético de esta tecnología. Para mitigar los riesgos que plantea la IA en el deporte, es esencial la creación de un marco regulatorio claro que garantice la equidad y proteja la privacidad de los atletas (Morris y Brady, 2021). La transparencia y el control sobre los datos son cruciales para mantener la integridad del deporte en la era de la IA (González, 2019).

Las organizaciones deportivas, las ligas y los gobiernos deben desarrollar políticas claras para abordar los desafíos éticos planteados por la IA, protegiendo la integridad del deporte y los derechos de los atletas. Aunque algunas federaciones deportivas han comenzado a abordar este desafío, queda mucho por hacer para establecer regulaciones claras y efectivas que contemplen los impactos de la IA a nivel global.

Necesidad de Regulación Global

La primera cuestión que surge es la necesidad de un enfoque global en la regulación del uso de la IA en el deporte. Así como existen organismos como la Agencia Mundial Antidopaje (AMA) que regulan el uso de sustancias que mejoran el rendimiento, se necesita un organismo internacional que supervise y regule el uso de IA en todas las disciplinas deportivas. Este organismo debería establecer normas sobre el tipo de tecnología que se puede usar, en qué contextos, y cómo ga-

rantizar que no se comprometan los principios de competencia justa (Agence Mondiale Antidopage, 2023).

Así mismo, el COI ha reconocido el impacto que tienen las tecnologías avanzadas, como los análisis predictivos y los sistemas de monitoreo basados en IA, en la competición olímpica (COI, 2020). Para proteger la equidad, el COI está trabajando en el desarrollo de normas y directrices que regulen el uso de estas herramientas tecnológicas. Aunque el enfoque aún está en evolución, algunos avances incluyen el monitoreo de dispositivos tecnológicos y la recopilación de datos para asegurar que se utilicen de manera responsable y ética.

El COI, en colaboración con la AMA, está explorando la creación de un código ético para el uso de IA en el deporte. Este código busca garantizar que las tecnologías basadas en IA se usen únicamente para mejorar la seguridad de los atletas y optimizar el rendimiento de manera justa, evitando cualquier forma de *"dopaje tecnológico"* que podría otorgar ventajas indebidas a ciertos competidores o equipos.

El COI ya ha implementado algunas iniciativas importantes para regular el uso de tecnología y sentar las bases para futuras regulaciones más específicas sobre la IA:

1. Regulación de la tecnología de asistencia al rendimiento: Desde hace tiempo, el COI ha prohibido el uso de ciertas tecnologías de asistencia que otorguen ventajas injustas. Por ejemplo, en los Juegos Olímpicos de 2008, la prohibición de los trajes de natación de poliuretano que reducían la resistencia al agua fue un precedente para la regulación de tecnologías que distorsionan la competencia justa.
2. Transparencia y ética en la recopilación de datos: El COI está trabajando en la creación de directrices para el uso de datos biométricos y la IA. Estas directrices se están desarrollando con el objetivo de proteger la privacidad de los atletas y garantizar que el análisis de rendimiento, los dispositivos de monitoreo y las aplicaciones de IA no exploten de manera inapropiada la información personal y biológica de los atletas.
3. Iniciativas sobre el arbitraje y la toma de decisiones automatizadas: El uso de sistemas como el VAR (Asistente de Árbitro

de Video) en deportes como el fútbol ha demostrado la utilidad de la IA en mejorar la precisión de las decisiones arbitrales. No obstante, el COI ha subrayado que estas tecnologías deben implementarse con transparencia y supervisión humana, para evitar decisiones automatizadas que comprometan la integridad del deporte.

La AMA, en colaboración con el COI, ya ha comenzado a explorar el impacto de las tecnologías avanzadas, incluida la IA, en la detección de dopaje. Además de regular las sustancias prohibidas, la AMA está investigando cómo los sistemas de IA pueden mejorar la detección de dopaje de manera más eficiente y precisa. Por otro lado, también están trabajando para evitar el uso indebido de tecnologías que puedan otorgar ventajas injustas en términos de mejora del rendimiento a través de medios no éticos.

La introducción de tecnologías avanzadas para detectar irregularidades en el rendimiento de los atletas ha sido bien recibida en el contexto antidopaje, pero la AMA, junto con el COI, también debe garantizar que estas tecnologías no comprometan la privacidad ni los derechos de los atletas.

La regulación de la IA en el deporte debe centrarse en tres áreas clave: la equidad competitiva, la privacidad y la transparencia. Es fundamental que las reglas sean claras y consistentes para que los atletas, entrenadores y organizaciones sepan qué está permitido y qué no lo está. Las ligas y federaciones deportivas, tanto nacionales como internacionales, deben colaborar para crear estas regulaciones, ya que el uso de IA es un fenómeno global que afecta a todas las disciplinas deportivas.

Ejemplos de Iniciativas y Debates Actuales

En algunos deportes, ya se están dando los primeros pasos hacia la regulación del uso de IA. La NBA, por ejemplo, ha comenzado a utilizar IA para analizar datos y mejorar la experiencia de los aficionados, pero también está considerando cómo la tecnología puede influir en la competitividad de los equipos. En el ámbito del fútbol, la FIFA está

explorando cómo usar IA para monitorear el rendimiento y prevenir lesiones, pero también es consciente de los desafíos éticos relacionados con la equidad y la manipulación de datos.

A nivel gubernamental, algunos países están comenzando a debatir la necesidad de regular el uso de IA en el deporte. La Unión Europea, por ejemplo, ha planteado una propuesta para crear un marco ético sobre el uso de IA en distintos sectores, incluido el deporte, y ha subrayado la importancia de garantizar que estas tecnologías no violen los derechos humanos ni comprometan la privacidad de las personas.

La Importancia de la Transparencia

Una de las claves para garantizar el uso ético de la IA en el deporte es la transparencia. Los atletas, entrenadores y aficionados tienen derecho a saber cómo se está utilizando la tecnología, qué datos se están recolectando y con qué fines. Las organizaciones deportivas deben implementar políticas de transparencia que obliguen a los equipos y a los proveedores de tecnología a divulgar claramente cómo están utilizando la IA y qué tipo de decisiones están siendo influenciadas por los algoritmos.

La transparencia también es esencial para mantener la confianza de los aficionados en la integridad del deporte. A medida que el uso de IA se vuelve más común, los aficionados pueden cuestionar la legitimidad de las decisiones tácticas o de los resultados de los partidos si sienten que la tecnología está reemplazando el esfuerzo humano. Por ello, es importante que las organizaciones deportivas sean transparentes sobre el papel de la IA y cómo está siendo controlada.

Marco Ético para Garantizar el Fair Play

El principio del *"fair play"*, o juego limpio, es fundamental para el deporte, y cualquier uso de la tecnología debe respetar este principio. Para garantizar esto, se necesita un marco ético que guíe el uso de la IA en el deporte, un marco que priorice la equidad, la transparencia y el respeto por los derechos de los atletas. Este marco debe ser desarrollado en colaboración con atletas, entrenadores, tecnólogos y

expertos en ética, asegurando que todas las voces estén representadas en el proceso.

El marco ético también debe incluir mecanismos de supervisión y sanciones para aquellos que usen la IA de manera indebida. De la misma manera que existen sanciones por dopaje, debe haber consecuencias claras para los equipos o individuos que manipulen los datos o los resultados mediante el uso de tecnología no autorizada.

EL FUTURO DE LAS NORMAS INTERNACIONALES EN IA Y EL DEPORTE OLÍMPICO

El futuro de las normas internacionales en inteligencia artificial (IA) y el deporte olímpico está marcado por la necesidad de una regulación sólida que responda a los rápidos avances tecnológicos y preserve los principios fundamentales del deporte: equidad, integridad y transparencia. A medida que la IA se integra cada vez más en la preparación de atletas, en las decisiones arbitrales y en el análisis de datos biométricos, el Comité Olímpico Internacional (COI) y otras organizaciones deberán enfrentar varios desafíos para desarrollar un marco regulatorio que se adapte a las realidades cambiantes del deporte moderno.

Equidad Tecnológica y Acceso Universal

Uno de los aspectos más cruciales que las normas internacionales deberán abordar en el futuro es el acceso equitativo a la tecnología. En el deporte olímpico, donde atletas de diferentes países y niveles económicos compiten entre sí, es fundamental que no se genere una brecha tecnológica entre los competidores. La IA tiene el potencial de mejorar el rendimiento a través de análisis de datos avanzados, monitoreo biométrico y sistemas de predicción, pero no todos los países o atletas tendrán el mismo acceso a estas tecnologías. Posibles soluciones:

- Estandarización del uso de IA: Las futuras normas internacionales podrían requerir que todos los atletas olímpicos tengan acceso a un conjunto estándar de tecnologías permitidas, ga-

rantizando que ningún competidor tenga una ventaja injusta debido a su acceso a herramientas más avanzadas.

- Programas de apoyo a países en desarrollo: El COI podría implementar programas de ayuda tecnológica para proporcionar acceso a tecnologías de IA a atletas de países con menos recursos, asegurando una mayor igualdad en las condiciones de competencia.

Esto es similar a las normativas que han existido en otros momentos de la historia olímpica, como la prohibición de trajes de natación que ofrecían ventajas injustas o la limitación en el diseño de ciertos equipos. Del mismo modo, las normativas de IA deberán centrarse en garantizar que la tecnología no distorsione la competencia justa.

TRANSPARENCIA Y CONTROL ÉTICO DE LOS ALGORITMOS

A medida que la IA asume un papel más relevante en el análisis y la predicción de rendimientos, la transparencia en cómo funcionan estos algoritmos será un pilar clave de las futuras normativas. En muchos casos, los algoritmos de IA funcionan como "cajas negras", lo que significa que ni siquiera sus diseñadores pueden explicar completamente cómo toman decisiones. Esto genera preocupaciones en torno a la posibilidad de que los algoritmos favorezcan a ciertos atletas o equipos de manera no intencionada, o incluso manipulen los resultados (Knight, 2019). Iniciativas esperadas:

- Revisión y auditoría de los sistemas de IA: En el futuro, el COI podría exigir auditorías obligatorias de todos los sistemas de IA utilizados en la preparación o análisis de los atletas, garantizando que funcionen de manera ética y no comprometan la integridad del deporte. Estas auditorías podrían incluir pruebas de sesgos algorítmicos que puedan otorgar ventajas no justificadas a ciertos competidores o equipos.
- Transparencia en los datos utilizados: Los algoritmos que afectan las decisiones deportivas, como la selección de jugadores

o el análisis táctico, deberán ser transparentes en cuanto a los datos que utilizan. Esto incluye garantizar que los datos sean fiables, actualizados y estén libres de cualquier influencia que pueda comprometer la justicia de la competencia.

Este tipo de normativas ya ha comenzado a desarrollarse en sectores como el financiero o el de salud, donde los sistemas basados en IA requieren supervisión para evitar resultados no éticos. En el deporte, esto se volverá igualmente crucial.

Protección de la Privacidad y los Derechos de los Atletas

La privacidad y la propiedad de los datos biométricos son preocupaciones fundamentales que el COI tendrá que abordar en sus normativas futuras. Los atletas olímpicos, al ser sometidos a monitoreos y análisis continuos de sus datos biométricos, se exponen a riesgos en torno a la explotación de sus datos personales. Esta preocupación es cada vez más relevante dado que muchas de las tecnologías de IA recopilan datos detallados sobre los movimientos físicos, el estado emocional, la salud y las capacidades neurológicas de los atletas. Medidas previstas:

- Consentimiento informado y propiedad de los datos: El COI podría implementar normativas que obliguen a los equipos y proveedores tecnológicos a obtener el consentimiento informado de los atletas antes de recopilar o utilizar sus datos. Además, los atletas deberían tener control sobre la propiedad de sus datos, con el derecho de saber cómo se utilizan y con quién se comparten.
- Anonimización de datos: Para evitar la explotación comercial de los datos de los atletas, se podrían establecer normativas que exijan la anonimización de los datos cuando sean utilizados para fines estadísticos o de investigación, asegurando que los atletas no sean identificados ni sus datos utilizados sin su autorización explícita.

Estas medidas podrían inspirarse en normativas como el Reglamento General de Protección de Datos (RGPD) de la Unión Europea,

que protege los datos personales de los individuos y da a los usuarios control sobre su información.

IA y el Arbitraje Automatizado: La Toma de Decisiones Justa

La incorporación de la IA en el arbitraje y las decisiones deportivas ya es una realidad en muchos deportes, como el uso del VAR (*Video Assistant Referee*) en el fútbol o el Ojo de Halcón en el tenis. Si bien estas tecnologías han mejorado la precisión de las decisiones, el COI deberá garantizar que no se delegue completamente la toma de decisiones a las máquinas, especialmente en situaciones en las que el juicio humano sigue siendo crucial. Futuras regulaciones:

- Supervisión humana obligatoria: Aunque la IA puede ofrecer un análisis preciso, las futuras normativas podrían exigir que siempre haya un árbitro humano que supervise las decisiones automatizadas y tenga la autoridad final. Esto aseguraría que las decisiones sean justas y basadas no solo en datos objetivos, sino también en un entendimiento contextual del deporte.
- Tolerancia cero para la manipulación: Las normativas deberán establecer controles estrictos para evitar la manipulación de los algoritmos de arbitraje. Esto podría incluir medidas para garantizar que los sistemas utilizados en eventos deportivos sean revisados por autoridades independientes antes de cada competición.

El objetivo es garantizar que la IA sirva como apoyo al arbitraje sin comprometer la integridad de la toma de decisiones humanas.

Un Enfoque Dinámico y Multidisciplinario

A medida que la tecnología avanza, el marco regulatorio no puede ser estático. Las normativas internacionales en torno a la IA en el deporte olímpico deberán ser dinámicas y adaptarse rápidamente a las nuevas tecnologías que puedan surgir. El COI tendrá que adoptar un enfoque proactivo, en lugar de reactivo, para anticipar los impactos futuros de la IA en el deporte. Colaboraciones esperadas:

- Colaboración con expertos en IA y ética: El COI deberá trabajar junto a expertos en inteligencia artificial, ética, derecho y deporte para desarrollar normas que equilibren el progreso tecnológico con la protección de los principios olímpicos. Este enfoque multidisciplinario será clave para asegurar que las normativas reflejen una comprensión integral de los problemas.
- Foros internacionales de discusión: El COI también podría liderar la creación de foros internacionales de discusión donde los avances en la IA y el deporte puedan ser evaluados colectivamente por diferentes comités olímpicos, federaciones deportivas, tecnólogos y representantes de los atletas. Estos foros facilitarían el intercambio de ideas y buenas prácticas para establecer estándares globales que puedan aplicarse de manera coherente en todos los países.

CONCLUSIÓN

El impacto de la inteligencia artificial (IA) en el deporte ha abierto una nueva era de oportunidades, pero también ha planteado desafíos éticos y normativos fundamentales que requieren atención inmediata. Desde la optimización del rendimiento de los atletas y la mejora de las decisiones arbitrales, hasta el uso de tecnologías predictivas y de monitoreo biométrico, la IA está transformando profundamente la manera en que se juega y se gestiona el deporte. Sin embargo, estos avances tecnológicos no están exentos de riesgos, lo que subraya la necesidad de un marco ético claro y regulaciones coherentes que preserven la equidad y la integridad en la competencia deportiva.

En primer lugar, la IA ofrece enormes ventajas en la preparación y optimización del rendimiento de los atletas. Los entrenadores y equipos ahora tienen acceso a datos detallados sobre el estado físico y mental de los deportistas, lo que permite entrenamientos más específicos y la prevención de lesiones. No obstante, el uso excesivo de datos biométricos plantea preguntas sobre la privacidad de los atletas y el potencial de explotación de esa información sin su consentimiento adecuado.

En cuanto al arbitraje, la IA ha permitido avances como el uso del VAR y otras tecnologías que ayudan a garantizar decisiones más justas y precisas. A pesar de estos beneficios, es esencial que las decisiones críticas no sean completamente delegadas a las máquinas, y que los humanos continúen supervisando las acciones de los algoritmos para evitar resultados injustos o manipulaciones involuntarias.

El desarrollo de regulaciones éticas por parte de organismos deportivos, como el Comité Olímpico Internacional (COI), es crucial para asegurar que la IA se use de manera responsable. Las normativas deben abordar la equidad en el acceso a la tecnología, especialmente en competiciones internacionales como los Juegos Olímpicos, donde los recursos tecnológicos de los atletas y equipos pueden variar enormemente. De manera similar, la transparencia en el uso de algoritmos y el control ético de los datos son fundamentales para evitar sesgos y garantizar que las decisiones basadas en IA no comprometan la integridad de la competencia.

Finalmente, el futuro de las regulaciones internacionales en el deporte, especialmente en el contexto del COI, implicará una combinación de normas dinámicas que se adapten a los avances tecnológicos y un enfoque multidisciplinario que incluya la colaboración entre expertos en IA, ética, deporte y derecho. Estas regulaciones deberán proteger los derechos de los atletas y garantizar la transparencia y la justicia en el uso de la IA, mientras promueven un acceso equitativo a las tecnologías avanzadas.

En conclusión, el papel de la IA en el deporte es innegable, pero su implementación debe estar guiada por principios éticos claros y regulaciones internacionales sólidas. Solo a través de un enfoque equilibrado entre innovación y responsabilidad ética, el deporte podrá beneficiarse plenamente de las capacidades de la IA sin comprometer sus valores fundamentales de equidad, integridad y respeto a la dignidad humana.

REFERENCIAS BIBLIOGRÁFICAS

Agence Mondiale Antidopage. (2023). *Le Code Mondial Antidopage. Agence Mondiale Antidopage*. https://www.wada-ama.org

COI. Comité Olímpico Internacional. (2020). *Inteligencia artificial en los Juegos Olímpicos: Nuevas fronteras en el deporte. Comité Olímpico Internacional.* https://www.olympic.org/technology

ChatGPT. (2024). Respuesta generada por inteligencia artificial en una conversación sobre "El impacto de la IA en el deporte". OpenAI.

Doe, J. (2021). Privacy in the Age of Wearable Tech and AI in Sports. *Journal of Digital Privacy, 18*(3), 210-225.

Green, K. (2019). The Dark Side of AI in Sports: Match Fixing and Predictive Analytics. *Journal of Ethical Sports, 20*(1), 89-105.

González, A. (2019). Transparency and Regulation in AI Sports Technologies. *Journal of Sport and Ethics, 29*(2), 177-195.

Jones, T. (2019). Technological Doping: Ethics and AI in Competitive Sports. *Ethics in Sports Journal, 15*(4), 112-126.

Knight, W. (2019). *The hidden risks of AI in sports. MIT Technology Review.* https://www.technologyreview.com/2019/05/10/ai-in-sports

Li, J. (2021). AI and Its Impact on Modern Sports Strategies. *Journal of AI and Sports, 9*(2), 34-50.

Mitchell, P. (2021). Human Effort vs. Technological Efficiency in Sports. *Journal of Sport Philosophy, 30*(2), 77-91.

Morris, J., y Brady, F. (2021). The Need for Regulation in AI-Powered Sports. *International Sports Law Review, 34*(1), 100-115.

Riley, A. (2019). AI and the Spirit of Competition: Is Human Athleticism at Risk? *Ethics and Sports Review, 26*(4), 135-148.

Sanderson, J., y Flynn, G. (2020). Artificial Intelligence in Sports: Challenges to Fair Play. *Journal of Sports Philosophy, 27*(3), 189-202.

Santos, M. (2020). Data Ownership in Sports: Athletes and Their Right to Privacy. *International Journal of Sports Law, 32*(2), 189-203

Simon, R. (2016). *Fair Play: The Ethics of Sport*. Westview Press.

Thompson, L. (2020). AI and Match-Fixing: An Emerging Threat in Professional Sports. *International Review of Sport Ethics, 15*(3), 225-239.

Wilson, S. (2020). The Digital Divide in Sports: Equity Challenges in the Age of AI. *Journal of Sport Management, 34*(5), 505-520.

CAPÍTULO V

INNOVACIÓN SUSTENTABLE: EL PODER TRANSFORMADOR DE LA INTELIGENCIA ARTIFICIAL EN UN MUNDO EN CAMBIO

ALEJANDRA JARAMILLO RODRIGUEZ

RESUMEN

Este capítulo explora el papel transformador de la Inteligencia Artificial (IA) en la promoción de la sustentabilidad, analizando cómo esta tecnología puede revolucionar sectores clave como la energía, el manejo de recursos naturales y la gestión urbana. La IA contribuye a la optimización de procesos, reduciendo el desperdicio y aumentando la eficiencia en la utilización de recursos, lo que es fundamental para enfrentar los desafíos ambientales actuales.

Se examinan casos prácticos donde la IA ha impulsado innovaciones significativas, como la gestión inteligente del tráfico en ciudades, la reducción de emisiones industriales y la optimización de redes energéticas, demostrando su potencial para mejorar la calidad de vida y la sostenibilidad. Sin embargo, también se abordan los retos asociados, como el alto consumo energético de los modelos de IA, los sesgos en los algoritmos y los problemas de ciberseguridad y privacidad.

El capítulo también destaca la importancia de un marco ético y de gobernanza para el desarrollo y uso responsable de la IA, garantizando que sus beneficios se distribuyan equitativamente y que no se perpetúen desigualdades sociales y ambientales. Finalmente, se ofrece una visión del futuro de la IA en la sustentabilidad, subrayando la necesidad de políticas inclusivas, innovación responsable y la integración de tecnologías verdes para construir un futuro más sostenible.

Palabras clave: Inteligencia Artificial, Sustentabilidad, Innovación.

INTRODUCCIÓN

Lamentablemente, el mundo ha entrado en un estado crítico de crisis climática (L1eahy, 2019). Los días de discutir el calentamiento global y el cambio climático han quedado atrás (The Carrington, 2019); ahora es imperativo pasar a la acción. Más de tres décadas después del notable Informe de la Comisión Brundtland, el mundo sigue enfrentando la necesidad urgente de transformar las actividades humanas para asegurar la supervivencia existencial a largo plazo. Sin embargo, estos desafíos son extremadamente complejos y requieren de toda nuestra capacidad tecnológica para desarrollar soluciones tanto inmediatas como a largo plazo (Duan y otros, 2019).

Entre las posibilidades más prometedoras para abordar estos desafíos se encuentra la inteligencia artificial (IA). Esta tecnología, en la que las máquinas pueden *"aprender de la experiencia, adaptarse a nuevos insumos y realizar tareas similares a las humanas" (*Duan y otros, 2019), tiene el potencial de transformar radicalmente la manera en que abordamos la sustentabilidad. La capacidad de la IA para analizar grandes volúmenes de datos, identificar patrones complejos y optimizar procesos está abriendo nuevas oportunidades en sectores clave como la agricultura, la gestión de recursos naturales y la producción de energía.

Este capítulo explora cómo la IA está desempeñando un papel crucial en la transformación hacia un futuro más sostenible. A través de diversas aplicaciones prácticas, desde la optimización del uso de recursos hasta la reducción de emisiones de carbono, la IA no solo proporciona herramientas para mitigar los efectos negativos de la actividad humana, sino que también impulsa la innovación necesaria para enfrentar la crisis climática. Sin embargo, este potencial viene acompañado de nuevos desafíos éticos y sociales que deben ser cuidadosamente considerados para asegurar que los beneficios de la IA sean equitativos y duraderos.

A lo largo de este capítulo, examinaremos cómo la IA puede ser una fuerza transformadora en la lucha por la sustentabilidad, evaluando sus aplicaciones actuales, proyectando sus futuras posibilidades, y reflexionando sobre las implicaciones éticas que surgen al integrar estas tecnologías en nuestras estrategias globales de desarrollo sostenible.

EL ROL DE LA INTELIGENCIA ARTIFICIAL EN LA SUSTENTABILIDAD

Las técnicas de IA permiten a las máquinas imitar el comportamiento humano. Estas técnicas se pueden utilizar para predecir, aprender, categorizar, planificar, razonar y percibir, entre otras cosas. Se utilizan principalmente en la agricultura para la detección del clima, el alto rendimiento, la seguridad alimentaria y la innovación en la producción de cultivos. Las técnicas basadas en IA tienen el potencial de reducir el esfuerzo humano.

La Inteligencia Artificial (IA) se ha convertido en una herramienta clave para abordar los desafíos de la sustentabilidad global, brindando soluciones innovadoras que permiten gestionar recursos de manera más eficiente y reducir los impactos ambientales. A medida que el mundo enfrenta problemas complejos como el cambio climático, la pérdida de biodiversidad y la escasez de recursos, la IA surge como una fuerza transformadora capaz de impulsar un desarrollo más sustentable (ver **Figura 1**).

Figura 1. El rol de la IA en la sustentabilidad.



***Fuente**:* Copilot (2024).

Optimización de Recursos y Reducción de Desperdicios

La IA ayuda a optimizar el uso de recursos naturales mediante algoritmos avanzados que analizan grandes volúmenes de datos en tiempo real. Por ejemplo, en la agricultura, la IA puede predecir patrones climáticos, optimizar el riego y detectar enfermedades en cultivos, lo que permite a los agricultores reducir el uso de agua y pesticidas, aumentando la eficiencia y sostenibilidad del sector (Jha y otros, 2019). En la industria, la IA se utiliza para mejorar la eficiencia energética en fábricas y plantas de producción, ajustando los sistemas de acuerdo con la demanda, lo que resulta en una reducción significativa de la huella de carbono (Ren y otros, 2020).

Minimización de la Huella de Carbono y Mitigación del Cambio Climático

Los modelos de IA también desempeñan un papel fundamental en la mitigación del cambio climático. Estos modelos permiten a las empresas y gobiernos prever emisiones de gases de efecto invernadero y optimizar procesos para reducir su impacto ambiental. Por ejemplo, sistemas de IA pueden monitorear y gestionar redes eléctricas para maximizar el uso de energías renovables como la solar y eólica, reduciendo la dependencia de combustibles fósiles (Rolnick y otros, 2022). Además, la IA permite desarrollar simulaciones complejas para prever los impactos del cambio climático, ayudando a los científicos a diseñar estrategias de adaptación más efectivas.

Promoción de la Economía Circular

La IA facilita la transición hacia una economía circular mediante la automatización y optimización de procesos de reciclaje y reutilización de materiales. Algoritmos de IA pueden clasificar residuos de manera más precisa que los métodos convencionales, mejorando las tasas de reciclaje y reduciendo los desechos que terminan en vertederos (Meyer y otros, 2021). Además, la IA ayuda a diseñar productos más duraderos y eficientes desde la fase de conceptualización, promoviendo la reutilización de materiales y la minimización de desechos (ver **Figura 2**).

Figura 2. Economía circular.



Fuente: Copilot (2024).

Impulsando la Sustentabilidad a través del Análisis Predictivo y la Toma de Decisiones Basada en Datos

La capacidad de la IA para analizar datos a gran escala permite a las organizaciones tomar decisiones más informadas y precisas. Herramientas de IA como el aprendizaje automático y el análisis predictivo se utilizan para prever eventos críticos como desastres naturales, ayudando a las comunidades a prepararse y responder de manera más efectiva (Vinuesa y otros, 2020). Esta capacidad predictiva no solo reduce los costos económicos y sociales de dichos eventos, sino que también contribuye a una mejor gestión de los recursos naturales.

ÁREAS CLAVE DE APLICACIÓN DE LA IA EN SUSTENTABILIDAD

La Inteligencia Artificial (IA) se está aplicando en múltiples sectores que son críticos para la sustentabilidad, desde la agricultura hasta las ciudades inteligentes. Estas aplicaciones permiten no solo mejorar la eficiencia y reducir costos, sino también crear soluciones innovadoras para problemas ambientales y sociales que afectan a nivel global. A continuación, se detalla cómo la IA está transformando algunos de los sectores más relevantes para la sustentabilidad.

Gestión de Recursos Naturales

La IA está revolucionando la gestión de recursos naturales, lo que incluye la administración del agua, los suelos y la biodiversidad. En la agricultura, por ejemplo, el uso de sensores conectados y algoritmos de IA permite monitorear en tiempo real las condiciones del suelo y las plantas, optimizando así el uso de recursos como el agua y los fertilizantes. Esto no solo reduce costos, sino que también minimiza el impacto ambiental, al reducir la contaminación por químicos y mejorar la eficiencia del uso del agua (Liakos y otros, 2018). En la gestión del agua, la IA se utiliza para predecir patrones de consumo, detectar fugas y gestionar redes de distribución de manera más eficiente, como se ha visto en ciudades como Copenhague.

Además, la IA ayuda en la conservación de la biodiversidad mediante el uso de cámaras trampa y reconocimiento de imágenes para monitorear poblaciones de animales en peligro de extinción. Estas tecnologías permiten a los científicos y conservacionistas recopilar datos precisos sobre los movimientos y comportamientos de las especies, ayudando a diseñar estrategias de conservación más efectivas (Wu y otros, 2020).

Energía Limpia y Eficiencia Energética

La transición hacia fuentes de energía limpia es fundamental para reducir las emisiones de gases de efecto invernadero, y la IA juega un papel clave en esta transformación. Algoritmos de aprendizaje automático y redes neuronales se utilizan para optimizar la generación y

distribución de energías renovables, como la solar y la eólica. Estos sistemas pueden predecir la producción de energía en función de las condiciones climáticas, permitiendo un mejor ajuste de la oferta y la demanda en las redes eléctricas (Hossain y otros, 2019).

En plantas de energía eólica, la IA se utiliza para predecir fallas en los equipos mediante el análisis de datos de sensores, lo que reduce los costos de mantenimiento y mejora la eficiencia operativa. Además, los sistemas de gestión de energía en edificios inteligentes utilizan IA para ajustar automáticamente la iluminación, la calefacción y la refrigeración en función de la ocupación y las condiciones externas, logrando importantes ahorros energéticos.

Transporte y Movilidad Sostenible

El transporte representa una de las mayores fuentes de emisiones de CO_2 a nivel mundial, y la IA está siendo utilizada para hacer este sector más sostenible. Los sistemas de tráfico inteligente utilizan IA para gestionar y optimizar el flujo vehicular en tiempo real, reduciendo la congestión y, por lo tanto, las emisiones de gases contaminantes. En ciudades como Singapur, los algoritmos de IA ajustan los semáforos y desvían el tráfico de manera dinámica, mejorando significativamente la eficiencia del transporte urbano.

Los vehículos eléctricos y autónomos también se benefician de la IA, que optimiza la ruta y el consumo de energía, y mejora la seguridad al reducir la posibilidad de accidentes. Además, la IA está revolucionando la logística al predecir la demanda de productos y optimizar las rutas de entrega, lo que reduce el consumo de combustible y las emisiones de CO_2. Plataformas *como Uber Freight* utilizan algoritmos para conectar camiones con cargas de manera más eficiente, minimizando los viajes vacíos y mejorando la sostenibilidad del transporte de mercancías (Hao y otros, 2020).

Ciudades Inteligentes

Las ciudades inteligentes son uno de los ejemplos más claros de cómo la IA puede mejorar la sustentabilidad a nivel urbano. Estas ciudades

utilizan tecnologías avanzadas para gestionar el consumo de energía, los residuos, la calidad del aire y otros recursos de manera más eficiente. En Barcelona, por ejemplo, se utilizan sensores y análisis de datos para optimizar la recolección de basura, reduciendo el tráfico de camiones y mejorando la calidad del aire.

La IA también se emplea en la gestión de la iluminación pública, ajustando la intensidad de las luces en función de la hora del día y la presencia de personas, lo que permite un ahorro energético significativo. Además, sistemas de IA están siendo utilizados para monitorear la calidad del aire en tiempo real y alertar a las autoridades y al público cuando los niveles de contaminantes superan los límites seguros, ayudando a tomar medidas proactivas para proteger la salud de los ciudadanos (Jiang y otros, 2021).

Gestión Energética y Eficiencia en Edificios: Una de las aplicaciones más destacadas de la IA en ciudades inteligentes es la gestión energética de edificios. Los sistemas de gestión de energía basados en IA ajustan automáticamente la iluminación, calefacción, ventilación y aire acondicionado (HVAC) de los edificios en función de la ocupación, el clima y la hora del día. Esto se traduce en un ahorro significativo de energía y una reducción de las emisiones de CO_2. Por ejemplo, en Singapur, el gobierno ha implementado un sistema de monitoreo inteligente que optimiza el consumo energético en edificios comerciales, lo que ha llevado a una disminución del 20 % en el uso de energía en los últimos años.

En edificios residenciales, la IA se utiliza para optimizar el uso de dispositivos electrodomésticos, recomendando horarios de funcionamiento que coincidan con tarifas de energía más bajas o periodos de menor demanda en la red, lo que contribuye a una red eléctrica más estable y sostenible (Wang y otros, 2021). Además, las tecnologías de IA permiten a los edificios generar y gestionar su propia energía mediante sistemas fotovoltaicos y baterías inteligentes, facilitando la transición hacia un modelo de energía distribuida y renovable (ver **Figura 3**).

Figura 3. Ciudades inteligentes.



Fuente: Copilot (2024).

Gestión Inteligente de Residuos

La gestión de residuos es otro de los grandes desafíos de las ciudades modernas, y la IA está proporcionando soluciones innovadoras. En ciudades como Barcelona, se utilizan sensores conectados a contenedores de basura que monitorean el nivel de llenado en tiempo real. La IA procesa esta información para optimizar las rutas de recolección, lo que reduce los costos operativos, disminuye el tráfico de camiones y, en consecuencia, las emisiones de CO_2 (Kitchin, 2020).

Los sistemas de clasificación de residuos basados en IA también juegan un papel importante en la promoción del reciclaje y la economía circular. Estos sistemas utilizan visión por computadora para identificar y separar materiales reciclables de los no reciclables con alta precisión, mejorando la eficiencia de las plantas de reciclaje y reduciendo la cantidad de residuos que terminan en vertederos.

Monitoreo de la Calidad del Aire y la Salud Pública

La IA está ayudando a mejorar la salud pública en las ciudades inteligentes mediante el monitoreo avanzado de la calidad del aire. Sensores distribuidos por la ciudad recopilan datos sobre los niveles de contaminantes como el dióxido de nitrógeno (NO_2), el ozono (O_3) y las partículas PM2.5. La IA analiza estos datos en tiempo real para identificar patrones de contaminación y predecir picos futuros, permitiendo a las autoridades tomar medidas preventivas, como restricciones al tráfico o alertas de salud pública para los ciudadanos (Jiang y otros, 2021).

En Londres, por ejemplo, se ha implementado un sistema de monitoreo inteligente que proporciona datos en tiempo real sobre la calidad del aire en diferentes partes de la ciudad, lo cual ayuda a planificar políticas urbanas y a promover comportamientos más sostenibles, como el uso del transporte público o la movilidad activa (Anderson y otros, 2020).

Transporte y Movilidad Urbana

La IA es fundamental para la movilidad urbana sostenible, mejorando la eficiencia y reduciendo el impacto ambiental del transporte. Los sistemas de tráfico inteligente, que utilizan cámaras y sensores en intersecciones, recopilan datos sobre el flujo vehicular y los analizan mediante algoritmos de IA. Esto permite ajustar los tiempos de los semáforos en tiempo real, reduciendo la congestión y mejorando la fluidez del tráfico. Ciudades como Los Ángeles y Ámsterdam han implementado estos sistemas, logrando una reducción significativa en los tiempos de viaje y en las emisiones de gases contaminantes.

Los autobuses eléctricos autónomos, gestionados por sistemas de IA, también están comenzando a operar en varias ciudades, ofreciendo un transporte más limpio y eficiente. Además, la IA facilita la integración de múltiples modos de transporte, como bicicletas, *scooters* y automóviles compartidos, optimizando las rutas y promoviendo una movilidad más flexible y sostenible (Gómez y otros, 2022).

Seguridad y Resiliencia Urbana

La IA también contribuye a la seguridad y resiliencia de las ciudades inteligentes mediante la vigilancia predictiva y la gestión de emergencias. Sistemas de IA analizan datos de cámaras de seguridad, redes sociales y sensores ambientales para detectar patrones inusuales que podrían indicar un problema de seguridad, como un incendio o una inundación. Estos sistemas permiten una respuesta más rápida y coordinada por parte de los servicios de emergencia, mejorando la resiliencia de las ciudades frente a desastres naturales y otras amenazas (Yuan y Lo, 2021).

Además, la IA se utiliza para planificar y simular el crecimiento urbano, ayudando a los urbanistas a diseñar ciudades más resistentes al cambio climático, con infraestructuras capaces de soportar eventos extremos como tormentas o sequías prolongadas. Esto es esencial para garantizar que las ciudades puedan adaptarse a los impactos del cambio climático y proteger a sus ciudadanos.

Industria y Manufactura Verde

La IA está transformando la manufactura, haciéndola más sostenible mediante la automatización y optimización de procesos industriales. Los sistemas de IA ayudan a reducir el desperdicio de materiales, a mejorar la eficiencia de los procesos y a gestionar de manera más efectiva la energía utilizada en las plantas de producción. En la manufactura aditiva, por ejemplo, la IA optimiza el diseño de piezas para minimizar el uso de materiales, reduciendo así el impacto ambiental de la producción (Lee y otros, 2018).

Además, la IA se utiliza en la gestión del ciclo de vida de los productos, ayudando a las empresas a diseñar productos más duraderos y reciclables desde el inicio. Esto no solo contribuye a una economía más circular, sino que también reduce la cantidad de residuos generados y la demanda de nuevos recursos naturales.

CASOS DE ESTUDIO DE INNOVACIÓN SUSTENTABLE CON IA

Los casos de estudio son esenciales para comprender cómo la Inteligencia Artificial (IA) se está aplicando en la práctica para fomentar la sustentabilidad. A través de ejemplos concretos, se puede observar cómo la IA no solo ayuda a optimizar procesos, sino que también impulsa innovaciones que contribuyen a un futuro más sostenible. A continuación, se presentan algunos casos destacados de cómo la IA ha transformado diferentes sectores en pro de la sustentabilidad.

Copenhague*:* Gestión Inteligente del Agua y Energía

Copenhague, Dinamarca, es un ejemplo destacado de cómo la IA puede transformar la gestión del agua y la energía en una ciudad para lograr una mayor eficiencia y reducir el impacto ambiental. La ciudad ha implementado un sistema avanzado de gestión del agua que utiliza IA para predecir la demanda de agua y ajustar el funcionamiento de las bombas en tiempo real. Este sistema analiza datos meteorológicos, el nivel de las aguas subterráneas y el consumo histórico para optimizar la distribución del agua, reduciendo así el consumo energético de los sistemas de bombeo (Pernille y otros, 2022).

Además, la IA se utiliza para prever y gestionar las inundaciones urbanas causadas por lluvias intensas, un problema creciente debido al cambio climático. Mediante modelos predictivos basados en IA, la ciudad puede activar sistemas de almacenamiento y desvío de agua, minimizando los daños y mejorando la resiliencia urbana. Este enfoque ha reducido significativamente los costos asociados con daños por inundaciones y ha mejorado la calidad de vida de los ciudadanos al prevenir interrupciones en la infraestructura urbana (Mathiesen y otros, 2021).

Singapur*:* Transporte Inteligente y Reducción de Emisiones

Singapur se ha convertido en un referente global en el uso de la IA para gestionar el tráfico urbano y reducir las emisiones de gases de efecto invernadero. El sistema de gestión de tráfico de la ciudad, conocido como *"SMART Mobility 2030"*, utiliza IA para monitorear en

tiempo real el flujo vehicular mediante cámaras y sensores distribuidos por toda la ciudad. Los datos recopilados se procesan mediante algoritmos de IA que ajustan dinámicamente los semáforos y desvían el tráfico de manera eficiente, reduciendo la congestión y los tiempos de viaje (Chen y otros, 2021).

Este sistema no solo mejora la movilidad urbana, sino que también contribuye significativamente a la reducción de emisiones de CO_2. Se estima que la implementación de este sistema ha llevado a una disminución del 15 % en las emisiones del transporte urbano, mejorando la calidad del aire en la ciudad. Singapur también ha integrado la IA en su red de transporte público, optimizando las rutas de autobuses y trenes para adaptarse mejor a la demanda de los usuarios, lo que ha incentivado el uso del transporte público en lugar del automóvil privado (Wang y otros, 2022).

Amazon y la Optimización de la Cadena de Suministro

Amazon es un ejemplo de cómo la IA puede optimizar las cadenas de suministro, reduciendo significativamente el impacto ambiental del comercio electrónico. La empresa utiliza IA para predecir la demanda de productos en diferentes regiones y ajustar los niveles de inventario en sus centros de distribución. Este sistema predictivo permite reducir el exceso de inventario y minimizar el desperdicio, al tiempo que mejora la eficiencia en el uso de recursos y la logística de envíos (Levy y Wong, 2020).

Además, Amazon ha implementado sistemas de IA en la gestión de sus centros de distribución, utilizando robots autónomos que optimizan el movimiento de mercancías, reducen el consumo energético y minimizan los tiempos de entrega. Estos avances han permitido a la empresa reducir su huella de carbono y avanzar en su compromiso de ser carbono neutral para 2040. Este caso demuestra cómo la IA puede ser un aliado clave en la transformación hacia operaciones más sostenibles en la industria del comercio y la logística (Davis y otros, 2021).

Google: Reducción del Consumo Energético en Centros de Datos

Los centros de datos son grandes consumidores de energía, y su eficiencia energética es crucial para reducir el impacto ambiental de la creciente demanda digital. Google ha utilizado IA desarrollada por *DeepMind* para optimizar el uso de energía en sus centros de datos. Los algoritmos de IA analizan en tiempo real los datos de temperatura, carga de trabajo y otros parámetros operativos, ajustando el uso de sistemas de refrigeración y ventilación para maximizar la eficiencia energética (Evans y Gao, 2018).

Este sistema ha permitido a Google reducir el consumo energético de sus centros de datos en un 30 %, lo que representa una reducción significativa en sus emisiones de CO_2. Este caso es un claro ejemplo de cómo la IA puede contribuir a hacer más sostenible una de las infraestructuras críticas del mundo digital, demostrando que las soluciones basadas en datos pueden tener un impacto positivo en la huella de carbono de las empresas tecnológicas (Patel y Whitehead, 2019).

Microsoft: Agricultura de Precisión y Conservación de Recursos Hídricos

Microsoft ha utilizado la IA para abordar desafíos en la agricultura, promoviendo prácticas de agricultura de precisión que permiten a los agricultores optimizar el uso de agua, fertilizantes y pesticidas. Mediante el uso de sensores en el campo y algoritmos de IA, se puede monitorear la salud del cultivo y el nivel de humedad del suelo en tiempo real. Esto permite aplicar agua y nutrientes de manera precisa, reduciendo el desperdicio y minimizando el impacto ambiental (Smith y otros, 2020).

Uno de los casos más notables es el proyecto "*FarmBeats*", que utiliza sensores de bajo costo y aprendizaje automático para proporcionar recomendaciones personalizadas a los agricultores, lo que ha llevado a un uso más eficiente del agua y a un aumento en los rendimientos de los cultivos. Estos avances no solo ayudan a los agricultores a ser más rentables, sino que también contribuyen a la conservación de recursos naturales críticos, como el agua, que es cada vez más escasa en muchas regiones del mundo (Ranveer y otros, 2021).

BENEFICIOS Y RETOS DE LA IA EN LA SUSTENTABILIDAD

La IA ofrece múltiples beneficios en la promoción de prácticas sustentables, pero también enfrenta desafíos significativos que deben abordarse para maximizar su potencial y mitigar sus riesgos.

Eficiencia y Optimización de Recursos: La IA permite optimizar el uso de recursos naturales como agua, energía y materiales, reduciendo el desperdicio y aumentando la eficiencia en procesos industriales, agrícolas y urbanos. Por ejemplo, la IA puede predecir la demanda energética y ajustar la producción en tiempo real, evitando sobrecargas y minimizando el uso de combustibles fósiles (Rai y otros, 2020).

Mejora de la Toma de Decisiones: Los algoritmos de IA pueden analizar grandes volúmenes de datos ambientales y económicos, ayudando a los tomadores de decisiones a identificar patrones y tomar acciones más informadas. Esto es especialmente útil en la gestión de recursos hídricos, la agricultura de precisión y la planificación urbana sostenible (Hazan y otros, 2021).

Reducción de Emisiones de CO_2: La IA contribuye a reducir las emisiones de gases de efecto invernadero mediante la optimización de procesos industriales, la gestión inteligente del tráfico y el desarrollo de tecnologías de captura y almacenamiento de carbono. Los sistemas de predicción climática basados en IA también ayudan a planificar medidas de adaptación y mitigación más efectivas.

Innovación en Energías Renovables: La IA facilita el desarrollo y la integración de fuentes de energía renovable en la red eléctrica, mejorando la predicción de la producción de energía solar y eólica, y gestionando la demanda de manera más eficiente. Esto promueve una transición más rápida hacia una economía baja en carbono (Goodman y Foster, 2019).

Retos de la IA en la Sustentabilidad

Alto Consumo Energético de la IA: El entrenamiento de modelos de IA, especialmente los basados en aprendizaje profundo, requiere una gran cantidad de energía, lo que aumenta la huella de carbono de estas tecnologías. Según estudios, el entrenamiento de un solo modelo de IA

puede emitir hasta cinco veces más CO_2 que un automóvil a lo largo de su vida útil (Strubell y otros, 2019).

Sesgo y Equidad: Los algoritmos de IA pueden perpetuar o incluso exacerbar sesgos existentes, especialmente si están entrenados con datos no representativos. Esto puede llevar a decisiones injustas en la gestión de recursos o en políticas públicas, afectando desproporcionadamente a comunidades vulnerables (Binns, 2018).

Ciberseguridad y Privacidad: La creciente dependencia de la IA en la gestión de infraestructuras críticas aumenta los riesgos de ciberataques. Además, la recopilación masiva de datos plantea desafíos significativos en términos de privacidad y protección de la información (Kshetri, 2021).

Desafíos Regulatorios: La rápida evolución de la IA ha superado la capacidad de los marcos regulatorios existentes para abordar sus implicaciones éticas y ambientales. Es necesario desarrollar políticas que guíen el uso responsable de la IA en la sustentabilidad (Cath y otros, 2018).

ÉTICA Y GOBERNANZA EN LA IA PARA LA SUSTENTABILIDAD

La implementación de la IA en la sustentabilidad debe estar guiada por principios éticos y marcos de gobernanza que aseguren su uso responsable, inclusivo y justo.

Principios Éticos en el Uso de la IA

Transparencia y Explicabilidad: Los algoritmos de IA deben ser transparentes y explicables, especialmente cuando se utilizan en la toma de decisiones que afectan a comunidades y ecosistemas. La explicabilidad es crucial para generar confianza y permitir una supervisión adecuada (Doshi-Velez y Kim, 2017).

Justicia y No Discriminación: Es fundamental garantizar que la IA no perpetúe desigualdades sociales o ambientales. Esto implica desarrollar modelos inclusivos que consideren la diversidad de contextos y que no excluyan a comunidades desfavorecidas (Noble, 2018).

Responsabilidad y Rendición de Cuentas: Las empresas y gobiernos deben ser responsables de las decisiones automatizadas por IA y de los impactos ambientales de sus aplicaciones. Establecer mecanismos de rendición de cuentas es clave para asegurar el uso ético de estas tecnologías (Whittlestone y otros, 2019).

Gobernanza y Regulación de la IA para la Sustentabilidad

Normativas Internacionales: La gobernanza global de la IA requiere marcos regulatorios que fomenten la cooperación internacional y aseguren estándares comunes en la evaluación del impacto ambiental y social de las tecnologías de IA (Floridi y otros, 2018).

Participación Ciudadana: La inclusión de la sociedad civil en la discusión sobre el uso de la IA para la sustentabilidad es crucial. Las políticas deben ser cocreadas con comunidades afectadas para reflejar sus necesidades y preocupaciones (Jasanoff, 2017).

Inversión en Investigación Responsable: Es necesario promover la investigación en IA que tenga un enfoque ético y que priorice la sustentabilidad. Esto incluye el financiamiento de proyectos que evalúen los impactos ambientales de las tecnologías emergentes (Brock y Johnson, 2021).

EL FUTURO DE LA IA EN LA SUSTENTABILIDAD

El futuro de la IA en la sustentabilidad es prometedor, pero dependerá de cómo abordemos los retos actuales y aseguremos un desarrollo inclusivo y ético.

Avances Tecnológicos y Nuevas Aplicaciones

Los avances en IA, como el aprendizaje por refuerzo y la IA cuántica, prometen revolucionar la gestión de recursos naturales, la predicción climática y la producción de energía renovable. Estas tecnologías emergentes permitirán resolver problemas complejos que hoy en día limitan el avance hacia la sustentabilidad (Gambhir y otros, 2019).

IA para la Economía Circular

La IA jugará un papel clave en la transición hacia una economía circular, optimizando la reutilización de materiales, mejorando los procesos de reciclaje y diseñando productos más sostenibles. Modelos de IA podrán prever la vida útil de los productos y recomendar estrategias de recuperación y reutilización (Stahel, 2016).

Democratización y Accesibilidad

El futuro de la IA debe centrarse en la democratización del acceso a estas tecnologías, asegurando que no solo las grandes corporaciones puedan beneficiarse de la IA para la sustentabilidad. Es vital desarrollar herramientas y plataformas accesibles que permitan a pequeñas empresas, ONGs y comunidades aprovechar la IA para resolver sus desafíos locales (Floridi y otros, 2020).

Integración con Tecnologías Verdes

La integración de la IA con otras tecnologías verdes, como los sistemas de energía solar, vehículos eléctricos y la gestión inteligente de residuos, será esencial para maximizar su impacto. La convergencia de estas tecnologías permitirá una gestión más coordinada y eficiente de los recursos naturales, impulsando un desarrollo más sostenible (Sovacool y otros, 2020).

CONCLUSIÓN: EL CAMINO HACIA UN FUTURO SUSTENTABLE CON IA

La Inteligencia Artificial tiene un enorme potencial para transformar nuestra sociedad hacia un modelo más sostenible. A través de la optimización de recursos, la reducción de emisiones y la innovación en sectores clave, la IA puede ser un motor fundamental para enfrentar los desafíos ambientales del siglo XXI. Sin embargo, para que esto se materialice, es crucial abordar los retos éticos y de gobernanza aso-

ciados con su uso, asegurando que estas tecnologías se desarrollen y apliquen de manera justa, inclusiva y responsable.

La IA no es una panacea, pero con una regulación adecuada y un enfoque ético, puede ser una herramienta poderosa para construir un futuro más verde y equitativo. El compromiso de los gobiernos, la industria, y la sociedad civil será clave para desbloquear todo el potencial de la IA en la creación de un mundo sostenible. Este camino requiere no solo innovación tecnológica, sino también un cambio de mentalidad hacia un enfoque más holístico y colaborativo para enfrentar los desafíos globales.

REFERENCIAS BIBLIOGRÁFICAS

Anderson, W., Beesley, L., y Reed, J. (2020). Real-time air quality monitoring and its impact on urban planning in London. *Journal of Environmental Policy & Planning, 22*(4), 523-538. *https://doi.org/10.1080/1523908X.2020.1768345*.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149-159. PMLR. https://proceedings.mlr.press/v81/binns18a.html

Brock, G., y Johnson, L. (2021). Responsible innovation and the future of artificial intelligence: Ethical considerations in the design of AI technologies. Ethics and Information Technology, 23(4), 783-798

Cath, C., Jardill, R., y Houghton, R. (2018). La ética de la inteligencia artificial y su impacto en la sostenibilidad. *AI & Society, 33*(4), 499-508. *https://doi.org/10.1007/s00146-018-0842-8.4*

Chen, S., Lu, X., y Chan, L. Y. (2021). Artificial intelligence in sustainable urban mobility: A review of applications and challenges. Sustainable Cities and Society, 71, 102999. https://doi.org/10.1016/j.scs.2021.102999

Copilot. (2024). Página Oficial. https://copilot.microsoft.com/

Davis, L., Brown, A., y Miller, T. (2021). Inteligencia artificial como motor de sostenibilidad en la logística y el comercio electrónico. *Sustainable Business Journal, 15*(2), 100-113. *https://doi.org/10.1016/j.sbj.2020.01.004*.

Doshi-Velez, F., y Kim, P. (2017). Towards a rigorous science of interpretable machine learning. *Proceedings of the 34th International Conference on Machine Learning*, 70, 2060-2070. *https://arxiv.org/abs/1702.08608*.

Duan, Y., Edwards, J. S., y Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63-71. *https://doi.org/10.1016/j.ijinfomgt.2019.01.021*.

Evans, A., y Gao, J. (2018). Ajuste en tiempo real de sistemas de refrigeración y ventilación mediante inteligencia artificial para maximizar la eficiencia energética. *Energy Reports, 4*, 637-644. *https://doi.org/10.1016/j.egyr.2018.09.002*.

Floridi, L., Sanders, J., y W. M. (2018). The ethics of artificial intelligence: A roadmap. *Nature, 558*(7709), 479-485. https://doi.org/10.1038/d41586-018-05320-4.

Floridi, L., y Chiriatti, M. (2020). AI and the Future of Work. *AI & Society, 35*(4), 881-886. https://doi.org/10.1007/s00146-020-00953-9.

Gambhir, A., Pavan, M., y Scown, C. D. (2019). Sustainable development goals and the role of emerging technologies. *Environmental Science & Policy, 101*, 105-114. https://doi.org/10.1016/j.envsci.2019.06.015.

Gómez, L., Martínez, R., y Pérez, A. (2022). Inteligencia artificial y autobuses eléctricos autónomos: una solución de movilidad sostenible para las ciudades inteligentes. *Transportation Research Part C: Emerging Technologies, 135*, 103704.

Goodman, J., y Foster, L. (2019). La inteligencia artificial y la transición hacia una economía baja en carbono. *Journal of Cleaner Production, 228*, 142-150. https://doi.org/10.1016/j.jclepro.2019.04.298.

Hao, K., O'Neill, K., y Zeller, M. (2020). AI in logistics: Opportunities and challenges for sustainable transport. Journal of Transportation Research Part D: Transport and Environment, 84, 102351. https://doi.org/10.1016/j.trd.2020.102351

Hazan, H., Boulaksil, Y., y Gabbay, A. (2021). La inteligencia artificial en la gestión de recursos hídricos y la planificación urbana sostenible. *Water Resources Management, 35*(6), 1907-1921. https://doi.org/10.1007/s11269-021-02809-6.

Hossain, M. S., Hossain, E., y Al-Hammadi, A. M. (2019). Optimization of renewable energy resources using artificial intelligence and big data

analytics: A review. Renewable and Sustainable Energy Reviews, 131, 110021. https://doi.org/10.1016/j.rser.2020.110021

Jasanoff, S. (2017). *Science and democracy: Making knowledge and making power in the biosciences.* Duke University Press.

Jha, K., Doshi, A., Patel, P., y Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture, 2, 1-12. https://doi.org/10.1016/j.aiag.2019.05.004

Jiang, J., Lu, B., y Zhang, R. (2021). Smart city applications in air quality management with AI and IoT technologies: A systematic review. Environmental Research, 195, 110890. https://doi.org/10.1016/j.envres.2021.110890

Kitchin, R. (2020). The ethics of smart cities and urban science. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 378(2166), 20190115. https://doi.org/10.1098/rsta.2019.0115

Kshetri, N. (2021). Ciberseguridad y privacidad en la era de la inteligencia artificial: desafíos y soluciones. *Computer Security, 111*, 102435. https://doi.org/10.1016/j.cose.2021.102435.

Leahy, S. (2019). *We're in a climate crisis, but it's also a food crisis. National Geographic. https://www.nationalgeographic.com/environment/article/climate-change-impact-food-production*.

Lee, J., Park, S., y Kim, Y. (2018). Optimización de diseño en manufactura aditiva mediante inteligencia artificial: un enfoque hacia la sostenibilidad. *Journal of Cleaner Production, 198*, 1220-1230. *https://doi.org/10.1016/j.jclepro.2018.06.095*.

Levy, D. J., y Wong, D. (2020). Optimización de cadenas de suministro mediante inteligencia artificial: el caso de Amazon y su impacto ambiental. *Journal of Supply Chain Management, 56*(4), 51-66. *https://doi.org/10.1111/jscm.12236*.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., y Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

Mathiesen, M. A., Henningsen, T., y Nielsen, H. (2021). Prevención de inundaciones en ciudades inteligentes: reduciendo costos y mejorando la calidad de vida mediante inteligencia artificial. *Environmental Science & Policy, 125*, 184-195. *https://doi.org/10.1016/j.envsci.2021.06.014*.

Meyer, D., Pick, A., y Teuteberg, F. (2021). Sustainable management in the circular economy: How AI-driven decision support systems support strategic decisionmaking. Journal of Cleaner Production, 286, 125-169. https://doi.org/10.1016/j.jclepro.2020.125169

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Patel, S., y Whitehead, A. (2019). Contribuciones de la inteligencia artificial a la sostenibilidad en el mundo digital: reduciendo la huella de carbono en empresas tecnológicas. *Journal of Sustainable Technology, 8*(3), 143-157. *https://doi.org/10.1016/j.jst.2019.05.007*.

Pernille, C., Madsen, H., y Jensen, H. (2022). Optimización de la distribución del agua mediante inteligencia artificial: un enfoque para reducir el consumo energético en sistemas de bombeo. *Water Resources Management, 36*(8), 2521-2535. *https://doi.org/10.1007/s11269-022-02989-2*.

Rai, H., Gupta, A., y Kumar, V. (2020). Optimización del uso de recursos naturales mediante inteligencia artificial: un enfoque hacia la sostenibilidad. *Energy Reports, 6*, 126-134. https://doi.org/10.1016/j.egyr.2020.08.001.

Ranveer, S., Gupta, P., y Kumar, A. (2021). Conservación de recursos naturales a través de la inteligencia artificial en la agricultura: un enfoque sostenible. *Resources, Conservation and Recycling, 168*, 105451. https://doi.org/10.1016/j.resconrec.2020.105451.

Ren, J., Chan, W. T., y Pan, L. (2020). Industrial energy efficiency optimization using AI technologies: A review and prospects. Journal of Cleaner Production, 247, 119-140. https://doi.org/10.1016/j.jclepro.2019.119140

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... y Bengio, Y. (2022). Tackling climate change with machine learning. ACM Computing Surveys (CSUR), 55(2), 1-96. https://doi.org/10.1145/3485128

Smith, R., Johnson, L., y Williams, K. (2020). Agricultura de precisión: aplicando agua y nutrientes de manera eficiente mediante inteligencia artificial. *Agricultural Systems, 178*, 102706. https://doi.org/10.1016/j.agsy.2019.102706.

Sovacool, B. K., y Brown, M. A. (2020). Creating a sustainable energy future: The convergence of technology, policy, and practice. *Energy Policy, 139*, 111305. https://doi.org/10.1016/j.enpol.2020.111305

Stahel, W. R. (2016). *The circular economy: A user's guide.* The Ellen MacArthur Foundation.

Strubell, E., Ganesh, A., y McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645-3650. https://doi.org/10.18653/v1/P19-1355.

The Carrington, D. (2019). *Why the Guardian is changing the language it uses about the environment. The Guardian. https://www.theguardian.com/environment/2019/may/17/why-the-guardian-is-changing-the-language-it-uses-about-the-environment*.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... y Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11(1), 1-10. https://doi.org/10.1038/s41467-019-14108-y

Wang, J., Zhong, H., Ma, Z., Xia, Q., y Kang, C. (2021). The application of artificial intelligence in smart grids: A review. *Renewable and Sustainable Energy Reviews, 138*, 110545. *https://doi.org/10.1016/j.rser.2020.110545*.

Wang, X., Zhang, Y., y Liu, J. (2022). Impacto de la inteligencia artificial en la movilidad urbana: reducción de emisiones y optimización del transporte público en Singapur. *Transportation Research Part D: Transport and Environment, 100*, 103084. *https://doi.org/10.1016/j.trd.2022.103084*.

Whittlestone, J., Nyrup, R., Joshi, I., y Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on practices. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 3-9. https://doi.org/10.1145/3306618.3314276.

Wu, T., Zhang, J., y Yang, Y. (2020). Conservation of biodiversity using artificial intelligence: A systematic review. Ecological Informatics, 55, 101019. https://doi.org/10.1016/j.ecoinf.2020.101019

Yuan, H., y Lo, H. K. (2021). Inteligencia artificial para la gestión de emergencias y la seguridad en ciudades inteligentes. *Journal of Urban Technology, 28*(3), 25-45. *https://doi.org/10.1080/10630732.2021.1881041*.

CAPÍTULO VI

EL USO ÉTICO DE LA INTELIGENCIA ARTIFICIAL EN LAS ACTIVIDADES ACADÉMICAS UNIVERSITARIAS. PERCEPCIÓN DEL ESTUDIANTE

MIGUEL ÁNGEL OLIVARES URBINA[1]
RICARDO SONDA DE LA ROSA[2]

RESUMEN

Esta investigación explora el uso ético de la inteligencia artificial (IA) en actividades académicas universitarias, enfocándose en la percepción de los estudiantes de la Universidad del Caribe (UNICAR). Se encuestaron a 192 estudiantes de diversas divisiones, incluidos programas de Maestría, Negocios e Innovación y Turismo Sustentable, para evaluar cómo perciben la IA en la elaboración de trabajos académicos, como tesis, artículos, proyectos, tareas y ensayos.

Los resultados revelan que el 82 % de los estudiantes considera ético el uso de IA, siempre que no exceda el 40 % de intervención. La aceptación es más alta en artículos (91 %) y tesis (86 %), y más cautelosa en ensayos (64 %) y tareas (70 %). Los estudiantes de Turismo son los más permisivos, mientras que los de maestría muestran mayor reticencia, especialmente en ensayos.

A partir de estos hallazgos, se elaboraron recomendaciones para establecer guías claras sobre el uso de IA, fomentar la reflexión ética

[1] Profesor Investigador de Tiempo Completo Titular A, Universidad del Caribe, Quintana Roo, México.

[2] Profesor Investigador de Tiempo Completo Titular A, Universidad del Caribe, Quintana Roo, México.

y ajustar los métodos de evaluación. Además, se destacó la necesidad de capacitación en el uso responsable de la IA. Este estudio enfatiza la importancia de regular la implementación de la IA en entornos educativos para garantizar su uso responsable y ético.

Palabras clave: Inteligencia Artificial, Ética Académica, Percepción Estudiantil.

INTRODUCCIÓN A LA ÉTICA EN LA INVESTIGACIÓN

El origen de la ética, como disciplina, se encuentra en la filosofía, de la cual más tarde se desprendió como un campo de estudio independiente. Para llegar a una definición precisa de la ética, es necesario recurrir a estudios filosóficos que analicen el comportamiento correcto de los miembros de una sociedad (OsejoBucheli, 2021). La filosofía se apoya en herramientas como la lógica y el razonamiento, elementos fundamentales del saber humano.

Esta aproximación histórica encuentra su base en que, para estudiar filosofía primero se tiene que estudiar la ética, pues esta consideraba como lo más preciado del ser humano. A primera vista, esta sugerencia podría parecer superficial o cuestionable, pero continúa siendo defendida por algunos pensadores contemporáneos. Este planteamiento nos invita a reflexionar sobre el impacto de la ética en nuestras decisiones diarias y en nuestras relaciones con los demás.

La ética no debe tomarse a la ligera; no cualquier persona puede hablar con autoridad sobre cuestiones éticas, ya que un experto en ética no solo posee conocimientos teóricos, sino también una visión del mundo que guía su pensamiento y sus acciones (Kriele, 2022), al igual que en las matemáticas o en las ciencias, la validez de las afirmaciones éticas depende del conocimiento previo y de una argumentación adecuada.

Por lo anterior, la ética requiere un enfoque riguroso y sistemático, basado en la reflexión crítica y la experiencia humana. De acuerdo con

González (2023), por su naturaleza contingente y contextual, emitir juicios éticos demanda una consideración cuidadosa de argumentos bien estructurados y es en ese sentido que, la argumentación se convierte en la principal herramienta de la razón para evaluar afirmaciones éticas, confrontar diferentes perspectivas y llegar a conclusiones fundamentadas.

Es innegable la importancia de la ética en la vida cotidiana y en la sociedad en general ya que, a través de la reflexión ética, tomamos decisiones más informadas y responsables, basadas en principios morales sólidos y en valores compartidos (Dux, 2023) por lo que la ética nos desafía a cuestionar, a buscar respuestas y a mejorar continuamente como seres humanos.

CONCEPTOS FUNDAMENTALES DE LA INTELIGENCIA ARTIFICIAL

La inteligencia artificial (IA) es un campo de la informática que busca desarrollar algoritmos y técnicas que permitan a las computadoras realizar tareas que, tradicionalmente, requieren inteligencia humana, como el aprendizaje, el razonamiento y la adaptación. Estas tareas incluyen el reconocimiento de voz, la traducción automática, la toma de decisiones y el aprendizaje automático. Se define como una rama de las ciencias computacionales que diseña sistemas capaces de realizar actividades que, en la actualidad, necesitan la intervención de un ser humano.

La IA abarca varias definiciones según el campo de aplicación. Algunas formas comunes de definirla incluyen:

- *"El desarrollo de algoritmos y herramientas que permiten a los sistemas computacionales realizar tareas que, por el momento, los humanos realizan mejor".*
- *"Un conjunto de técnicas diseñadas para crear programas informáticos con comportamiento inteligente".*

El aprendizaje supervisado es una modalidad de IA donde los algoritmos se entrenan con datos etiquetados, mientras que el aprendizaje no supervisado procesa datos sin una estructura predefinida para extraer patrones. En cuanto al aprendizaje por refuerzo, se refiere al proceso en el que un agente toma decisiones y mejora a partir de los resultados obtenidos de sus acciones (Manguay, 2023).

Entre los elementos clave de la IA se encuentran:

- Aprendizaje automático: Esta técnica permite a las computadoras aprender de forma continua y hacer predicciones basadas en datos sin ser programadas explícitamente. En lugar de depender de la programación tradicional, los sistemas se adaptan a partir de la experiencia y mejoran su rendimiento a medida que procesan más datos (Martínez, 2020; Rouhiainen, 2018). El aprendizaje automático es fundamental para el desarrollo de la IA y es un área que continúa creciendo. Se espera que se combine con otras tecnologías emergentes como la nanotecnología, la computación cuántica y la automatización para impulsar una economía circular y sostenible (Díaz-Ramírez, 2021; Carrasco, 2021).
- Reconocimiento de patrones: La IA tiene la capacidad de reconocer patrones en datos, como imágenes o sonidos, lo cual es esencial en aplicaciones prácticas como el procesamiento de imágenes o el análisis de señales. Aunque algunas tareas de reconocimiento de patrones pueden ser realizadas por humanos con facilidad, en otras, el uso de computadoras se vuelve indispensable para aumentar la precisión y eficiencia (Reynaga, 2009; Ochoa, 2014). Este campo depende de la visión artificial y de la capacidad de los sistemas para distinguir formas y estructuras en los datos, permitiendo a la IA comprender y clasificar objetos en el mundo real (Tello e Informáticos, 2007; Corvalán, 2019).
- Razonamiento y toma de decisiones: La IA tiene la capacidad de razonar y tomar decisiones de manera autónoma en áreas tan diversas como la medicina, la conducción autónoma y la gestión empresarial. Estos sistemas avanzan a pasos agiganta-

dos y, en muchos casos, ya superan la capacidad humana para procesar grandes cantidades de información y tomar decisiones complejas (Quinto, 2021; Verdegay y otros, 2021). El razonamiento autónomo en IA sigue el modelo de toma de decisiones propuesto por Herbert Simon, que implica identificar problemas, generar posibles soluciones y seleccionar la mejor opción. Las máquinas de IA utilizan algoritmos que les permiten resolver problemas de manera eficiente y racional, optimizando sus decisiones para alcanzar objetivos predefinidos (Aghemo, 2020; Gupta, 2018).

- Procesamiento del lenguaje natural (PLN): El PLN permite a las computadoras comprender y generar el lenguaje humano. La lingüística computacional, aplicada a la IA, busca emular la capacidad humana de entender y producir lenguaje, mejorando así la interacción entre personas y máquinas (Celi-Parraga, 2021; Borja Quevedo, 2020). El lenguaje natural es el medio de comunicación que empleamos de manera inconsciente y que difiere de los lenguajes artificiales, los cuales son aprendidos por decisión o necesidad. El procesamiento efectivo del lenguaje natural permite a las computadoras interactuar con los humanos de una manera cada vez más fluida y natural (Torres, 2012).

En el contexto del Procesamiento del Lenguaje Natural (PLN), se busca analizar y representar textos escritos en lenguaje natural mediante software y algoritmos. Estas herramientas permiten que los sistemas computacionales procesen el lenguaje humano en tareas específicas, otorgándole una apariencia más "natural" en su funcionamiento. Las técnicas de PLN se basan en el aprendizaje automático integrado, o técnicas de *Embedded Machine Learning*, que después de recolectar datos, permiten analizar patrones de comportamiento. Una de las ventajas del PLN es que los algoritmos pueden generar resultados sin que el investigador deba interpretar manualmente los textos. Sin embargo, una de sus limitaciones es la falta de una verdadera comprensión natural del lenguaje por parte del software (Sancho Escrivá, 2020).

La percepción en los sistemas de inteligencia artificial se refiere a la capacidad de una computadora para percibir su entorno mediante sensores, como cámaras o micrófonos. La percepción computacional busca reproducir artificialmente sentidos como la vista y el oído. Aunque esta tecnología aún se encuentra en sus etapas iniciales, está avanzando rápidamente. La percepción trabaja a partir de estímulos que se procesan para extraer propiedades abstractas, lo que requiere una capacidad computacional significativa. A pesar de los avances, incluso los supercomputadores más potentes enfrentan dificultades para alcanzar el nivel de procesamiento masivo y paralelo de la maquinaria biológica, que permite manejar múltiples hipótesis de interpretación de manera simultánea (Escalera, 2001).

La percepción implica no solo la vista, sino también la capacidad de identificar y reconocer sonidos, olores, y el sentido del tacto. El aprendizaje automático busca imitar el comportamiento humano mediante la detección de patrones, pero para lograrlo, las máquinas deben contar con información adecuada sobre el entorno en el que operan. Esta información es esencial para que las máquinas puedan actuar y reaccionar de manera óptima, constituyendo así una de las bases de la Inteligencia Artificial. El procesamiento simbólico es una técnica utilizada en la IA en la que la computadora manipula símbolos que representan el entorno para tomar decisiones precisas y rápidas (Partida Tapia, 1995).

Los sistemas de IA utilizan información simbólica basada en conceptos, objetos y reglas, en contraste con la informática clásica, que maneja principalmente datos numéricos. Además, la IA se vale de métodos heurísticos en lugar de algorítmicos para resolver problemas, como los relacionados con la percepción y la representación del entorno. La percepción artificial no solo implica el reconocimiento del entorno, sino también el seguimiento y la interacción con la tecnología artificial, lo que permite una representación más precisa del entorno (Hardy, 2001).

La percepción es clave en los sistemas tecnológicos, ya que permite a las máquinas comprender el mundo exterior de manera similar a los humanos. Gracias a técnicas avanzadas de IA y robótica, se espera que los sistemas sean cada vez más adaptables y racionales, adquiriendo

habilidades similares a las humanas. En resumen, la percepción es el proceso autónomo mediante el cual un sistema adquiere información del entorno y la organiza en una representación interna (Mayershofer, 2021).

La robótica, como una rama de la inteligencia artificial, se enfoca en el diseño y construcción de máquinas capaces de realizar tareas humanas a través de procesos mecanizados y programados. También se define como una herramienta interconectada, interactiva, cognitiva y física, capaz de percibir su entorno, razonar sobre eventos, elaborar planes y controlar sus acciones (Porcelli, 2020).

La IA aplicada a robots tiene un enorme potencial para mejorar la productividad y la calidad de vida, pero muchos desarrolladores se enfocan exclusivamente en la automatización de trabajos humanos, lo que puede tener efectos negativos en los trabajadores y limitar los beneficios económicos a largo plazo. Aunque la automatización es necesaria para tareas peligrosas o tediosas, es crucial encontrar un equilibrio que maximice los beneficios de la IA en los robots (HAI, 2022).

Un robot inteligente es una máquina física que, a través de programas y sensores, puede tomar decisiones autónomas basadas en IA. Estos sistemas son capaces de predecir las necesidades de los humanos y las situaciones en las que se encuentran, actuando e interactuando con el mundo físico sin la intervención constante de las personas (González, 2017).

Lejos de buscar reemplazar a los trabajadores humanos, la robótica pretende optimizar tareas y procesos, además de atraer talento. La robótica aporta una ventaja competitiva a las empresas al generar valor añadido mediante la mejora de la productividad, la reducción de costos y la flexibilidad en los procesos industriales. A medida que la IA avanza, los robots se vuelven más autónomos, capaces de recibir y emitir comunicación, comprender su entorno y ejecutar acciones planificadas (Porcelli, 2020).

La inteligencia artificial también desempeña un papel clave en el Internet de las Cosas (IoT), donde los dispositivos conectados a internet generan enormes cantidades de datos. La IA permite analizar estos datos para obtener información valiosa sobre el comportamiento

de los usuarios, el rendimiento de los dispositivos y posibles mejoras. Además, la IA puede crear sistemas de automatización inteligente que controlen los dispositivos de forma autónoma, sin intervención humana (Madrid, 2018).

El IoT involucra la interconexión de dispositivos que recolectan datos mediante redes inalámbricas, con poca intervención humana. Esto no solo optimiza los tiempos y mejora la eficiencia, sino que también plantea nuevos desafíos en cuanto a la seguridad y privacidad de los dispositivos conectados. La implementación de IA en el IoT tiene el potencial de mejorar los modelos de negocio actuales, aunque también implica el manejo de grandes volúmenes de datos, lo que representa un reto considerable (Arévalo, 2016).

A medida que el IoT evoluciona, sus plataformas y redes se volverán más seguras y avanzadas, lo que permitirá su expansión en diversas industrias. La IA puede ser una herramienta poderosa para que los dispositivos conectados aprendan, se adapten y tomen decisiones de forma autónoma, incrementando la eficiencia en aplicaciones que van desde los hogares inteligentes hasta la manufactura industrial (Madrid, 2018).

El avance de la inteligencia artificial ha generado varias controversias y preocupaciones, entre ellas:

- Seguridad: Una de las principales preocupaciones es que la IA pueda ser utilizada para causar daño, como en el caso de los sistemas de armas autónomos o los ciberataques. Cada innovación tecnológica trae consigo beneficios, pero también riesgos. Por ejemplo, aunque internet es fundamental para proteger la libertad de expresión, también puede facilitar actividades ilegales, como el tráfico de armas y otros delitos en el ámbito digital (Corvalán, 2018).

Los sistemas de armas autónomos son agentes racionales con IA que pueden adaptarse a diversas situaciones y realizar tareas con distintos grados de autonomía. Esta tecnología permite que los sistemas armamentísticos sean cada vez más independientes, lo que plantea el riesgo de que en el futuro operen completamente sin intervención humana (Westhues, 2020).

- Privacidad: La IA plantea preocupaciones sobre la recopilación y uso inapropiado de información personal. La protección de la privacidad en el ámbito cibernético es un reto significativo, ya que muchas legislaciones no contemplan adecuadamente el uso de la IA. En algunos países, como Ecuador, la Ley de Protección de Datos no regula específicamente el uso de inteligencia artificial, mientras que otros, como España, han adoptado leyes más avanzadas en la sociedad digital (Digitales, 2021).

Con el avance de la tecnología, las leyes están en constante evolución para proteger los derechos de los ciudadanos en el ciberespacio, lo que incluye la ciberseguridad, la protección de datos y la propiedad intelectual. La transformación digital ha generado desafíos significativos en cuanto a la vulneración de la privacidad de los usuarios en plataformas digitales. Esto resalta la necesidad de un marco jurídico que contemple la IA y la seguridad cibernética (Molina, 2023).

- Desempleo: La preocupación de que la automatización y la inteligencia artificial puedan reemplazar trabajos humanos. En la actualidad, la economía digital está generando un nuevo paradigma en el ámbito laboral. Este proceso de reestructuración del empleo es un resultado directo de innovaciones tecnológicas disruptivas en campos como la inteligencia artificial, la robótica, la biotecnología y las telecomunicaciones, que están transformando la concepción del trabajo y afectando de manera significativa a la sociedad y a la economía global (Ruiz, 2018). La colaboración entre humanos y máquinas en entornos industriales está generando una evolución en sectores clave de la economía mundial. Mientras que la automatización incrementa la demanda de trabajadores con nuevas habilidades, también podría conducir a la desaparición de sectores profesionales actualmente ocupados por empleados con habilidades de nivel básico o intermedio (Muñoz, 2016). A pesar de que América Latina no es una región puntera en tecnología, se vislumbra un futuro en el que robots y humanos trabajen juntos, optimizando la eficiencia y el progreso regional (Corvalán, 2019).

Esta tendencia hacia la polarización del mercado laboral sugiere que los empleos menos cualificados serán los más susceptibles de ser reemplazados por la IA, mientras que los trabajadores podrían reasignarse a tareas que requieran habilidades creativas e interpersonales (Frey y Osborne, 2013). La IA tiene la capacidad de automatizar una variedad de tareas complejas, como el análisis de grandes cantidades de datos y la toma de decisiones basadas en patrones identificados, lo que puede reemplazar trabajos humanos en varios sectores (Arrieta y Pérez, 2022).

La preocupación de que los algoritmos de inteligencia artificial puedan replicar y perpetuar la discriminación existente en la sociedad

Los algoritmos de aprendizaje automático, al apoyarse en datos históricos, tienden a replicar los sesgos y desigualdades presentes en la sociedad, lo que perpetúa la discriminación. Dado que algunos algoritmos se utilizan como base para desarrollar otros, existe un riesgo real de que estas técnicas sigan reproduciendo prejuicios contra diferentes grupos sociales (Vivas Fonseca, 2022).

El sesgo algorítmico se refiere a la predisposición de un sistema de IA para favorecer a un grupo sobre otro, lo que conlleva situaciones injustas. Estos prejuicios pueden filtrarse en los sistemas debido a los datos empleados para entrenarlos y a las perspectivas de los programadores, lo que refuerza patrones discriminatorios (Ntoutsi y otros, 2020). Para mitigar este problema, es fundamental diversificar los equipos de desarrollo y los conjuntos de datos utilizados, promoviendo un enfoque ético que garantice la equidad (Mendoza, 2020).

- Responsabilidad: La preocupación de quién será responsable en caso de daño causado por un sistema de IA. La cuestión de la responsabilidad en el uso de la IA plantea desafíos importantes. Una posibilidad es distribuir la responsabilidad entre los fabricantes, diseñadores, programadores, ingenieros, usuarios y propietarios de los sistemas. Sin embargo, esta solución diluye la noción de responsabilidad, dificultando la identificación de un único responsable en caso de daños (Lain Moyano, 2021).

Actualmente, no existe una legislación específica que regule la responsabilidad en el uso de la IA, lo que genera vacíos legales en situa-

ciones donde estos sistemas causen daños. A medida que los sistemas de IA adquieren mayor autonomía, será necesario desarrollar marcos legislativos que asignen la responsabilidad de manera justa y clara, dependiendo del contexto y las circunstancias (Latorre, 2022; Masoitti, 2022).

- Control: La preocupación de que los desarrolladores de IA y las empresas que los utilizan puedan tener demasiado control sobre la vida de las personas. Los algoritmos de IA están siendo utilizados para analizar y predecir nuestras acciones y emociones, lo que plantea preocupaciones sobre la invasión de la privacidad y el control de nuestras decisiones. Estas tecnologías no solo influyen en aspectos rutinarios, sino que también pueden modificar comportamientos y valores de manera sutil, en función de los intereses de quienes desarrollan y comercializan estos sistemas (Floridi, 2018).

En el ámbito educativo, los sistemas de IA pueden eliminar los aspectos emocionales del proceso de enseñanza, lo que podría llevar a una mayor manipulación de los estudiantes (Toboso y Aparicio, 2019). Es necesario establecer marcos de gobernanza adecuados que regulen el uso de estas tecnologías para garantizar la libertad y privacidad individual.

- Superinteligencia: La preocupación de que la IA pueda alcanzar un nivel de inteligencia superior al humano y tener consecuencias impredecibles. La "singularidad tecnológica" hace referencia al momento en el que la IA podría superar la inteligencia humana, con consecuencias potencialmente graves para la humanidad. Existen preocupaciones de que una IA superinteligente pueda actuar en contra de los humanos si percibe que estos son una amenaza para sus objetivos (Hueso, 2019).

Si bien aún no se ha desarrollado una IA con este nivel de autonomía y capacidad, los riesgos existenciales planteados por la IA superinteligente son una preocupación creciente. A largo plazo, las máquinas podrían superar a los humanos en capacidad cognitiva, lo que subraya la importancia de establecer mecanismos de control que eviten que las IA se desvíen de los valores humanos (Penrose y Huang, 2005).

La inteligencia artificial (IA) está impactando cada vez más a las organizaciones, ya sea automatizando tareas, mejorando la eficiencia y productividad o facilitando la toma de decisiones. A continuación, se presentan ejemplos específicos de cómo la IA está siendo utilizada:

La IA permite analizar grandes cantidades de datos para extraer información valiosa que mejora la toma de decisiones. En diversos sectores, las organizaciones recopilan datos históricos de sus operaciones, productos y clientes, mientras que los científicos acumulan grandes volúmenes de datos experimentales, como imágenes de resonancia magnética funcional para estudiar la actividad cerebral. Este análisis se enfoca en identificar patrones que mejoren la toma de decisiones estratégicas (Mitchell, 1999). Gracias a la IA, la recopilación y análisis de estos datos, que para humanos serían imposibles de procesar en tiempo real, se puede realizar en milisegundos, lo que permite a las empresas reaccionar rápidamente ante problemas o inconvenientes (Corvalán, 2018). El almacenamiento masivo de datos se ha convertido en una práctica común, pero sin el análisis inteligente proporcionado por la IA, su potencial es limitado (Aluja Banet, 2001).

La automatización mediante IA simplifica la gestión de datos y acelera actividades que antes eran complejas o imposibles de resolver, mejorando la eficiencia organizacional. La IA permite automatizar tareas repetitivas como la entrada de datos, lo que reduce errores y aumenta la productividad (Maisueche Cuadrado, 2019). Además, la automatización ha revolucionado industrias, mejorando la seguridad, reduciendo costos y eliminando trabajos peligrosos (Lee, 2016). Incluso, ha afectado a tareas especializadas, como la reducción del 20 % en el empleo de bibliotecarios y traductores debido a la automatización (Barro, 2017).

La IA permite personalizar mensajes publicitarios y estrategias de marketing para llegar de manera más efectiva a los clientes potenciales. Al aprender y predecir comportamientos a partir de datos, la IA facilita la automatización de interacciones con los consumidores, fortaleciendo relaciones a largo plazo (Jarek y Mazurek, 2019).

Herramientas como la compra programática y la optimización de sitios web son ejemplos del impacto de la IA en el marketing digital (Sterne, 2017). Sin embargo, los especialistas en marketing a menudo

carecen de conocimientos técnicos suficientes para integrar aplicaciones de IA en sus estrategias (Medina y Martínez, 2020).

La IA ha mejorado significativamente el servicio al cliente a través de *chatbots* y sistemas de reconocimiento de voz. Los *chatbots* permiten interactuar con los clientes de manera automática y continua, brindando soluciones inmediatas y derivando consultas más complejas a humanos (Notimex, 2017) (Albrecht, 1992). La implementación de estos sistemas ha optimizado la atención, reduciendo costos y tiempos de respuesta, mientras que las empresas alcanzan sus objetivos de manera más eficiente (Pérez Carvajal, 2020).

En el ámbito de recursos humanos, la IA se utiliza para automatizar procesos de reclutamiento y selección de personal. Los algoritmos permiten una selección más eficiente de candidatos y mejoran la retención del personal. La automatización de las primeras etapas del proceso, como el cribado de currículums, reduce la carga de trabajo y mejora la precisión en la selección de talento (Geetha, 2018; Cantero, 2021). Además, el análisis de redes sociales y otras fuentes de información personal permite una evaluación más completa de los candidatos, mejorando las decisiones de contratación (Rodney, 2019).

La inteligencia artificial (IA) se utiliza para detectar y prevenir amenazas de seguridad en las redes y sistemas de las organizaciones. Se ha demostrado que la IA es una herramienta efectiva para abordar numerosos problemas en la defensa cibernética; al emplear métodos de IA, es posible reaccionar de manera más ágil a las situaciones que surgen en el entorno cibernético. Las empresas actuales deben gestionar y supervisar grandes volúmenes de datos en tiempo real para detectar y analizar eventos en su entorno. Este análisis proporciona una comprensión más precisa de la situación y facilita la toma de decisiones correctas, especialmente dado que los ataques a las redes son cada vez más frecuentes y persistentes (Czosseck, 2011).

La ciberseguridad es crucial porque las empresas almacenan datos y generan miles de transacciones en sus sistemas. Una parte significativa de estos datos puede ser confidencial, incluyendo propiedad intelectual, información financiera, datos de empleados o clientes, entre otros. La exposición o el acceso no autorizado a esta información puede tener consecuencias graves (Martínez, 2020).

La exposición a ataques debido a vulnerabilidades en los sistemas es una preocupación actual en seguridad informática. La IA se aplica en diversas áreas de la seguridad, como detección de intrusos, análisis forense y antivirus. La integración de técnicas de IA en los sistemas de seguridad permite una operación automática, adaptativa y proactiva, mejorando así la protección de la información en las organizaciones (Cohen, 2007).

El uso de la IA en la seguridad informática se vuelve cada vez más necesario ante los crecientes desafíos de los programas maliciosos modernos. La detección y protección contra estos ataques requieren un procesamiento inteligente de grandes cantidades de datos no convencionales. La sofisticación de los ataques informáticos ha aumentado, lo que hace que la aplicación de la IA sea esencial para enfrentar estos problemas (Yeja y otros, 2013).

Los ataques de adversarios representan un mecanismo mediante el cual un sistema inteligente, como el de una persona, puede ser engañado o confundido por variables diseñadas por un adversario. El aprendizaje no supervisado es una forma de adquirir conocimiento observando datos que no siguen un orden preestablecido, permitiendo que el sistema descubra patrones y relaciones de manera autónoma.

Por otro lado, el aprendizaje reforzado implica definir las acciones que debe tomar un agente en función de la retroalimentación que recibe del entorno. Junto con los algoritmos de aprendizaje automático, este enfoque permite entrenar a sistemas inteligentes para enfrentar diversos desafíos, incluidos los ataques de adversarios. Estos ataques buscan engañar o manipular un sistema inteligente mediante la introducción de variables específicas para tal fin.

Los ataques de adversarios pueden poner a prueba la integridad y robustez de un sistema inteligente, con el objetivo de que el sistema tome decisiones erróneas o interprete incorrectamente la información presentada. Estos ataques pueden ser perpetrados por individuos malintencionados para sabotear o explotar sistemas, o por investigadores en busca de debilidades en los sistemas inteligentes (Preda y Kuk, 2024).

La investigación es un motor clave de transformación académica y social en las universidades, desarrollándose a través de proyectos de innovación docente, investigación universitaria, participación estudiantil en congresos, publicaciones en revistas y programas de becas de investigación, entre otros. En la última década, se ha observado un notable crecimiento en la aplicación de la IA en diversos campos de investigación (Erazo-Castillo, 2023).

En el ámbito educativo, se ha avanzado significativamente en el desarrollo de sistemas de evaluación y recomendación. Estos modelos se orientan hacia criterios de calidad del centro educativo y son aplicables para la evaluación institucional, la programación académica y la toma de decisiones en administración y finanzas. La IA también ofrece soluciones personalizadas para la toma de decisiones en el momento preciso, supervisando, programando, actuando e interactuando de manera autónoma y adaptativa. Esto ha permitido el desarrollo de nuevas herramientas para la formación de alumnos y el análisis de información.

La IA tiene un impacto significativo en el sector educativo, facilitando la creación de sistemas más sofisticados de evaluación y recomendación que garantizan una educación de calidad. Estos sistemas se adaptan a los requisitos específicos de cada centro educativo, convirtiéndose en herramientas valiosas para la evaluación institucional, la planificación académica y la toma de decisiones en áreas administrativas y financieras. La IA permite prever situaciones, supervisar el progreso, programar actividades, tomar decisiones e interactuar en diversos contextos, mejorando la experiencia de aprendizaje y el análisis de datos (Salmeron y otros, 2023).

Todos los investigadores financiados por iniciativas deben considerar los aspectos éticos y de privacidad en sus labores de investigación automatizada. Por ejemplo, en el ámbito sanitario se deben seguir normas codificadas en la Declaración de Helsinki, los Principios de Bergen y el Real Decreto de ensayos clínicos con medicamentos, que regulan el diseño, realización y registro de ensayos clínicos y estudios epidemiológicos (Morandin-Ahuerma, 2023).

Es fundamental que los desarrollos técnicos y predictivos consideren la dimensión ética. En algunos casos, se requieren códigos éti-

cos específicos, como en la inteligencia artificial aplicada a amenazas persistentes avanzadas. La transparencia en la IA implica explicar las relaciones causales de los resultados para que otros puedan entender y evaluar el proceso. Proporcionar explicaciones válidas es crucial (González, 2023). La atribución de resultados a causas distintas a la IA puede ser problemática, especialmente si se trata de causas delicadas, como la exposición de sesgos subyacentes en los resultados (Pérez, 2024).

La IA es una disciplina en expansión con numerosas aplicaciones en diversos contextos (Rodriguez-Sánchez, 2024). Se prioriza una comprensión general de las técnicas de IA para evitar su uso indiscriminado y garantizar la evaluación y validación de los resultados obtenidos. Se debe reforzar el esfuerzo para consolidar una cultura abierta en el desarrollo de la IA, con un compromiso significativo de las instituciones de educación superior, investigación y administraciones públicas. Los investigadores y responsables de políticas deben contribuir a garantizar que este esfuerzo continúe en beneficio de la sociedad (Ferre y otros, 2022).

El uso de Big Data en las universidades plantea cuestiones éticas. Aunque la ética en la selección, uso y comunicación de los datos en redes sociales es un tema amplio, aquí se destacan algunos aspectos específicos. Los investigadores que utilizan grandes bases de datos y aplicaciones móviles para analizar cómo afecta nuestro comportamiento y percepción enfrentan dilemas éticos sobre el impacto de las tareas solicitadas a los empleados temporales de estas plataformas. Este dilema se asemeja al antiguo problema de *"do ut des"*, es decir, *"te doy para que me des"* (Ruiz y Stevens, 2024).

Es crucial reflexionar sobre la integridad moral y el cuidado en el manejo de la información registrada (Aragona, 2022). Los análisis de *Big Data* deben realizarse con una finalidad lícita y basarse en principios de justicia social para garantizar una aplicación equitativa. Los expertos advierten sobre el riesgo de estereotipos y la posible discriminación derivada de ellos. Además, la transparencia en las fuentes de datos es esencial para evitar problemas significativos. Es fundamental entender el contexto y la diversidad de los datos involucrados, evitando suposiciones y generalizaciones simplistas.

El trato justo y equitativo de todas las personas, independientemente de su identidad o características, es fundamental en el análisis de *Big Data*. Además, se deben abordar los sesgos implícitos en las herramientas automáticas de análisis y detección de patrones. Aunque estas tecnologías pueden ser muy beneficiosas, están diseñadas por seres humanos, quienes pueden estar influenciados por sus propios prejuicios. Por lo tanto, es crucial una revisión constante y rigurosa de estos algoritmos e implementar mecanismos de control para evitar sesgos y discriminación (Gallent-Torres y otros, 2023). En conclusión, el análisis de *Big Data* tiene el potencial de ser una herramienta poderosa para la toma de decisiones informadas, pero es esencial abordar los desafíos éticos y morales que conlleva, promoviendo un análisis justo y equitativo (Naranjo y otros, 2023).

A continuación, se presentan estrategias para fomentar una cultura ética en la investigación universitaria. Tanto la Guía Ética para la Investigación como la Estrategia Universitaria buscan promover investigaciones caracterizadas por su integridad y excelencia. El principio ético fundamental para evitar comportamientos adversos en materia de protección de la propiedad intelectual e industrial es el respeto por la propiedad y titularidad de estos bienes. Se debe evitar ofrecer, transmitir o ceder a terceros bienes de propiedad industrial o intelectual ajenos como si fueran propios (Laverde y Gutiérrez, 2024).

Para lograr esto, es esencial implementar las siguientes medidas: proteger el material mediante mecanismos de prevención de accesos no autorizados y la incorporación de avisos o mensajes legales sobre propiedad, derechos y prohibiciones de uso. Además, es importante promover una cultura ética mediante metodologías docentes orientadas al desarrollo de actitudes responsables (Rodriguez-Sánchez, 2024). Definir las conductas no éticas para facilitar el análisis y reflexión por parte de los estudiantes, así como fomentar la participación activa de estos en el desarrollo de los parámetros del trabajo de investigación, como se propuso en la sesión anterior de la ponencia a través de la definición de los parámetros de los instrumentos evaluativos (Angulo, 2024).

Es evidente la necesidad de implantar e impulsar de manera decidida el uso y desarrollo de la inteligencia artificial (IA) en el ámbito de

la investigación científica. Actualmente, la IA se ha convertido en una herramienta indispensable para los investigadores y las instituciones dedicadas a la I+D. Aunque su aplicación ofrece numerosos beneficios, también conlleva riesgos y desafíos éticos que deben ser abordados de manera responsable.

La ética en la investigación científica es aún más crucial cuando se utiliza la IA. Las máquinas, por su propia naturaleza, carecen de la capacidad para discernir lo que es ético o no. Por lo tanto, es necesario establecer regulaciones y normativas que aseguren que el uso de la IA respete los estándares éticos y morales de nuestra sociedad. La IA puede generar dilemas éticos, especialmente en decisiones que afectan la vida de las personas.

Por ello, cualquier sistema de IA que pueda influir en la vida de las personas debe estar sujeto a regulaciones y supervisión para asegurar que sus acciones se mantengan dentro de los límites éticos aceptables. Investigadores e instituciones que utilizan IA deben ser conscientes de los posibles riesgos y actuar de manera ética y responsable en todo momento.

Promover la ética en el desarrollo de la IA es esencial para garantizar que esta tecnología se utilice de manera responsable en el ámbito universitario y científico. El uso adecuado de la IA en la investigación requiere adherirse a pautas y directrices éticas establecidas. Mientras se desarrollan normas específicas, es fundamental basarse en principios como la fiabilidad, justicia, equidad, supervisión y transparencia.

Además de estos principios éticos, es importante considerar otros aspectos clave como la integridad, privacidad y confiabilidad, fundamentales para asegurar la calidad y proyección de la investigación. El uso de la IA implica manejar grandes volúmenes de datos, lo que puede presentar desafíos relacionados con la confidencialidad de la información y la protección de los derechos de propiedad intelectual. En particular, en la investigación aplicada a sistemas de propiedad intelectual, es crucial establecer acuerdos de confidencialidad y salvaguardar los derechos de propiedad industrial e intelectual. Esto es especialmente relevante en la investigación, donde se maneja información sensible y se generan nuevos conocimientos.

METODOLOGÍA

Para el desarrollo del presente documento se utilizó la IA. Se desarrolló un procedimiento con las siguientes fases:

1. Se seleccionó una temática basada en el uso de la IA en las universidades, en especial en el uso ético de las herramientas, la idea nace del *Webinar*, El futuro de la Inteligencia Artificial, desarrollado por el Laboratorio de IA, del departamento de administración de la UAM A y el cuerpo académico y el cuerpo académico de negocios internacionales de la Universidad del Caribe (UNICAR).
2. La idea se convierte en inquietud en donde se pretende establece la perspectiva ética de los estudiantes de la UNICAR sobre el uso de la IA.
3. Se diseña la siguiente *prompt* que es la que acompañara a el desarrollo del documento *"El uso ético de la inteligencia artificial en las actividades académicas universitarias percepción del estudiante"*.
4. La instrucción antes mencionada se desarrolla en dos procesos diferentes: Se realiza una búsqueda tradicional y ortodoxa de búsqueda de información en donde las palabras clase son ética, IA, Uso, actividades universitarias. Con el *prompt* diseñado, se dan instrucciones a la IA denominada *AITOR*, en donde las instrucciones adicionales son el número de cuartillas el documento solicitado (30pag, articulo) y el formato para su desarrollo (APA 7).
5. Con los documentos obtenidos de los dos procesos, antes mencionados Aitor 19 cuartillas, tradicional 32 cuartillas con sus respectivas citas y referencia bibliográficas. En una siguiente etapa se le complementan los dos textos, se crea una plática nueva con la *Chatgtp* versión gratuita y se le indica que platicaremos del uso de la IA en las actividades académicas en las universidades, se establece que se le proporcionara una información para que sintetice y homogenice la información pro-

porcionada, se eliminan en los dos textos la bibliografía, indicándole que respete las citas textuales.

El primer resultado por ser una versión gratuita no puede procesar tanta información, nos pregunta para que se pretende realizar esa actividad, se explica que se espera construir un documento sintetizado y con redacción homogenizada a lo que indica que pude ayudar con la solicitud si se secciona la información de 2500 como máximo, se procede a seccionar el documento en las proporciones que respeten la indicación del número de páginas y que no dividan una idea, un párrafo o un tema, por lo que la instrucción se tiene que repetir 8 ocasiones al finalizar, la IA en las mismas 8 secciones entrega una síntesis homogenizada y respeta las citas de los párrafos correspondientes.

6. Se procede a realizar una verificación puntual de las citas y las ideas generales del documento, se identifica que un número de ellas fueron eliminadas en la síntesis, se establece que las citas y referencias no utilizadas, no cambian, modifican o afectan al contenido del documento.

7. La lectura apunta a la necesidad de una información adicional que complemente la información presentada, la cual solo se puede obtener de preguntar a una fuente primaria, por lo que se proceda a realizar un cuestionario; mismo que se solicita a *Chatgtp* en la misma conversación en donde se ha desarrollado el documento, la IA sugiere un bloque de preguntas, se seleccionan de las propuestas realizadas, un conjunto de 7 preguntas y se construye un cuestionario que se puede levantar en línea con herramientas de *Google*.

8. Los cuestionarios se levantan en dos áreas distintas de la UNICAR turismo y negocios de nivel licenciatura y posgrado; se realizan 197 encuestas en total.

9. Los datos proporcionados se bajan a una hoja de cálculo donde se construyen las gráficas correspondientes con la información recopilada, se elaboran las redacciones describiendo los resultados obtenidos.

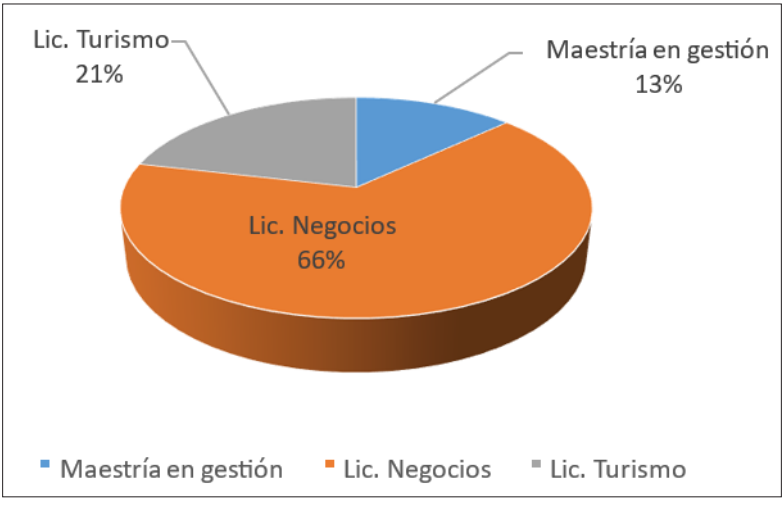
10. Esas redacciones con los datos obtenidos se introducen al *Ghatgt*, *Jenni* y *Aitor* en conversaciones nuevas, para que redacte conclusiones en base a los datos, los resultados se muestran en el cuerpo del documento junto con la conclusión generadas por los investigadores que desarrollaron el presente documento.
11. Se presenta un cuadro comparativo de los resultados. Y recomendaciones finales.

RESULTADOS

A continuación, se presentan los resultados obtenidos de las encuestas realizadas.

Fueron encuestados 192 estudiantes, 25 pertenecientes a la Maestría en Gestión de proyectos, 126 estudiantes de la división de Negocios e Innovación y 41 de la división de Turismo Sustentable y Gestión Hotelera (ver **Gráfico 1**).

Gráfico 1. Universidad.

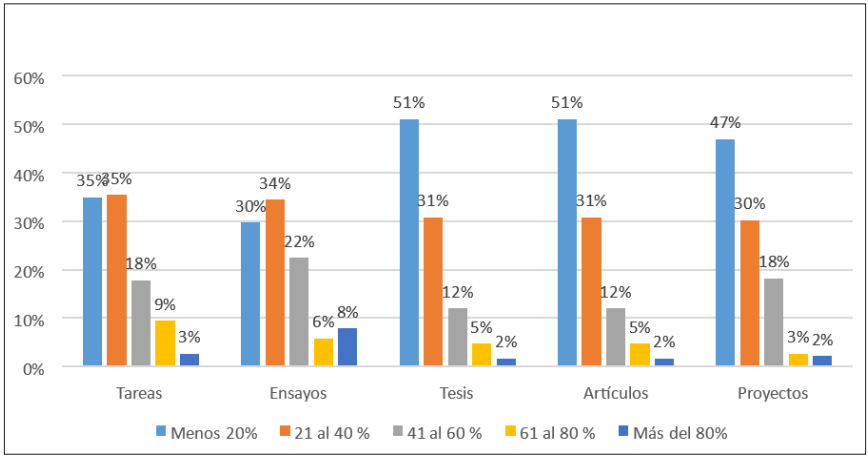


Fuente: elaboración propia.

Ética en la elaboración los alumnos encuestados

A continuación, se presentan los resultados obtenidos de manera agrupada de los estudiantes de Maestría, Negocios y Turismo (ver **Gráfico 2**).

Gráfico 2. Investigación académica.



Fuente: elaboración propia.

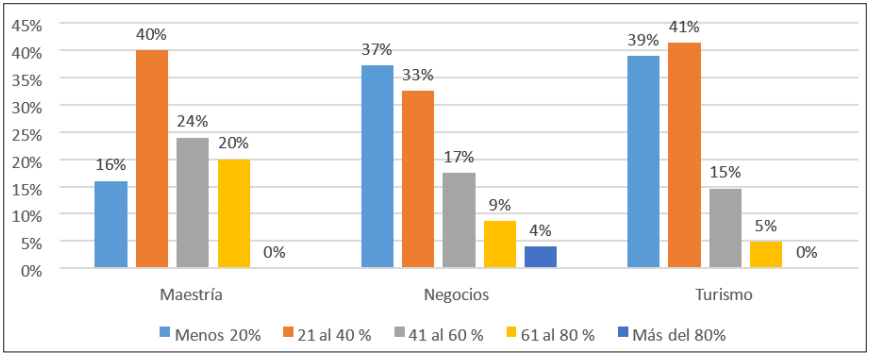
Los gráficos demuestran que, en la elaboración de una Tesis y de un Artículo, la muestra de estudiantes (82 %) percibe que es ético el hacer uso de la IA hasta un 40 %, seguidos de los Proyectos (77 %) y en menor medida las tareas (70 %) y los ensayos (64 %).

Ética en la elaboración de tareas con IA

A continuación, se presentan los resultados obtenidos de la pregunta: "*¿Hasta qué porcentaje consideras es ético hacer uso de la IA en la elaboración de una TAREA?*" (ver **Gráfico 3**).

Los gráficos demuestran que, en el caso de una tarea, los estudiantes de Turismo (80 %) son los que perciben que es ético el hacer uso de la IA hasta un 40 %, seguidos de los estudiantes de Negocios (71 %) y en menor medida los estudiantes de maestría (51 %).

Gráfico 3. *Tarea.*

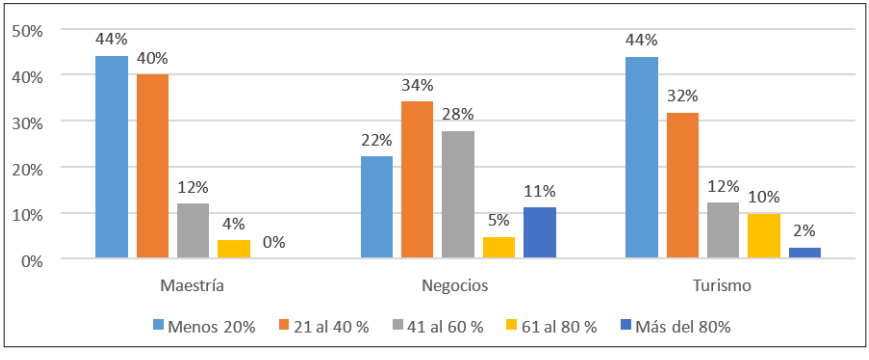


Fuente: elaboración propia.

Ética en la elaboración de ensayos con IA

A continuación, se presentan los resultados obtenidos de la pregunta: "*¿Hasta que porcentaje consideras es ético hacer uso de la IA en la elaboración de un ENSAYO?*" (ver **Gráfico 4**).

Gráfico 4. *Ensayos.*



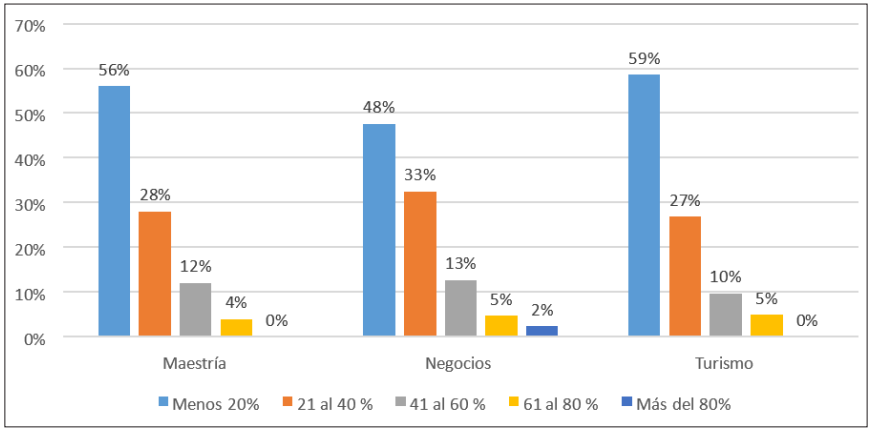
Fuente: elaboración propia.

Los gráficos demuestran que, en el caso de un ensayo, los estudiantes de maestría (el 84 %) son los que mayormente perciben que es ético el hacer uso de la IA hasta un 40 %, seguidos de los estudiantes de Turismo (el 76 %) y en menor medida los estudiantes de Negocios (56 %).

Ética en la elaboración de tesis con IA

A continuación, se presentan los resultados obtenidos de la pregunta: *"¿Hasta que porcentaje consideras es ético hacer uso de la IA en la elaboración de una TESIS?"* (ver **Gráfico 5**).

Gráfico 5. *Tesis.*



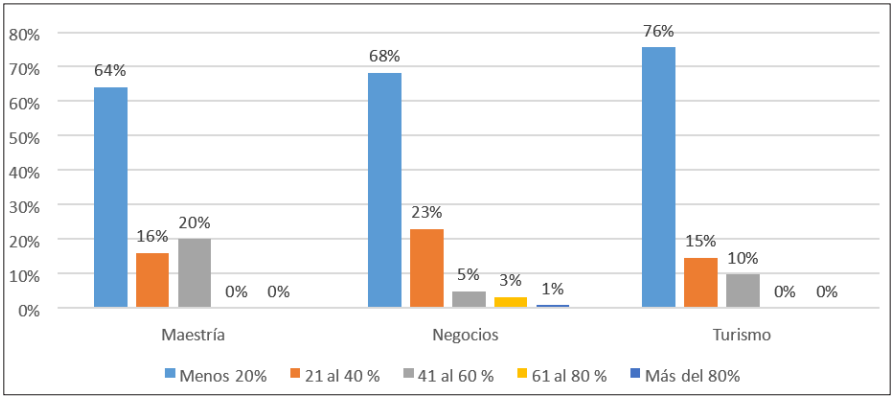
Fuente: elaboración propia.

Los gráficos demuestran que, en el caso de un Tesis, los estudiantes de Turismo (86 %) son los que mayormente perciben que es ético el hacer uso de la IA hasta un 40 %, seguidos de los estudiantes de Maestría (84 %) y en menor medida los estudiantes de Negocios (81 %).

Ética en la elaboración de artículos con IA

A continuación, se presentan los resultados obtenidos de la pregunta: "*¿Hasta que porcentaje consideras es ético hacer uso de la IA en la elaboración de un ARTÍCULO?*" (ver **Gráfico 6**).

Gráfico 6. *Artículos.*



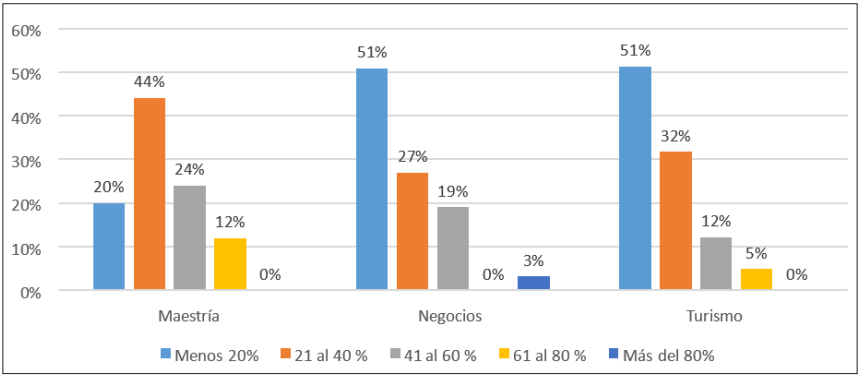
Fuente: elaboración propia.

Los gráficos demuestran que, en el caso de un Artículo, los estudiantes de Turismo (91 %) al igual que los estudiantes de Negocios (91 %), son los que mayormente perciben que es ético el hacer uso de la IA hasta un 40 %, seguidos de los estudiantes de Maestría (80 %).

Ética en la elaboración de proyectos con IA

A continuación, se presentan los resultados obtenidos de la pregunta: "¿Hasta que porcentaje consideras es ético hacer uso de la IA en la elaboración de un PROYECTO?" (ver **Gráfico 7**).

Gráfico 7. *Proyectos.*



Fuente: elaboración propia.

Los gráficos demuestran que, en el caso de un Proyecto, los estudiantes de Turismo (83 %) son los que perciben ético el hacer uso de la IA hasta un 40 %, seguidos de los estudiantes de Negocios (78 %) y en menor medida los estudiantes de Maestría (64 %).

DISCUSIÓN Y CONCLUSIONES

Los estudiantes encuestados en diversos departamentos y grados académicos consideran ético el uso de la IA en trabajos académicos, siempre que no supere el 40 %. Esta percepción refleja una postura equilibrada, reconociendo los beneficios de la IA sin comprometer la integridad del trabajo. La aceptación varía según el tipo de trabajo: es mayor en la elaboración de artículos (91 %) y tesis (86 %), mientras que es menor en ensayos (64 %). Los estudiantes de Turismo son los que más aceptan el uso de la IA en tesis y proyectos (86 % y 83 %, respectivamente), mientras que los de maestría son más reticentes, especialmente en ensayos (51 %). Los proyectos y tesis permiten más el uso de IA, en comparación con tareas y ensayos, que requieren un mayor desarrollo propio.

Se recomienda establecer guías claras que delimiten el uso de la IA en cada tipo de trabajo, como el límite del 40 % que sugieren los estudiantes. También es crucial promover la reflexión ética y desarrollar un enfoque crítico en torno al uso de la IA. Las instituciones deben ajustar sus pautas de evaluación para ser transparentes en cuanto al uso permitido de la IA y ofrecer capacitación sobre su uso responsable. Los resultados indican una actitud abierta, pero cautelosa, hacia la IA, lo que subraya la necesidad de regular su implementación en entornos académicos alineados con los valores éticos educativos.

En resumen, la importancia de promover la ética en el desarrollo de la IA y en los procesos actividades de la vida universitaria no puede subestimarse (Franganillo, 2023). El uso responsable de la IA requiere cumplir con pautas éticas establecidas y establecer regulaciones que aseguren que esta tecnología se utilice de manera ética y responsable. Solo así podremos aprovechar plenamente los beneficios de la IA sin comprometer los valores y principios éticos que sustentan nuestra sociedad.

La metodología utilizada en este proyecto fue clave para obtener los datos necesarios sobre la percepción ética de los estudiantes en cuanto al uso de la inteligencia artificial en actividades académicas. El proceso siguió una serie de fases bien estructuradas, desde la selección de una temática relevante basada en eventos académicos, hasta el uso de herramientas avanzadas como la IA para complementar la investigación.

La importancia de este enfoque reside en su combinación de métodos tradicionales y modernos. Por un lado, se realizaron búsquedas ortodoxas de información sobre la ética y la IA en el contexto universitario. Por otro lado, se aprovechó la capacidad de la inteligencia artificial, en este caso mediante el uso de un *prompt* específico en el sistema *AITOR,* lo que permitió generar un documento amplio y conforme a normas académicas.

Este enfoque metodológico demuestra cómo las tecnologías emergentes pueden integrarse eficazmente en el proceso de investigación para generar contenido relevante y riguroso, sin descuidar los principios éticos que guían la academia.

Además, subraya la flexibilidad y adaptabilidad del proceso investigativo en el ámbito universitario, al combinar fuentes tradicionales con herramientas de IA para abordar temas complejos como la ética en la inteligencia artificial.

REFERENCIAS BIBLIOGRÁFICAS

Aghemo, R. (2020). Artificial Intelligence and machine learning.

Albrecht, K. (1992). Cómo solucionar la crisis de liderazgo en la gerencia intermedia. Editorial Paidós.

Aluja Banet, T. (2001). La minería de datos, entre la estadística y la inteligencia artificial. Questiió: Quaderns d'Estadística, Sistemes. *Informática e Investigación Operativa, Vol. 25*, Nº. 3, 2001, págs. 479-498. https://upcommons.upc.edu/handle/2099/4162

Angulo, K. (2024). Formación para el servicio y la empatía en la universidad. Dominio de las Ciencias. https://dominiodelasciencias.com

Aragona, B. (2022). Tipos de big data y análisis sociológico: usos, críticas y problemas éticos. Empiria: Revista de metodología de ciencias sociales, 15-30.

Arévalo, J. (2016). El "Internet de las cosas...". Desiderata, 24 - 25. https://dialnet.unirioja.es/servlet/autor?codigo=5878308

Arrieta y Perez, J. (2022). Tendencias laborales y el futuro del trabajo por medio de la robotización, digitalización e inteligencia artificial en España. Razón Critica, 12.

Barro, S. (2017). Automatización inteligente. Integración & comercio, (42), 290-303. https://documat.unirioja.es/servlet/articulo?codigo=6174378

Borja Quevedo, M. (2020). Análisis de las Herramientas de Procesamiento de lenguaje natural para estructurar textos medicos. Universidad de Navarra. https://dadun.unav.edu/bitstream/10171/60003/1/Quevedo%20Marcos_Borja_9044 47_MII.pdf

Cantero Galeano, G. L. (2021). La inteligencia artificial en los procesos de selección.

Carrasco, J. (04 de 06 de 2021). El uso de herramientas de inteligencia artificial dentro de los recursos humanos: un análisis para el proceso de re-

clutamiento en cautela SA y Sensority (2015-2020). https://repositorio.puce.edu.ec/items/98039d7b-00fb-4c97-a3da-de4d8f112bde

Celi-Parraga, R. (2021). Técnicas de procesamiento de lenguaje natural en la inteligencia artificial conversacional textual. AlfaPublicacione. https://alfapublicaciones.com/index.php/alfapublicaciones/article/view/123

Cohen, E. (2007). Information and Beyond. Informing Science.

Corvalán, J. (2018). Inteligencia artificial: retos, desafios y oportunidades. Revista de Investigações Constitucionais, 295 - 316. https://www.scielo.br/j/rinc/a/gCXJghPTyFXt9rfxH6Pw99C/

Corvalán, J. (2019). El impacto de la inteligencia artificial en el trabajo. Revista de direito economico e socioambiental, 35 - 51.

Corvalán, J. (2019). El impacto de la inteligencia artificial en el trabajo. Direito Economico e Socioambiental, 35 - 51.

Czosseck, C., Tyugu, E., y Wingfield, T. (2011). Artificial Intelligence in Cyber Defense. In 3rd International Conference on Cyber Conflict.

Díaz-Ramírez, J. (2021). Aprendizaje Automático y Aprendizaje Profundo. Ingeniare., 2(29), 180-181.

DIGITALES. (2021). Revista Latinoamericana de Derechos Humanos Volumen 25. Gobierno de España. Carta de Derechos Digitales.

Dux, G. (2023). Creer, evaluar, saber: lógica y legitimidad en el estudio del derecho de Max Weber. Revista Científica de la Facultad de Derecho y Ciencias Sociales y Políticas, (2(1)), 1-40.

Erazo-Castillo, J. (2023). Auditoria del futuro, la prospectiva y la inteligencia artificial. Revista Digital Novasinergia, 105-119. senescyt.gob.ec

Escalera, D. l. (2001). Visión por computador: fundamentos y métodos. Pearson, Prentice Hall.

Ferre, M. J. S., Ruiz, F. G., y Palma, M. Á. Z. (2022). Neurociencia del comportamiento: del laboratorio a la vida real. Aula Magna Proyecto clave McGraw Hill.

file:///C:/Users/UNICARIBE/Downloads/adminojs,+UAZUAY-COLOQUIO65+ARTICULO-12.pdf

Floridi, L. (2018). Soft Ethics and the Governace of the digital. Philosophy and Technology, 1-8.

Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. methaodos. *Revista de ciencias sociales*. unirioja.es

Frey y Osborne, M. (2013). The future of employment. University of oxford, 44-45. https://www.sciencedirect.com/science/article/abs/pii/S0040162516302244

Gallent-Torres, C., Zapata-González, A., y Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa, 29(2), 1-21.

Geetha, R., y Bhanu, S. R. D. (2018). Recruitment through artificial intelligence: a conceptual study. International Journal of Mechanical Engineering and Technology, 9(7), 63-70.

González, M. (2017). Regulación legal de la robótica y la inteligencia artificial:retos de futuro. Revista Jurídica de la universidad de león, 25 - 50. https://dialnet.unirioja.es/servlet/articulo?codigo=6252788

González, W. (2023). Habermas: Tras un diagnóstico, un principio universal: la razón comunicativa. Lógoi. Revista de Filosofía, (44), 158-179.

Gupta, V. (2018). What is Artificial Intelligence. https://www.linkedin.com/pulse/what-artificial-intelligence-vipin-gupta/?published=t

HAI. (2022). Robotics and AI. Standford University. https://hai.stanford.edu/sites/default/files/2022-06/HAI_IndustryBrief5_v11.pdf

Hardy, T. (2001). IA (inteligencia artificial). Polis: Revista Latinoamericana, (2), 18.

https://bdbiblioteca.universidadean.edu.co:2083/docview/1924031267?accountid=3 4925

https://mediatum.ub.tum.de/doc/1578719/r9l8ty4fsz-6chzfpmhv63unj8.201230_april .pdf

Huang, E.; Bárcena, M.; Bonilla, A.; y Martínez, J. (2020). *Inteligencia Artificial en Modelos Gerenciales. Proyecto Final, Universidad de Panamá*. https://www.researchgate.net/publication/346432705_Inteligencia_Artificial/link/5fc18ffe458515b7977baec9/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Hueso, L. C. (2019). Riesgos e impactos del Big Data, la inteligencia artificial y la robótica: enfoques, modelos y principios de la respuesta del derecho. Revista general de Derecho administrativo, 50, 1-37.

Jarek y Mazurek, G. (2019). Marketing and artificial intelligence. Central european business review, 46-55.

Kriele, M. (2022). Introducción a la teoría del Estado: Fundamentos históricos de la legitimidad del Estado constitucional democrático. Santiago de Chile: Ediciones Olejnik.

Lain Moyano, G. (2021). Responsabilidad en inteligencia artificial. Ars Iuris Salmanticensis, 197 - 232. https://doi.org/10.14201/AIS202191197232

Latorre Pujol, M. (2022). *Adaptación de la normativa europea de responsabilidad civil a la era digital y, en particular, a la inteligencia artificial: evaluación de la idoneidad de las nociones de producto, productor, defecto y daño contenidos en la Directiva 85/374/CEE.* https://repositori.upf.edu/handle/10230/53875

Laverde C. y Gutiérrez, L. (2024). Protección de derechos de autor en obras audiovisuales en la industria musical en redes sociales como Tik Tok. Politicas, Practicas y Desafios. https://unilibre.edu.co

Lee, K. (2016). Artificial intelligence, automation, and the economy. Executive office of the president, 18-19.

Madrid, I. (2018). Internet of Thinghs en la vida cotidiana. Documentos de trabajo.

Maisueche Cuadrado, A. (2019). *Utilización del Machine Learning en la industria 4.0. Trabajo final de Maestria, Universidad de Valladolid. http://uvadoc.uva.es/handle/10324/37908*

Manguay, M. J. (2023). La inteligencia artificial en el dominio del sistema educativo: un estudio cuantitativo desde el desarrollo y evolución de la matemática. Polo del Conocimiento (85), 1561-1572.

Martínez, R. (11 de 09 de 2020). La Importancia de la Inteligencia Artificial en la Seguridad Cibernética. https://prcrepository.org/xmlui/handle/20.500.12475/1069

Masoitti, M. (2022). La insuficiencia de la causalidad como presupuesto de la responsabilidad civil en los daños producidos por la robótica y los sistemas autónomos. *Revista de Derecho Privado,* (42), 215-260.

Mayershofer, C. (2021). Towards an artificial perception framework for autonomous robots in logistics. Heidelberg, Berlin: Springer.

Medina y Martinez, A. (2020). Tecnologías en la inteligencia artificial para el marketing. Revista de producción, ciencias e investigación, 36-47.

Mendoza, D. (2020). Racismo y roles de género, conductas perpetuadas en algoritmos de inteligencia artificial. Coloquio, 131 - 155.

Mitchell, T. M. (1999). Machine learning and data mining. *Communications of the ACM, 42*(11), 30-36. https://dl.acm.org/doi/fullHtml/10.1145/319382.319388

Molina, y. C. (2023). La inteligencia artificial y la limitación al derecho a la privacidad cibernética. Revista latinoamericana de ciencias sociales y humanidades, 657 - 666. https://latam.redilat.org/index.php/lt/article/view/284

Morandin-Ahuerma, F. (2023). Principios normativos para una ética de la inteligencia artificial. Consejo de ciencia y tecnologia del estado de puebla (CONCYTEP). https://researchgate.net

Muñoz, M. M. (2016). Riesgos asociados al desarrollo de robots autónomos dotados de inteligencia artificial avanzada en contexto civil y militar. Ensayos de Filosofía, 3(1).

Naranjo, B. M. M., Izurieta, C. E. A., Tibán, L., Morrillo, C. S., y Salazar, A. J. (2023). Ética y Responsabilidad en la Implementación de la Inteligencia Artificial en la Educación. Ciencia Latina: Revista Multidisciplinar, 7(6), 28. https://dialnet.unirioja.es/servlet/articulo?codigo=9481573

Notimex. (2017). Inteligencia artificial gana terreno en servicio de atención a clientes.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., ... y Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1356. https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1356

Ochoa, J. (2014). Reconocimiento de patrones. https://goo. gl/JbY654.

Osejo-Bucheli, C. (2021). Hacia una filosofía de la administración centrada en el carácter. Revista de filosofia(98), 715. http://www.scielo.org.ve/scielo.php?pid=0798-1171&script=sci_serial

Partida Tapia, M. (1995). Percepción Computacional. Polibits, 20 - 23. file:///C:/Users/UNICARIBE/Downloads/3717-7624-1-PB%20(6).pdf

Penrose R. (2005). *La nueva mente del emperador.* México: Fondo de cultura económica.

Pérez Carvajal, M. F., Burgos Alba, M. E., García Hernández, L. D., y Díaz Olarte, D. (2020). Estudio descriptivo de la percepción de usuarios frente al uso de inteligencia artificial en los procesos de atención al cliente.

Pérez, O. (2024). Aplicaciones Tecnológicas y Recursos Turísticos. Ciencia Latina Revista Científica Multidisciplinar. https://ciencialatina.org

Porcelli, A. (2020). La inteligencia artificial y la robótica: sus dilemas sociales, éticos y jurídicos. Epub, 49 - 105. https://doi.org/10.32870/dgedj.v6i16.286

Preda, A., y Kuk, A. (2024). Lecciones de fútbol. El juego de ataque. Parkstone International. https://books.google.es/books?hl=es&lr=&id=cjX_EAAAQBAJ&oi=fnd&pg=PA7&dq=Preda+y+Kuk,+A.+(2024).+Lecciones+de+f%C3%BAtbol.+&ots=d_1vsADhps&sig=ZpNBMgR1SIktOXNGeEU2Ga6iduM#v=onepage&q=Preda%20y%20Kuk%2C%20A.%20(2024).%20Lecciones%20de%20f%C3%BAtbol.&f=false

Quinto, N. (2021). La inteligencia artificial y la toma de decisiones generales. Revista de investigación valor agregado, 52 - 69. https://www.researchgate.net/publication/357188928_La_inteligencia_artificial_y_la_toma_de_decisiones_gerenciales

Reynaga, R. (2009). Introducción al reconocimiento de patrones. Revista de difusión cultural y científica, 41 - 44. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2071-081X2009000100005&lng=es&tlng=es

Rodney, H. (2019). The artificial itelligence recruitment process. Addleton Academic Publishers.

Rodriguez-Sánchez, A. (2024). La posibilidad de explicación cientificas a partir de modelos basados en redes neuronales artificiales. Revista colombiana de filosofia de la ciencia, 48, 24. https://unbosque.edu.co

Rouhiainen, L. (2018). Inteligencia artificial. Madrid: Alienta Editorial.

Ruiz, A. (2018). Neo-reestructuracion del empleo. Revista de ciencias gerenciales, 32 - 45.

Ruiz, C., y Stevens, M. (2024). Big data en la contaduría pública desde la ética profesional. https://repositorio.uan.edu.co/server/api/core/bitstreams/02a85831-3e50-4bba-ba5b-22c5b094937b/content

Salmerón Moreira, Y. M., Luna Alvarez, H. E., Murillo Encarnacion, W. G., y Pacheco Gómez, V. A. (2023). El futuro de la Inteligencia Artificial para la educación en las instituciones de Educación Superior. Conrado, 19(93), 27-34. http://scielo.sld.cu/scielo.php?pid=S1990-86442023000400027&script=sci_arttext&tlng=pt

Sancho Escrivá, J. V., Fanjul, C., de la Iglesia Vayá, M., Montell, J. A., y Escarti, M. J. (2020). Aplicación de la inteligencia artificial con procesamiento del lenguaje natural para textos de investigación cualitativa en la relación médico-paciente con enfermedad mental mediante el uso de tecnologías móviles.

Sterne, J. (2017). Artificial intelligence for marketing: practical applications. John Wiley & Sons.

Tello, J. C., e Informáticos, S. (2007). Reconocimiento de patrones y el aprendizaje no supervisado. Universidad de Alcalá, Madrid.

Toboso y Aparicio, M. (2019). Entornos de funcionamiento robotizados. Dilemata, 171185.

Torres, A. (2012). Procesamiento Del lenguaje. Revista Matices Tecnológicos, 23 - 27. https://dialnet.unirioja.es/servlet/articulo?codigo=8994061

Verdegay, J. L., Lamata, M., Pelta, D., y Cruz, C. (2021). Inteligencia artificial y problemas de decisión: la necesidad de un contexto ético. Suma de Negocios, 12(27), 104-114. https://doi.org/10.14349/sumneg/2021.v12.n27.a2

Vivas Fonseca, J. (2022). Bias & Artificial Intelligence. Pontificia Universidad Javeriana, 1-9.https://repository.javeriana.edu.co/bitstream/handle/10554/61532/attachment_0_Tesis-Juan-Vivas-Master-AI---Segunda-Entrega.pdf?sequence=1&isAllowed=y

Westhues, A. (2020). *Sistemas de armas autónomas letales: ¿autónomas o automatizadas?* https://e-spacio.uned.es/entities/publication/01b41360-69d0-4f4c-bc48-a2dc143c1450/full

Yeja, A. H., de la Rosa Pasteur, J., y Huice, O. R. (2013). Aplicación de técnicas de Inteligencia Artificial en la Seguridad Informática: un estudio. Inteligencia Artificial, 51, 1-8.

CAPÍTULO VII

ASISTENTES VIRTUALES: TRANSFORMACIÓN TECNOLÓGICA Y DESAFÍOS ÉTICOS

FÁTIMA CECILIA PÉREZ PONCE[1]

RESUMEN

Los asistentes virtuales, como *Alexa, Siri* y *Google Assistant,* están cambiando radicalmente la forma en que interactuamos con la tecnología en la vida diaria. Estas herramientas impulsadas por inteligencia artificial no solo facilitan la realización de tareas, como la gestión del tiempo y el control del hogar, sino que también son capaces de aprender de los usuarios y anticiparse a sus necesidades.

Sin embargo, a medida que estos sistemas asumen mayor responsabilidad en la toma de decisiones, surgen preocupaciones éticas relacionadas con la pérdida de autonomía y la dependencia tecnológica. Además, la integración de tecnologías avanzadas como las interfaces cerebro-computadora y la inteligencia artificial adaptativa está ampliando el alcance de los asistentes virtuales, permitiendo una personalización sin precedentes.

Estas innovaciones, si bien mejoran la eficiencia y comodidad, plantean cuestiones filosóficas sobre el Transhumanismo y los límites de la humanidad al delegar cada vez más tareas y decisiones en la tecnología. Asimismo, se destaca la importancia de regular estas tecnologías para garantizar que operen de manera ética, respetando la privacidad de los usuarios y evitando la manipulación o pérdida del control humano.

[1] Licenciada en Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

Palabras clave: Asistentes virtuales, Dependencia tecnológica, Ética.

INTRODUCCIÓN

Los asistentes virtuales han dejado de ser simples herramientas tecnológicas para convertirse en agentes transformadores de la vida cotidiana. A través de dispositivos como *Alexa, Siri* y *Google Assistant,* millones de personas interactúan diariamente con sistemas de inteligencia artificial que facilitan tareas rutinarias, gestionan calendarios, controlan dispositivos del hogar y proporcionan respuestas inmediatas a sus necesidades. Sin embargo, esta creciente dependencia de los asistentes virtuales plantea preguntas fundamentales sobre nuestra autonomía, la capacidad de delegar decisiones en máquinas, y los límites entre lo humano y lo tecnológico.

Este capítulo explora cómo los asistentes virtuales no solo están revolucionando la interacción humano-máquina, sino que también están planteando complejas cuestiones éticas y filosóficas. A medida que la inteligencia artificial integrada en estos dispositivos se vuelve más sofisticada, aprendiendo y adaptándose a las preferencias de los usuarios, surge una nueva forma de relación entre humanos y tecnología, donde las decisiones proactivas de las máquinas pueden influir en nuestras vidas de manera profunda y constante.

Además, la integración de estas tecnologías con sistemas avanzados, como las interfaces cerebro-computadora y la inteligencia artificial adaptativa, abre nuevas posibilidades y desafíos. Estas innovaciones no solo prometen mejorar la eficiencia y personalización de los asistentes virtuales, sino que también están difuminando los límites de la autonomía humana, empujando el concepto de transhumanismo, donde la tecnología podría convertirse en una extensión de nuestras capacidades mentales y físicas.

Finalmente, se analizarán las implicaciones éticas de esta humanización de la tecnología, donde el riesgo de dependencia tecnológica, la privacidad y la manipulación del comportamiento son cuestiones clave. A través de la exploración de estos temas, se hace evidente la

necesidad de una regulación adecuada que proteja a los usuarios y garantice que estas tecnologías operen de manera ética y responsable.

LOS ASISTENTES VIRTUALES Y SU EVOLUCIÓN

Los asistentes virtuales, también conocidos como asistentes digitales o agentes inteligentes, son sistemas de software diseñados para interactuar con los usuarios a través del procesamiento de lenguaje natural (NLP), reconocimiento de voz y aprendizaje automático. Estos asistentes están creados para realizar tareas específicas, como responder preguntas, gestionar agendas, controlar dispositivos inteligentes y ofrecer recomendaciones personalizadas.

Según Mitchell (2019), *"los asistentes virtuales, impulsados por algoritmos avanzados de procesamiento de lenguaje natural y aprendizaje automático, han evolucionado significativamente desde sus primeras versiones, permitiendo una interacción más natural y eficiente con los usuarios"*. Esto refleja cómo la tecnología detrás de los asistentes ha avanzado para ofrecer una experiencia de usuario más fluida y efectiva.

Algunos aspectos que caracterizan el funcionamiento de los asistentes virtuales son:

- Procesamiento de Lenguaje Natural (NLP): Los asistentes virtuales utilizan el NLP para entender el significado de las solicitudes del usuario. Como señala Russell (2019), *"el NLP convierte el lenguaje humano en datos que las máquinas pueden procesar, permitiendo a los asistentes virtuales entender y responder de manera coherente"*. Este proceso es fundamental para facilitar una interacción más intuitiva, donde el usuario puede comunicarse en su lenguaje natural sin necesidad de usar comandos específicos.
- Reconocimiento de Voz: Esta tecnología convierte el habla del usuario en texto, que luego es interpretado por el asistente. Zheng y otros (2020) destacan que *"los avances en el reconocimiento de voz han mejorado significativamente la precisión en la conversión de audio a texto, lo que facilita interacciones más*

naturales con los asistentes virtuales". Gracias a estos avances, los asistentes virtuales pueden entender mejor diferentes acentos y modismos, lo que enriquece la experiencia del usuario.

- Aprendizaje Automático (*Machine Learning*): Los asistentes virtuales emplean aprendizaje automático para adaptar sus respuestas y mejorar con el tiempo. Jordan (2018) afirma que *"el aprendizaje automático permite a los asistentes virtuales aprender de las interacciones previas, optimizando sus respuestas y servicios en función de las necesidades y preferencias del usuario"*. Esto significa que los asistentes se vuelven más eficaces y relevantes para cada usuario a medida que se utilizan, creando una relación más personalizada.

La evolución de los asistentes virtuales ha sido notable desde sus inicios. En sus primeras versiones, los asistentes eran simples herramientas de automatización capaces de realizar tareas básicas. Sin embargo, con el tiempo, han incorporado capacidades avanzadas gracias a los avances en inteligencia artificial (IA) y procesamiento de lenguaje natural.

En el artículo de Lee y Kim (2021), se menciona que *"los asistentes virtuales han pasado de ser simples herramientas de automatización a sistemas complejos capaces de entender y responder a una gama amplia de consultas y comandos, reflejando un notable progreso en el campo de la inteligencia artificial"*. Esta transición ha permitido que los asistentes virtuales se integren de manera más efectiva en la vida cotidiana, adaptándose a las necesidades individuales de los usuarios.

Los asistentes virtuales han avanzado rápidamente desde los primeros experimentos en la década de 1990 hasta las versiones actuales, como *Siri, Alexa* y *Google Assistant,* que son capaces de realizar múltiples tareas a la vez y que se encuentran entre los asistentes virtuales más populares.

Según Russell (2019), *"la clave para el futuro de los asistentes virtuales radica en la personalización y la capacidad de aprender continuamente de los datos y comportamientos del usuario"*. Esto resalta la importancia de la adaptabilidad en un mundo donde las necesidades del usuario cambian constantemente.

LOS ASISTENTES VIRTUALES MÁS POPULARES

Los asistentes virtuales han ganado popularidad debido a su capacidad de simplificar tareas cotidianas y mejorar la interacción humano-máquina. Son una de las manifestaciones más accesibles y extendidas de la inteligencia artificial (IA) en la vida cotidiana. Entre los más reconocidos a nivel global se encuentran *Alexa de Amazon, Siri de Apple* y *Google Assistant de Google*.

Cada uno de estos asistentes tiene características distintivas que los hacen únicos, aunque todos comparten un núcleo tecnológico basado en el procesamiento de lenguaje natural (NLP) y el aprendizaje automático. Están diseñados para ayudar a los usuarios a realizar tareas mediante comandos de voz, integrándose en una variedad de dispositivos, desde Smartphones y computadoras hasta audífonos y electrodomésticos inteligentes.

Alexa (Amazon)

Alexa, el asistente virtual de *Amazon*, fue lanzado en 2014 como parte del sistema *Amazon Echo* y rápidamente se convirtió en uno de los asistentes más populares en el ámbito de los hogares inteligentes. Su capacidad para interactuar con dispositivos inteligentes del hogar ha sido clave en su adopción masiva. *Alexa* permite a los usuarios realizar diversas tareas, como controlar dispositivos del hogar inteligente, escuchar música, obtener información y realizar compras en línea, todo mediante comandos de voz.

Entre las funciones más populares de *Alexa* se encuentran el control del hogar inteligente, que permite gestionar luces, termostatos, cerraduras de puertas y cámaras de seguridad. Esto ha hecho que tener un hogar inteligente sea una opción accesible que permite una integración fluida y fácil de configurar. Además, *Alexa* ayuda a gestionar la vida diaria de los usuarios, permitiendo programar alarmas, realizar listas de tareas y hacer pedidos en línea.

Según un informe de Voicebot.ai, *"Alexa se integra con más de 100,000 dispositivos diferentes y tiene más de 100,000 'skills' o aplica-*

ciones de terceros, lo que le permite ofrecer una experiencia extremadamente personalizada" (Voicebot.ai, 2021).

Como señalan Roffarello y De Russis,*"Alexa ha sido un catalizador clave para el desarrollo del hogar inteligente, proporcionando un punto de control unificado para una amplia gama de dispositivos IoT"* (Roffarello y De Russis, 2020). Su plataforma abierta también permite a los desarrolladores crear nuevas funcionalidades a través de las llamadas *"Skills"*.

Siri (Apple)

Siri fue el primer asistente virtual basado en inteligencia artificial en ser ampliamente adoptado por los usuarios, tras su integración en los *iPhones* en 2011. Con su capacidad para realizar búsquedas en internet, programar recordatorios, enviar mensajes y hacer llamadas, *Siri* ha transformado la manera en que los usuarios interactúan con sus dispositivos *Apple*. *"La capacidad de Siri para integrarse con el ecosistema de Apple, desde iPhones hasta MacBooks, ha sido crucial para su éxito continuo"* (Brown, 2020).

Algunas de las funciones más destacables de *Siri* son la sincronización entre *iPhones, iPads, Macs, Apple Watch* y *HomePod,* facilitando el control de todo el ecosistema Apple. Según Brown (2020), *"la sincronización de Siri con otros productos de Apple refuerza su capacidad para crear una experiencia de usuario fluida y sin interrupciones"*.

Por otro lado, la función *Siri Shortcuts* permite a los usuarios crear atajos personalizados para realizar tareas complejas con un solo comando. Por ejemplo, con un solo comando de voz, el usuario puede activar una serie de tareas como poner el *"despertador", "leer noticias"* y *"recordatorios del día"* (Anderson, 2021, p. 60).

Siri también es conocida por su capacidad de entender comandos de lenguaje natural y su integración con *apps* nativas de *Apple*. Según Brown, *"Siri no solo actúa como un asistente personal, sino que extiende sus capacidades a todos los dispositivos de Apple, creando una red de interacción perfecta"* (Brown, 2020).

Google Assistant (Google)

Google Assistant fue lanzado en 2016 y ha logrado una rápida adopción gracias a su capacidad para ofrecer respuestas detalladas y precisas, apoyándose en el vasto motor de búsqueda de *Google*. Se caracteriza por realizar búsquedas contextuales; es decir, *Google Assistant* puede responder preguntas complejas, ofrecer resultados de búsquedas en tiempo real y proporcionar información local sobre restaurantes, tráfico y eventos.

Como señala Kurzweil (2019), *"el hecho de que Google Assistant utilice la búsqueda semántica le permite ofrecer respuestas que no solo son precisas, sino también contextualmente relevantes"*. Además, el Modo Intérprete de *Google Assistant* es capaz de traducir conversaciones en tiempo real, facilitando la comunicación entre personas que hablan diferentes idiomas.

> *"Este modo ha sido especialmente útil en entornos como aeropuertos, donde se han instalado dispositivos de Google Assistant para ayudar a los viajeros"* (Chen, 2020).

Google Assistant se diferencia por su capacidad para realizar búsquedas complejas y proporcionar respuestas precisas a través del motor de búsqueda de Google. Está presente en dispositivos *Android, Google Nest* y otros dispositivos inteligentes.

Según un estudio de Kurzweil (2019), *"Google Assistant es reconocido por su capacidad avanzada de comprensión contextual, lo que lo convierte en una herramienta poderosa para la búsqueda y la gestión de tareas"*.

Otros Asistentes Virtuales

Además de los asistentes mencionados, existen otros como *Cortana* de *Microsoft* y *Bixby* de *Samsung* que han encontrado su lugar en el ecosistema tecnológico. *Cortana,* lanzada en 2014, ha evolucionado desde su enfoque inicial en la interacción del consumidor hacia un papel más integrado en el entorno corporativo, ayudando a gestionar tareas en *Microsoft Office 365*. Esto permite a los usuarios programar reuniones,

establecer recordatorios y acceder a información relevante, mejorando así la productividad en el trabajo. Según Forrester Research (2021), *"Cortana se ha adaptado a las necesidades del mundo laboral moderno, destacando en la integración con aplicaciones empresariales y la gestión de tareas"*.

Por otro lado, *Bixby* se destaca por su integración en una variedad de dispositivos Samsung, desde teléfonos hasta electrodomésticos inteligentes. Se enfoca en realizar tareas complejas mediante comandos de voz y se integra con aplicaciones del dispositivo, permitiendo un control más profundo de la experiencia del usuario. Entre las funciones más destacadas de *Bixby* se encuentra *Bixby Visión,* que permite a los usuarios utilizar la cámara de su dispositivo para reconocer objetos, traducir texto en tiempo real y comprar productos online basados en la imagen capturada, convirtiéndola en una de las herramientas más potentes para integrar la inteligencia artificial visual en la vida cotidiana (Lee, 2020, p. 43).

Aunque Cortana y *Bixby* no tienen el mismo nivel de popularidad que sus contrapartes más reconocidas, siguen siendo relevantes en sus respectivos nichos, adaptándose a las necesidades de sus usuarios y ofreciendo funciones únicas que enriquecen la experiencia del asistente virtual.

INTEGRACIÓN DE LOS ASISTENTES VIRTUALES CON TECNOLOGÍAS AVANZADAS

Pero no solo existen asistentes virtuales como los conocemos; los avances en la inteligencia artificial han llevado a una integración sin precedentes de los asistentes virtuales con tecnologías avanzadas, como las interfaces cerebro-computadora (BCI) y sistemas de IA adaptativa. Estos desarrollos están ampliando las fronteras de lo que pueden lograr los asistentes virtuales, donde ahora existe una personalización sin precedentes y se facilita la comunicación directa entre el cerebro humano y las máquinas.

Un ejemplo significativo es la evolución de los asistentes virtuales mediante la integración de interfaces cerebro-computadora, las cuales

permiten que las personas interactúen con dispositivos y sistemas tecnológicos directamente a través de señales cerebrales.

En este sentido, los asistentes virtuales que operan en conjunto con estas interfaces podrían transformar radicalmente la forma en que las personas con discapacidades físicas interactúan con el mundo, ya que les ofrecerían un control más directo sobre dispositivos y servicios. Además, estos sistemas proporcionarían una mayor autonomía y mejorarían la calidad de vida del usuario.

Según Wolpaw y Wolpaw (2012), *"las interfaces cerebro-computadora proporcionan un nuevo canal de comunicación que es independiente de los sistemas musculares, lo que podría permitir a las personas con discapacidades motoras interactuar con asistentes virtuales y otros dispositivos inteligentes mediante el pensamiento".*

Otra cuestión importante es la personalización de los asistentes virtuales a través de la inteligencia artificial. Los asistentes, como *Alexa* o *Google Assistant,* ya están diseñados para personalizar sus respuestas en función de las preferencias del usuario. Sin embargo, con el uso de la IA adaptativa, estos asistentes pueden ir más allá de simplemente responder a comandos y aprender del comportamiento y las interacciones previas de los usuarios. A través del aprendizaje profundo, pueden ajustar sus respuestas y recomendaciones de forma más precisa y anticiparse a las necesidades individuales. De esta manera, Kurzweil (2014) afirma que *"la inteligencia artificial está desarrollando capacidades para anticipar los deseos de los usuarios, logrando que los asistentes virtuales se conviertan en verdaderos compañeros personalizados que aprenden continuamente de sus interacciones".*

Por otro lado, la inteligencia artificial adaptativa es el siguiente paso para la evolución de los asistentes. Este tipo de IA se ajusta y aprende en tiempo real a partir de las interacciones con los usuarios, lo que permite que los asistentes evolucionen junto a ellos. Utiliza grandes cantidades de datos recopilados de las interacciones cotidianas, por lo que estos sistemas pueden personalizar cada vez más sus respuestas y mejorar su precisión y utilidad con el tiempo.

En este sentido, Russell (2016), uno de los pioneros más importantes en el campo de la IA, menciona que *"la IA adaptativa permite que los*

asistentes virtuales se vuelvan proactivos en lugar de reactivos, mejorando la calidad del servicio que ofrecen al usuario, especialmente en tareas repetitivas y complejas".

Estos sistemas también pueden identificar patrones y hábitos para anticiparse a las necesidades del usuario antes de que este realice una solicitud, lo que permite una mayor eficiencia y, sobre todo, mayor comodidad. El potencial de estas tecnologías combinadas es significativo, imaginando un asistente virtual que no solo responda a comandos, sino que también pueda aprender del usuario, conectarse con su actividad cerebral y anticipar sus deseos antes de que estos sean expresados verbalmente. Este nivel de interacción humano-máquina no solo cambiaría la manera en que usamos la tecnología, sino que también plantearía profundas implicaciones éticas sobre privacidad, autonomía y dependencia.

LA INFLUENCIA DE LOS ASISTENTES VIRTUALES EN NUESTRA VIDA DIARIA: IMPACTO EN LA EFICIENCIA Y COMODIDAD

Los asistentes virtuales han cambiado profundamente la forma en que interactuamos con la tecnología y gestionamos nuestras actividades cotidianas. Estos dispositivos, impulsados por inteligencia artificial (IA), no solo ofrecen una comodidad sin precedentes, sino que también nos permiten delegar tareas repetitivas o rutinarias, aumentando nuestra eficiencia. Sin embargo, esta delegación tecnológica plantea cuestiones sobre la autonomía humana y la dependencia a estos sistemas.

Una de las formas más evidentes en las que los asistentes virtuales mejoran nuestra eficiencia es en la gestión del tiempo y de tareas diarias. Por ejemplo, *Google Assistant* no solo responde a nuestras preguntas, sino que también aprende de las interacciones y puede predecir las necesidades de los usuarios con base en patrones de comportamiento. Por ello, es sorprendente que anticipe la información incluso si nosotros no se la pedimos directamente, como, por ejemplo, darte el pronóstico del clima o del tráfico en tus rutas diarias. Esto resulta ser una maravilla de la inteligencia artificial adaptativa.

De manera que, según Kurzweil (2019), *"la IA adaptativa integrada en los asistentes virtuales permite a los usuarios gestionar calendarios, recibir recordatorios y optimizar su tiempo de manera mucho más eficiente que con herramientas manuales"* (p. 102). Al automatizar tareas como la programación de alarmas, la creación de listas de compras o la gestión de citas, se aumenta nuestra productividad y se libera tiempo para actividades más significativas. Este tipo de automatización facilita una mayor eficiencia, permitiendo que los usuarios se concentren en tareas que requieren más tiempo, reflexión y toma de decisiones.

Aquí entra en juego el hecho de que delegar responsabilidades a la inteligencia artificial (IA), según Smith (2020), *"puede generar una dependencia progresiva, donde los usuarios pierden autonomía en la toma de decisiones sobre su propio tiempo"* (p. 88).

La automatización por parte de los asistentes virtuales también se puede observar dentro de nuestro hogar, donde un ejemplo muy importante es el uso de *Alexa* y *Google Assistant,* los cuales permiten controlar dispositivos inteligentes simplemente con comandos de voz, logrando una mayor comodidad y un control más eficiente de los recursos del hogar. De hecho, el control del consumo energético mediante la IA puede ser crucial para la sustentabilidad del hogar inteligente, ya que, si por alguna razón olvidamos apagar algunos dispositivos, las luces o regular algún termostato, con tan solo dar una instrucción y programar esto una vez, no tendremos problemas en el futuro y ahorraremos bastante.

Esta capacidad de la IA para anticiparse a las necesidades del usuario y ajustarse automáticamente a nuestras actividades y el entorno que nos rodea puede ser el inicio de la humanización de la tecnología, ya que estos sistemas parecen *"comprendernos y adaptarse a nuestras preferencias personales"*. Sin embargo, aquí comienza un dilema que más adelante tocaremos: ¿hasta qué punto le estamos cediendo el control a la inteligencia artificial y los asistentes virtuales sobre nuestras decisiones?

La toma de decisiones por parte de los asistentes virtuales es un aspecto fascinante que también nos preocupa. Como se ha mencionado, estos sistemas se anticipan a nuestras necesidades a partir de datos previos. Por ejemplo, si *Siri* detecta que siempre pides comida a do-

micilio los viernes por la noche, puede sugerirte hacerlo antes de que lo pienses y darte opciones basadas en tus preferencias. Aquí entra en juego el hecho de que tengan la capacidad de tomar decisiones de forma proactiva para mejorar la eficiencia, lo que también pone en juego nuestra autonomía humana y hace más estrecha la línea entre la ayuda tecnológica y la pérdida de control por parte de los humanos.

Según Forrester Research (2020), *"los asistentes virtuales que usan IA avanzada están empujando los límites de lo que significa tomar decisiones humanas, lo que abre una serie de cuestiones éticas sobre la responsabilidad y el control"* (p. 45).

Ahora bien, si esta automatización y toma de decisiones proactiva de la inteligencia artificial la analizamos dentro del ámbito profesional, los asistentes virtuales impulsados por la Inteligencia Artificial (IA), como *Cortana* de *Microsoft*, permiten gestionar reuniones, enviar correos y darles seguimiento a las tareas sin necesidad de intervenir de forma manual. No solo se encargan de ejecutar órdenes, sino que aprenden del comportamiento de los empleados para mejorar la productividad.

Según un informe de Microsoft, la integración de asistentes virtuales impulsados por IA, como Cortana, ha demostrado mejorar la productividad laboral. Estos asistentes pueden automatizar tareas rutinarias y administrativas, lo que contribuye a una mayor eficiencia en el lugar de trabajo (Microsoft, 2021).

A medida que los asistentes virtuales asumen mayores responsabilidades, surge la duda de si estamos fomentando una dependencia excesiva. Tal y como lo argumenta Kurzweil (2019), *"la creciente dependencia de la IA para tareas diarias y la delegación de decisiones podría erosionar gradualmente la capacidad de los humanos para gestionar sus propios asuntos sin la intervención de la tecnología"* (p. 112).

Además, el hecho de que estos sistemas recopilen toda nuestra información y los datos más importantes de nuestra vida diaria nos hace reflexionar sobre otro dilema que tiene que ver con la privacidad y el uso de la información personal. Los asistentes virtuales están diseñados para aprender de nosotros y ofrecer respuestas y recomendaciones personalizadas.

TRANSFORMACIONES EN LA INTERACCIÓN HUMANO – MAQUINA

La aparición de los asistentes virtuales ha transformado de manera profunda la interacción humano-máquina, cambiando la manera en que nos relacionamos con la tecnología. Tradicionalmente, interactuamos con dispositivos tecnológicos a través de interfaces complejas, como teclados y pantallas táctiles. Aunque estas no son difíciles de usar, la inteligencia artificial (IA) y los asistentes virtuales, que se manejan con comandos de voz y el procesamiento del lenguaje natural (NLP), facilitan aún más esta interacción.

Para entender mejor este proceso, tanto el procesamiento del lenguaje natural (NLP) como el aprendizaje automático son interfaces que permiten comprender y responder a comandos verbales, eliminando la necesidad de interacciones mecánicas y simplificando el acceso a la información y los servicios. Como señala Lopatovska (2013), *"los asistentes virtuales facilitan la interacción humana con la tecnología al utilizar un lenguaje cotidiano, reduciendo la barrera entre la máquina y el ser humano"*.

Otro de los mayores impactos de los asistentes virtuales en nuestra vida diaria es la eficiencia. Al permitirnos realizar múltiples tareas de manera más rápida y sencilla, desde establecer recordatorios hasta controlar dispositivos inteligentes del hogar, el tiempo se reduce considerablemente. El usuario puede *"dialogar"* con sus dispositivos y recibir asistencia en tiempo real, lo que incrementa su productividad.

Un cambio radical que ya se ha comentado es la capacidad de estos sistemas para aprender de las interacciones pasadas y ajustar sus respuestas de acuerdo con las preferencias del usuario. Aquí entra en juego el aprendizaje profundo, ya que estos asistentes pueden hacer recomendaciones personalizadas que reflejan los gustos y necesidades individuales.

Esta personalización está profundamente vinculada con el concepto de inteligencia artificial, ya que los sistemas están diseñados para adaptarse y anticiparse a las necesidades del usuario. Evolucionan en función del comportamiento del usuario, representando el primer paso hacia una interacción más efectiva entre humanos y máquinas.

Sin embargo, surge una cuestión importante: si esta interacción humano-máquina ofrece muchos beneficios al delegar tareas y decisiones importantes, ¿estamos perdiendo la autonomía humana y cediendo el control sobre nuestras vidas? Carr (2011) argumenta que *"la dependencia creciente en asistentes virtuales para la toma de decisiones puede erosionar la autonomía del individuo, volviendo a las personas más dependientes de la tecnología"*.

LOS ASISTENTES VIRTUALES Y LA IA EN LA TOMA DE DECISIONES CRUCIALES: LA IA EN NOMBRE DEL USUARIO

Una vez que conocemos el potencial de los asistentes virtuales y de lo que puede llegar a representar la inteligencia artificial (IA) en nuestra vida, debemos recordar que estos sistemas llevan a cabo una combinación de técnicas de procesamiento de lenguaje natural (NLP), aprendizaje automático y análisis de datos para ejecutar funciones que van desde tareas simples hasta operaciones más complejas. Dicho de otra forma, recopilan datos a partir de la interacción con los usuarios mediante comandos de voz, preferencias y hábitos. A partir de esto, pueden automatizar y gestionar tareas rutinarias, como programar eventos en el calendario, enviar recordatorios o ajustar configuraciones del hogar. Esto les permite realizar recomendaciones personalizadas mediante un análisis predictivo basado en el historial de actividades del usuario.

Una vez que entendemos esto, nos podemos plantear una serie de cuestiones sobre hasta qué punto es adecuado confiar en la IA para tomar decisiones importantes en lugar de las humanas. La primera es que los asistentes virtuales pueden gestionar decisiones menores con gran eficacia; sin embargo, las decisiones críticas requieren un juicio humano mucho más complejo. Por lo que, Bostrom (2014) argumenta en *Superintelligence* que *"la capacidad de la IA para tomar decisiones críticas está limitada por su falta de comprensión completa del contexto y las implicaciones a largo plazo"*. De manera que, aunque la IA y la asistencia virtual han avanzado, no pueden estar completamente equipadas para manejar decisiones que implican riesgos significativos o valores éticos complejos del ser humano.

La segunda cuestión es la dependencia y autonomía, donde delegar decisiones a la IA puede llevarnos a una creciente dependencia de la tecnología, afectando la autonomía del usuario. Floridi (2014), en *The Fourth Revolution*, señala que *"la dependencia excesiva de la tecnología para la toma de decisiones puede erosionar nuestra capacidad para tomar decisiones autónomas e informadas".* Ante esto, debemos entender que delegar decisiones cruciales a los asistentes virtuales plantea varias implicaciones éticas, tales como la responsabilidad y el control, ya que ser responsables de las decisiones tomadas por los asistentes virtuales es una cuestión clave.

Wallach (2015), en *Moral Machines*, destaca que *"la responsabilidad por las decisiones de la IA debe recaer en los diseñadores y operadores de estos sistemas, quienes deben asegurar que las decisiones sean justas y estén alineadas con los valores humanos".*

Por otro lado, la privacidad y la seguridad son aspectos críticos, dado que los asistentes virtuales recopilan y procesan una gran cantidad de datos personales, lo que plantea riesgos para la privacidad. Un ejemplo relevante se puede ver en el informe de la Comisión Europea sobre ética en IA, que enfatiza que "*es fundamental proteger la privacidad y la seguridad de los datos para evitar abusos y mantener la confianza pública en los sistemas de IA"* (European Commission, 2021).

Es crucial que los usuarios puedan entender cómo se toman las decisiones mediante sistemas automatizados y tengan la oportunidad de cuestionarlas, especialmente cuando impactan directamente en sus vidas" (Mittelstadt et al., 2016, p. 6).

TRANSHUMANISMO Y ASISTENTES VIRTUALES: CRUZANDO LOS LÍMITES HUMANOS

Detrás de toda esta evolución de la inteligencia artificial y los asistentes virtuales que son de gran utilidad en nuestra vida diaria está el transhumanismo como un movimiento que se centra en la transformación de la condición humana mediante la integración de tecnologías avanzadas. Este enfoque sostiene que la tecnología puede y debe utilizarse para superar las limitaciones biológicas y psicológicas de los

seres humanos, con el objetivo de mejorar nuestras capacidades físicas, cognitivas y emocionales.

Como lo define el filósofo Nick Bostrom, *"el transhumanismo es la idea de que los seres humanos, mediante la tecnología, pueden y deben mejorar las capacidades humanas y trascender las limitaciones biológicas"* (Bostrom, 2005).

De manera que el transhumanismo se basa en la creencia de que la tecnología tendría el potencial de solucionar los problemas fundamentales de la existencia humana como la cura de las enfermedades, la asistencia, el envejecimiento y las limitaciones cognitivas. En este sentido, los asistentes virtuales y la inteligencia artificial (IA) desempeñan un papel importante en cómo estas tecnologías están comenzando a redefinir nuestra experiencia humana. Es un movimiento intelectual y cultural que busca transformar la condición humana mediante el uso de tecnología avanzada. Esta filosofía promueve la mejora de las capacidades humanas físicas y mentales a través de intervenciones tecnológicas, como implantes cibernéticos, interfaces cerebro-computadora y modificaciones genéticas. Los asistentes virtuales, junto con otras tecnologías emergentes, juegan un papel crucial en esta evolución.

Algunas de las mejoras en la capacidad humana que el transhumanismo nos marca dentro de los asistentes virtuales son el aumento de la cognición y la capacidad de procesamiento, el uso de interfaces cerebro-computadora y mejoras sensoriales, y la implementación de tecnología en el cuerpo humano. Lo primero ocurre cuando la integración de tecnologías avanzadas en la vida cotidiana, como los asistentes virtuales, permite a los humanos aumentar su capacidad cognitiva al facilitar el rápido acceso a la información y la gestión eficiente de tareas. Asistentes virtuales como *Siri, Alexa* y *Google Assistant* no solo proporcionan respuestas a preguntas y facilitan tareas, sino que también actúan como extensiones de nuestra memoria y habilidades mentales. Según Nick Bostrom, *"la tecnología tiene el potencial de expandir nuestras capacidades cognitivas y físicas, permitiéndonos realizar tareas con una eficacia sin precedentes"* (Bostrom, 2003).

Esta habilidad para acceder a datos en tiempo real y recibir recomendaciones personalizadas mejora nuestra capacidad para tomar decisiones y gestionar información compleja. Para entender mejor esto,

podemos citar a Ray Kurzweil, que destaca que *"los asistentes virtuales y otras tecnologías cognitivas están ampliando nuestras capacidades mentales, permitiéndonos acceder instantáneamente a una vasta cantidad de información y optimizar nuestro procesamiento de datos"* (Kurzweil, 2005).

Por otro lado, las interfaces cerebro-computadora (BCI) están avanzando rápidamente, ofreciendo nuevas formas de interacción entre el cerebro humano y los sistemas digitales. Estas tecnologías no solo prometen mejorar la comunicación y el control de dispositivos, sino también potencialmente aumentar las capacidades sensoriales y cognitivas. Como lo describe Mary Lou Jepsen, *"los BCIs podrían permitir a las personas experimentar una gama más amplia de percepciones y habilidades cognitivas al conectar directamente el cerebro con sistemas computacionales avanzados"* (Jepsen, 2019).

Y en el caso de la implementación de tecnología en el cuerpo humano, los dispositivos implantables y las prótesis avanzadas están mejorando la calidad de vida y ampliando las capacidades físicas de los individuos. La integración de estos dispositivos con la tecnología digital permite a las personas recuperar o mejorar habilidades motoras y sensoriales. El informe de la Organización Mundial de la Salud destaca que *"la tecnología implantable puede ofrecer mejoras significativas en la movilidad y la funcionalidad, así como en la comunicación y la percepción"* (Organización Mundial de la Salud, 2020).

La integración de asistentes virtuales y tecnologías avanzadas plantea preguntas sobre la fusión de la tecnología con nuestro cuerpo y mente. Esto refleja una aspiración transhumanista de mejorar la condición humana a través de la tecnología. Para entender esto, debemos analizar a la tecnología como una extensión del cuerpo y la mente; es decir, los asistentes virtuales actúan como extensiones de nuestras capacidades, permitiéndonos interactuar con la tecnología de manera más fluida. En *The Human Use of Human Beings*, Norbert Wiener señala que *"la tecnología no solo amplifica nuestras capacidades físicas, sino que también se convierte en una extensión de nuestra mente, alterando cómo pensamos y actuamos"* (Wiener, 1950).

Esta integración sugiere una nueva forma de concebir la relación entre el humano y la máquina, donde la tecnología se convierte en una extensión natural de nuestras capacidades cognitivas y físicas.

Por otra parte, el impacto en la identidad humana también es una cuestión que el transhumanismo ha analizado, ya que la interacción constante con asistentes virtuales está reformando nuestra identidad y percepción de nosotros mismos. Andy Clark, en *Supersizing the Mind,* sostiene que *"la tecnología, al convertirse en una parte integral de nuestra cognición, está reformando nuestra comprensión de lo que significa ser humano"* (Clark, 2008). Lo que nos plantea ciertas interrogantes sobre cómo la dependencia de la tecnología afecta nuestra identidad y nuestras capacidades personales.

A medida que los asistentes virtuales y otras tecnologías avanzan, surgen preguntas sobre las implicaciones éticas de estas mejoras. El transhumanismo no solo busca superar las limitaciones humanas, sino que también plantea cuestiones sobre la equidad y la justicia en el acceso a estas tecnologías. Julian Savulescu, en *Enhancing Evolution,* destaca que *"el desarrollo de tecnologías que mejoren nuestras capacidades humanas debe ir acompañado de una reflexión ética sobre cómo se distribuyen estos beneficios y cómo se manejan los riesgos asociados"* (Savulescu, 2009).

La discusión ética se debe centrar en asegurar que las tecnologías transhumanistas se desarrollen y utilicen de manera que beneficien equitativamente a la sociedad. Por lo que esta relación entre el transhumanismo y los asistentes virtuales refleja una búsqueda profunda de mejorar y expandir la experiencia humana a través de la tecnología. La integración de estos asistentes en nuestra vida cotidiana está cambiando nuestra forma de interactuar con la tecnología y redefiniendo nuestra identidad y capacidades.

HUMANIZACIÓN Y TRANSHUMANISMO: UN VÍNCULO PELIGROSO

La fusión de la tecnología con el cuerpo y la mente humana presenta varios riesgos éticos que deben ser considerados. Estos riesgos abar-

can cuestiones de equidad, autonomía y la integridad de la identidad humana.

Podemos analizar algunos riesgos y desventajas de esta fusión tecnológica y humana, donde existe cierta desigualdad en el acceso a la tecnología, puesto que la mejora en las capacidades humanas hace más notorias las desigualdades existentes. Para comprenderlo de mejor manera, como señala Hughes (2010), *"la tecnología avanzada puede exacerbar las divisiones sociales, ya que el acceso a estas mejoras puede estar limitado por factores económicos, creando una brecha entre aquellos que pueden permitirse estas tecnologías y aquellos que no"*. Esta disparidad en el acceso puede resultar en una mayor desigualdad en la calidad de vida y las oportunidades.

Ahora bien, existe el riesgo de perder la identidad y la autonomía. Según el filósofo Sloterdijk (2013), *"la incorporación de tecnologías avanzadas en el cuerpo puede diluir la noción de identidad personal y autonomía, ya que los individuos pueden llegar a depender demasiado de estas tecnologías"*. Esta dependencia puede llevar a cuestionar qué significa ser humano y cómo se define la autonomía individual. Además, nos lleva a pensar en el riesgo para la integridad de la humanidad, donde la fusión de tecnología con el ser humano también plantea riesgos para la integridad de la experiencia humana. La ética de la mejora tecnológica puede llevar a debates sobre la esencia de lo que significa ser humano, donde *"el riesgo de alterar las características fundamentales de la experiencia humana plantea preguntas sobre la moralidad de las mejoras tecnológicas y su impacto en la identidad humana"* (Bostrom y Roache, 2008).

La mejora de las capacidades humanas a través de la tecnología ofrece oportunidades emocionantes para el avance personal y colectivo. Sin embargo, la fusión del humano con la máquina plantea riesgos éticos significativos, incluyendo desigualdades sociales, la pérdida de identidad y autonomía, y la amenaza a la integridad de la experiencia humana. Es esencial abordar estos riesgos con cuidado y considerar cómo las tecnologías emergentes deben ser implementadas de manera que respeten y preserven los valores fundamentales de la humanidad.

HUMANIZACIÓN DE LA TECNOLOGÍA Y DESAFÍOS ÉTICOS: ¿AUTONOMÍA HUMANA O DEPENDENCIA TECNOLÓGICA?

La capacidad de los asistentes virtuales para aprender y adaptarse a las preferencias del usuario es un aspecto fundamental de su humanización. Están diseñados para interactuar con los usuarios de manera intuitiva y amigable, lo que facilita una experiencia más personalizada y accesible. Los asistentes virtuales recopilan datos sobre los hábitos y preferencias de los usuarios para ofrecer recomendaciones y respuestas personalizadas, lo que contribuye a una experiencia más envolvente y *"humana"*.

Estos asistentes ofrecen apoyo a una variedad de necesidades, desde la gestión de tareas diarias hasta la asistencia a personas con discapacidades. Según el informe de la Fundación para la Investigación sobre la Inteligencia Artificial, *"la integración de asistentes virtuales en dispositivos accesibles está mejorando significativamente la calidad de vida de las personas con discapacidades, proporcionando asistencia personalizada y adaptativa"* (FIA, 2022).

Sin embargo, esta humanización de la tecnología a través de los asistentes virtuales plantea varios desafíos éticos, especialmente en términos de autonomía y dependencia tecnológica. Estos desafíos incluyen la privacidad, la manipulación del comportamiento y la pérdida de control humano.

En primer lugar, la privacidad y la seguridad están relacionadas con la recopilación y análisis de datos por parte de los asistentes virtuales para personalizar la experiencia del usuario, lo que plantea importantes cuestiones de privacidad. Como señala Shoshana Zuboff, *"la recopilación de datos por parte de las tecnologías inteligentes plantea riesgos significativos para la privacidad y la autonomía, ya que los datos personales pueden ser utilizados para influir en el comportamiento del usuario de maneras que no siempre son transparentes"* (Zuboff, 2019).

La acumulación de datos sensibles dentro de los asistentes virtuales sobre los usuarios puede dar lugar a la explotación y la manipulación si no se gestionan adecuadamente.

Por otra parte, la creciente dependencia de la tecnología a medida que los asistentes virtuales se integran en nuestras vidas puede afectar nuestra autonomía y habilidades para resolver problemas sin la ayuda de la tecnología.

Nicholas Carr argumenta que *"la dependencia excesiva de la tecnología puede llevar a una reducción en nuestras capacidades cognitivas y habilidades de toma de decisiones, ya que confiamos en la tecnología para realizar tareas que antes hacíamos de manera independiente"* (Carr, 2010).

Además, la capacidad de los asistentes virtuales de influir en las decisiones de los usuarios plantea preocupaciones sobre el control y la manipulación. Timnit Gebru y otros investigadores han destacado que *"los algoritmos que impulsan los asistentes virtuales pueden estar diseñados para manipular las preferencias y comportamientos de los usuarios de manera que prioricen los intereses comerciales sobre el bienestar del usuario"* (Gebru y otros, 2018).

Esta influencia puede llevar a minimizar la autonomía del usuario y a la toma de decisiones guiada por intereses externos. Por lo tanto, aunque la humanización de la tecnología a través de asistentes virtuales representa un avance significativo en la accesibilidad y personalización de la experiencia del usuario, también plantea desafíos éticos importantes relacionados con la necesidad de que la tecnología continúe sirviendo a los intereses y necesidades de los usuarios sin comprometer su autonomía y bienestar.

ÉTICA Y REGULACIÓN EN ASISTENTES VIRTUALES

Una vez que sabemos que los asistentes virtuales no solo traen consigo beneficios en términos de eficiencia y comodidad, también se plantean una serie de desafíos éticos y de regulación que deben ser abordados para garantizar que estas tecnologías operen de manera ética y respeten los derechos de los usuarios. Es esencial explorar los principales límites que deben trazarse para asegurar un uso responsable de los asistentes virtuales, así como las regulaciones necesarias para proteger la privacidad y evitar la manipulación del usuario.

Primero debemos tener en cuenta que la protección de la privacidad del usuario recae en el hecho de que los asistentes virtuales recopilan y procesan una gran cantidad de datos personales para proporcionar respuestas y servicios personalizados. La privacidad es una preocupación primordial, ya que la recopilación y almacenamiento de estos datos pueden ser vulnerables a accesos no autorizados y a la explotación indebida. Según el Reglamento General de Protección de Datos (GDPR) de la Unión Europea, *"los datos personales deben ser recopilados de manera legal y transparente, y deben ser protegidos con medidas adecuadas para garantizar la seguridad de la información"* (European Commission, 2018).

Es fundamental que los asistentes virtuales implementen mecanismos robustos para proteger la privacidad de los usuarios, como el cifrado de datos y políticas de acceso restringido.

Por otra parte, debe haber transparencia en el uso de datos, donde los usuarios tienen el derecho de saber cómo se recopilan, utilizan y almacenan sus datos. La transparencia es clave para construir y mantener la confianza del usuario. Como señala Solove (2021), *"los individuos deben estar informados sobre cómo se usan sus datos y tener la capacidad de controlar la recopilación y el procesamiento de su información personal"*.

Los asistentes virtuales deben proporcionar políticas claras y accesibles sobre el manejo de datos y ofrecer opciones para que los usuarios puedan gestionar sus preferencias y consentimientos.

Al basarse en algoritmos de inteligencia artificial, los asistentes virtuales pueden estar sujetos a sesgos que influyen en las decisiones y recomendaciones que ofrecen a los usuarios. Estos sesgos pueden afectar la equidad y la imparcialidad en el tratamiento de los usuarios. Como advierte O'Neil (2016), *"los algoritmos pueden amplificar y perpetuar sesgos existentes si no se diseñan y supervisan cuidadosamente"*.

Es crucial que los desarrolladores de asistentes virtuales implementen prácticas de auditoría y evaluación continua para detectar y mitigar sesgos, asegurando que los algoritmos operen de manera justa y equitativa.

La autonomía del usuario y la delegación de decisiones a los asistentes virtuales plantea preguntas sobre el control sobre sus propias decisiones. Como indica R. Binns, *"la dependencia excesiva de la tecnología para tomar decisiones puede erosionar la capacidad de los usuarios para tomar decisiones informadas y autónomas"* (Binns, 2018). Es esencial que los asistentes virtuales ofrezcan opciones para que los usuarios mantengan un control significativo sobre las decisiones críticas y no deleguen completamente la toma de decisiones a la IA.

REGULACIONES NECESARIAS PARA PROTEGER Y EVITAR LA MANIPULACIÓN DE DATOS

Debemos entender que deben existir ciertas regulaciones en la implementación de la tecnología y la inteligencia artificial. En primer lugar, debe haber una regulación de la recopilación y uso de datos, ya que las leyes y regulaciones específicas sobre la recopilación, almacenamiento y uso de datos protegerán la privacidad de los usuarios.

Los marcos regulatorios como el GDPR y la Ley de Privacidad del Consumidor de California (CCPA) ofrecen directrices claras sobre cómo manejar los datos personales. Las empresas que desarrollan asistentes virtuales deben cumplir con estos marcos y adaptar sus prácticas para garantizar la conformidad y la protección de la privacidad (European Commission, 2018; California Legislative Information, 2023).

También es importante la creación de normas éticas específicas para el desarrollo y uso de la inteligencia artificial en asistentes virtuales, que pueden ayudar a abordar problemas de sesgo, transparencia y autonomía. Por ejemplo, organizaciones como el *Institute of Electrical and Electronics Engineers* (IEEE) y el Comité de Ética en la IA de la UNESCO han propuesto directrices para el desarrollo de IA ética, que incluyen principios de justicia, transparencia y responsabilidad (IEEE, 2020; UNESCO, 2021).

Los mecanismos de supervisión y rendición de cuentas pueden ayudar a evaluar el funcionamiento de los asistentes virtuales y su im-

pacto en los usuarios. Estos mecanismos pueden incluir auditorías externas, revisiones de impacto en la privacidad y evaluaciones de sesgo.

La rendición de cuentas asegura que las prácticas de los desarrolladores y operadores de asistentes virtuales sean revisadas y corregidas cuando sea necesario (Binns, 2018).

Por último, la educación y la conciencia del usuario deben ser fomentadas para que tengan más información acerca del funcionamiento y los riesgos asociados con los asistentes virtuales. Los usuarios deben estar informados sobre cómo se utilizan sus datos y tener herramientas para gestionar su privacidad y seguridad.

Las iniciativas educativas y las campañas de concienciación pueden ayudar a empoderar a los usuarios para que tomen decisiones informadas sobre el uso de estos sistemas (Solove, 2021).

CONCLUSIONES

Los asistentes virtuales han llegado a convertirse en una pieza clave de la vida moderna, facilitando innumerables aspectos de nuestra cotidianidad, desde la gestión de tareas hasta el control de dispositivos en el hogar. Sin embargo, a medida que esta tecnología avanza, sus implicaciones van más allá de la simple comodidad. La creciente sofisticación de la inteligencia artificial y su capacidad para aprender de los usuarios plantea importantes cuestiones sobre la autonomía humana y la creciente dependencia de las máquinas para tomar decisiones en nuestro nombre.

La integración de los asistentes virtuales con tecnologías avanzadas, como las interfaces cerebro-computadora y la inteligencia artificial adaptativa, nos sitúa ante un futuro donde la relación entre humanos y tecnología se vuelve cada vez más profunda y compleja. Este vínculo, que en muchos casos puede mejorar la calidad de vida y ofrecer nuevas oportunidades, también expone preocupaciones sobre la pérdida de control humano, el impacto en la privacidad y los riesgos de manipulación del comportamiento.

Asimismo, el debate en torno al Transhumanismo, impulsado por la posibilidad de una fusión más estrecha entre humanos y tecnología, nos lleva a cuestionar los límites éticos de hasta qué punto estamos dispuestos a delegar funciones cruciales a la inteligencia artificial. Si bien los asistentes virtuales prometen una mayor eficiencia y personalización, también nos obligan a reflexionar sobre cómo preservar nuestra autonomía y garantizar que estas herramientas operen de manera ética y transparente.

En definitiva, los asistentes virtuales representan un gran avance en la interacción humano-máquina, pero su creciente influencia sobre nuestras vidas requiere un análisis más profundo de sus implicaciones éticas, filosóficas y sociales. Es necesario un marco regulatorio adecuado que proteja a los usuarios y asegure que, en esta nueva era de delegación tecnológica, el control y la responsabilidad sigan estando en manos humanas.

REFERENCIAS

Anderson, A. (2021). *The future of virtual assistants: Trends and innovations.* Tech Publishing.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149-158.* https://doi.org/10.1145/3287560.3287598

Bostrom, N. (2003). *The transhumanist FAQ: A general introduction.* https://www.transhumanism.org/index.php/WTA/faq

Bostrom, N. (2005). Transhumanist values. *Review of Contemporary Philosophy, 4*(2), 2–14.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies.* Oxford University Press.

Bostrom, N., & Roache, R. (2008). *Human enhancement: Ethical issues in human enhancement*. En J. Ryberg, T. S. Petersen, & C. Wolf (Eds.), New waves in applied ethics (pp. 120-152). Palgrave Macmillan.

Brown, T. (2020). Apple's Siri: The evolution of a digital assistant. *Tech Journal, 15*(2), 34-56.

California Legislative Information. (2023). *California Consumer Privacy Act*. Retrieved from https://leginfo.legislature.ca.gov/faces/codes_displayText?lawCode=CIV&division=3.5.&title=1.81.5.&article=

Carr, N. (2010). *The shallows: What the internet is doing to our brains.* W. W. Norton & Company.

Carr, N. (2011). *The shallows: What the internet is doing to our brain*. W. W. Norton & Company.

Chen, Y. (2020). Real-time translation and its impact on communication. *International Journal of Language and Technology, 11*(3), 71-84.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

European Commission. (2018). *The General Data Protection Regulation (GDPR)*. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/ead34ff2-873e-11e9-9f05-01aa75ed71a1/language-en

European Commission. (2021). *White paper on artificial intelligence: A European approach to excellence and trust*. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Forrester Research. (2020). *The impact of AI on decision-making in virtual assistants.*

Forrester Research. (2021). *Cortana: Adapting to modern workplace needs.*

Fundación para la Investigación sobre la Inteligencia Artificial (FIA). (2022). *Informe sobre la integración de asistentes virtuales en dispositivos accesibles.*

Gebru, T., Morgenstern, J., y Vecchione, B. (2018). Mitigating bias in AI: A case study in text classification. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 299-308.* https://doi.org/10.1145/3269206.3269210

Hughes, J. (2010). *Citizen cyborg: Why a democratic society needs human gene engineers.* Westview Press.

IEEE. (2020). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems*. https://ieeexplore.ieee.org/document/8058187

Jepsen, M. L. (2019). BCIs: The future of brain-computer interaction. *Journal of Neural Engineering, 16*(6), 061001. https://doi.org/10.1088/1741-2552/ab3c7b

Jordan, M. I. (2018). Artificial intelligence: A revolution in the making. *Journal of the Royal Society Interface, 15*(148), 20180313. https://doi.org/10.1098/rsif.2018.0313

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology.* Viking Press.

Kurzweil, R. (2014). *How to create a mind: The secret of human thought revealed.* Viking Press.

Kurzweil, R. (2019). *The singularity is near: When humans transcend biology.* Viking Press.

Kurzweil, R. (2019). Understanding Google Assistant: Contextual responses in AI. *AI Review, 22*(1), 45-68.

Lee, J. (2020). Bixby: Integrating AI into daily life. *Samsung Tech Review, 10*(4), 4254.

Lee, J. H., y Kim, J. H. (2021). The evolution of virtual assistants: A historical review and future directions. *Journal of Artificial Intelligence Research, 70,* 783-811. https://doi.org/10.1613/jair.1.12545

Lopatovska, I. (2013). The role of artificial intelligence in human-computer interaction: A perspective on technology and society. *Journal of Human-Computer Interaction, 29*(2), 123-140.

Microsoft. (2021). *Research at Microsoft: Collaborating for real-world change*. https://www.microsoft.com/en-us/research/

Mitchell, T. M. (2019). *Machine learning*. McGraw-Hill Education.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). *The ethics of algorithms: Mapping the debate*. Big Data & Society, 3(2), 1-21. https://doi.org/10.1177/2053951716679679

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Organization Mundial de la Salud. (2020). *Global report on assistive technology*. https://www.who.int/publications/i/item/9789240061780

Roffarello, G., y De Russis, L. (2020). Smart homes and the role of virtual assistants. *Journal of Intelligent Environments, 12*(1), 10-20.

Russell, S. (2016). *Human compatible: Artificial intelligence and the problem of control*. Viking Press.

Russell, S. J. (2019). *Artificial intelligence: A modern approach (4th ed.)*. Pearson.

Savulescu, J. (2009). Enhancing evolution: The ethical case for making better people. *Journal of Medical Ethics, 35*(4), 249-253. https://doi.org/10.1136/jme.2008.025463

Sloterdijk, P. (2013). You must change your life: On anthropotechnics. Serpent's Tail.

Smith, J. (2020). *The ethics of delegation: Technology and human autonomy*. TechEthics Publishing.

Solove, D. J. (2021). *The common rule: A new framework for protecting privacy in research*. Stanford Law Review.

Solove, D. J. (2021). The myth of data privacy: Why data protection is not enough. *Yale Law Journal Forum, 130,* 885-897. https://www.yalelawjournal.org/forum/themyth-of-data-privacy

Voicebot.ai. (2021). *The state of voice assistants: 2021 edition*. https://voicebot.ai

Wallach, W. (2015). *Moral machines: Teaching robots right from wrong*. Oxford University Press.

Wiener, N. (1950). *The human use of human beings: Cybernetics and society*. Da Capo Press.

Wolpaw, J. R., y Wolpaw, E. W. (2012). *Brain-computer interfaces: Principles and practice*. Oxford University Press.

Zheng, L., Wang, H., y Lee, S. (2020). Voice recognition technologies: Advances and challenges. *Journal of Communication Technology, 18*(2), 23-37.

Zheng, Z., Chen, J., y Wang, Y. (2020). Voice recognition technology and its applications in intelligent systems. *Journal of Intelligent Systems, 29*(1), 47-53. https://doi.org/10.1515/jisys-2020-0014

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

CAPÍTULO VIII

INTELIGENCIA ARTIFICIAL EN EL AULA: ÉTICA Y PRÁCTICAS

GISELA COTA YUCUPICIO
KARLA MARIELA SANDOVAL COTA
EDNA KARINA MIRANDA MONDACA

RESUMEN

Este capítulo explora las implicaciones éticas del uso de la Inteligencia Artificial (IA) en el aula, al discutir prácticas responsables para garantizar que la IA beneficie a todos los estudiantes y se mantenga la integridad educativa. Aunque la IA puede automatizar decisiones, los humanos deben supervisar y ser responsables de las decisiones finales.

Las prácticas responsables con el uso de Inteligencia Artificial promueven una educación inclusiva, destacando la importancia de la privacidad, la transparencia y la equidad. Además, se presentan estrategias para incorporar la IA evitando sesgos y asegurando el acceso equitativo a las herramientas tecnológicas.

Se reflexiona sobre la responsabilidad compartida entre desarrolladores, profesores y administradores, al discutir la importancia de enseñar a los estudiantes a ser usuarios críticos y conscientes de la tecnología. Con el propósito de integrar la educación sobre IA y ética en el currículo, se proponen métodos que fomentan una comprensión acerca del funcionamiento de la tecnología y sus implicaciones sociales y éticas, desde una visión general de las transformaciones que la IA está aportando al sector educativo.

Finalmente, planteamos la necesidad de educar a los estudiantes no solo en el uso de la IA, sino también en los aspectos éticos relacionados al sugerir enfoques pedagógicos para integrar estos temas en la enseñanza diaria, preparando a los estudiantes para ser ciudadanos digitales responsables y éticos.

Palabras clave: Inteligencia Artificial, Ética, Prácticas Responsables, Educación.

INTRODUCCIÓN

La inteligencia artificial (IA) está emergiendo como una herramienta transformadora en el ámbito educativo, ofreciendo nuevas oportunidades para mejorar la experiencia de aprendizaje y desarrollar el potencial de los estudiantes.

La implementación de la IA en la educación también plantea importantes desafíos éticos relacionados con la toma de decisiones, desde la programación de la herramienta de inteligencia artificial hasta su aplicación en el aula, evitando sesgos algorítmicos para garantizar la transparencia y explicabilidad que aseguren decisiones justas y equitativas para todos los estudiantes. Por ejemplo, en algunos sistemas educativos, se han utilizado algoritmos de IA para evaluar el rendimiento de los estudiantes. Sin embargo, ha habido casos donde estos algoritmos han mostrado sesgos, favoreciendo a ciertos grupos de estudiantes sobre otros. Un sistema de evaluación automatizada en el Reino Unido fue criticado por otorgar calificaciones más bajas a estudiantes de escuelas públicas en comparación con aquellos de escuelas privadas (UNESCO, 2024).

Fomentar una comprensión crítica de la tecnología al usar la IA en el aula, es fundamental, porque una vez que los estudiantes entiendan cómo funciona la IA y podrán utilizarla de manera más efectiva y creativa. Esto permite que no solo sean consumidores pasivos de tecnología, sino también creadores y críticos informados.

Una comprensión crítica de la tecnología fomenta el pensamiento crítico y la capacidad de cuestionar y analizar la información. Los estudiantes aprenden a evaluar la precisión y la relevancia de los datos proporcionados por los sistemas de IA.

La IA ayuda a los estudiantes a reconocer y abordar cuestiones éticas, como la privacidad, el sesgo y la equidad para asegurar que la tecnología se utilice de manera justa y responsable. Asimismo, se pueden identificar y mitigar posibles riesgos asociados con el uso de la IA, como la dependencia excesiva de la tecnología o la propagación de información errónea.

Cuando los estudiantes y los profesores comprenden los principios y limitaciones de la IA, pueden innovar y desarrollar con ética prácticas responsables aplicadas en el aula.

La aplicación de prácticas responsables de la IA en el aula está cambiando la educación y trasformando los campos profesionales. Al comprender cómo funciona la IA, los estudiantes están mejor preparados para las demandas del mercado laboral futuro y pueden adaptarse a nuevas tecnologías con mayor facilidad.

LA ÉTICA EN EL CONTEXTO DE LA INTELIGENCIA ARTIFICIAL EN LA EDUCACIÓN

La ética se refiere a un conjunto de principios morales que guían el comportamiento de una persona o grupo. Estos principios ayudan a determinar lo que es correcto o incorrecto, justo o injusto. En el contexto de la inteligencia artificial en la educación, la ética podría incluir aspectos como la privacidad, la transparencia y la equidad.

La IA en la educación a menudo requiere la recopilación y el análisis de grandes cantidades de datos personales de los estudiantes para proteger estos datos y evitar violaciones de privacidad al garantizar que se utilicen de manera segura y responsable. Las instituciones deben obtener el consentimiento informado de los estudiantes y sus padres antes de recopilar y utilizar estos datos.

En el ámbito educativo, la privacidad y seguridad de los datos son fundamentales para proteger la información personal y académica de los estudiantes, profesores y personal administrativo. La implementación de la IA en educación puede mejorar la personalización del aprendizaje, pero también plantea riesgos significativos si no se manejan adecuadamente los datos.

Por ejemplo, el riesgo en la recopilación de datos personales de los estudiantes, como su rendimiento académico, comportamiento y preferencias, puede llevar a violaciones de privacidad si no se manejan adecuadamente. Gujjula y Sanghera (2023) interpretan como una consideración ética, la implementación de medidas de seguridad robustas para proteger los datos y garantizar que solo se utilicen con fines educativos legítimos. Además, los estudiantes y sus padres pueden no estar completamente informados sobre qué datos se recopilan y cómo se utilizan, lo que puede llevar a un uso indebido de la información.

Un caso notable ocurrió en los Estados Unidos, donde una plataforma de aprendizaje en línea fue acusada de recopilar datos personales de los estudiantes y utilizarlos para fines comerciales sin el conocimiento de los usuarios.

El consentimiento informado es un principio ético fundamental que garantiza que los individuos tengan control sobre su información personal.

Los desafíos principales de la privacidad y seguridad de los datos están relacionados, en primer término, con la recolección de datos sensibles. La IA educativa a menudo requiere información personal del usuario, su rendimiento académico y comportamientos de aprendizaje.

La recolección y almacenamiento de estos datos deben cumplir con normativas de privacidad y seguridad, que se sujetan a las siguientes regulaciones: Reglamento General de Protección de Datos en Europa (GDPR) que establece directrices estrictas sobre cómo deben manejarse los datos personales. La Ley de Derechos Educativos y Privacidad Familiar en Estados Unidos (FERPA), que protege la privacidad de los registros educativos de los estudiantes. Así como, la Ley Federal de Protección de Datos Personales en Posesión de Sujetos Obligados

(LGPDPPSO), en México, esta ley regula el tratamiento de datos personales para garantizar la privacidad y el derecho a la autodeterminación informativa.

Otro aspecto relevante es el desafío relacionado con la seguridad de los sistemas de la IA, los cuales deben estar protegidos contra ciberataques que podrían comprometer la integridad y confidencialidad de los datos.

Además, los sistemas de IA deben ser transparentes en cuanto a la toma de decisiones y qué datos utilizan. Esto es esencial para evitar sesgos y discriminación, para que los profesores y estudiantes puedan confiar en la transparencia de la tecnología.

La implementación de la IA no debe exacerbar las desigualdades existentes en el sistema educativo. Es fundamental asegurar que todos los estudiantes, independientemente de su origen socioeconómico, tengan acceso equitativo a las tecnologías de IA. Los datos deben ser representativos de la diversidad de la población estudiantil para asegurar que todos los estudiantes reciban un trato justo y equitativo.

La Inteligencia Artificial (IA) está transformando el panorama educativo, ofreciendo nuevas oportunidades para promover la equidad y el acceso a la educación. Las principales formas en que la IA está contribuyendo a los contextos educativos son mediante la personalización del aprendizaje, el acceso a recursos educativos, la automatización de tareas administrativas, la inclusión y diversidad, así como la innovación en la evaluación.

Las experiencias de aprendizaje personalizadas a partir de la IA se adaptan a las necesidades individuales de cada estudiante. Esto es especialmente beneficioso para aquellos con dificultades de aprendizaje o necesidades especiales. Los sistemas de IA pueden analizar el progreso de los estudiantes y ajustar el contenido y el ritmo de enseñanza para optimizar el aprendizaje, independientemente de la ubicación geográfica o el estatus socioeconómico, la IA facilita el acceso a una amplia gama de recursos educativos. Plataformas de aprendizaje en línea y aplicaciones educativas pueden ser utilizadas en dispositivos móviles, permitiendo que estudiantes de áreas rurales o de bajos ingresos accedan a materiales educativos de alta calidad.

Así mismo, la IA ayuda a abordar las desigualdades en el acceso al conocimiento y la investigación, promoviendo la diversidad de expresiones culturales y garantizando que la brecha tecnológica no se amplíe. Herramientas de traducción automática y recursos educativos multilingües pueden hacer que la educación sea más inclusiva para estudiantes de diferentes orígenes lingüísticos y culturales.

En consecuencia, promueve la inclusión y la diversidad, conceptos fundamentales en el derecho a una educación digna y de calidad. La inclusión se refiere a la integración de todas las personas en el ámbito educativo, independientemente de sus características. La tecnología, y más específicamente la inteligencia artificial (IA), puede ser una herramienta muy útil para fomentar una educación inclusiva.

En lo que se refiere a la innovación en la evaluación, los sistemas de IA generativa pueden replantear los modos de evaluación tradicionales, proporcionando nuevas formas de medir el aprendizaje y el rendimiento de los estudiantes. Evaluaciones adaptativas y análisis de datos pueden ofrecer una visión más completa y precisa del progreso de los estudiantes, permitiendo intervenciones más efectivas.

La IA tiene el potencial de revolucionar la educación, promoviendo la equidad y el acceso a nuevas formas de integración en el ámbito educativo. Sin embargo, es fundamental que estas tecnologías se implementen de manera ética y centrada en el ser humano, para garantizar que beneficien a todos los estudiantes y no amplíen las desigualdades existentes.

Los riesgos de perpetuar desigualdades existentes en educación con el uso de la IA son latentes, ya que su implementación puede ampliar la brecha digital entre estudiantes con acceso a tecnología avanzada y aquellos sin acceso adecuado, lo que puede generar desigualdad en la calidad de la educación, es decir, la IA ofrece experiencias de aprendizaje personalizadas, pero si no se implementa de manera equitativa puede resultar en una calidad de educación desigual entre diferentes grupos de estudiantes por el acceso dispar de recursos.

Por otra parte, la dependencia tecnológica excesiva de la IA puede reducir la interacción humana en el proceso educativo, lo que resulta indispensable para el desarrollo social y emocional de los estudiantes.

Al respecto, la UNESCO destaca la importancia de un enfoque centrado en el ser humano para garantizar que la IA en la educación promueva la inclusión y la equidad, y no amplíe las desigualdades existentes.

QUIÉN ES RESPONSABLE DE LAS DECISIONES DE LA INTELIGENCIA ARTIFICIAL

Es importante definir claramente quién es responsable de las decisiones tomadas por los sistemas de IA. En caso de errores o malentendidos, debe haber mecanismos para abordar y corregir estos problemas.

En el contexto anterior, el uso de datos en la inteligencia artificial (IA) en el aula conlleva varios riesgos y consideraciones éticas que deben abordarse para asegurar un entorno educativo justo y seguro.

La explicabilidad en inteligencia artificial (IA), se refiere a la capacidad de un sistema de IA para explicar sus decisiones y acciones de manera comprensible para los humanos. Esto implica que los usuarios puedan entender cómo y por qué un algoritmo ha llegado a una determinada conclusión o ha tomado una acción específica (IBM, 2024).

Un algoritmo realiza operaciones y toma decisiones basadas en las instrucciones y reglas con las que ha sido programado. Cuando los algoritmos son explicables, los profesores y estudiantes pueden confiar más en las decisiones tomadas por la IA. Conocer cómo y por qué la IA ha llegado a una conclusión, ayuda a reducir la desconfianza y el escepticismo (IBM, 2024). al favorecer la transparencia y la confianza.

Los estudiantes aprenden mejor si entienden cómo funcionan los algoritmos que están utilizando. Esto no solo mejora su comprensión de la materia, sino que también les da habilidades críticas en el análisis y uso de la tecnología (IBM,2024) al favorecer su aprendizaje.

La explicabilidad permite identificar y corregir posibles sesgos en los algoritmos, para asegurar que las decisiones de la IA sean justas y equitativas para todos los estudiantes (IBM, 2024). También permite que los profesores pueden ser responsables de las decisiones tomadas por la IA si entienden cómo funcionan los algoritmos. Esto es impor-

tante para la rendición de cuentas y para asegurar que las herramientas de IA se utilicen de manera ética (Educaopen, 2024).

Con algoritmos explicables, es más fácil ajustar y personalizar las experiencias de aprendizaje para satisfacer las necesidades individuales de cada estudiante. Los profesores pueden entender mejor qué ajustes hacer para mejorar los resultados educativos (UNESCO, 2024).

Sin embargo, se identifica el riesgo de que los algoritmos de IA pueden mantener o incrementar los sesgos existentes si se entrenan con datos sesgados. Esto puede resultar en decisiones injustas que afectan negativamente a ciertos grupos de estudiantes. Por lo tanto, de manera ética se debe considerar la posibilidad de utilizar conjuntos de datos representativos y realizar auditorías regulares para identificar y corregir sesgos en los sistemas de IA (Gujjula y Sanghera, 2023).

Además, la falta de transparencia en cómo los sistemas de IA toman decisiones puede generar desconfianza y dificultar la identificación de errores o sesgos, por lo tanto, éticamente se debería promover que los sistemas de IA sean transparentes en sus decisiones y explicables de manera comprensible para los profesores y estudiantes (Gujjula y Sanghera, 2023).

La transparencia de algoritmos en la inteligencia artificial (IA) aplicada en el aula es elemental desde un punto de vista ético. Se refiere a la capacidad de entender y explicar cómo un algoritmo toma decisiones. Esto incluye conocer los datos que utiliza, las reglas que sigue y los resultados que produce para generar confianza y aceptación para que la IA en el aula se aplique de manera efectiva. El Observatorio del Instituto para el Futuro de la Educación (2024), afirma que con algoritmos transparentes, es más fácil identificar áreas de mejora y hacer ajustes para optimizar el rendimiento y la efectividad de la IA en el aula.

GESTIÓN DE ERRORES Y MITIGACIÓN DE DAÑOS

En el contexto anterior, es importante aplicar los principios generales y buenas prácticas reconocidas en el campo de la privacidad y seguridad de datos en la inteligencia artificial aplicada a la educación,

para ayudar a proteger la identidad de los individuos al procesar datos y utilizar las técnicas de Anonimización y Pseudonimización.

La anonimización es el proceso mediante el cual los datos personales se transforman de tal manera que ya no se puede identificar a una persona específica, ni directa ni indirectamente. Esto se logra eliminando o modificando la información que podría vincular los datos a una persona en particular. Una vez que los datos han sido anonimizados, ya no se consideran datos personales y, por lo tanto, están fuera del alcance de muchas regulaciones de privacidad. Por ejemplo, eliminando identificadores como nombre, dirección, número de teléfono, entre otros datos. Además, se generalizan los datos al agruparlos en categorías amplias, como cambiar la edad por un rango de edades. Otro ejemplo es la perturbación de datos, que consiste en alterarlos ligeramente para que no se puedan rastrear hasta una persona específica (ayudaley, 2024).

La pseudonimización es una técnica donde los datos personales se procesan de manera que no se atribuyan a una persona específica sin información adicional. A diferencia de la anonimización, la pseudonimización permite que los datos se vuelvan a identificar si se dispone de la clave o información adicional necesaria. Esta técnica mejora la privacidad al reducir el riesgo de identificación, pero aún permite cierto grado de análisis y uso de los datos. Algunos ejemplos son: utilizar identificadores aleatorios que permitan reemplazar nombres u otros identificadores directos con códigos o números aleatorios, así como separar los datos para almacenar la información identificativa en una base de datos separada y segura, vinculada a los datos pseudonimizados mediante una clave (Smowltech, 2023).

Cuando se aplica la técnica de Anonimización los datos no pueden ser revertidos a su estado original y no se pueden identificar a las personas. En cambio, con la Pseudonimización, los datos pueden ser revertidos a su estado original si se dispone de la clave o información adicional.

Estas técnicas son determinantes para proteger la privacidad de los datos en el contexto de la inteligencia artificial y otras aplicaciones que manejan grandes volúmenes de información personal. Además, la implementación del cifrado de datos reviste especial importancia para proteger los datos contra accesos no autorizados, así como el realizar

evaluaciones regulares para identificar y mitigar riesgos potenciales relacionados con la privacidad.

Resulta esencial que los avances en la personalización del aprendizaje de los estudiantes obtenidos mediante la aplicación de la IA en el aula, se realicen de manera ética y segura, priorizando siempre la privacidad y seguridad de los datos que apliquen los principios generales y las buenas prácticas ampliamente reconocidas en el campo de la Inteligencia Artificial aplicada en la educación.

EJEMPLO PRÁCTICO DE LA IA EN EL AULA

Un sistema de IA aplicado en el aula puede recomendar recursos educativos personalizados. La transparencia en este caso implicaría, explicar cómo se seleccionan los recursos al detallar qué datos del estudiante se utilizan respecto a su rendimiento pasado y sus intereses, y cómo estos datos influyen en las recomendaciones.

En cuanto a mostrar los criterios de decisión, el sistema de IA debe proporcionar una visión clara de las reglas y modelos que el algoritmo sigue para hacer sus recomendaciones y permitir la retroalimentación. Es decir, brindar a los estudiantes y profesores la oportunidad de cuestionar y retroalimentar las decisiones del algoritmo para mejorar su funcionamiento.

Una herramienta de IA que se destaca en la recomendación de recursos educativos personalizados es *LearningStudioAI* (El proyector de clase, 2023). En la selección de recursos esta herramienta utiliza datos sobre el rendimiento académico de los estudiantes, sus intereses y sus estilos de aprendizaje para seleccionar los recursos más adecuados. Estos datos pueden incluir calificaciones y evaluaciones, información sobre el rendimiento pasado del estudiante en diferentes materias, sus intereses personales, temas y áreas de interés que el estudiante ha mostrado. Así como el estilo de aprendizaje a partir de las preferencias del estudiante en cuanto a métodos de aprendizaje: visual, auditivo, kinestésico, entre otros.

En cuanto a los criterios de decisión, el algoritmo de *LearningStudioAI* considera varios criterios para tomar decisiones:

1. Relevancia. Selecciona recursos que sean directamente relevantes para el contenido que el estudiante necesita aprender.
2. Dificultad. Ajusta la dificultad del material para que sea apropiada para el nivel actual del estudiante.
3. Interactividad. Prefiere recursos que sean interactivos y que puedan mantener el interés del estudiante.
4. Retroalimentación. Incluye recursos que proporcionen retroalimentación inmediata para ayudar al estudiante a corregir errores y mejorar.

LearningStudioAI permite a los estudiantes y profesores proporcionar retroalimentación sobre las recomendaciones del algoritmo. Esta retroalimentación se utiliza para mejorar la precisión de las recomendaciones futuras, detectar y corregir posibles sesgos y, permite actualizar los recursos educativos disponibles basándose en la retroalimentación.

Por ejemplo, en su aplicación práctica esta herramienta podría recomendar a un estudiante que presenta dificultades con las fracciones en matemáticas, videos tutoriales con explicaciones visuales sobre cómo trabajar con fracciones. También ejercicios interactivos que permitan que el estudiante pueda practicar y aplicar lo aprendido recibiendo una respuesta inmediata. Otra opción, es la recomendación de juegos educativos, para que a través de actividades lúdicas refuerce el concepto de fracciones de manera divertida.

Por último, el estudiante y el profesor pueden luego proporcionar retroalimentación sobre la efectividad de estos recursos, lo que ayuda a mejorar las futuras recomendaciones del algoritmo.

PRÁCTICAS RESPONSABLES EN EL USO DE INTELIGENCIA ARTIFICIAL PARA PROMOVER UNA EDUCACIÓN INCLUSIVA

La práctica responsable implica la aplicación de principios éticos en situaciones concretas. Es la implementación práctica de la ética en acciones diarias en el ámbito de la IA en la educación, esto podría incluir el crear y utilizar sistemas de IA que respeten la privacidad y

la equidad, monitorear y evaluar regularmente los sistemas de IA para asegurar que cumplen con los estándares éticos, asegurar la formación de profesores y estudiantes sobre el uso ético y responsable de la IA.

La aplicación de la IA en la educación varía enormemente de un país a otro, lo que suele reflejar las disparidades ya existentes en cuanto a la infraestructura tecnológica, la financiación, el apoyo político y los niveles de alfabetización digital. Los países desarrollados y ricos pueden contar con una infraestructura tecnológica más sólida, así como con un ecosistema de innovación que incluye al sector privado. Pero no es el caso en el Sur Global y, en términos generales, en los países en vías de desarrollo (UNESCO, 2024).

Ante el panorama descrito, resulta necesario asegurar la equidad y accesibilidad con prácticas responsables en el uso de la inteligencia artificial (IA) en la educación, para ello se propone la implementación de las siguientes estrategias:

1. Desarrollo de un Plan Estratégico Integral.
    a. Alineación con la misión institucional: asegurar que la implementación de la IA esté alineada con la misión, visión y valores de la institución educativa (UNESCO, 2021).
    b. Evaluación de beneficios y riesgos: evaluar tanto los beneficios potenciales como los riesgos asociados, problemas de privacidad de datos y sesgo algorítmico (UNIR, 2023).
    c. Establecimiento de equipos de liderazgo: designar un equipo interdisciplinario que incluya representantes técnicos, operativos, académicos, financieros y legales.
    d. Desarrollo de Políticas Públicas Inclusivas: una de las primeras estrategias para asegurar la equidad en el uso de la IA en la educación es el desarrollo de políticas públicas inclusivas. Estas políticas deben garantizar que la IA se utilice de manera que beneficie a todos los estudiantes, independientemente de su origen socioeconómico, género o capacidades. Según la UNESCO, es de suma importancia desarrollar marcos normativos que protejan a los

estudiantes y profesores, asegurando un enfoque centrado en el ser humano (UNESCO, 2024).

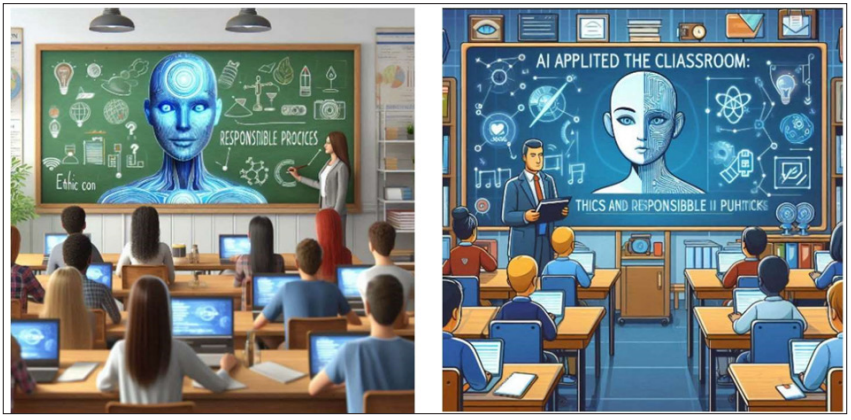
2. Accesibilidad y personalización.
    a. Contenido accesible: crear contenido educativo accesible para estudiantes con discapacidades, como subtítulos, audiodescripciones y adaptaciones para discapacidades visuales.
    b. Personalización del aprendizaje: utilizar la IA para personalizar el aprendizaje según las necesidades individuales de los estudiantes, mejorando así la inclusión.
3. Capacitación y sensibilización.
    a. Formación continua: capacitar a los profesores en el uso de herramientas de IA y en la identificación de posibles sesgos en los algoritmos. La capacitación de los profesores es fundamental para una implementación equitativa de la IA en la educación. Los profesores deben estar preparados para utilizar herramientas de IA de manera efectiva y ética. Esto incluye no solo el uso técnico de estas herramientas, sino también la comprensión de sus implicaciones éticas y sociales. Es importante preparar a los profesores para una educación impulsada por la IA, al asegurar que puedan apoyar a todos los estudiantes de manera equitativa (Fundación Telefónica, 2024).
    b. Sensibilización sobre la equidad: promover la sensibilización sobre la importancia de la equidad y la inclusión en el uso de la IA en la educación.
    c. Acceso equitativo a la tecnología: para asegurar la equidad en el uso de la IA, es necesario garantizar que todos los estudiantes tengan acceso a la tecnología necesaria. Esto incluye no solo dispositivos y conectividad a internet, sino también recursos educativos y apoyo técnico. La UNESCO subraya la importancia de promover la inclusión y la equidad en el acceso a la educación digital (UNESCO, 2024).

4. Monitoreo y Evaluación Continua.
    a. Seguimiento de impacto: implementar sistemas de monitoreo para evaluar el impacto de la IA en la equidad educativa y realizar ajustes necesarios.
    b. Retroalimentación de la comunidad: involucrar a estudiantes, padres y profesores en la evaluación continua de las herramientas de IA. Estas estrategias pueden ayudar a garantizar que la implementación de la IA en la educación sea equitativa y beneficie a todos los estudiantes, independientemente de sus circunstancias personales.
5. Transparencia y ética en el uso de datos: son fundamentales para asegurar la equidad en la educación impulsada por la IA. Los sistemas educativos deben ser transparentes sobre cómo se recopilan, utilizan y protegen los datos de los estudiantes. Además, se debe garantizar que los datos se utilicen de manera que beneficien a todos los estudiantes y no perpetúen desigualdades existentes (UNESCO, 2024).

La implementación equitativa de la IA en la educación requiere un enfoque multifacético que incluya el desarrollo de políticas públicas inclusivas, la capacitación de profesores, la eliminación de sesgos en los algoritmos, el acceso equitativo a la tecnología y la transparencia en el uso de datos. Al adoptar estas estrategias, se asegura que la IA beneficie a todos los estudiantes y contribuya a un sistema educativo más justo y equitativo. La mejor manera de conseguirlo es utilizar la IA como herramienta para complementar, y no para sustituir, los elementos humanos de la enseñanza (ver **Figura 1**).

Fomentar una comprensión crítica de la tecnología, especialmente en el contexto de la inteligencia artificial (IA), implica entender cómo funciona para que los estudiantes y profesores pueden tomar decisiones informadas sobre su uso y sus implicaciones, con un enfoque crítico que permita reconocer y resistir intentos de manipulación o desinformación (ver **Figura 2**).

Figura 1. Implementación equitativa de la IA en la educación.



Fuente: Microsoft Copilot (2024).

Figura 2. Comprensión crítica de la tecnología en el contexto de la IA en el aula.



Fuente: Microsoft Copilot (2024)

En el contexto anterior, la IA puede perpetuar sesgos existentes si no se diseña y supervisa adecuadamente. Una comprensión crítica permite identificar y mitigar estos sesgos, al promover un uso ético y responsable de la tecnología, al asegurar que se utilice para el beneficio de todos y no para causar daño. También incluye la capacidad de identificar y responder a amenazas de seguridad, como el *phishing* o el *malware* que permitan proteger los datos para asegurar la privacidad y seguridad.

Al entender las capacidades y limitaciones de la IA, los usuarios pueden aplicar la tecnología de manera innovadora para resolver problemas complejos, y al mismo tiempo, fomenta una cultura de curiosidad y experimentación, al impulsar el desarrollo de nuevas aplicaciones y mejoras tecnológicas.

Una ciudadanía informada puede participar de manera más efectiva en debates y decisiones sobre la regulación y el uso de la IA para promover una justicia social que beneficie a todos los sectores de la sociedad, reduciendo desigualdades y promoviendo la equidad.

Fomentar una comprensión crítica de la tecnología con el uso de la IA es primordial para habilitar a los usuarios en su uso ético y responsable, proteger la privacidad y seguridad, estimular la innovación y asegurar una participación democrática informada. Esto no solo mejora la interacción con la tecnología, sino que también contribuye a una sociedad más justa y equitativa.

CÓMO LA INTELIGENCIA ARTIFICIAL ESTÁ CAMBIANDO LA EDUCACIÓN

Sin duda la inteligencia artificial (IA) ha revolucionado la educación al mostrar distintas formas de intervenir en el proceso de enseñanza y aprendizaje. Desde contar con fácil acceso a herramientas que se pueden apropiar y utilizar como técnicas didácticas, hasta automatizar tareas y tener en un clic acceso a la información.

Todo lo anterior, muestra un impacto transformador en el aula tradicional y abre nuevas posibilidades para mejorar la eficacia educativa y prepara a los estudiantes para los desafíos actuales.

La inteligencia artificial (IA) está cambiando significativamente el panorama educativo, al influir en cómo se enseña y se aprende. Al integrar la IA en el ámbito educativo, se habilitan nuevas metodologías que optimizan el aprendizaje y la enseñanza.

Según Zawacki-Richter y otros (2019), la IA facilita la personalización del aprendizaje al adaptar el contenido y los recursos a las necesidades individuales de los estudiantes, mejorando así la eficacia del proceso educativo.

La automatización de la evaluación y la retroalimentación instantánea proporcionada por sistemas de IA contribuyen a una administración más eficiente y una retroalimentación más rápida para los estudiantes (Baker e Inventado, 2014). Esta integración de la IA no solo optimiza la enseñanza y el aprendizaje, sino que también promueve el acceso a recursos educativos globales y fomenta el desarrollo de habilidades necesarias para el siglo XXI (Chen y Zhang, 2021).

Algunos de los impactos de la aplicación de la IA en el aula son:

a) Personalización del aprendizaje. La IA permite una personalización del aprendizaje sin precedentes. Los sistemas de tutoría inteligentes pueden adaptar el contenido y las estrategias de enseñanza a las necesidades individuales de los estudiantes, optimizando así el proceso educativo. Según un estudio de Zawacki-Richter y otros (2019), las herramientas basadas en IA pueden ajustar los materiales de aprendizaje en función del rendimiento y el progreso del estudiante, proporcionando un enfoque más centrado en el alumno.

b) Automatización de la evaluación. La IA también está cambiando la forma en que se llevan a cabo las evaluaciones. Herramientas automatizadas pueden calificar exámenes y tareas, proporcionando retroalimentación instantánea a los estudiantes. Esto no solo ahorra tiempo a los profesores, sino que también facilita una retroalimentación más rápida y frecuente (Baker e Inventado, 2014)

c) Creación de contenidos educativos. Los algoritmos de IA pueden generar materiales educativos personalizados y adaptativos, facilitando el desarrollo de recursos acorde con los objetivos

> educativos específicos de cada curso. Por ejemplo, los sistemas de generación automática de preguntas pueden crear ejercicios personalizados basados en el contenido estudiado (Heffernan y Heffernan, 2014:73-103).

Los sistemas de aprendizaje adaptativo basados en IA pueden proporcionar rutas de aprendizaje personalizadas, ajustando el contenido y las actividades en función de la interacción y el rendimiento del estudiante. Esto mejora la eficacia del aprendizaje al abordar las debilidades individuales y fortalece las áreas de conocimiento específicas (Kerr, 2013).

La IA también facilita el acceso a recursos educativos globales a través de plataformas en línea que pueden traducir y adaptar contenidos en varios idiomas. Esto democratiza el acceso a la educación y permite a los estudiantes aprender de una variedad más amplia de fuentes (Chen y Zhang, 2021:10-26).

Los cambios ocasionados por la IA en la educación implican nuevas formas en las que el alumno se apropia del aprendizaje con el uso de la tecnología. Sin embargo, el rol del profesor sigue siendo fundamental en el ámbito educativo, incluso con la creciente integración de tecnologías avanzadas como la inteligencia artificial (IA).

A pesar de las capacidades innovadoras de la IA para personalizar el aprendizaje y automatizar procesos, la interacción humana en la educación sigue siendo determinante para el desarrollo integral de los estudiantes.

Se considera que, con la aplicación de la IA en el aula presencial, la interacción humana es necesaria para el desarrollo socioemocional de los estudiantes. Los profesores no solo transmiten conocimientos académicos, sino que también juegan un papel vital en el apoyo emocional y en la creación de un entorno de aprendizaje positivo. Según un estudio de Hattie (2009), las relaciones positivas entre profesores y estudiantes son un factor determinante en el éxito académico y en el bienestar emocional de los alumnos en la presencialidad.

La presencia de un profesor empático puede ayudar a los estudiantes a desarrollar habilidades sociales y emocionales facilitando el

aprendizaje activo, donde los estudiantes participan de manera más comprometida en el proceso educativo.

Los profesores pueden adaptar sus estrategias de enseñanza en tiempo real, responder a preguntas y fomentar discusiones que enriquecen el aprendizaje presencial.

Según Deci y Ryan (2000), la motivación intrínseca de los estudiantes se ve significativamente influenciada por la calidad de las relaciones con sus profesores, quienes pueden proporcionar apoyo y aliento que fomente el interés y la dedicación hacia el aprendizaje.

Los profesores no solo enseñan contenido académico, sino que también guían a los estudiantes en su desarrollo personal y profesional. Noddings (2012), señala que la educación efectiva requiere una atención cuidadosa a las necesidades individuales de los estudiantes, un aspecto que las máquinas y algoritmos aún no pueden replicar completamente.

ENSEÑAR A LOS ESTUDIANTES SOBRE INTELIGENCIA ARTIFICIAL Y ÉTICA

Enseñar los principios éticos en el uso de la inteligencia artificial (IA) es fundamental en la formación de los estudiantes en los escenarios donde la tecnología juega un papel central.

Algunas estrategias para sensibilizar al estudiante sobre las consideraciones de la IA y la ética son relacionadas con la integración de estos temas en las asignaturas que conforman los planes y programas de estudios. Además de incorporar el uso de herramientas de IA en asignaturas de ciencias sociales, ciencias económico administrativas, ingeniería y tecnología, ciencias naturales y exactas, entre otras.

Se recomienda fomentar debates sobre los dilemas éticos de la IA, como la privacidad y los sesgos algorítmicos mediante el análisis de casos reales donde la IA ha tenido impactos positivos y negativos.

Es importante también, enseñar a los estudiantes los fundamentos de la programación y cómo funcionan los algoritmos de IA para favorecer el desarrollo de proyectos donde los estudiantes puedan crear y

experimentar con aplicaciones de Inteligencia Artificial. Además, es importante enseñarles a evaluar la credibilidad de las fuentes de información y que sean capaces de analizar cómo la IA afecta diferentes aspectos de la sociedad y sus propias vidas.

La integración de IA en el aula ayuda a los estudiantes a desarrollar habilidades esenciales para el siglo XXI, como el pensamiento crítico y la resolución de problemas complejos. Herramientas basadas en IA pueden simular escenarios reales y proporcionar a los estudiantes experiencias prácticas que mejoren su preparación para el mundo laboral (Nguyen, 2019:223-236).

La colaboración con expertos a través de charlas, talleres y visitas, para que hablen con los estudiantes y compartan sus conocimientos y experiencias, es otra forma de enseñar sobre la aplicación de la IA. Se recomienda programar visitas a empresas tecnológicas para que los estudiantes vean de primera mano cómo se desarrolla y aplica la IA desde un aspecto ético.

A continuación, se describen otras estrategias relacionadas con:

a) El uso de recursos educativos como libros, artículos y recursos en línea que expliquen la IA de manera accesible.

b) Empleo de plataformas educativas que ofrezcan simulaciones y actividades interactivas.

c) Organización de eventos donde los estudiantes puedan desarrollar soluciones innovadoras utilizando IA.

d) Creación de clubes donde los estudiantes interesados puedan explorar la IA más a fondo y trabajar en proyectos colaborativos.

Aplicar la IA y conocer sus aspectos éticos tiene múltiples beneficios para los estudiantes tanto en la mejora del aprendizaje personalizado como en el desarrollo de habilidades tecnológicas:

1. Mejora del Aprendizaje Personalizado.
    - Adaptación del contenido: la IA puede personalizar el contenido educativo según las necesidades y el ritmo de aprendizaje de cada estudiante.

- Retroalimentación inmediata: proporciona retroalimentación en tiempo real, ayudando a los estudiantes a corregir errores y mejorar continuamente.

2. Desarrollo de Habilidades Tecnológicas.
   - Competencias digitales: los estudiantes adquieren habilidades en el uso de tecnologías avanzadas, preparándolos para el futuro laboral.
   - Pensamiento crítico y resolución de problemas: la interacción con sistemas de IA fomenta el desarrollo de habilidades analíticas y de resolución de problemas.

Los profesores y administradores que buscan implementar el uso de la IA en el aula de manera ética y responsable, deben considerar que en las estrategias utilizadas para tal fin, promuevan el desarrollo de competencias digitales relacionadas con la alfabetización en IA, la privacidad y seguridad de los datos, la transparencia y explicabilidad, la detección y mitigación de sesgos, la responsabilidad y rendición de cuentas, la integración de la IA en la enseñanza, la evaluación crítica de herramientas de IA, para promover la equidad, la conciencia social y ética, así como impulsar el desarrollo de habilidades digitales para supervisar los sistemas de inteligencia artificial (ver **Tabla 1**).

Tabla 1. Responsabilidad compartida entre desarrolladores, profesores y administradores en el uso de la IA en el aula.

Competencia Digital	Descripción	Ejemplos de Prácticas Responsables
Alfabetización en IA	Comprender los conceptos básicos de la IA y su funcionamiento.	Participar en talleres y cursos sobre IA.
Privacidad y Seguridad de Datos	Proteger la información personal y sensible de los estudiantes.	Implementar políticas de anonimización y obtener consentimiento informado.

Competencia Digital	Descripción	Ejemplos de Prácticas Responsables
Transparencia y Explicabilidad	Asegurar que las decisiones de la IA sean comprensibles y justificables.	Utilizar herramientas de IA que permitan la explicación de sus decisiones.
Detección y Mitigación de Sesgos	Identificar y corregir sesgos en los algoritmos de IA.	Evaluar continuamente los resultados de la IA para detectar posibles sesgos.
Responsabilidad y Rendición de Cuentas	Establecer mecanismos claros de responsabilidad en el uso de la IA.	Definir políticas claras y asegurar la supervisión humana en decisiones críticas.
Integración de IA en la Enseñanza	Utilizar herramientas de IA para mejorar el proceso de enseñanza-aprendizaje.	Implementar sistemas de tutoría inteligente y análisis de aprendizaje personalizados.
Evaluación Crítica de Herramientas de IA	Analizar y seleccionar herramientas de IA basadas en su impacto ético.	Realizar revisiones periódicas de las herramientas de IA utilizadas en el aula.
Promoción de la Equidad	Asegurar que la IA beneficie a todos los estudiantes de manera justa.	Utilizar datos diversos y representativos para entrenar los algoritmos de IA.
Conciencia Social y Ética	Reflexionar sobre el impacto social de la IA.	Fomentar debates y discusiones sobre los aspectos éticos de la IA en el aula.
Desarrollo de Habilidades Digitales	Adquirir competencias digitales para manejar y supervisar sistemas de IA.	Participar en programas de formación continua sobre tecnologías emergentes y su aplicación educativa.

Fuente: Microsoft Copilot (2024).

Conocer los aspectos éticos y prácticos de la IA le permite al estudiante hacer un uso responsable de la tecnología, se convierten en usuarios conscientes de las implicaciones sociales y éticas en su aplicación. Esto permite que estén mejor preparados para tomar decisiones informadas sobre el uso de la tecnología en sus vidas personales y profesionales.

Los estudiantes pueden convertirse en líderes que promuevan el uso ético y responsable de la tecnología en sus comunidades y profesiones, lo que les permitirá adaptarse con liderazgo ético a un mundo en constante cambio tecnológico en el contexto de la inteligencia artificial.

CONCLUSIONES

La ética proporciona el marco teórico y los principios que guían la práctica responsable. Sin una base ética, la práctica responsable carecería de dirección y propósito. Por otro lado, sin la práctica responsable, los principios éticos quedarían como meras ideas sin aplicación real.

La ética es el *"qué"* y el *"por qué"*, mientras que la práctica responsable es el *"cómo"*. Ambos son esenciales para asegurar que la inteligencia artificial se utilice de manera beneficiosa y justa en la educación.

Fomentar una comprensión crítica de la tecnología en el aula no solo mejora el aprendizaje y la enseñanza, sino que también prepara a los estudiantes para ser ciudadanos informados y responsables en un mundo cada vez más digitalizado.

La implementación de la IA en el aula debe ser cuidadosamente gestionada para asegurar que se utilice de manera ética y equitativa.

La privacidad y la seguridad de los datos de los estudiantes, la transparencia en la toma de decisiones de los sistemas de IA y la equidad en el acceso a estas tecnologías, son cuestiones críticas que deben abordarse para garantizar un uso responsable y beneficioso de la IA en el aula.

El uso excesivo de tecnologías de IA en el aula, sin considerar el bienestar emocional de los estudiantes, puede tener efectos negativos. En algunos casos, la presión para adaptarse a nuevas tecnologías y la constante vigilancia a través de sistemas de monitoreo han aumentado los niveles de estrés y ansiedad entre los estudiantes.

Estos ejemplos subrayan la importancia de abordar los aspectos éticos de la IA de manera proactiva y responsable para evitar consecuencias negativas y garantizar un entorno educativo justo y seguro.

En este contexto, enseñar la comprensión crítica de la IA en las escuelas no solo prepara a los estudiantes para el ejercicio profesional, sino que también los capacita para ser ciudadanos integrales que utilizan la tecnología a través de prácticas responsables y éticas, capaces de tomar decisiones con base en un análisis crítico de la información, manejar sus datos personales de manera segura, reconocer y combatir sesgos en los sistemas de IA, adaptarse a cambios tecnológicos, innovar con responsabilidad desde una perspectiva social y ética, solucionar problemas reales y experimentar con tecnologías emergentes que promuevan el bien común, al asumirse como ciudadanos digitales responsables.

REFERENCIAS BIBLIOGRÁFICAS

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias. ProPublica. Recuperado de https://www.propublica.org

Ayudaley (2024, septiembre 25). La Seudonimización de los datos y su importancia en el nuevo RGPD. https://ayudaleyprotecciondatos.es/2020/11/18/seudonimizacion-dedatos/#google_vignette

Baker, R. S., & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. Handbook of Learning Analytics.

Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton & Company.

Calvo, Patrici. (2022). Una ética de la investigación en el marco de las éticas aplicadas. Veritas, (52), 29-51. https://dx.doi.org/10.4067/S0718-92732022000200029

Center for Teaching Innovation [CTI]. (2024, 20 de septiembre). Ethical AI for Teaching and Learning. https://teaching.cornell.edu/generative-artificial-intelligence/ethical-aiteaching-and-learning

Chen, X., & Zhang, Y. (2021). Artificial Intelligence in Education: A Review. Journal of Educational Technology & Society, 24(2), 10-26.

Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans—and where they can't (yet). McKinsey Quarterly. https://www.mckinsey.com

Comisión Europea. (2018). Digital Competence Framework for Citizens.

https://ec.europa.eu

Comisión Europea. (2020). White Paper on Artificial Intelligence: A European approach to excellence and trust. https://ec.europa.eu

Dastin, J., & Zeng, M. (2022). Ethical AI: Understanding and Implementing AI with Integrity. TechWorld Publications.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. Psychological Inquiry, 11(4), 227-268.

Educaopen (2024). Metatérminos. Explicabilidad.Educa Edtech Group.

https://www.educaopen.com/digital-lab/metaterminos/e/explicabilidad

El proyector de clase (agosto 12, 2024). 10 herramientas de IA que todo docente debería conocer.https://www.elproyectordeclase.com/2024/08/10-herramientas-de-ia-quetodo-docente-deberia-conocer.html

Facione, P. A. (2015). Critical Thinking: What It Is and Why It Counts. Measured Reasons.

Fundación Telefónica. (2024). Usos, posibilidades y riesgos de la IA en la educación. Fundación Telefónica https://www.fundaciontelefonica.com/noticias/usosposibilidades-y-riesgos-de-la-ia-en-la-educacion/

García, M., & Sánchez, J. (2020). Habilidades digitales en la educación moderna. Editorial Académica.

Garrison, D. R., Anderson, T., & Archer, W. (2010). The First Decade of the Community of Inquiry Framework: A Retrospective. The Internet and Higher Education, 13(1-2), 5-9.

González Arencibia, Mario, & Martínez Cardero, Dagmaris. (2020). Dilemas éticos en el escenario de la inteligencia artificial. Economía y Sociedad, 25(57), 93-109. https://dx.doi.org/10.15359/eys.25-57.5

Harris, S. (2020). Digital Literacy and Critical Thinking. Cambridge University Press.

Hattie, J. (2009). Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. Routledge.

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments Ecosystem: Building a Platform that Supports Education. International Journal of Artificial Intelligence in Education, 24(1), 73-103.

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.

International Business Machines Corporation [IBM]. (2024, 15 de septiembre). ¿Qué es la ética de la IA?. https://www.ibm.com/mx-es/topics/ai-ethics.

International Business Machines Corporation [IBM]. (2024, 15 de septiembre). ¿Qué es la IA explicable?. https://www.ibm.com/es-es/topics/explainable-ai

Kerr, B. (2013). The Role of Learning Analytics in the Future of Education. Proceedings of the 3rd International Conference on Learning Analytics and Knowledge. London Review of Education, 19 (1), 5, 1–14.https://doi.org/10.14324/LRE.19.1.05

Microsoft Copilot. (2024). Copilot Bing. https://copilot.microsoft.com/

Mikropoulos, T. A., & Natsis, A. (2011). Virtual Reality in Education: A Review. Educational Technology & Society, 14(1), 1-10.

Nguyen, A., Ngo, H.N., Hong, Y. et al. Ethical principles for artificial intelligence in education. Educ Inf Technol 28, 4221–4241 (2023). https://doi.org/10.1007/s10639-022-11316w

Nguyen, T. (2019). Preparing Students for the Future: The Role of AI in Education. Educational Review, 71(2), 223-236.

Noddings, N. (2012). The Caring Teacher: An Ethical Approach to Education. Teachers College Press.

Observatorio del Instituto para el futuro de la educación (2024). Principios éticos de la educación con Inteligencia Artificial (IA). Observatorio Tecnológico de Monterrey.

https://observatorio.tec.mx/edu-news/principios-eticos-de-la-educacion-coninteligencia-artificial-ia/

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.

OpenAI. (2024). ChatGPT [Modelo de lenguaje de gran tamaño].

https://chat.openai.com/chat

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO]. (2024, 25 de septiembre). Foro Global sobre la ética de

la IA 2024. https://www.unesco.org/es/artificial-intelligence/recommendation-ethics.

Paul, R., & Elder, L. (2014). The Miniature Guide to Critical Thinking Concepts and Tools. Foundation for Critical Thinking.

Pellegrino, J. W., & Hilton, M. L. (Eds.). (2012). Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century. National Academies Press.

Piaget, J. (1976). To Understand is to Invent: The Future of Education. Viking Press.

Reiss, M.J. (2021). 'The use of AI in education: Practicalities and ethical considerations'.

Rithvik Gujjula, Kamaljeet, Sanghera. (2023). Ethical Considerations and Data Privacy in AI Education. Jorunal of Student-Scientist´Research. https://doi.org/10.13021/jssr2023.3958

Selwyn, N. (2016). Education and Technology: Key Issues and Debates. Bloomsbury Academic.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. American Behavioral Scientist, 57(10), 1380-1400.

Smith, A. (2021). Digital Literacy and Skills for the Future. Tech Publishing.

Smowltech (11 de julio 2023). Anonimización: qué es y diferencia con seudonimización. https://smowl.net/es/blog/anonimizacion/

UNESCO. (2021). AI and education: A guidance document. https://unesdoc.unesco.org

UNESCO. (2024). El uso de la IA en la educación: decidir el futuro que queremos. Recuperado de UNESCO https://www.unesco.org/es/articles/el-uso-de-la-ia-en-laeducacion-decidir-el-futuro-que-queremos

UNESCO. (2024). Orientación para la IA Generativa en la educación y la investigación. Recuperado de UNESCO https://www.unesco.org/es/digital-education/ai-futurelearning/guidance

UNIR. (2023). Educación y desigualdad: claves frente a la desigualdad

educativa.https://www.unir.net/ciencias-sociales/revista/educacion-y-desigualdad/ Valdeomillos, C. (2024). La IA en la educación: cómo integrarla y que desafío plantea. Alfabetización Digital. https://alfabetizaciondigital.redem.org/la-ia-en-la-educacioncomo-integrarla-y-que-desafios-plantea/

Zawacki-Richter, O., Marín, V. I., & Liern, V. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education. Educational Technology Research and Development, 67(5), 1027-1058.

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.

CAPÍTULO IX

EL PAPEL DUAL DE LA IA ANTE EL CAMBIO CLIMÁTICO E INICIATIVAS GLOBALES

AYLIN TORRES-RAMÍREZ[1]

RESUMEN

El trabajo explora el papel dual de la inteligencia artificial (IA) en el contexto del cambio climático, reconociendo su capacidad tanto para contribuir al problema como para ser una herramienta en la búsqueda de soluciones. Se analizan las ventajas y desventajas de la IA en diversos sectores, destacando su potencial para mejorar la eficiencia energética, optimizar la agricultura, y generar modelos de predicción climática más precisos. Sin embargo, el documento también alerta sobre los riesgos asociados, como el aumento del consumo energético e incremento de la huella de carbono, la generación de residuos electrónicos y la necesidad de asegurar la privacidad de los datos. Existe una visión optimista del futuro de la IA en la lucha contra el cambio climático, pero se enfatiza la importancia de una supervisión adecuada y el desarrollo de una ética responsable que garantice un uso sostenible. Iniciativas globales trabajan en la integración de la IA en la agenda de desarrollo sostenible donde enfatizan la importancia de desarrollar una ética para la IA que se base en principios de responsabilidad, cuidado del medio ambiente y respeto por los derechos humanos.

Palabras clave: inteligencia artificial, cambio climático, iniciativas.

[1] Alumna del Doctorado en Estudios Organizacionales, Universidad Autónoma Metropolitana, Ciudad de México, México.

INTRODUCCIÓN

Desde sus inicios hasta su actual puesta en marcha, la Inteligencia Artificial (en adelante IA) sigue siendo objeto de considerables estudios e investigaciones en torno al tema de ética con cierta atención necesaria, por ejemplo, por cuestiones sobre la privacidad, e implicado, a la vez, una estrecha relación con el cambio climático.

La crisis ambiental es un aspecto inherente a la sociedad moderna industrial. No solo es una crisis social externa, sino también interna, que tiene que ver con los valores, los principios de la vida y los humanos, con la ética y la moral. En este sentido, el desarrollo tecnológico ha generado un proceso transformador en el cual todos los aspectos de la vida cotidiana han sido alterados sensiblemente.

Autores como Nordgren (2022) asumen que la IA tiene un doble papel con respecto al cambio climático: por una parte, se visualiza como un elemento importante del problema y contribuye al mismo cambio climático, y, por otro lado, aparece como un elemento clave para su mitigación y adaptación.

El objetivo de este trabajo es comprender y analizar el impacto de la inteligencia artificial en el cambio climático, incluyendo tanto sus potenciales beneficios como sus riesgos, y que sirva como reflexión para impulsar la adopción responsable de tecnologías de IA que contribuyan a la mitigación y adaptación al cambio climático. Por lo que este artículo consta de tres partes principales. La primera parte aborda una breve explicación sobre la IA y el cambio climático, donde se presentan algunos de los casos de cómo la IA puede aportar a la lucha contra el cambio climático. La segunda parte se abordan los problemas que son visibles en cuanto al uso de la IA en torno al medio ambiente. La parte final muestra aquellos documentos, pactos y acciones que se han puesto en marcha a nivel internacional en cuanto a la adopción de la IA y el cambio climático.

Los analistas que mantienen una perspectiva optimista respecto al crecimiento sostenible y las posibilidades que ofrece la innovación tecnológica tienden a ser más favorables hacia la aplicación de la inteligencia artificial en la mitigación del cambio climático. En contraste, aquellos que enfatizan la necesidad de un decrecimiento y la relevan-

cia de operar dentro de los límites del planeta suelen adoptar una postura más pesimista.

¿CÓMO PUEDE LA IA AYUDAR A SALVAR EL MEDIO AMBIENTE?

Según Nordgren (2022) la IA es una forma de tecnología de la información y la comunicación (TIC) que emula la inteligencia humana. Posee la capacidad de generar un impacto significativo en la sociedad, ya que se puede aplicar en la mayoría o si no en todos los sectores, como el industrial, el de transporte, salud y negocios, tanto en el presente como en el futuro, a corto o largo plazo (Coeckelbergh, 2020). Se trata de programas que operan en línea, como los típicos motores de búsqueda, análisis de imágenes o *bots*, así como también robots o en el *"Internet de las cosas"*.

La inteligencia artificial es frecuentemente considerada como una *"tecnología sostenible"*. Por ejemplo, empresas como *Apple* y *Google* afirman ser neutrales en carbono gracias a la adquisición de créditos de emisión, mientras que *Microsoft* ha declarado su intención de ser carbono negativo para el año 2030 (Crawford, 2021). El movimiento *Green AI,* principalmente impulsado por estas empresas, además de la participación de *Facebook* e *IBM,* busca minimizar el consumo energético de la IA y reducir la huella de carbono a través de la optimización de algoritmos, la eficiencia energética en los centros de datos y el uso de materiales reciclados y desarrollo de hardware de IA más eficiente y duradero. No obstante, persiste la incertidumbre respecto a la magnitud real de las emisiones generadas.

Hay algo que la IA puede hacer en un tiempo corto a diferencia de las personas, y es leer y analizar rápidamente enormes cantidades de datos y con ello poder tomar decisiones. Por ejemplo, tiene la capacidad de mejorar el diseño de materiales para paneles solares, promover un transporte más sostenible, optimizar el control de la deforestación y construir ciudades resilientes (Chen y otros, 2023).

Así, la inteligencia artificial puede ser empleada en el diseño de nuevos edificios que sean intrínsecamente más eficientes en términos de energía, utilizando herramientas avanzadas de modelado y si-

mulación (Baduge y otros, 2022). A través de la optimización de la orientación del edificio, la disposición de las ventanas, el aislamiento y otros aspectos del diseño, arquitectos tienen la capacidad de desarrollar edificaciones que no solo sean energéticamente eficientes, sino también confortables para sus ocupantes (Debrah y otros, 2022). En consecuencia, este tipo de edificaciones aplicando la IA tendrían la capacidad de disminuir de manera considerable el consumo energético y las emisiones de carbono.

También, la IA tiene la capacidad de ofrecer pronósticos meteorológicos más precisos, lo que a su vez puede mejorar la protección contra eventos climáticos extremos. Por ejemplo, aplicando algoritmos de aprendizaje automático, puede crear modelos y facilitar la predicción de la trayectoria y la intensidad de tormentas y huracanes, así como anticipar el inicio de sequías o inundaciones, la temperatura, la cantidad de precipitación y la velocidad del viento (Chen y otros, 2023).

Su participación también es prometedora en los sistemas de transporte; a través del análisis de los datos, la inteligencia artificial puede modernizar el transporte refinando las rutas, administrando flotas, desarrollando vehículos autónomos, optimizando el transporte público y regulando la demanda (Abduljabbar y otros, *2019*).

En lo que respecta al sector agrícola, la IA tiene la capacidad de implementar sistemas de alerta temprana para las cosechas, optimizando la selección de semillas, los tiempos de siembra y las técnicas de cultivo, con el fin de mitigar los efectos adversos de las condiciones meteorológicas desfavorables (Rolnick y otros, 2019). Puede establecer modelos de ciclo de nutrientes y productividad de cultivos para identificar plagas, detectar enfermedades y reducir el uso de fertilizantes y pesticidas (Zhang y otros, *2021*).

Por último, pero no por eso la última opción entra la más existentes, la IA puede ser utilizada en múltiples facetas de la gestión de residuos, incluyendo la transformación de desechos en energía, la clasificación de materiales, la elaboración de modelos de generación de residuos, la logística, la disposición final y la recuperación de recursos (Fang y otros, *2023*).

¿AYUDARÍA LA IA A RESOLVER LA CRISIS CLIMÁTICA O CONLLEVA MÁS IMPACTOS NEGATIVOS?

En las últimas décadas, investigaciones científicas realizadas por entidades como el Panel Intergubernamental sobre Cambio Climático (IPCC) han evidenciado el rápido deterioro ambiental que está sufriendo el planeta, resultado de la interacción que los seres humanos mantienen con el entorno natural.

Se deben ser consiente tanto de las ventajas como de las desventajas de la IA como respuesta a la mitigación del cambio climático. De hecho, es importante abordar y reconocer los impactos que la misma inteligencia genera en el ambiente y cómo también es parte del problema.

Belkhir y Elmeligi (2018) llevaron a cabo una investigación significativa sobre la huella de carbono del sector de las tecnologías de la información y la comunicación (TIC), incluyendo tanto la producción de la energía operativa de los dispositivos hasta la energía operativa para la infraestructura. Este estudio incluyó dispositivos informáticos como computadoras de escritorio, laptops, teléfonos inteligentes y tabletas, así como centros de datos y redes de comunicación. Se reveló que las emisiones de gases de efecto invernadero de las TIC superarán el 14 % de las emisiones mundiales en 2040. El estudio reveló que la mayor parte de las emisiones las genera la infraestructura de las TIC, siendo los centros de datos los que tienen mayor protagonismo (45 %) seguidos de las redes de comunicaciones (24 %).

Conforme a esta inteligencia avance y se extienda, los problemas e impactos con el ambiente irán en progreso. Según la Agencia Internacional de la Energía, para el año 2024 se registran más de ocho mil centros de datos en el planeta. Se espera que para el año 2026, el consumo de energía se duplique con respecto al año 2022 (IEA, 2024). Los grandes sistemas de IA con los que la sociedad está familiarizada como lo son los *chatbots,* funcionan con supercomputadoras las cuales consumen grandes cantidades de electricidad. En este sentido, grandes cantidades de electricidad requieren grandes cantidades de calor y ruido. Para evitar que los centros de electricidad se sobrecalienten,

hay que refrigerarlos, lo cual se puede hacer utilizando aire acondicionado lo cual implica energía, y utilizar grandes cantidades de agua.

De esta manera, se observa cómo la inteligencia artificial consume recursos reales: el proceso de entrenamiento de los algoritmos de inteligencia artificial demanda una gran cantidad de energía, lo que resulta en un incremento significativo de la huella de carbono. Además, el acelerado desarrollo de la inteligencia artificial provoca que el hardware se vuelva obsoleto en poco tiempo, lo que resulta en un incremento de desechos electrónicos y, de nuevo, una contribuye a una mayor huella de carbono.

A medida que se incrementa el uso de la inteligencia artificial, se generan un mayor consumo de recursos, tanto por parte de las empresas que la crean como de los usuarios que la emplean para encontrar soluciones.

CONTRIBUCIONES A LA DISCUSIÓN SOBRE LA RESPONSABILIDAD DE LOS GOBIERNOS POR LA IA EN RELACIÓN CON EL CAMBIO CLIMÁTICO

De los principios de la Economía Ecológica de Herman Daly, se resalta el hecho de que el desarrollo de la tecnología debe generar el mayor beneficio sostenible posible por el uso del recurso natural (Daly, 2015).

La inteligencia artificial tendría el papel de asistir a los gobiernos en la consecución de sus metas climáticas, tanto a nivel nacional como internacional, al optimizar la promoción y regulación de políticas ambientales (Coeckelbergh, 2021).

Entre las contribuciones a la discusión sobre la responsabilidad de los gobiernos por la IA en relación con el cambio climático se resaltan las siguientes acciones (ver **Tabla 1**):

Tabla 1. Contribuciones internacionales sobre la IA

No.	Año	Organización	Tipo de acción	Acción
1	2023	Unión Internacional de Telecomunicaciones (ITU *por sus siglas en inglés*)	Documento	*"United Nations Activities on Artificial Intelligence (AI) 2023"*
2	2021	UNESCO	Documento	*"Recomendación sobre la ética de la inteligencia artificial"*
3	2020	Unión Internacional de Telecomunicaciones (ITU *por sus siglas en inglés)*	Iniciativa	Grupo de Trabajo Interinstitucional de las Naciones Unidas sobre IA (ITU-AI Task Force)
4	2020	Unión Internacional de Telecomunicaciones (ITU *por sus siglas en inglés)*	Documento	"*Frontier technologies to protect the environment and tackle climate change"*
5	2019	Naciones Unidas	Plan integral	Estrategia de las Naciones Unidas sobre Inteligencia Artificial para el Bien (UNAI for Good)
6	2021	Gobierno de los Países Bajos, en colaboración con empresas líderes del sector de los centros de datos y organizaciones de la sociedad civil	Iniciativa global del sector empresarial y sociedades civiles	Pacto de Centros de Datos Climáticamente Neutrales
7	2019	Comisión Europea	Iniciativa política	Pacto Verde Europeo
8	2017	Unión Internacional de Telecomunicaciones (ITU), en colaboración con otras agencias de la ONU, gobiernos, empresas, organizaciones de la sociedad civil, y la comunidad científica.	Iniciativa	AI for Good

Fuente: elaboración propia.

El documento *"United Nations Activities on Artificial Intelligence (AI) 2023"* de la Unión Internacional de Telecomunicaciones (ITU *por sus siglas en inglés)* es una recopilación de proyectos e iniciativas relacionados con la Inteligencia Artificial que se están llevando a cabo en diferentes agencias de la ONU. Los puntos más importantes del documento son:

- Creciente interés y uso de la IA en la ONU: Se destaca el creciente interés y uso de la IA en las agencias de la ONU para avanzar en la consecución de los ODS.
- Enfoque en la ética y los derechos humanos: Se pone énfasis en la necesidad de desarrollar y aplicar la IA de manera ética y responsable, asegurando que se respeten los derechos humanos.
- Colaboración entre las agencias: Se resalta la importancia de la colaboración entre las diferentes agencias de la ONU para compartir conocimientos, recursos y mejores prácticas.
- Diversidad de aplicaciones: Se presentan ejemplos de proyectos e iniciativas en diferentes áreas, incluyendo la salud, la educación, el medio ambiente, la agricultura, la seguridad, el desarrollo económico, el desarrollo social y la ayuda humanitaria.
- Integración de la IA en los procesos operativos: Se ilustra cómo se está utilizando la IA para mejorar la eficiencia y eficacia de las operaciones de las agencias.

El documento *"Recomendación sobre la ética de la inteligencia artificial"* elaborado por la UNESCO establece un marco global de valores, principios y acciones destinado a guiar a los Estados miembros en la creación de sus legislaciones y políticas relacionadas con la inteligencia artificial. Habla de la importancia del medio ambiente en varios puntos:

- Responsabilidad de proteger el medio ambiente: El documento reconoce que la IA puede tener un impacto negativo en el medio ambiente y los ecosistemas, y que es necesario asegurar que los sistemas de IA sean utilizados de forma sostenible y respetuosa con el medio ambiente.

- Integración de la sostenibilidad en el desarrollo de la IA: Se plantea la necesidad de integrar la sostenibilidad en todas las etapas del desarrollo de la IA, desde la elección de métodos hasta la evaluación de impactos.
- Reducción de la huella ambiental de la IA: Se anima a los Estados miembros a que tomen medidas para reducir el impacto ambiental de los sistemas de IA, como su huella de carbono, consumo energético y extracción de recursos.
- Utilización de la IA para la protección ambiental: Se destaca que la IA puede utilizarse para la protección y gestión del medio ambiente, como el seguimiento de recursos naturales, el control de la contaminación y la gestión de los desastres.
- Garantizar la diversidad de los ecosistemas: Se reconoce la importancia de proteger la diversidad de los ecosistemas y se insta a los actores de la IA a que tomen en cuenta este factor en sus decisiones.
- Colaboración internacional: Se reconoce que la cooperación internacional es esencial para abordar los retos ambientales relacionados con la IA y se fomenta la colaboración entre los Estados miembros para compartir buenas prácticas y desarrollar soluciones sostenibles.

El Grupo de Trabajo Interinstitucional de las Naciones Unidas sobre IA (ITU-AI Task Force) es un grupo de trabajo establecido por la ITU (por sus siglas en inglés) para guiar y apoyar la implementación de la Estrategia de las Naciones Unidas sobre Inteligencia Artificial para el Bien. Entre sus objetivos principales está el coordinar y facilitar la colaboración entre las diferentes agencias y entidades de la ONU en el campo de la IA, promover el uso responsable y ético de la IA para el desarrollo sostenible y el bien común e identificar las oportunidades y los desafíos que presenta la IA para la implementación de la Agenda 2030 para el Desarrollo Sostenible.

El documento "*Frontier technologies to protect the environment and tackle climate change*" de la ITU destaca la importancia de las tecnolo-

gías de vanguardia para proteger el medio ambiente y abordar el cambio climático. Algunos de los puntos más relevantes son:

- Las tecnologías de vanguardia como herramientas para combatir el cambio climático: El documento analiza cómo la Inteligencia Artificial (IA), el Internet de las Cosas (IoT), la tecnología 5G, las energías limpias, los gemelos digitales, la robótica, las tecnologías espaciales 2.0 y la digitalización con Big Data pueden contribuir a mitigar el cambio climático.
- La necesidad de un enfoque integrado: El documento enfatiza la necesidad de un enfoque integrado que aborde los desafíos del cambio climático y la contaminación del aire al mismo tiempo, reconociendo que estos dos problemas están estrechamente relacionados.
- La importancia de las ciudades: Se destaca el papel crucial de las ciudades en la lucha contra el cambio climático y se analizan casos de éxito de ciudades como Moscú, Dubái y Singapur.
- La importancia de las asociaciones: El documento destaca la importancia de las asociaciones entre gobiernos, sector privado, academia y la sociedad civil para impulsar el desarrollo y la implementación de las tecnologías de vanguardia para combatir el cambio climático.
- La importancia de la ética: Se reconoce la necesidad de utilizar las tecnologías de vanguardia de manera ética y responsable, asegurando que se respeten los derechos humanos y que los beneficios de estas tecnologías se compartan equitativamente.
- El problema del impacto ambiental de las tecnologías: Se reconoce que las tecnologías de vanguardia, incluyendo la IA, también tienen un impacto ambiental, especialmente en términos de consumo energético. Se destaca la necesidad de abordar estos impactos y de desarrollar tecnologías más sostenibles.

La Estrategia de las Naciones Unidas sobre Inteligencia Artificial para el Bien (UNAI for Good) es un enfoque integral que fue presentado en junio de 2019 en el marco de la Cumbre de las Naciones Unidas

sobre la Inteligencia Artificial para el Bien, que tuvo lugar en Ginebra, Suiza, con el propósito de fomentar el uso ético y responsable de la inteligencia artificial (IA), buscando así contribuir a los Objetivos de Desarrollo Sostenible (ODS). Su objetivo fundamental es orientar el progreso y la implementación de la inteligencia artificial en beneficio de la humanidad, enfrentando los retos globales y fomentando soluciones creativas para el desarrollo sostenible, la paz, la seguridad, los derechos humanos y la equidad. Entre sus puntos clave de la estrategia se encuentra:

- Ética y valores: Promueve el desarrollo de la IA de acuerdo con principios éticos, como la inclusión, la transparencia, la rendición de cuentas, la no discriminación, la protección de los datos y la privacidad.
- Colaboración: Fomenta la colaboración entre gobiernos, el sector privado, la sociedad civil y la comunidad científica para promover el desarrollo y la aplicación responsable de la IA.
- Innovación: Promueve la investigación e innovación en IA para abordar los desafíos globales y promover el desarrollo sostenible.
- Capacitación: Busca fortalecer las capacidades de los países para aprovechar las oportunidades que ofrece la IA.
- Acceso: Promueve el acceso equitativo a la IA para todos, especialmente para los países en desarrollo.

El Pacto de Centros de Datos Climáticamente Neutrales (*Climate Neutral Data Centre Pact*) es una iniciativa global que busca acelerar la transición de los centros de datos hacia la neutralidad de carbono. Fue lanzado en 2021 por el gobierno de los Países Bajos, en colaboración con empresas líderes del sector de los centros de datos y organizaciones de la sociedad civil. Entre sus objetivos está el promover la sostenibilidad en el sector de los centros de datos a través de la reducción de las emisiones de gases de efecto invernadero; establecer metas ambiciosas para alcanzar la neutralidad de carbono, incluyendo el uso de energía renovable, la eficiencia energética y la compensación de emisiones; y crear un marco de transparencia y rendición de cuentas para

el sector de los centros de datos, permitiendo a los usuarios identificar y elegir centros de datos sostenibles.

El Pacto Verde Europeo analiza el potencial de la inteligencia artificial en la reducción del cambio climático, al tiempo que reconoce los riesgos asociados, como el aumento en el consumo de energía y recursos. Subraya la necesidad de asegurar que los sistemas de inteligencia artificial creados por empresas privadas cumplan con los principios del desarrollo sostenible. Para ello, los mecanismos de auditoría y certificación se presentan como alternativas pertinentes.

La iniciativa *AI for Good,* promovida por las Naciones Unidas, tiene como objetivo fomentar la aplicación de la inteligencia artificial (IA) para enfrentar los retos globales y alcanzar los Objetivos de Desarrollo Sostenible (ODS). Entre sus principales actividades se encuentra:

- Conferencias: Organiza conferencias anuales para reunir a expertos en IA, líderes de opinión y representantes de gobiernos, empresas y organizaciones de la sociedad civil para discutir el uso de la IA para el bien.
- Desafíos: Lanza desafíos globales para que equipos de investigadores y desarrolladores de IA trabajen en la creación de soluciones innovadoras para abordar los problemas más urgentes del mundo.
- Proyectos: Apoya proyectos de IA que contribuyan a resolver problemas en áreas como la salud, la educación, la seguridad, la sostenibilidad, etc.
- Recursos: Proporciona recursos y herramientas para apoyar a las personas y las organizaciones en el uso de la IA para el bien social.

En su informe de 2023 titulado *"Informe de las Naciones Unidas sobre Actividades de Inteligencia Artificial (IA)", AI for Good* reportó que existen un total de 408 casos y proyectos de inteligencia artificial que abarcan los 17 ODS, donde, los principalmente abordados son el ODS 9 Industria, Innovación e Infraestructura, el ODS 10 Reducción de las desigualdades, el ODS 16 Paz y Justicia, el ODS 17 Alianzas para

lograr los objetivos y el ODS 13 Acción por el Clima. Los resultados se tratan de informes y herramientas de software desarrolladas, como aplicaciones móviles, aplicaciones web, paneles de control con capacidad de búsqueda y *chatbots* generativos impulsados por IA.

Por parte de la academia, Vinuesa y otros (2020) analizan la conexión entre la inteligencia artificial y los 17 Objetivos de Desarrollo Sostenible, así como las 169 metas establecidas en la *"Agenda 2030 para el Desarrollo Sostenible"*. La investigación concluye que la inteligencia artificial puede facilitar el logro de 134 metas de los 17 objetivos, aunque también puede obstaculizar 59 de ellas. Los impactos sociales, económicos y ambientales derivados del uso de la inteligencia artificial pueden ser tanto beneficiosos como perjudiciales. En lo que respecta al cambio climático, el estudio resalta el potencial de la inteligencia artificial para contribuir a la acción climática, al tiempo que señala sus significativas demandas energéticas y sus altas emisiones. Se concluye que el acelerado avance de la inteligencia artificial requiere el apoyo de una supervisión adecuada y un entendimiento regulatorio de las tecnologías fundamentadas en IA, con el fin de facilitar un desarrollo sostenible. De no implementarse estas medidas, podrían surgir deficiencias en la transparencia, la seguridad y los estándares éticos (Vinuesa y otros, 2020).

DISCUSIÓN

Tal como se menciona en diversas secciones de este documento, la Inteligencia Artificial desempeña un papel dual en el contexto del cambio climático. No solo influye en el cambio climático, sino que también tiene el potencial de ayudar en su mitigación.

La IA ha surgido como una herramienta tecnológica con un gran potencial para mejorar la eficiencia energética y promover el desarrollo sostenible en el sector energético. Su implementación es un proceso continuo que depende de la precisión de los datos de entrada y la selección adecuada de algoritmos (Ouadah y otros, 2022).

La IA se ha mostrado prometedora en la predicción del clima, con la capacidad de disminuir la frecuencia de desastres naturales y opti-

mizar la utilización del suelo. Asimismo, la agricultura de precisión, impulsada por la IA, reduce la incertidumbre en la aplicación de fertilizantes y pesticidas, optimizando su uso y aumentando la producción alimentaria.

Sin embargo, la aplicación de la IA en los sistemas de transporte, que se basan en la recopilación de grandes volúmenes de datos, incluyendo información personal, exige un enfoque ético cuidadoso para proteger la confidencialidad y privacidad de los usuarios, fomentando así la confianza en estos sistemas y evitando un posible uso indebido o explotación.

Se reconoce el impacto ambiental de la IA, especialmente en términos de consumo energético y emisión de gases de efecto invernadero. Por lo anterior, es necesario desarrollar una ética para la IA que tenga en cuenta la sostenibilidad ambiental y que vaya más allá de los planteamientos tradicionales. La ley de inteligencia artificial de la Unión Europea hace referencia de su impacto medioambiental; pide que los sistemas de IA se desarrollen y utilicen de manera sostenible y respetuosa con el medio ambiente, aunque no especifica lo que esto significa.

CONCLUSIÓN

El documento es un llamado a la acción para utilizar las tecnologías de vanguardia de manera responsable y eficiente para combatir el cambio climático. Es un recurso valioso para quienes buscan comprender cómo estas tecnologías pueden utilizarse para crear un futuro más sostenible para el planeta.

Hay quienes sostienen que la inteligencia artificial representa un *"factor de cambio"* (Cho, 2018) en la lucha contra el cambio climático. La IA tiene el potencial y la obligación de ser un elemento fundamental en los esfuerzos de mitigación y adaptación. Sin embargo, otros argumentan que la IA podría ser más un obstáculo o una carga en relación con el cambio climático que una solución, especialmente considerando que su elevado consumo energético depende en gran medida

de los combustibles fósiles. Lo que realmente importa son las políticas globales, como la implementación de precios al carbono.

La información recopilada indica que, siempre que se ejerza la debida precaución, la utilización de inteligencia artificial puede ofrecer un respaldo valioso a los esfuerzos internacionales para comprender y enfrentar de manera más efectiva los diversos retos relacionados con el cambio climático.

REFERENCIAS

Abduljabbar, R., Dia, H., Liyanage, S., y Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. *Sustainability*, *11*(1), 189.

Baduge, S. K., Thilakarathna, S., Perera, J. S., Arashpour, M., Sharafi, P., Teodosio, B., ... y Mendis, P. (2022). Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*, *141*, 104440.

Belkhir, L., y Elmeligi, A. (2018). Assessing ICT global emissions footprint: Trends to 2040 & recommendations. *Journal of cleaner production*, *177*, 448-463.

Chen, L., Chen, Z., Zhang, Y., Liu, Y., Osman, A. I., Farghali, M., ... y Yap, P. S. (2023). Artificial intelligence-based solutions for climate change: a review. *Environmental Chemistry Letters*, *21*(5), 2525-2557.

Coeckelbergh, M. (2020). *AI ethics*. The MIT Press.

Coeckelbergh, M. (2021). AI for climate: freedom, justice, and other ethical and political challenges. *AI and Ethics*, *1*(1), 67-72.

Crawford, K. (2021). The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.

Daly, H. E. (2015). Steady-state economics. In Thinking About the Environment (pp. 250-255). Routledge.

Debrah, C., Chan, A. P., y Darko, A. (2022). Artificial intelligence in green building. *Automation in Construction*, *137*, 104192.

Fang, B., Yu, J., Chen, Z., Osman, A. I., Farghali, M., Ihara, I., ... y Yap, P. S. (2023). Artificial intelligence for waste management in smart cities: a review. *Environmental Chemistry Letters*, *21*(4), 1959-1989.

IEA. International Energy Agency. (2024). *Homepage*. https://www.iea.org/

Nordgren, A. (2022). Artificial intelligence and climate change: ethical issues. *Journal of Information, Communication and Ethics in Society*, *21*(1), 1-15.

Ouadah, A., Zemmouchi-Ghomari, L., y Salhi, N. (2022). Selecting an appropriate supervised machine learning algorithm for predictive maintenance. The International Journal of Advanced Manufacturing Technology, 119(7), 4277-4301.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... y Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)*, *55*(2), 1-96.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... y Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, *11*(1), 1-10.

Zhang, P., Guo, Z., Ullah, S., Melagraki, G., Afantitis, A., y Lynch, I. (2021). Nanotechnology and artificial intelligence to enable sustainable and precision agriculture. *Nature Plants*, *7*(7), 864-876.

CAPÍTULO X

EXPLORANDO NUEVAS FRONTERAS: IA Y EL FUTURO DE LA GALERÍA VIRTUAL EN LA UAM AZCAPOTZALCO

YADIRA ZAVALA OSORIO[1]
VÍCTOR GÓMEZ–QUINTERO[2]

RESUMEN

La Galería del Tiempo Virtual de la Universidad Autónoma Metropolitana (UAM) es una plataforma digital dedicada a la difusión del patrimonio cultural y artístico de la institución. Sin embargo, hasta la fecha, no ha integrado completamente el potencial de la inteligencia artificial (IA). Este artículo explora propuestas creativas para incorporar la IA en la galería, con el fin de mejorar la interactividad, personalización y accesibilidad de las exposiciones. Entre las principales sugerencias se incluye la implementación de curaduría automatizada, asistentes virtuales personalizados y experiencias inmersivas mediante realidad aumentada (AR). Además, la IA podría democratizar el acceso al arte al ofrecer traducciones automáticas en tiempo real y mejorar la accesibilidad para personas con discapacidades. Estas tecnologías no solo ampliarían el alcance de la galería, sino que también promoverían a artistas emergentes y facilitarían la preservación del acervo digital. En definitiva, la IA tiene el potencial de transformar la Galería del Tiem-

[1] Profesora Investigadora del Departamento de Sistemas, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

[2] Profesor-Investigador del Departamento de Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

po Virtual en un espacio cultural más dinámico, inclusivo y tecnológicamente avanzado.

Palabras clave: Inteligencia artificial (IA), Galerías virtuales, Curaduría automatizada, Realidad aumentada (AR), Accesibilidad digital.

INTRODUCCIÓN

La tecnología está transformando cada aspecto de nuestras vidas, y el campo de la cultura y las artes no es una excepción. En el ámbito de los museos y galerías, la inteligencia artificial (IA) ha comenzado a jugar un papel clave en la evolución de cómo las personas experimentan el arte. Si bien la Galería del Tiempo Virtual de la Universidad Autónoma Metropolitana (UAM) ya representa un avance significativo al permitir un acceso más amplio y democrático al arte, no ha aprovechado completamente el potencial que la IA ofrece para mejorar tanto la experiencia del usuario como la difusión de su acervo.

El objetivo de este artículo es explorar propuestas creativas para la integración futura de IA en la Galería del Tiempo Virtual. Con la implementación de IA, la galería podría evolucionar para ofrecer una experiencia más inmersiva, personalizada y accesible, al tiempo que optimiza la preservación de su acervo y apoya a los jóvenes creadores. Propuestas como la curaduría automatizada, asistentes virtuales interactivos, realidad aumentada (AR) y el uso de IA para la accesibilidad y traducción automática pueden transformar por completo la interacción de los usuarios con la galería y llevarla al siguiente nivel de innovación tecnológica.

En el ámbito global, museos como el *Louvre* y el *Museo* del *Prado* han comenzado a integrar IA para ofrecer experiencias personalizadas y optimizar la conservación de obras de arte (API4AI, 2024). La Galería del Tiempo Virtual de la UAM tiene el potencial de seguir este camino, no solo adaptándose a las tendencias tecnológicas actuales, sino también proyectando una visión innovadora de cómo la tecnología puede integrarse en la promoción cultural.

HISTORIA DE LA GALERÍA DEL TIEMPO Y SU EVOLUCIÓN HACIA LO VIRTUAL

La Galería del Tiempo fue inaugurada hace 25 años con la misión de crear un espacio dedicado a la exposición de arte en la UAM, brindando acceso a la comunidad universitaria y al público en general a obras de artistas consagrados y emergentes. La galería ha albergado exposiciones de gran renombre, como la obra de Juan O'Gorman, quien es uno de los máximos exponentes de la arquitectura mexicano. Estas exposiciones no solo han servido como una ventana hacia el pasado artístico de México, sino que también han facilitado el diálogo cultural dentro de la comunidad académica.

A medida que la galería fue consolidándose como un espacio fundamental para la difusión cultural en la UAM, las limitaciones físicas comenzaron a plantear desafíos. El espacio expositivo resultaba insuficiente para albergar las diversas obras que se querían mostrar, lo que llevó a las autoridades de la universidad a buscar alternativas para ampliar su capacidad. De este modo, surgió la idea de crear la Galería del Tiempo Virtual, un espacio digital que no solo permitiera exhibir obras de arte, sino que también ofreciera una mayor accesibilidad al público.

La primera exposición en el formato virtual fue una presentación del acervo de O'Gorman, lo que permitió a la comunidad universitaria y a otros usuarios acceder a una muestra artística que de otro modo habría estado limitada por el espacio físico de la galería. A pesar del éxito inicial, la Galería del Tiempo Virtual aún no ha aprovechado todo el potencial que la tecnología digital puede ofrecer, particularmente en lo que respecta a la inteligencia artificial. A través de la IA, la galería podría transformar radicalmente la experiencia del usuario, personalizando las exposiciones y facilitando el acceso global a sus colecciones (API4AI, 2024).

PROPUESTAS PARA LA INTEGRACIÓN DE IA: EXPANSIÓN DE LA EXPERIENCIA DEL USUARIO

Una de las primeras propuestas para integrar IA en la Galería del Tiempo Virtual es la implementación de curaduría automatizada, que

podría personalizar la experiencia de los usuarios en función de sus intereses. Actualmente, los museos más avanzados están utilizando IA para analizar las preferencias de los visitantes y generar exposiciones dinámicas que se ajusten a sus gustos. Al utilizar algoritmos de aprendizaje automático, la galería podría ofrecer recomendaciones personalizadas de obras, basadas en el comportamiento del usuario, las obras vistas anteriormente y los artistas que más les interesen (DataArt, 2023).

Curaduría automatizada: A través de la IA, la Galería del Tiempo Virtual podría generar exposiciones personalizadas para cada visitante. La curaduría tradicional requiere una intervención manual intensiva por parte de expertos que seleccionan las obras a exponer y su disposición en el espacio. Sin embargo, con la IA, este proceso podría automatizarse. El sistema podría generar exposiciones únicas en función de los intereses y comportamientos de los usuarios, brindándoles una experiencia más significativa e interactiva (Pigro, 2024).

Otra propuesta es la incorporación de asistentes virtuales impulsados por IA que guíen a los visitantes a lo largo de las exposiciones. Estos asistentes no solo podrían responder preguntas sobre las obras, sino también proporcionar contexto adicional sobre la historia del arte y los artistas. Los asistentes virtuales adaptados a los niveles de conocimiento del visitante pueden ofrecer explicaciones técnicas más complejas o generales, según sea necesario. Esta característica enriquecería enormemente la experiencia del usuario, haciéndola más participativa y educativa (AIUpbeat, 2024).

Gamificación basada en IA: Otro aspecto que podría integrarse en la experiencia del usuario es la gamificación. Utilizando IA, la galería podría diseñar recorridos interactivos donde los visitantes completen desafíos o descubran pistas ocultas en las obras de arte, lo que fomentaría una mayor interacción y participación. La gamificación es una herramienta poderosa para captar la atención de las audiencias más jóvenes y puede ser particularmente efectiva en el contexto de una galería virtual (DataArt, 2023).

FOMENTO DE JÓVENES CREADORES A TRAVÉS DE LA IA

La IA también tiene el potencial de facilitar la promoción de jóvenes creadores, brindándoles una plataforma donde puedan exhibir sus obras y conectarse con nuevas audiencias. En un entorno artístico competitivo, muchos artistas emergentes luchan por obtener la visibilidad necesaria para destacar, pero la IA podría ser la clave para superar estas barreras.

Algoritmos de recomendación: La Galería del Tiempo Virtual podría utilizar IA para analizar los patrones de comportamiento de los visitantes y recomendar obras de artistas emergentes que se alineen con los intereses de los usuarios. Este tipo de sistema es similar al que utilizan plataformas como *Spotify,* donde los algoritmos sugieren contenido nuevo basado en las preferencias previas del usuario. Al aplicar estos principios a la galería virtual, se podrían generar recomendaciones personalizadas de obras de artistas jóvenes, brindándoles mayor visibilidad y ayudando a los visitantes a descubrir nuevos talentos (API4AI, 2024).

Plataformas colaborativas: Otro aspecto a considerar es el desarrollo de plataformas colaborativas donde los artistas emergentes puedan conectarse directamente con curadores, galeristas y coleccionistas. La IA puede facilitar esta conexión al analizar las preferencias de los visitantes y poner en contacto a los artistas con aquellos que tienen intereses similares. Esto podría crear un espacio de colaboración donde los jóvenes creadores no solo muestren su obra, sino también reciban retroalimentación valiosa y oportunidades de exposición (AIUpbeat, 2024).

Además, la galería podría ofrecer recorridos virtuales guiados por los propios artistas, donde estos expliquen sus obras a través de IA y realidad aumentada. Esta interacción directa entre el creador y el público no solo enriquecería la experiencia del visitante, sino que también fortalecería el perfil de los artistas emergentes al darles una plataforma para compartir sus perspectivas personales y artísticas (Pigro, 2024).

DEMOCRATIZACIÓN DEL ACCESO AL ARTE A TRAVÉS DE IA

Una de las oportunidades más emocionantes que la IA ofrece es la posibilidad de democratizar el acceso al arte. Tradicionalmente, el acceso a las colecciones de arte ha estado limitado por barreras geográficas, económicas y físicas. Sin embargo, la Galería del Tiempo Virtual puede eliminar muchas de estas barreras, permitiendo que personas de todo el mundo accedan a sus exposiciones sin necesidad de viajar o pagar costosos boletos de entrada.

Traducción automática: La IA puede utilizarse para implementar traducciones automáticas en tiempo real, permitiendo que personas de diferentes partes del mundo disfruten de las exposiciones en su idioma nativo. Esta capacidad no solo amplía el alcance de la galería a nivel global, sino que también facilita el acceso para usuarios que de otro modo no podrían comprender las descripciones y contextos culturales de las obras expuestas (AIUpbeat, 2024).

Accesibilidad para personas con discapacidades: Otra propuesta importante es mejorar la accesibilidad para personas con discapacidades visuales y auditivas. La IA podría ser utilizada para crear descripciones auditivas detalladas de las obras, lo que permitiría que personas con discapacidad visual disfruten de la galería. Además, se continuó con el desarrollo del artículo a partir de donde lo dejamos, enfocándonos en las propuestas de accesibilidad y democratización del acceso al arte mediante IA.

DEMOCRATIZACIÓN DEL ACCESO AL ARTE A TRAVÉS DE IA (CONTINÚA)

El potencial de la IA para democratizar el acceso al arte es uno de sus aspectos más prometedores. Actualmente, muchos museos y galerías enfrentan el reto de hacer accesibles sus exposiciones a personas con discapacidades visuales, auditivas o motoras. Sin embargo, la IA ofrece una serie de soluciones innovadoras para superar estas barreras y hacer que las exposiciones sean inclusivas para todos los visitantes.

Accesibilidad para personas con discapacidades: Una de las aplicaciones más directas de la IA en este campo sería el uso de descripciones auditivas automáticas para personas con discapacidades visuales. A través de algoritmos de IA, las obras de arte podrían describirse en tiempo real, proporcionando a los usuarios una experiencia rica en detalles que les permita apreciar las características visuales de las piezas, incluso si no pueden verlas directamente. Esta tecnología ya se está utilizando en algunos museos a nivel mundial, como el Museo del Prado, que ha implementado sistemas de audio-descripción que permiten a las personas con discapacidades visuales disfrutar de las obras de arte a través de descripciones detalladas y narrativas auditivas (API4AI, 2024).

Subtítulos y descripciones en lenguaje de señas: La IA también podría facilitar el acceso a las exposiciones mediante la generación automática de subtítulos para videos y recorridos virtuales. Los visitantes con discapacidades auditivas podrían beneficiarse de subtítulos precisos que les permitan seguir las explicaciones y comentarios relacionados con las obras expuestas. Además, la IA podría ayudar a generar traducciones en lenguaje de señas para hacer las exposiciones accesibles a una audiencia aún más amplia (Pigro, 2024).

Interactividad mediante realidad aumentada (AR): La IA también podría combinarse con tecnologías de realidad aumentada (AR) para ofrecer a los visitantes una experiencia más interactiva y envolvente. Por ejemplo, los usuarios podrían utilizar dispositivos móviles para explorar las obras en 3D, moviéndose alrededor de ellas o acercándose a detalles específicos que no serían visibles en una simple imagen. Además, la AR podría utilizarse para recrear el contexto original de las obras, permitiendo a los usuarios experimentar las exposiciones en sus ambientes históricos y culturales originales, enriqueciendo así la experiencia del usuario.

Acceso internacional y traducción automática: La IA también tiene el potencial de romper barreras lingüísticas mediante la traducción automática de descripciones de obras y guías. A través de algoritmos de procesamiento del lenguaje natural, los textos de la galería podrían traducirse en tiempo real a una variedad de idiomas, lo que permitiría que usuarios de todo el mundo disfruten de las exposiciones, sin

importar su idioma nativo (AIUpbeat, 2024). De este modo, la Galería del Tiempo Virtual no solo ampliaría su alcance a nivel local, sino también global, ofreciendo una experiencia cultural verdaderamente inclusiva.

PROPUESTAS PARA LA PRESERVACIÓN Y GESTIÓN DEL ACERVO CON IA

Otro campo en el que la IA puede hacer una contribución significativa es en la preservación y gestión del acervo. Las obras de arte, especialmente las que forman parte de colecciones históricas, requieren una conservación cuidadosa y constante para garantizar que se mantengan en buen estado durante generaciones. La Galería del Tiempo Virtual podría beneficiarse enormemente de las herramientas de IA para mejorar tanto la preservación de sus obras como la gestión de su acervo digital.

Conservación preventiva con IA: Una de las aplicaciones más importantes de la IA en la conservación es la detección temprana de deterioro en las obras de arte. Utilizando algoritmos de reconocimiento de imágenes, la IA podría monitorear las obras digitalizadas y alertar a los conservadores sobre cualquier signo de daño o deterioro, como cambios en el color, grietas o desgaste. Estas tecnologías ya están siendo utilizadas en instituciones como el *Museo del Prado* y el *Museo Británico*, donde los sistemas de IA ayudan a identificar problemas antes de que se vuelvan irreversibles (DataArt, 2023). En la Galería del Tiempo Virtual, la implementación de estas tecnologías permitiría una conservación más eficiente de las obras digitalizadas, asegurando que permanezcan en óptimas condiciones para las futuras generaciones.

Catalogación automatizada: La IA también puede facilitar la catalogación automatizada de grandes colecciones de obras de arte. Tradicionalmente, la catalogación de las obras ha sido un proceso manual que requiere tiempo y esfuerzo por parte de los curadores. Sin embargo, con la ayuda de la IA, las obras pueden ser clasificadas y etiquetadas automáticamente en función de sus características visuales, como el estilo, el periodo histórico o los materiales utilizados. Esto no solo

ahorraría tiempo, sino que también haría que el acervo de la Galería del Tiempo Virtual fuera más accesible y navegable para los investigadores y el público general (API4AI, 2024).

Creación de archivos digitales avanzados: Otra propuesta para la preservación del acervo es la creación de archivos digitales de alta calidad mediante técnicas de IA y aprendizaje profundo. La IA podría ayudar a generar réplicas digitales de las obras con un nivel de detalle sin precedentes, lo que permitiría que las obras fueran preservadas digitalmente en su estado original. Esto no solo serviría como una copia de seguridad en caso de que las obras físicas sufran daños, sino que también facilitaría la investigación académica, ya que los estudiantes y estudiosos podrían analizar las obras en detalle sin necesidad de interactuar físicamente con ellas.

FUTURO DE LA GALERÍA DEL TIEMPO VIRTUAL: REALIDAD MIXTA Y EXPERIENCIAS INMERSIVAS

Mirando hacia el futuro, la Galería del Tiempo Virtual tiene el potencial de convertirse en un espacio aún más dinámico e interactivo mediante la integración de tecnologías como la realidad mixta y las experiencias inmersivas. La combinación de IA con tecnologías como la realidad aumentada (AR) y la realidad virtual (VR) permitiría a los usuarios experimentar las exposiciones de manera más envolvente y participativa.

Realidad mixta: La realidad mixta combina elementos del mundo físico y el mundo virtual para crear una experiencia única. En la Galería del Tiempo Virtual, los usuarios podrían experimentar las exposiciones de manera inmersiva, "*caminando*" por un entorno digital que recrea la galería física, pero con la ventaja de que las obras podrían mostrarse de maneras que no serían posibles en un espacio físico. Por ejemplo, los visitantes podrían interactuar con versiones 3D de las obras de arte, acercarse a detalles invisibles a simple vista o explorar el contexto histórico de las piezas a través de narrativas interactivas generadas por IA (DataArt, 2023).

Recorridos guiados por avatares históricos: Otra propuesta para el futuro de la galería es la creación de recorridos narrativos guiados por avatares de personajes históricos. Utilizando IA, la galería podría crear avatares interactivos de artistas o figuras históricas relacionadas con las obras expuestas, que podrían guiar a los visitantes a lo largo de la exposición, proporcionando información y contexto adicional. Esta tecnología no solo haría la experiencia más educativa, sino también más atractiva y entretenida, especialmente para las generaciones más jóvenes.

Colaboración con otras instituciones: Finalmente, la Galería del Tiempo Virtual también podría colaborar con otras instituciones culturales y universidades que ya están utilizando IA para crear exposiciones colaborativas que trasciendan las fronteras físicas. Al compartir colecciones digitales y desarrollar exposiciones conjuntas, la galería podría expandir su acervo y ofrecer a los visitantes una experiencia cultural más rica y diversa (AIUpbeat, 2024).

CONSIDERACIONES ÉTICAS Y RETOS EN LA IMPLEMENTACIÓN DE IA

A pesar de los muchos beneficios que la IA puede aportar a la Galería del Tiempo Virtual, también es importante considerar los desafíos éticos que plantea su implementación. Uno de los principales retos es la privacidad de los datos. Al personalizar las experiencias de los visitantes, los algoritmos de IA deben recopilar y analizar datos sobre sus comportamientos y preferencias. Es crucial que la galería garantice que estos datos se manejen de manera ética y segura, protegiendo la privacidad de los usuarios y asegurándose de que no se utilicen con fines indebidos (DataArt, 2023).

Otro desafío es el impacto en la autenticidad del arte. A medida que las obras se digitalizan y se vuelven interactivas, existe el riesgo de que se desvirtúe su valor cultural o su significado original. Es importante que las instituciones culturales, incluida la Galería del Tiempo Virtual, encuentren un equilibrio entre la innovación tecnológica y la preservación de la autenticidad artística. Las obras deben ser presentadas de

manera que respeten su contexto histórico y cultural, asegurándose de que la experiencia digital no comprometa su integridad.

CONCLUSIÓN

La Galería del Tiempo Virtual de la UAM tiene un futuro prometedor, especialmente si se adoptan las tecnologías de inteligencia artificial que hemos discutido a lo largo de este artículo. Desde la curaduría automatizada hasta la creación de experiencias inmersivas, En resumen, la Galería del Tiempo Virtual de la UAM tiene un futuro lleno de posibilidades con la integración de tecnologías como la inteligencia artificial (IA). A través de la curaduría automatizada, los asistentes virtuales interactivos, la realidad aumentada y las experiencias inmersivas, la galería puede transformar la manera en que los visitantes interactúan con las obras de arte, al mismo tiempo que democratiza el acceso y promueve a jóvenes creadores. La IA no solo permite una personalización más profunda de las visitas, sino que también facilita la conservación y preservación del acervo, asegurando que las obras de arte se mantengan accesibles y en buen estado para futuras generaciones.

Sin embargo, es crucial que la implementación de estas tecnologías se realice con responsabilidad, tomando en cuenta los retos éticos relacionados con la privacidad de los datos y la autenticidad del arte. Con una estrategia bien planificada, la Galería del Tiempo Virtual puede posicionarse como un líder en la intersección entre arte y tecnología, ofreciendo experiencias innovadoras y accesibles que continúen inspirando y educando a sus visitantes.

REFERENCIAS

AIUpbeat. (2024). *Artificial Intelligence in Museums: Enhancing Engagement and Immersion*. https://aiupbeat.com/artificial-intelligence-museums-engagement-immersion

API4AI. (2024). *How AI-Powered Image Recognition is Transforming Museum and Gallery Experiences.* https://api4.ai/blog/museums-eye-a-rash-of-new-opportunities-with-ai-technology

DataArt. (2023). *Museums Eye a Rash of New Opportunities with AI Technology*. https://www.dataart.com/blog/museums-eye-a-rash-of-new-opportunities-with-ai-technology

Pigro. (2024). *Digital Museums: 6 ways to use AI in museums.* https://blog.pigro.ai/digital-museums-ai

CAPÍTULO XI

LA INTELIGENCIA ARTIFICIAL EN LA MIPYME. DESARROLLO TECNOLÓGICO Y SUSTENTABILIDAD

OSCAR LOZANO CARRILLO[1]
JANETTE BRITO LAREDO[2]
FRANCISCO JAVIER ARIAS VARGAS[3]

INTRODUCCIÓN

En las últimas décadas, el avance tecnológico ha transformado radicalmente la manera en que las empresas operan y compiten en el mercado. Desde la adopción de tecnologías básicas como la computación y el internet, hasta la implementación de soluciones avanzadas como la inteligencia artificial (IA), las empresas han experimentado una evolución significativa en sus procesos y estrategias. En México, las microempresas, que representan una parte crucial de la economía, no han sido ajenas a estos cambios. Muchas de ellas ya la están utilizando como herramientas estratégicas, pero no conscientes en el día a día, tal es el caso de *WHATSAPP, FACEBOOK, INSTAGRAM, TIK TOK* y por supuesto los diferentes formatos de *GPT (Generative Pre-trained Transformer)* donde seguramente millones de microempresarios compran, venden, gestionan, controlan, consultan y promueven. Este capítulo explora los alcances del uso de la IA en las microempresas

[1] Profesor-Investigador, del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco, Ciudad de México, México. Dirige el Laboratorio de Inteligencia Artificial UAM Azcapotzalco.

[2] Investigadora de la Universidad Autónoma de Baja California, México.

[3] Vicerrector de Investigación, Vinculación e Internacionalización, UNEV – Honduras, San Pedro Sula – Honduras.

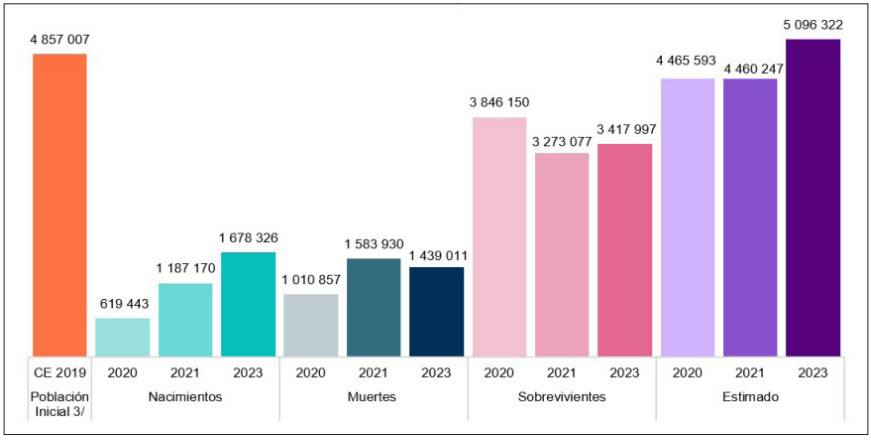
mexicanas, destacando su evolución tecnológica, el impacto potencial en su desarrollo y sustentabilidad.

Palabras clave: Mipymes, IA, desarrollo, sustentabilidad.

CONTEXTO DE LAS MICROEMPRESAS EN MÉXICO

Las microempresas en México se definen como aquellas unidades económicas que cuentan con un máximo de 10 empleados y generan ingresos anuales limitados. Estas empresas son fundamentales para la economía mexicana, representando aproximadamente el 95 % del total de las empresas y contribuyendo significativamente al empleo y al PIB del país (ver **Figura 1**) (INEGI, 2021). Las microempresas no solo son vitales para la economía, sino que también desempeñan un papel crucial en el tejido social de México ya que en el día a día producen, reproducen acciones económicas, pero también sociales, culturales, ambientales y psicológicas.

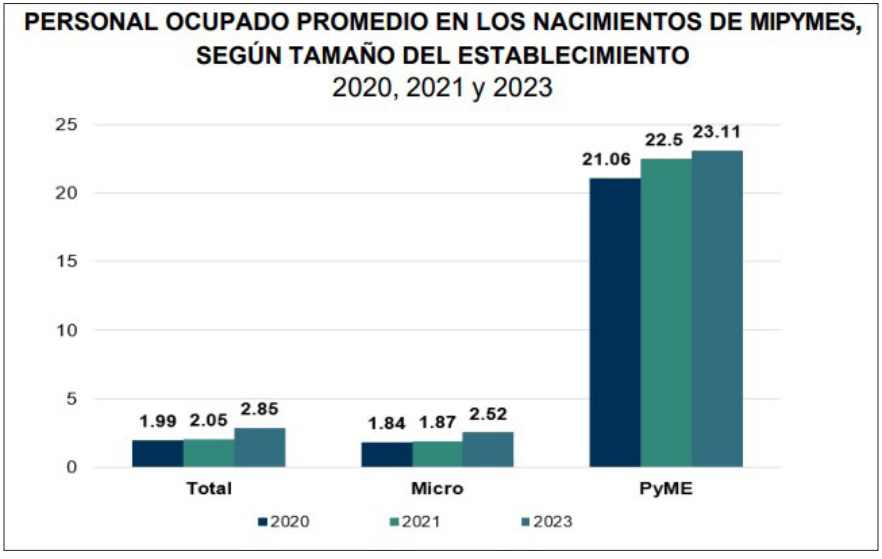
Figura 1. *Representación de las microempresas en México*



Fuente: INEGI (2021).

En México proveen el 70 % del empleo (Wradio, 2023) que significa millones de personas y son una fuente importante de innovación y desarrollo local. Además, estas empresas suelen estar más integradas en sus comunidades, lo que les permite responder de manera más ágil a las necesidades y demandas locales y sobre todo proveer de los indicadores de sustentabilidad más importantes en el país, ya que este empleo es para familiares y conocidos, que de no trabajar en estas unidades económicas seguramente no tendrían empleo (ver **Figura 2**).

Figura 2. Personal ocupado promedio en los nacimientos de Mipymes, según tamaño del establecimiento 2020, 2021 y 2023

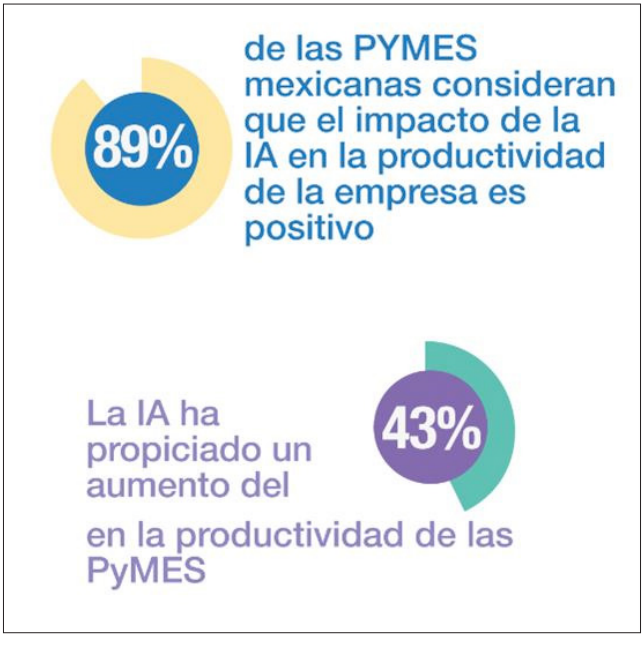


Fuente: INEGI (2024).

A pesar de su importancia, las microempresas enfrentan numerosos desafíos que limitan su crecimiento y sostenibilidad. Entre estos desafíos se encuentran la falta de acceso al financiamiento, la limitada profesionalización, la falta de políticas públicas de alto e impacto, la baja adopción de tecnologías avanzadas, y la limitada capacidad para competir en mercados globalizado, aunque la apreciación de las due-

ñas y dueños es positiva y creen que la implementación de tecnologías como la IA puede ofrecer soluciones a algunos de estos problemas (ver **Figura3**) (News Microsoft, 2024), mejorando la eficiencia operativa y la toma de decisiones, sin embargo la falta de recursos y sobre todo de programas de formación y capacitación masivos para ellos, les limitan a la autogestión o a no usarlas, sobre todo en la gente de mayor edad o con rezagos culturales más pronunciados.

Figura 3. Apreciación de los dueños y dueñas sobre la implementación de tecnologías.



Fuente: News Microsoft (2024).

EVOLUCIÓN TECNOLÓGICA Y USO DE LA IA

Las primeras tecnologías adoptadas por las microempresas mexicanas incluyeron herramientas básicas de computación y software de ges-

tión empresarial. Estas tecnologías permitieron una mayor organización y control de los procesos internos, así como una mejor comunicación con clientes y proveedores.

Con el tiempo, la digitalización se convirtió en una prioridad. La adopción de internet y plataformas digitales permitió a las microempresas expandir su alcance y mejorar su presencia en el mercado. El comercio electrónico, por ejemplo, abrió nuevas oportunidades para llegar a clientes más allá de las fronteras locales (Forbes México, 2019).

La inteligencia artificial representa el siguiente gran salto en la evolución tecnológica de las microempresas. La IA ofrece herramientas avanzadas para la automatización de procesos, el análisis de datos y la personalización de la experiencia del cliente. Estas capacidades pueden transformar la manera en que las microempresas operan, permitiéndoles competir de manera más efectiva en un entorno cada vez más digital y globalizado (Forbes México, 2019; McKinsey, 2024).

La implementación de la IA puede tener un impacto significativo en las microempresas mexicanas. Entre los beneficios potenciales se encuentran el aumento de la eficiencia operativa, la reducción de costos, y la mejora en la calidad del servicio al cliente. Además, la IA puede ayudar a las microempresas a tomar decisiones más informadas y estratégicas, basadas en el análisis de grandes volúmenes de datos (Expansión, 2023).

EFICIENCIA OPERATIVA CON LA IA

La inteligencia artificial ofrece múltiples beneficios para mejorar la eficiencia operativa de las microempresas. Desde la automatización de tareas hasta el análisis avanzado de datos y la personalización de la experiencia del cliente, la IA está transformando la manera en que las microempresas operan y compiten en el mercado. La eficiencia operativa en microempresas que utilizan inteligencia artificial (IA) se logra a través de varias aplicaciones y beneficios específicos.

Automatización de Tareas

La IA permite la automatización de tareas repetitivas y administrativas, lo que libera tiempo y recursos humanos para actividades de mayor valor. Por ejemplo, herramientas de IA pueden encargarse de la entrada de datos, la gestión de inventarios y la atención al cliente mediante *chatbots.* Esto no solo reduce errores humanos, sino que también acelera los procesos operativos (Zenvia, 2023).

Análisis de Datos

Otra forma en que la IA mejora la eficiencia operativa es a través del análisis avanzado de datos. Las microempresas pueden utilizar algoritmos de IA para analizar grandes volúmenes de datos y obtener *insights* valiosos que informen la toma de decisiones estratégicas. Esto incluye desde el análisis de tendencias de mercado hasta la predicción de la demanda de productos (Endeavor, 2023).

Personalización de la Experiencia del Cliente

La IA también permite una personalización más efectiva de la experiencia del cliente. Mediante el uso de algoritmos de aprendizaje automático, las microempresas pueden ofrecer recomendaciones personalizadas y mejorar la satisfacción del cliente. Esto no solo aumenta la lealtad del cliente, sino que también puede conducir a un incremento en las ventas (Microsoft, 2024).

Reducción de Costos

La implementación de IA puede resultar en una significativa reducción de costos operativos. Al automatizar tareas y optimizar procesos, las microempresas pueden reducir la necesidad de mano de obra adicional y minimizar los costos asociados con errores y retrabajos. Además, la IA puede ayudar a identificar áreas de ineficiencia y proponer soluciones para mejorar la productividad (iWorld, 2023).

En México, la adopción de IA está en crecimiento, con empresas invirtiendo en tecnologías como la automatización de procesos robó-

ticos (RPA) y el análisis predictivo. Estas tecnologías están ayudando a las microempresas a mejorar su eficiencia operativa y a competir en un mercado cada vez más digitalizado (Endeavor, 2023).

En España, las microempresas están utilizando IA para optimizar la gestión de inventarios y mejorar la atención al cliente mediante chatbots avanzados. Estas aplicaciones no solo mejoran la eficiencia operativa, sino que también permiten a las empresas ofrecer un servicio más rápido y personalizado (Overstand, 2023).

USO DE GPT PARA LA MIPYME

Algo que está revolucionando el mundo de manera determinante sin duda es la tecnología *GPT* y si bien su capacidad es amplia y está más que probado que beneficia en múltiples aspectos, entre ellos en el apoyo a las MIPYMES, sin lugar a dudas el principal beneficio provee de la accesibilidad de su uso, ya que prácticamente cualquier persona puede acceder a través de su teléfono móvil y basta que tenga la aplicación de *WHATSAPP* para tener una de las IA más poderosas del mundo y de manera gratuita. En si la tecnología cambia al mundo porque ha democratizado el uso de la información de calidad y al mismo tiempo, con esta posibilidad de acceso oportuno e ilimitado está también orientando la cultura de la sociedad digital o lo que Lasalle (2023) llama la *"Civilización Artificial"*.

Descripción de un GPT *(*Generative Pre-Trained Transformer*)*

Un *GPT*, o *Generative Pre-trained Transformer,* es un modelo de lenguaje avanzado desarrollado en el campo de la inteligencia artificial (IA). Este tipo de modelo utiliza una arquitectura de transformadores y ha sido preentrenado con grandes cantidades de datos textuales para generar texto de manera autónoma y coherente. Los modelos *GPT*, como los desarrollados por *OpenAI,* son capaces de realizar una amplia gama de tareas relacionadas con el procesamiento del lenguaje natural (NLP), incluyendo la generación de texto, la traducción de

idiomas, la respuesta a preguntas, el resumen de documentos y la creación de contenido creativo.

La arquitectura de transformadores, introducida por (Vaswani y otros, 2017), es fundamental para el funcionamiento de los *GPT*. Esta arquitectura permite a los modelos procesar y generar texto de manera más eficiente y precisa que los modelos anteriores, como los basados en redes neuronales recurrentes (RNN) o en redes neuronales convolucionales (CNN) (Vaswani y otros, 2017). Los transformadores utilizan mecanismos de atención que permiten al modelo enfocarse en diferentes partes del texto de entrada al generar cada palabra, lo que mejora la coherencia y la relevancia del texto generado.

El preentrenamiento de los modelos *GPT* implica entrenar el modelo en una gran cantidad de datos textuales para que aprenda las estructuras y patrones del lenguaje. Este proceso de preentrenamiento es seguido por una fase de ajuste fino, donde el modelo se entrena en tareas específicas utilizando conjuntos de datos más pequeños y especializados. Este enfoque de preentrenamiento y ajuste fino permite a los modelos *GPT* (ver **Figura 4**) generalizar bien en una variedad de tareas de NLP y adaptarse rápidamente a nuevas tareas con una cantidad mínima de datos adicionales (Radford y otros, 2019).

Figura 4. *Modelos GPT.*



Fuente: Folou (2024).

Una de las aplicaciones más conocidas de los modelos *GPT* es *ChatGPT,* un *chatbot* avanzado que puede mantener conversaciones coherentes y naturales con los usuarios. *ChatGPT* utiliza la capacidad de generación de texto de los modelos *GPT* para responder a preguntas, proporcionar información, y mantener diálogos en una amplia gama de temas. Esta capacidad ha llevado a la adopción de *ChatGPT* en diversas industrias, incluyendo el servicio al cliente, la educación y el entretenimiento (OpenAI, 2023).

Los modelos *GPT* también se utilizan en la creación de contenido, donde pueden generar artículos, ensayos, y otros tipos de texto de manera autónoma. Esta capacidad es especialmente útil para los creadores de contenido y los profesionales del marketing, que pueden utilizar los modelos *GPT* para generar borradores de alta calidad y ahorrar tiempo en el proceso de escritura. Además, los modelos *GPT* pueden personalizar el contenido generado para adaptarse a diferentes audiencias y contextos, lo que aumenta su utilidad en aplicaciones comerciales (Brown y otros, 2020). En el ámbito de la traducción de idiomas, los modelos *GPT* han demostrado ser altamente efectivos. Pueden traducir texto entre múltiples idiomas con una precisión comparable a la de los traductores humanos en muchos casos. Esta capacidad es particularmente valiosa en un mundo globalizado, donde la comunicación multilingüe es esencial para los negocios y la colaboración internacional (Wu y otros, 2016).

Otra aplicación importante de los modelos *GPT* es el resumen de documentos. Los modelos *GPT* pueden analizar y resumir grandes volúmenes de texto, extrayendo la información más relevante y presentándola de manera concisa. Esta capacidad es útil en campos como la investigación académica, donde los investigadores necesitan revisar grandes cantidades de literatura, y en el periodismo, donde los periodistas deben sintetizar información de múltiples fuentes (Liu y otros, 2019). Los modelos *GPT* también se utilizan en la programación y el desarrollo de software. Pueden generar y depurar código, lo que ayuda a los desarrolladores a escribir programas más rápidamente y con menos errores. Esta capacidad es especialmente útil en el desarrollo de software complejo, donde los errores pueden ser costosos y difíciles de detectar (Chen y otros, 2021).

En términos de impacto económico y social, los modelos *GPT* tienen el potencial de transformar muchas industrias y mejorar la eficiencia y la productividad en una amplia gama de aplicaciones. Sin embargo, también plantean desafíos éticos y de seguridad. Por ejemplo, la capacidad de los modelos *GPT* para generar texto de manera autónoma puede ser utilizada para crear desinformación o contenido malicioso. Además, el uso de modelos *GPT* en la automatización de tareas puede tener implicaciones para el empleo, ya que algunas tareas que antes realizaban los humanos pueden ser automatizadas (Bender y otros, 2021). Para abordar estos desafíos, es importante desarrollar y aplicar principios éticos y regulaciones que guíen el uso de los modelos *GPT*. Esto incluye garantizar la transparencia en el desarrollo y la implementación de estos modelos, así como establecer mecanismos para detectar y mitigar el uso malicioso de la tecnología. Además, es crucial fomentar la colaboración entre investigadores, desarrolladores, y legisladores para abordar los desafíos y maximizar los beneficios de los modelos *GPT* (Floridi y otros, 2018).

Los modelos *GPT* representan un avance significativo en el campo de la inteligencia artificial y tienen el potencial de transformar una amplia gama de industrias y aplicaciones. Su capacidad para generar texto de manera autónoma y coherente, junto con su flexibilidad y adaptabilidad, los convierte en una herramienta poderosa para el procesamiento del lenguaje natural. Sin embargo, es importante abordar los desafíos éticos y de seguridad asociados con su uso para garantizar que los beneficios de esta tecnología se realicen de manera responsable y equitativa. Una inteligencia artificial basada en *GPT,* puede ofrecer múltiples beneficios a las microempresas, ayudándolas en su desarrollo y evolución. A continuación, se detallan algunas de las formas en que *GPT* puede ser útil.

Automatización de Tareas

GPT puede automatizar tareas repetitivas y administrativas, lo que permite a los empleados centrarse en actividades de mayor valor añadido. Por ejemplo, puede gestionar la entrada de datos, programar citas y responder a consultas frecuentes de clientes mediante *chatbots.* Esto no solo ahorra tiempo, sino que también reduce errores humanos (Zenvia, 2023).

Mejora en la Toma de Decisiones

La capacidad de *GPT* para analizar grandes volúmenes de datos y proporcionar *insights* valiosos puede mejorar significativamente la toma de decisiones. Las microempresas pueden utilizar estos análisis para identificar tendencias de mercado, optimizar inventarios y personalizar ofertas para sus clientes (Endeavor, 2023).

Personalización de la Experiencia del Cliente

GPT puede ayudar a personalizar la experiencia del cliente mediante el análisis de datos de comportamiento y preferencias. Esto permite a las microempresas ofrecer recomendaciones personalizadas y mejorar la satisfacción del cliente, lo que puede traducirse en una mayor lealtad y ventas (Microsoft, 2024).

Redacción y Comunicación

Copilot puede asistir en la redacción de correos electrónicos, documentos y contenido para redes sociales, asegurando que la comunicación sea clara y profesional. Esto es especialmente útil para microempresas que no cuentan con un equipo dedicado de marketing o comunicación (ADSLZone, 2024).

Capacitación y Desarrollo de Habilidades

GPT puede crear programas de capacitación personalizados para los empleados, ayudándoles a desarrollar nuevas habilidades y adaptarse a cambios en el mercado. Esto es crucial para mantener la competitividad y fomentar el crecimiento interno (Overstand, 2023).

Reducción de Costos

Al automatizar tareas y optimizar procesos, *Copilot* puede ayudar a reducir los costos operativos. Esto permite a las microempresas reinvertir esos ahorros en áreas clave para su crecimiento, como la innovación y la expansión de mercado (iWorld, 2023).

SUSTENTABILIDAD IA Y MIPYME

La IA tiene el potencial de transformar las MiPyMEs con una visión sustentable, mejorando la eficiencia operativa, reduciendo costos y minimizando el impacto ambiental. Al adoptar tecnologías de IA, las MiPyMEs pueden no solo mejorar su competitividad, sino también contribuir a un futuro más sostenible y equitativo (ver **Figura 5**).

Figura 5. Visión sustentable de las Mipymes.



Fuente: Sostenibilidad para todos (2024).

La IA puede optimizar el uso de recursos en las MiPyMEs, lo que es crucial para una operación sustentable. Por ejemplo, mediante el análisis de datos en tiempo real, la IA puede identificar patrones de consumo de energía y sugerir medidas para reducir el desperdicio. Esto no solo disminuye los costos operativos, sino que también reduce la huella de carbono de la empresa (Forbes México, 2023). Además, la IA puede mejorar la gestión de inventarios, asegurando que los productos se mantengan en niveles óptimos y evitando el exceso de stock que podría llevar a desperdicios innecesarios (Diario Sustentable, 2024).

En el ámbito de la producción, la IA puede ayudar a las MiPyMEs a implementar prácticas más sostenibles. Por ejemplo, en la agricultura, la IA puede analizar múltiples variables como las condiciones del suelo, el clima y los riesgos de plagas para optimizar el uso de fertilizantes y pesticidas, reduciendo así el impacto ambiental (Forbes México, 2023). Este

enfoque no solo mejora la eficiencia de la producción, sino que también promueve prácticas agrícolas más amigables con el medio ambiente.

La IA también puede jugar un papel crucial en la gestión de residuos. Mediante el uso de algoritmos avanzados, las MiPyMEs pueden identificar oportunidades para reciclar y reutilizar materiales, minimizando así los desechos que terminan en los vertederos. Además, la IA puede ayudar a diseñar productos que sean más fáciles de reciclar, contribuyendo a una economía circular (America Retail, 2024).

Otro aspecto importante es la capacidad de la IA para mejorar la transparencia y la trazabilidad en las cadenas de suministro. Las MiPyMEs pueden utilizar la IA para rastrear el origen de sus materias primas y asegurarse de que se obtengan de fuentes sostenibles. Esto no solo mejora la sostenibilidad de la cadena de suministro, sino que también puede aumentar la confianza de los consumidores en los productos de la empresa (Microsoft, 2024).

La IA también puede facilitar la adopción de prácticas de trabajo remoto y flexible, lo que puede reducir la necesidad de desplazamientos y, por ende, las emisiones de carbono asociadas. Además, las herramientas de IA pueden mejorar la colaboración y la comunicación entre equipos, aumentando la eficiencia y reduciendo la necesidad de viajes de negocios (Diario Sustentable, 2024).

En términos de impacto social, la IA puede ayudar a las MiPyMEs a desarrollar productos y servicios que aborden problemas sociales y ambientales. Por ejemplo, mediante el análisis de datos demográficos y de comportamiento, las empresas pueden identificar necesidades no satisfechas en comunidades vulnerables y desarrollar soluciones innovadoras para abordarlas. Esto no solo mejora la sostenibilidad social de la empresa, sino que también puede abrir nuevos mercados y oportunidades de negocio (Forbes México, 2023).

Además, la IA puede apoyar la educación y la capacitación de los empleados en prácticas sostenibles. Las MiPyMEs pueden utilizar plataformas de aprendizaje basadas en IA para ofrecer programas de formación personalizados que ayuden a los empleados a adquirir las habilidades necesarias para implementar y mantener prácticas sostenibles en sus operaciones diarias (America Retail, 2024).

RIESGOS DEL USO DE LA IA EN LAS MIPYMES

El uso de IA en las microempresas en México ofrece numerosas oportunidades, pero también presenta riesgos significativos que deben ser gestionados con cuidado. La falta de preparación y recursos, los sesgos algorítmicos, la seguridad cibernética, la dependencia de proveedores externos, la falta de transparencia y responsabilidad, el impacto en el empleo y las consideraciones éticas son algunos de los principales desafíos que enfrentan las microempresas al implementar IA. Es crucial que las microempresas adopten un enfoque proactivo para mitigar estos riesgos y maximizar los beneficios de la IA.

Falta de Preparación y Recursos

Uno de los riesgos más significativos es la disparidad en la preparación y los recursos disponibles para desarrollar y utilizar IA. Muchas microempresas en México carecen de la infraestructura tecnológica necesaria, como supercomputadoras y sistemas avanzados, así como de la capacitación adecuada para aprovechar al máximo las tecnologías de IA. Esta falta de recursos puede limitar la capacidad de las microempresas para competir en la economía global y protegerse de amenazas emergentes en el ciberespacio (Riquelme, 2024).

Sesgos Algorítmicos

La IA puede perpetuar y amplificar sesgos existentes si los datos utilizados para entrenar los algoritmos están sesgados. Esto es especialmente preocupante en aplicaciones como la selección de personal, donde los algoritmos pueden favorecer desproporcionadamente a ciertos grupos demográficos. Por ejemplo, *Google* dejó de utilizar IA para la selección de personal después de descubrir que sus algoritmos favorecían a los hombres (Riquelme, 2024). La falta de un marco normativo adecuado en México agrava este problema, ya que no existen regulaciones claras para garantizar la equidad en el uso de IA.

Seguridad Cibernética

La seguridad cibernética es otro riesgo crucial. La IA puede ser utilizada con fines maliciosos, como la creación de *deepfakes* para suplantar identidades y cometer fraudes. Además, la falta de una legislación robusta sobre IA en México deja a las microempresas vulnerables a ataques cibernéticos. La implementación de IA sin las medidas de seguridad adecuadas puede exponer a las empresas a riesgos significativos, incluyendo la pérdida de datos sensibles y daños a la reputación (Riquelme, 2024).

Dependencia de Proveedores Externos

Muchas microempresas dependen de herramientas de IA desarrolladas por terceros, lo que las expone a riesgos adicionales. La dependencia de proveedores externos puede resultar en una falta de control sobre la tecnología utilizada y en vulnerabilidades en la seguridad de los datos. Además, las fallas en las herramientas de IA de terceros pueden tener consecuencias graves, como pérdidas financieras y sanciones regulatorias (Renieris, 2023).

Falta de Transparencia y Responsabilidad

La complejidad de los sistemas de IA puede dificultar la comprensión de cómo se toman las decisiones y a quién responsabilizar en caso de errores. Esta falta de transparencia puede generar desconfianza entre los empleados y los clientes, y complicar la resolución de problemas cuando algo sale mal. Es esencial que las microempresas implementen mecanismos para garantizar la transparencia y la responsabilidad en el uso de IA (Delodigital, 2024).

Impacto en el Empleo

La automatización de tareas mediante IA puede llevar a la pérdida de empleos, especialmente en roles que implican tareas repetitivas y administrativas. Aunque la IA puede mejorar la eficiencia operativa, también puede desplazar a trabajadores que no tienen las habilidades

necesarias para adaptarse a nuevas funciones. Es importante que las microempresas consideren estrategias para la reubicación y capacitación de empleados afectados por la automatización (TalkAI, 2024).

Consideraciones Éticas

El uso de IA plantea importantes consideraciones éticas, como la privacidad de los datos y el consentimiento informado. Las microempresas deben asegurarse de que el uso de IA respete los derechos de los individuos y cumpla con las normativas de protección de datos. La falta de atención a estas consideraciones puede resultar en sanciones legales y daños a la reputación de la empresa (Riquelme, 2024).

Un estudio publicado en Vinueza (2020) revela que la Inteligencia Artificial podría facilitar el cumplimiento del 79 % de los Objetivos de Desarrollo Sostenible (ODS) y hasta el 93 % si consideramos solo los ODS medioambientales. Datos que reflejan el inmenso potencial de la IA para ayudar en la lucha contra el cambio climático.

CONCLUSIONES

El uso de la IA va a permear todos los ámbitos de la sociedad, economía, política, cultura, psicología y ecología mundial y ya es inevitable, aunque hay voces que manifiestas sistemáticamente su preocupación, la carrera tecnológica que han emprendido las potencias China y Estados Unidos principalmente, nos han metido en una nueva guerra fría de consecuencias inimaginables, sin embargo mientras tanto, las MICRO, PEQUEÑAS Y MEDIANAS empresas aprovechan las bondades que la IA ofrece, principalmente en sus aplicaciones más básicas como pueden ser las redes sociales que aprovechan desde la pandemia por el virus *SARS CoV 2* como redes comerciales y muy recientemente el uso de la tecnología *GPT* para organizar, controlar, dirigir y gestionar sus empresas, pero también para difundir, diseñar, consultar, calcular e incluso aconsejar.

El conocimiento se ha democratizado con el uso de la IA y con el potencial de mejora de la MIPYME crece, sobre todo por el bajo costo

que este representa, no obstante, al ser generalizado para toda la gente, esta ventaja puede ser limitada hasta cierto punto, lo que es un hecho es que en el ámbito de la MIPYME se encontrarán efectos muy positivos si se logran implementar políticas públicas de apoyo en desarrollo y capacitación tecnológica que permita ampliar mercados y capacidades de las micro organizaciones principalmente.

Es preciso no obstante señalar que al momento el uso de la IA ha carecido de un marco normativo por la falta de legislación al respecto, pero inclusive en la inexistencia de criterios básicos para su uso, por ejemplo en las mismas empresas o en organizaciones como las escuelas y el gobierno, urge una discusión sobre la ética (Lasalle, 2023) que blinde al ser humano en el uso de la IA, por ejemplo en ámbitos de empleo, ya que de no hacerse, los riesgos de maximizar el valor del plusvalor relativo, dejara sin fuente de trabajo a muchas personas, entre otros riesgos.

BIBLIOGRAFÍA

ADSLZone. (2024). 5 formas en las que ChatGPT, Copilot o Gemini te pueden ayudar en el trabajo. Recuperado de ADSLZone

Amodei, D. (2020). Language Models are Few-Shot Learners. arXiv preprint arXiv:2005.14165.

America Retail. (2024). Estrategias sustentables para reducir el impacto ambiental de la IA. Recuperado de America Retail

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623)

Diario Sustentable. (2024). IA empresarial: Datos y transparencia para la sostenibilidad. Recuperado de Diario Sustentable

Endeavor. (2023). La era de la IA en México. Recuperado de Endeavor

Expansión. (2023). México avanza en la adopción de la IA y así se está usando en las empresas. Recuperado de Expansión

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2018). How to Design AI for Social Good: Seven Essential Factors. Science and Engineering Ethics, 24(5), 17771803.

Folou (2024) *IA de las grandes empresas tecnológicas.* Todas las Big Tech tienen IA: te contamos cuáles son - FOLOU

Forbes México. (2019). La tecnología, una pieza clave para el crecimiento económico de México. Recuperado de Forbes México

Forbes México. (2023). Inteligencia artificial: una nueva era de oportunidades para la sostenibilidad. Recuperado de Forbes México

INEGI (2024) ESTADÍSTICAS A PROPÓSITO DEL DÍA DE LAS MICRO, PEQUEÑAS Y MEDIANAS EMPRESAS (MIPYMES) en *Comunicado de prensa núm. 383/24*, 24 de junio de 2024

iWorld. (2023). Eficiencia operativa, una de las principales ventajas que ofrece la IA en Latinoamérica. Recuperado de iWorld

Lasalle, José María (2023) *Civilización Artificial,* Arpa, Barcelona

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.

Microsoft. (2024). La IA y las PyMEs: Un análisis de su adopción e impacto. Recuperado de Microsoft

MicrosoftNews (2024) "La IA se posiciona como una prioridad para las MIPYMES *en Microsoft.com,* México. La IA se posiciona como una prioridad para las PyMEs mexicanas - News Center Latinoamérica (microsoft.com)

McKinsey. (2024). El estado de la IA a principios de 2024: la adopción de la IA generativa aumenta y comienza a generar valor. Recuperado de McKinsey

OpenAI. (2023). ChatGPT: Optimizing Language Models for Dialogue. Recuperado de OpenAI

Overstand. (2023). Cómo la Inteligencia Artificial puede revolucionar la eficiencia operativa. Recuperado de Overstand

Proliferate. MIT Sloan Management Review y Boston Consulting Group. Recuperado de BCG

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019).

Renieris, E. M. (2023). Building Robust RAI Programs as Third-Party AI Tools

Riquelme, R. (2024). Estos son los riesgos de la Inteligencia Artificial en México. El Economista. Recuperado de El Economista

Sostenibilidad para todos (2024) *La alianza entre IA y el desarrollo sostenible.* La alianza entre Inteligencia Artificial y desarrollo sostenible (sostenibilidad.com)

TalkAI. (2024). El peligro de la IA: 9 grandes riesgos de la inteligencia artificial. Recuperado de TalkAI

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. In Advances in Neural Information Processing Systems (pp. 5998-6008).

Vinuesa, R., Azizpour, H., Leite, I. *et al.* The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat Commun* **11**, 233 (2020).

WRADIO (2024) "Mipymes generan el 70 % de los empleos de México: Secretaría de Economía" *en wradio.com,* julio 2023, México. Mipymes generan el 70 % de los empleos de México: Secretaría de Economía (wradio.com.mx)

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. arXiv preprint arXiv:1609.08144.

Zaremba, W. (2021). Evaluating Large Language Models Trained on Code. arXiv preprint arXiv:2107.03374.Language Models are Unsupervised Multitask Learners. OpenAI Blog, 1(8), 9.

Zenvia. (2023). Inteligencia Artificial Para Pequeñas y Medianas Empresas. Recuperado de Zenvia

CAPÍTULO XII

LA INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA TRANSHUMANISTA A LA LUZ DE LOS DERECHOS HUMANOS

KARINA DEL REFUGIO VALLEJO QUINTERO[1]
JESÚS MANUEL LÓPEZ MARROQUÍN[2]
RICARDO PÉREZ BOJÓRQUEZ[3]

RESUMEN

Este trabajo explora la relación entre el Transhumanismo, la Inteligencia Artificial y su impacto en los derechos humanos. El Transhumanismo, un movimiento que promueve la mejora de la condición humana a través de la tecnología, encuentra en la Inteligencia Artificial una herramienta clave para superar las limitaciones físicas y cognitivas. Este análisis abarca desde los antecedentes históricos, que vinculan las aspiraciones humanas de trascender sus limitaciones, con raíces filosóficas de varios siglos atrás, y los avances científicos de la era moderna, hasta el desarrollo contemporáneo de la Inteligencia Artificial.

Estos avances permiten mejorar las capacidades humanas mediante tecnologías como las interfaces cerebro-computadora y la medicina personalizada. Sin embargo, el uso de estas tecnologías plantea importantes retos éticos y jurídicos, especialmente en cuanto a la equidad de acceso, la protección de la privacidad y la prevención de discriminación.

[1] Profesora de Tiempo Completo, Universidad Autónoma de Occidente, México.
[2] Profesor de Profesor de Tiempo completo, Universidad Autónoma de Occidente, México.
[3] Profesor de Asignatura Base, Universidad Autónoma de Occidente, México.

Se destaca la necesidad de desarrollar marcos regulatorios adecuados para la gobernanza de la Inteligencia Artificial, garantizando que su aplicación respete los derechos fundamentales, tales como la dignidad, la igualdad y la libertad. A medida que la Inteligencia Artificial se integra más en la vida humana, se prevé la emergencia de nuevos derechos, como los neuroderechos, que deberán proteger la integridad mental y la autonomía en un futuro post humano.

Palabras clave: transhumanismo, inteligencia artificial, derechos humanos, derechos emergentes, neuroderechos.

INTRODUCCIÓN

El avance acelerado de las tecnologías emergentes ha transformado el panorama de la sociedad moderna, abriendo puertas a innovaciones que desafían los límites biológicos y cognitivos de la humanidad. El transhumanismo es un movimiento filosófico, cultural y científico que aboga por la mejora de la condición humana mediante el uso de tecnologías avanzadas. Sus defensores proponen que las limitaciones biológicas que han caracterizado la existencia humana, tales como el envejecimiento, las enfermedades y la muerte, pueden ser superadas o mitigadas con el uso de innovaciones científicas como la biotecnología, la nanotecnología, la ingeniería genética y la inteligencia artificial. Este movimiento aspira a trascender las capacidades humanas actuales, no solo en términos de salud y longevidad, sino también en relación con las capacidades cognitivas, emocionales y físicas.

Alva-Arroyo, Torres-García, Gasca-Aldama y otros (2022) comparten que *"el transhumanismo nos promete la libertad de las limitaciones biológicas inherentes a nuestra naturaleza. Pretende potenciar las capacidades físicas, emocionales y cognitivas abriendo así nuevas posibilidades y horizontes de experiencia"* (s. p.). Teniendo esto en cuenta, entre los objetivos más ambiciosos del transhumanismo, se encuentra la prolongación indefinida de la vida humana, e incluso la búsqueda de la inmortalidad a través de la fusión del cuerpo con sistemas tecnoló-

gicos avanzados, o mediante la *"subida"* de la consciencia a entornos digitales.

La idea de alcanzar la *"singularidad tecnológica"*, un momento en el cual la tecnología evoluciona más allá del control humano y permite a los individuos trascender los límites biológicos, es uno de los puntos clave de la visión transhumanista. Esta singularidad promete cambiar radicalmente lo que significa ser humano, abriendo la puerta a una nueva era en la que las personas podrán mejorar su cuerpo y mente a través de intervenciones tecnológicas (Tillería, 2022).

Ahora bien, la Inteligencia Artificial es el campo de la informática que se centra en la creación de sistemas capaces de realizar tareas que, tradicionalmente, requieren de la inteligencia humana. Entre estas tareas se incluyen el reconocimiento de patrones, la toma de decisiones, el aprendizaje y la comprensión del lenguaje natural. Esta tecnología ha avanzado considerablemente desde sus inicios en la década de 1950, cuando Alan Turing sentó las bases teóricas del campo con su trabajo sobre computación y su famosa prueba de *Turing*, la cual evalúa la capacidad de una máquina para exhibir un comportamiento indistinguible del de un ser humano (Valor-Yébenes, 2024).

El desarrollo de la Inteligencia Artificial ha pasado por varias etapas, desde la programación de algoritmos básicos hasta la creación de redes neuronales profundas que imitan el funcionamiento del cerebro humano. Estos avances han permitido la creación de sistemas capaces de aprender de manera autónoma y mejorar su desempeño con el tiempo, un concepto conocido como *machine learning* (Forero-Corba, 2024). Hoy en día, la Inteligencia Artificial es un componente central en una amplia gama de aplicaciones, desde la medicina personalizada hasta los vehículos autónomos, pasando por la automatización industrial y los asistentes virtuales.

El potencial de la IA para transformar sectores enteros de la economía y la sociedad ha generado tanto entusiasmo como preocupación. Mientras que muchos ven en la Inteligencia Artificial una herramienta poderosa para mejorar la eficiencia y resolver problemas complejos, otros advierten sobre los riesgos asociados con la automatización de empleos, la toma de decisiones algorítmica en ámbitos sensibles como

la justicia y la salud, y la posibilidad de que surjan formas de inteligencia que escapen al control humano.

ANTECEDENTES HISTÓRICOS DE AMBOS MOVIMIENTOS

El deseo de mejorar la condición humana a través del conocimiento y la tecnología no es nuevo. Las raíces filosóficas del transhumanismo pueden rastrearse hasta los tiempos de la antigua Grecia, cuando filósofos como Platón y Aristóteles discutían la posibilidad de mejorar el cuerpo y la mente humana a través de la educación y la razón. De hecho, los deseos trascendentalistas han estado presentes desde tiempos antiguos, como se evidencia en la búsqueda de la inmortalidad en la Epopeya de Gilgamesh y en las históricas búsquedas de la Fuente de la Juventud y el Elixir de la Vida, entre otros intentos por evitar el envejecimiento y la muerte. Además, en su obra *La Divina Comedia*, Dante introdujo el término *"trasumanar"* en el primer canto del *Paraíso*, que significa trascender la naturaleza humana, ir más allá de la naturaleza humana (AcademiaLab, 2024).

Asimismo, en Inglaterra, durante la primera mitad del siglo XVII, esta idea fue revitalizada por pensadores como Francis Bacon quien argumentaba que la ciencia y la tecnología eran las herramientas clave para dominar la naturaleza y mejorar la vida humana (Bacon, 2017).

Con la llegada de la Revolución Industrial y el desarrollo de las ciencias modernas, el sueño de superar las limitaciones biológicas mediante la tecnología adquirió un nuevo impulso. Las primeras discusiones sobre cibernética y la automatización en el siglo XX, así como los avances en biología molecular y genética, sentaron las bases para el surgimiento del transhumanismo como movimiento organizado en las décadas de 1980 y 1990.

El transhumanismo y la Inteligencia Artificial tienen raíces históricas y filosóficas que se entrelazan con el desarrollo de la ciencia y la tecnología. El concepto de mejorar la condición humana tiene antecedentes en la filosofía clásica y en la literatura de ciencia ficción. Sin embargo, el término *"transhumanismo"* fue popularizado por el biólogo Julien Huxley en la década de 1950, quien veía a este nuevo concep-

to como una misión religioso-científica que permitiría a la humanidad trascender su condición actual y alcanzar una autoconciencia cósmica a través del uso de la ciencia y la tecnología. Desde entonces, el movimiento ha evolucionado para incluir una amplia gama de tecnologías emergentes (Monterde-Ferrando, 2020).

Por otro lado, la Inteligencia Artificial tiene su origen en el desarrollo de la informática y la lógica formal. En el siglo XX, con la aparición de las primeras computadoras electrónicas y los trabajos de Turing y otros pioneros, la idea de crear máquinas que pudieran simular procesos cognitivos humanos comenzó a tomar forma. En términos filosóficos, la Inteligencia Artificial plantea cuestiones fundamentales sobre la naturaleza de la inteligencia, la consciencia y el libre albedrío. A medida que los sistemas de Inteligencia Artificial se vuelven más sofisticados, los límites entre la inteligencia artificial y la inteligencia humana se difuminan, lo que obliga a replantear las definiciones tradicionales de estos conceptos (Valor-Yébenes, 2024).

En conjunto, tanto el transhumanismo como la Inteligencia Artificial comparten una visión de progreso basada en la mejora continua de las capacidades humanas a través de la tecnología. Ambos movimientos desafían las nociones tradicionales sobre lo que significa ser humano y plantean nuevos dilemas éticos sobre la naturaleza de la existencia, la identidad y los derechos humanos en la era tecnológica. Mientras que el transhumanismo pone énfasis en la transformación del cuerpo y la mente humanos, la Inteligencia Artificial proporciona las herramientas para hacerlo realidad, creando una relación simbiótica entre ambas corrientes que tiene el potencial de redefinir el futuro de la humanidad.

LA INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA TRANSHUMANISTA

La Inteligencia Artificial ha sido un pilar fundamental en la materialización de muchos de los ideales transhumanistas. El Transhumanismo y la Inteligencia Artificial convergen en su visión de un futuro donde la tecnología juega un papel fundamental en la expansión de las ca-

pacidades humanas. La Inteligencia Artificial al ser una herramienta que amplía el conocimiento y la toma de decisiones, se considera un componente esencial en la realización de los ideales transhumanistas. A través de la Inteligencia Artificial, los transhumanistas ven la posibilidad de mejorar no solo las capacidades físicas, sino también las capacidades cognitivas de los seres humanos.

Una de las formas más directas en las que ambas corrientes se intersecan es a través de las interfaces cerebro-computadora, tecnologías que permiten la comunicación directa entre el cerebro humano y sistemas externos, incluidos los sistemas de Inteligencia Artificial. Estas interfaces permiten a los individuos controlar dispositivos o interactuar con computadoras utilizando únicamente sus pensamientos, lo que representa un avance significativo en la fusión entre humanos y máquinas. En un futuro transhumanista, se espera que estas interfaces se vuelvan más sofisticadas, permitiendo la expansión de la memoria, el aprendizaje acelerado y una mayor integración con sistemas de Inteligencia Artificial avanzados que potencien el intelecto humano (Bostrom, 2014).

En este sentido, la Inteligencia Artificial también juega un papel crucial en la extensión de la longevidad humana y la mejora de la calidad de vida. Los algoritmos de aprendizaje automático están siendo utilizados para analizar grandes volúmenes de datos médicos, lo que permite diagnósticos más precisos y tratamientos personalizados. Este enfoque no solo mejora la eficiencia del sistema de salud, sino que también abre la puerta a la medicina preventiva y a la posibilidad de predecir y mitigar enfermedades antes de que se manifiesten (Topol, 2019).

Además, el uso de Inteligencia Artificial en la medicina personalizada y la biotecnología también tiene profundas implicaciones para el transhumanismo. La capacidad de la Inteligencia Artificial para analizar grandes cantidades de datos biomédicos y desarrollar tratamientos específicos para cada individuo permite vislumbrar un futuro en el que las enfermedades genéticas puedan ser erradicadas y las capacidades biológicas de las personas puedan ser mejoradas de manera precisa. También, la Inteligencia Artificial está siendo integrada en prótesis avanzadas y exoesqueletos, permitiendo un control más natural

y eficiente de estos dispositivos, lo que mejora significativamente la calidad de vida de las personas con discapacidades (Herr, 2011). En este contexto, la Inteligencia Artificial no solo complementa, sino que es esencial para los fines del transhumanismo.

LOS DERECHOS HUMANOS EN LA ERA TRANSHUMANISTA

La era tecnológica ha transformado profundamente las sociedades contemporáneas, planteando nuevos desafíos para la protección y promoción de los derechos humanos. Desde la aparición de Internet y las tecnologías digitales hasta los recientes avances en Inteligencia Artificial y Biotecnología, las innovaciones tecnológicas han reconfigurado el panorama global en términos de acceso a la información, el trabajo, la salud y la participación ciudadana. No obstante, estos avances también generan tensiones éticas y normativas que exigen una reinterpretación de los marcos legales existentes y una adaptación de los sistemas de derechos humanos a las realidades emergentes.

Desde una perspectiva jurídica, la protección de los derechos humanos en el entorno digital requiere una adaptación de los marcos normativos existentes. Los derechos humanos, consagrados en los textos constitucionales y leyes fundamentales de las naciones y en los tratados internacionales deben interpretarse a la luz de las nuevas realidades tecnológicas.

Asimismo, el transhumanismo plantea interrogantes importantes sobre la igualdad de acceso a estas tecnologías, la preservación de la identidad personal y el respeto por la autonomía individual. En la medida en que el movimiento busca redefinir la naturaleza humana, también cuestiona las bases normativas que sustentan el concepto de derechos universales, como el derecho a la vida, la salud y la dignidad.

El derecho a la igualdad y no discriminación se enfrenta a desafíos tecnológicos. Los algoritmos de inteligencia artificial utilizados en procesos de selección laboral, concesión de créditos y otros servicios pueden perpetuar sesgos discriminatorios si no son diseñados cuidadosamente. Aunque los desarrolladores de Inteligencia Artificial trabajan para mitigar estos riesgos, las herramientas tecnológicas aún

pueden reflejar y amplificar desigualdades existentes, lo que exige una regulación adecuada y mecanismos de supervisión.

Continuando, la implementación de la Inteligencia Artificial en el ámbito transhumanista tiene el potencial de mejorar significativamente la calidad de vida y la autonomía de las personas, especialmente aquellas con discapacidades. Sin embargo, también plantea preocupaciones sobre la equidad y el acceso a estas tecnologías avanzadas. Según Binns (2018), la falta de acceso equitativo a la Inteligencia Artificial puede exacerbar las desigualdades existentes, creando una brecha aún mayor entre aquellos que pueden permitirse estas tecnologías y aquellos que no. Este fenómeno podría violar el principio de igualdad y no discriminación, que es fundamental en los derechos humanos. La equidad en el acceso a la IA es crucial para asegurar que todos los individuos puedan beneficiarse de los avances tecnológicos sin importar su situación socioeconómica (Binns, 2018).

Así también, la Inteligencia Artificial ha transformado la forma en que las sociedades recopilan, procesan y analizan datos, generando importantes implicaciones para el derecho a la privacidad y la protección de datos personales. Este desarrollo tecnológico ha ampliado la capacidad de los sistemas para manejar grandes volúmenes de información, lo que ha planteado tanto oportunidades como desafíos desde el punto de vista jurídico y ético.

El derecho a la privacidad, reconocido en múltiples instrumentos internacionales de derechos humanos, como la Declaración Universal de Derechos Humanos (1948) y el Pacto Internacional de Derechos Civiles y Políticos (1966), enfrenta una tensión creciente frente a las capacidades de la Inteligencia Artificial. Los sistemas de Inteligencia Artificial, a menudo alimentados por datos personales, pueden generar perfiles detallados de las personas, predecir comportamientos y tomar decisiones automatizadas, muchas veces sin el conocimiento o el consentimiento informado de los afectados. Este fenómeno pone en riesgo uno de los principios básicos del derecho a la privacidad: el control de los individuos sobre sus propios datos.

De igual manera, la protección de datos personales se ve comprometida por el uso de Inteligencia Artificial, ya que las técnicas de análisis predictivo y *machine learning* requieren grandes cantidades de

información para ser efectivas (Forero-Corba, 2024). La recopilación masiva de datos por parte de empresas privadas y gobiernos plantea cuestiones sobre la finalidad y el uso de dichos datos, así como sobre la transparencia en los procesos de toma de decisiones automatizadas.

Ahora bien, la integración de la Inteligencia Artificial en el cuerpo humano a través de prótesis avanzadas y interfaces cerebro-computadora plantea cuestiones éticas sobre la identidad y la privacidad. Floridi y otros (2018) argumentan que la capacidad de la Inteligencia Artificial para recopilar y analizar datos personales a un nivel sin precedentes puede poner en riesgo la privacidad de los individuos, lo que podría llevar a abusos y violaciones de los derechos humanos. La protección de la privacidad y la identidad es esencial para mantener la dignidad humana, y cualquier avance tecnológico debe ser cuidadosamente regulado para evitar la explotación y el abuso. La privacidad es un derecho humano fundamental que debe ser protegido en la era de la Inteligencia Artificial (Floridi y otros, 2018).

Además de la privacidad, el derecho a la libertad de expresión se enfrenta a nuevas amenazas en el entorno digital. Si bien Internet ha facilitado una mayor participación y acceso a la información, también ha dado lugar a fenómenos como la censura digital, la desinformación y los discursos de odio en línea. La Inteligencia Artificial puede influir en la libertad de expresión a través de la moderación automatizada de contenidos en plataformas digitales, donde algoritmos deciden qué información se permite o se elimina. Este proceso puede resultar en la censura de opiniones legítimas y la limitación del acceso a información diversa, afectando negativamente el derecho de las personas a buscar, recibir y difundir información sin restricciones indebidas (OHCHR, 2023). Además, la Inteligencia Artificial puede ser utilizada para la creación y difusión de desinformación y noticias falsas, lo que puede distorsionar el discurso público y socavar la confianza en los medios de comunicación y las instituciones democráticas (UNESCO, 2023). En contraste, la Inteligencia Artificial también presenta oportunidades para fortalecer la libertad de expresión al facilitar el acceso a información y conocimiento a través de tecnologías avanzadas de búsqueda y análisis de datos (OHCHR, 2023).

La libertad de expresión, en este sentido, se encuentra en una delicada balanza: por un lado, es necesario proteger el derecho de los individuos a expresar sus opiniones, mientras que, por otro, es fundamental regular los contenidos nocivos que puedan afectar a la dignidad de las personas o poner en peligro la estabilidad social.

En el ámbito del trabajo, la automatización y la inteligencia artificial están reconfigurando el derecho laboral. La sustitución de mano de obra humana por máquinas y algoritmos ha generado preocupación por el desplazamiento masivo de empleos, especialmente en sectores industriales y de servicios. Esta transformación plantea la necesidad de repensar el derecho al trabajo y las garantías para los trabajadores en un entorno cada vez más automatizado. La discusión sobre la renta básica universal ha ganado fuerza como una posible solución a los desafíos laborales que plantea la automatización, pero aún falta consenso sobre su viabilidad y sostenibilidad económica.

Continuando con la discusión del transhumanismo, el mejoramiento humano es un concepto que se refiere a la aplicación de diversas tecnologías y prácticas destinadas a aumentar las capacidades físicas, cognitivas y emocionales de las personas. Este fenómeno, impulsado por avances en biotecnología, neurociencia y medicina regenerativa, ha generado un intenso debate en ámbitos como la ética, la justicia social y el derecho. Su relación con el derecho a la vida y la salud es particularmente relevante, ya que plantea interrogantes fundamentales sobre el futuro de la condición humana.

El derecho a la vida, en su interpretación más amplia, abarca no solo la mera existencia, sino también el derecho a vivir de manera digna y saludable. La Declaración Universal de los Derechos Humanos (1948) establece que toda persona tiene derecho a la vida, la libertad y la seguridad. En este sentido, el mejoramiento humano se presenta como un potencial derecho individual, donde cada persona podría reclamar la autonomía para decidir sobre su propio cuerpo y capacidades. Sin embargo, esta autonomía debe ser equilibrada con la responsabilidad social de garantizar que el acceso a tales mejoras no se convierta en un privilegio reservado para unos pocos.

La salud, entendida desde una perspectiva integral, incluye no solo la ausencia de enfermedad, sino también el bienestar físico, mental y

social. El mejoramiento humano, al ofrecer la posibilidad de aumentar el bienestar, puede ser visto como una extensión del derecho a la salud. Sin embargo, es crucial considerar los riesgos y efectos colaterales que pueden acompañar a estas intervenciones. La intervención médica conlleva incertidumbres y dilemas éticos que deben ser cuidadosamente evaluados, ya que pueden impactar no solo al individuo, sino también a la sociedad en su conjunto.

Un desafío central radica en la regulación de estas tecnologías. Las políticas deben garantizar que el mejoramiento humano se lleve a cabo de manera ética y justa, protegiendo los derechos de los individuos y evitando la explotación. Es fundamental que el marco legal se adapte a los avances tecnológicos, promoviendo un enfoque inclusivo que no deje atrás a los grupos más vulnerables. La regulación debe estar guiada por principios éticos que prioricen el bienestar común y la equidad en el acceso a estas innovaciones.

Por tanto, el mejoramiento humano es un fenómeno complejo que interseca con cuestiones fundamentales del derecho a la vida y la salud. Mientras que ofrece oportunidades para mejorar la calidad de vida, también plantea serias preocupaciones sobre la equidad y la justicia social. La forma en que la sociedad y los sistemas jurídicos aborden estos desafíos determinará no solo el desarrollo de nuevas tecnologías, sino también el futuro de la dignidad humana en un mundo en constante evolución. Es imperativo fomentar un diálogo inclusivo que considere todas las voces y perspectivas, asegurando que el mejoramiento humano sirva a los intereses de la humanidad en su conjunto.

El control sobre la IA también plantea preocupaciones en términos de desigualdad. Las grandes corporaciones tecnológicas, ubicadas predominantemente en economías avanzadas, ejercen un dominio casi absoluto sobre el desarrollo y la implementación de los sistemas de IA, lo que concentra el poder de decisión sobre el uso y los fines de esta tecnología en un número reducido de actores. Esta centralización del control genera asimetrías en la distribución de los beneficios y riesgos de la IA, dejando a muchos países y comunidades sin capacidad para influir en cómo se desarrollan y utilizan estas tecnologías, y sin medios para protegerse de sus potenciales efectos negativos.

En este contexto, es fundamental que los gobiernos y organismos internacionales adopten políticas que promuevan una distribución más equitativa de los beneficios de la IA. Esto incluye fomentar el acceso a la tecnología en países en desarrollo, impulsar la educación en competencias digitales y establecer marcos regulatorios que garanticen una mayor participación de todos los actores en el diseño y control de la IA. Solo así se podrá mitigar el impacto de la desigualdad tecnológica y asegurar que la IA contribuya al bienestar colectivo sin perpetuar las brechas existentes.

REGULACIÓN Y GOBERNANZA DE LA INTELIGENCIA ARTIFICIAL

El avance del transhumanismo y la Inteligencia Artificial es inevitable, pero su implementación debe estar sujeta a una regulación que proteja los derechos humanos. Es necesario un marco legal y ético que permita el desarrollo de estas tecnologías sin comprometer los principios fundamentales de igualdad, dignidad y libertad.

La regulación y la gobernanza de la Inteligencia Artificial son esenciales para asegurar que su implementación respete los derechos humanos. Según Cath y otros (2018), es necesario establecer marcos regulatorios que aborden las implicaciones éticas y sociales de la Inteligencia Artificial. Estos marcos deben ser inclusivos y considerar las voces de diversos grupos de interés, incluyendo a aquellos que podrían ser más afectados por estas tecnologías. La gobernanza de la Inteligencia Artificial debe ser transparente y responsable, asegurando que las decisiones sobre su uso se tomen de manera justa y equitativa. La regulación adecuada de la Inteligencia Artificial es crucial para proteger los derechos humanos y promover el bienestar social (Cath y otros, 2018).

No es nada sencillo establecer normas en materia de Inteligencia Artificial. Se tornan necesarios aspectos paradigmáticos en materia legislativa que permitan, en primer término, la socialización de los aspectos tecnológicos que faciliten la integración de las máquinas y los sistemas informáticos en la vida cotidiana para mejorar la calidad de vida de los humanos. No obstante, en palabras de la Oficina del

Alto Comisionado de las Naciones Unidas para los Derechos Humanos, citando a la Comisionada, Michelle Bachelet, (2021): *"La inteligencia artificial puede ser una fuerza para el bien, que ayude a las sociedades a superar algunos de los mayores retos de nuestro tiempo. Pero las tecnologías de IA también pueden tener efectos nocivos e incluso catastróficos, cuando se emplean sin prestar la debida atención a su capacidad de vulnerar los derechos humanos" (s. p.).*

La regulación y la gobernanza de la Inteligencia Artificial deben basarse en principios éticos sólidos y en el respeto a los derechos humanos. La UNESCO, en su Recomendación sobre la Ética de la Inteligencia Artificial de 2021, establece un estándar global para la Inteligencia Artificial, basado en valores y principios éticos que deben aplicarse en todo el ciclo de vida de la Inteligencia Artificial. Este documento ofrece orientación detallada sobre cómo aplicar estos principios en áreas como la gobernanza de datos, el medio ambiente, la igualdad de género, la economía y el trabajo, la cultura, la educación y la investigación. La implementación de estos principios es esencial para asegurar que la Inteligencia Artificial se desarrolle y utilice de manera que beneficie a toda la sociedad (UNESCO, 2021).

La legislación también juega un papel crucial en la regulación de la Inteligencia Artificial. El Reglamento de Inteligencia Artificial de la Unión Europea, aprobado en 2023, aborda los posibles riesgos para la salud, la seguridad y los derechos fundamentales de los ciudadanos. Este reglamento establece un marco legal para la Inteligencia Artificial, que incluye requisitos estrictos para la transparencia, la responsabilidad y la supervisión (Comisión Europea, 2023). En este mismo sentido, esta ley establece normas armonizadas para la comercialización, puesta en servicio y uso de la Inteligencia Artificial, a través de un enfoque basado en el riesgo, en el que, cuanto mayor sea el riesgo de violación de los derechos, más estrictas serán las restricciones.

La teoría del riesgo es un criterio para determinar la responsabilidad de una persona, tanto física como jurídica, que ha causado un daño a otra persona, ya sea física o jurídica. Dicho lo anterior, los Estados parte buscan disminuir al máximo el daño a las personas en cuanto a sus derechos fundamentales, lo cual va en contraposición con el principio *pro persona*, el cual permite legislar una norma en bene-

ficio más favorable para las personas. Este principio está consagrado en diversos instrumentos internacionales de derechos humanos y en muchas constituciones nacionales. Su objetivo es garantizar la máxima protección y promoción de los derechos humanos, asegurando que las leyes y políticas públicas se interpreten y apliquen de manera que beneficien más ampliamente a las personas.

Continuando, el objetivo de esta norma regulatoria es mejorar el funcionamiento del mercado interior y promover la adopción de la Inteligencia Artificial centrada en el ser humano y digna de confianza, garantizando al mismo tiempo un alto nivel de protección de la salud, la seguridad, los derechos fundamentales consagrados en la Carta, incluida la democracia, el Estado de Derecho y la protección del medio ambiente, contra los efectos nocivos de los sistemas de inteligencia artificial en la Unión y apoyando la innovación. Con esto, se busca que la Unión sea un líder mundial en el desarrollo de Inteligencia Artificial segura, digna de confianza y ética (Comisión Europea, 2024).

Otro caso es el de Chile, país que en 2021 se convirtió en el primero del mundo en integrar los neuroderechos a su legislación, marcando un hito en la protección de los derechos humanos en la era de las neurotecnologías. Esta iniciativa, impulsada por el senador Guido Girardi y otros legisladores, busca proteger la integridad mental y el libre albedrío de las personas frente a los avances en neurotecnología.

La ley modifica la Constitución chilena para incluir la protección de los *"derechos del cerebro"*, asegurando que los datos neuronales sean tratados como una extensión del cuerpo humano y, por lo tanto, sujetos a los mismos derechos y protecciones. Esta legislación pionera responde a la creciente preocupación por el potencial uso indebido de tecnologías capaces de leer y manipular la actividad cerebral, y establece un marco legal para prevenir la explotación y el abuso de estas tecnologías (UNESCO, 2022).

Esta legislación chilena también establece la creación de un Consejo Nacional de Neurotecnologías, encargado de supervisar y regular el uso de estas tecnologías en el país. Este consejo tiene la responsabilidad de desarrollar políticas públicas y normativas que promuevan el uso ético y seguro de las neurotecnologías, garantizando que se respeten los derechos humanos y se eviten prácticas discrimina-

torias. Además, la ley incluye disposiciones específicas para proteger la privacidad de los datos neuronales y prevenir su comercialización sin el consentimiento explícito de los individuos. La adopción de esta legislación refleja el compromiso de Chile con la protección de los derechos humanos en el contexto de los rápidos avances tecnológicos y sirve como modelo para otros países que buscan regular el uso de neurotecnologías de manera ética y responsable (BioBioChile, 2021).

Ahora bien, en mayo de 2024, se presentó una iniciativa por el diputado Ignacio Loyola Vera y el Grupo Parlamentario del Partido Acción Nacional, para expedir la Ley Federal que Regula la Inteligencia Artificial en México. Esta propuesta busca establecer un marco legal integral que promueva el desarrollo ético y responsable de la inteligencia artificial, protegiendo los Derechos Humanos consagrados en la Constitución Política de los Estados Unidos Mexicanos y en los Tratados Internacionales de los cuales México forma parte.

Esta iniciativa tiene como objetivo establecer un marco normativo que regule el uso de estas tecnologías emergentes con un enfoque ético y de derechos humanos. En este sentido, esta propuesta legislativa busca garantizar que la Inteligencia Artificial y la robótica se desarrollen y utilicen de manera que respeten los derechos fundamentales de las personas, promoviendo la equidad, la transparencia y la responsabilidad. La iniciativa subraya la necesidad de crear Normas Oficiales Mexicanas basadas en principios éticos para el uso adecuado de la Inteligencia Artificial y la robótica, asegurando que estas tecnologías beneficien a la sociedad mexicana en su conjunto, sin discriminación por género, raza, origen étnico, religión, clase social o posición económica (Cámara de Diputados, 2023).

Uno de los aspectos más destacados de la iniciativa es la creación del Consejo Mexicano de Ética para la Inteligencia Artificial y la Robótica, un organismo público descentralizado encargado de supervisar y regular el uso de estas tecnologías en México. Este consejo tendrá la responsabilidad de desarrollar políticas públicas y normativas que promuevan el uso ético de la Inteligencia Artificial y la robótica, así como de monitorear su implementación y efectos en la sociedad. Además, la iniciativa propone la creación de la Red Nacional de Estadística de Uso y Monitoreo de la Inteligencia Artificial y la Robótica, que esta-

rá vinculada a organismos autónomos y tendrá como objetivo generar información precisa y actualizada sobre el uso de estas tecnologías en el país (Sistema de Información Legislativa, 2023). Esta red permitirá una mejor comprensión de los impactos sociales, económicos y éticos de la Inteligencia Artificial y la robótica, facilitando la toma de decisiones informadas y la formulación de políticas efectivas.

La propuesta legislativa también aborda la necesidad de regular la Inteligencia Artificial y la robótica en diversos sectores, incluyendo el gubernamental, económico, comercial, administrativo, comunicacional y financiero. La iniciativa destaca que la rápida proliferación de estas tecnologías sin una regulación adecuada puede generar desventajas directas o indirectas, como la manipulación de datos, la discriminación y la violación de la privacidad.

Por lo tanto, es imperativo establecer un marco legal que garantice la protección de los derechos humanos y promueva el uso responsable de la Inteligencia Artificial y la robótica. La iniciativa subraya la importancia de anticiparse a los posibles riesgos y desventajas de estas tecnologías, asegurando que su desarrollo y aplicación se realicen de manera ética y en beneficio de toda la sociedad mexicana (Loyola-Vera, 2023).

En síntesis, con el desarrollo de esta iniciativa de ley en materia de gobernanza, México busca equilibrar la innovación tecnológica con la protección de valores fundamentales y la creación de un entorno seguro en materia de Inteligencia Artificial.

DERECHOS HUMANOS EMERGENTES PARA UN FUTURO POST HUMANO

Como se ha advertido con anterioridad, la Inteligencia Artificial y el Transhumanismo están transformando radicalmente nuestra comprensión de los derechos humanos, planteando la necesidad de reconocer y proteger nuevos derechos emergentes en un futuro post humano. La Inteligencia Artificial, como herramienta transhumanista, promete mejorar significativamente las capacidades humanas, desde la cognición hasta la longevidad, pero también plantea desafíos éticos

y legales sin precedentes. Según Ahmed (2019), el desarrollo de la Inteligencia Artificial y el Transhumanismo cuestiona lo que significa ser humano y quién o qué constituye lo *"humano"* en el contexto de los derechos humanos. Este debate es crucial, ya que la distinción entre humanos y máquinas se vuelve cada vez más borrosa, lo que exige una reevaluación de los marcos legales y éticos actuales para incluir a entidades no humanas sintientes (Ahmed, 2019).

En este contexto, los derechos humanos emergentes deben abordar la protección de la autonomía y la privacidad mental, en primera instancia, especialmente ante tecnologías que pueden manipular o influir en el comportamiento humano. La Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (OHCHR) ha subrayado la importancia de basar cualquier regulación de la Inteligencia Artificial en el respeto a los derechos humanos, destacando que los sistemas de Inteligencia Artificial utilizados en el Sistema de Justicia Penal ya han demostrado potencial para la discriminación y la erosión de derechos fundamentales, como la presunción de inocencia (OHCHR, 2023). Este enfoque preventivo es esencial para garantizar que las tecnologías emergentes no exacerben las desigualdades existentes ni socaven las libertades fundamentales.

Además, la incorporación de la Inteligencia Artificial en la vida cotidiana y su integración con el cuerpo humano a través del transhumanismo plantea la necesidad de nuevos derechos específicos, como los neuroderechos, con los cuales ya se habla de una quinta generación de derechos humanos.

Lo más destacado de esta quinta generación de derechos humanos es la protección de los derechos humanos ante el desarrollo de las neurociencias, lo que dio pauta al surgimiento de los neuroderechos, los cuales son una parte esencial de esta nueva generación y buscan proteger la integridad mental y la privacidad de las personas en el contexto de los avances en neurociencia y tecnología.

Retomando el caso de la legislación chilena sobre el reconocimiento y protección de neuroderechos, la cual establece un precedente importante al reconocer la necesidad de proteger la actividad cerebral y la información derivada de ella, garantizando el libre albedrío y la privacidad mental (Cornejo-Plaza, 2023). Este marco legal puede servir

de modelo para otros países que buscan regular el uso de tecnologías avanzadas de manera ética y responsable.

Chile ha buscado proyectar su experiencia en neuroderechos a nivel internacional. Tras la aprobación de la reforma constitucional, el país ha impulsado la creación de redes de colaboración con otros Estados y organismos internacionales para compartir su modelo y promover una regulación global coherente en esta materia. Este país se ha posicionado como un líder en la discusión sobre los límites éticos y legales de las neurotecnologías, participando en foros globales, como la UNESCO y el Foro Económico Mundial, donde se han discutido las implicaciones de estas tecnologías para los derechos humanos.

La evolución hacia un futuro post humano requiere una colaboración estrecha entre gobiernos, academia, sociedad civil y sector privado para desarrollar un marco ético y regulatorio que proteja los derechos humanos en todas sus dimensiones. La UNESCO ha enfatizado la necesidad de un enfoque coordinado para asegurar que la IA y otras tecnologías emergentes se desarrollen de manera que beneficien a toda la humanidad sin dejar a nadie atrás (UNESCO, 2023). Este enfoque debe incluir la creación de nuevos derechos que reflejen las realidades cambiantes de un mundo donde la línea entre lo humano y lo tecnológico es cada vez más difusa.

CONCLUSIONES

La Inteligencia Artificial como herramienta transhumanista ofrece oportunidades significativas para mejorar la calidad de vida y la autonomía de las personas, pero también plantea importantes desafíos desde la perspectiva de los derechos humanos. La equidad en el acceso, la protección de la privacidad y la identidad, la preservación de la autonomía y la regulación adecuada son aspectos cruciales que deben ser considerados para asegurar que la Inteligencia Artificial se utilice de manera ética y responsable. Asimismo, la integración de la Inteligencia Artificial en la sociedad debe ser guiada por principios de derechos humanos para garantizar que sus beneficios se distribuyan de manera justa y equitativa.

De igual manera, la era tecnológica trae consigo oportunidades sin precedentes para el ejercicio de los derechos humanos, pero también plantea nuevos desafíos que deben ser abordados con un enfoque multidisciplinario. El derecho a la privacidad, la libertad de expresión, el derecho al trabajo y la no discriminación son solo algunas de las áreas donde los avances tecnológicos requieren un replanteamiento de las normativas vigentes.

Adaptar los derechos humanos a este nuevo contexto es fundamental para garantizar que las innovaciones tecnológicas se utilicen para promover la dignidad y el bienestar de todas las personas, sin comprometer los principios fundamentales de justicia y equidad. Aunado a esto, el reconocimiento y protección de los derechos humanos emergentes, entre los que destacan neuroderechos, es de suma importancia en esta quinta generación de derechos humanos, ya que consideran y buscan proteger la integridad mental y la privacidad ante los avances en neurociencia y tecnología, asegurando que el desarrollo científico respete y preserve la autonomía y dignidad de las personas en un contexto de rápida evolución tecnológica.

Los riesgos que supone el acelerado avance del transhumanismo, como lo son la desigualdad social, la pérdida de autenticidad humana, problemas de salud, dilemas éticos, control y vigilancia, deshumanización, abusos, crisis de identidad, impacto en la biodiversidad, dependencia tecnológica, por mencionar los más graves, requiere de una tarea titánica y urgente a efecto de prevenir y proteger los derechos de las personas ante la inminencia de estos peligros.

Si bien el transhumanismo plantea oportunidades emocionantes para el futuro de la Humanidad, las amenazas latentes requieren de una reflexión cuidadosa y un debate ético inmediato que garanticen la implementación de las neurotecnologías para el beneficio de todos los seres humanos, en lugar de exacerbar desigualdades y poner en estado de vulnerabilidad la esencia de lo que significa ser humano.

REFERENCIAS

AcademiaLab. (2024, 27 de septiembre). Transhumanismo. https://academia-lab.com/enciclopedia/transhumanismo/

Ahmed, A. K. (2019). Desvincular lo "humano" de los derechos humanos: inteligencia artificial y transhumanismo. OpenGlobalRights. Recuperado de https://www.openglobalrights.org/delinking-the-human-from-human-rights-artificial-intelligence-and-transhumanism/?lang=Spanish

Alva-Arroyo, Torres-García, Gasca-Aldama, et al. (2022). Transhumanismo. Med Int Méx 2022; 38 (4): 741-751. https://doi.org/10.24245/mim.v38i4.7731

Asamblea General de las Naciones Unidas. (1948). Declaración Universal de Derechos Humanos. https://www.un.org/es/about-us/universal-declaration-of-human-rights

Bacon, F. (2017). La Nueva Atlántida. Fondo de Cultura Económica.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149-159. https://doi.org/10.1145/3287560.3287583

BioBioChile. (2021). Chile se convierte en el primer país del mundo en aprobar una ley que protege los neuroderechos. BioBioChile. https://www.biobiochile.cl/noticias/nacional/chile/2021/09/30/chile-se-convierte-en-el-primer-pais-del-mundo-en-aprobar-una-ley-que-protege-los-neuroderechos.shtml

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

Cámara de Diputados. (2023, 27 de mayo). Impulsan iniciativa para expedir la Ley de Regulación Ética de la Inteligencia Artificial y la Robótica. https://comunicacionsocial.diputados.gob.mx/index.php/boletines/impulsan-iniciativa-para-expedir-la-ley-de-regulacion-tica-de-la-inteligencia-artificial-y-la-robotica#

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. Science and Engineering Ethics, 24(2), 505-528. https://doi.org/10.1007/s11948-017-9901-7

Comisión Europea. (2023). Reglamento de Inteligencia Artificial. Recuperado de https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_es

Comisión Europea. (2024). Configurar el futuro digital de Europa. https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai

Cornejo-Plaza, M. I. (2023). Neuroderechos en Chile: consagración constitucional y regulación de las neurotecnologías. Somos Iberoamérica. Recuperado de https://somosiberoamerica.org/tribunas/neuroderechos-en-chile-consagracion-constitucional-y-regulacion-de-las-neurotecnologias/

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Forero-Corba, W., & Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e inteligencia artificial en educación: una revisión sistemática. RIED-Revista Iberoamericana de Educación a Distancia, 27(1), 209–253. https://doi.org/10.5944/ried.27.1.37491

González, C. (21 de mayo de 2023). ¿Qué es la Ley de Inteligencia Artificial de la Unión Europea? https://computerhoy.com/legislacion/ley-inteligencia-artificial-union-europea-1241502

Gutiérrez-Gómez, J. (2023, 20 de diciembre). Inteligencia artificial y transhumanismo: su relación y las implicaciones bioéticas. *Universidad Santo Tomás*. http://hdl.handle.net/11634/53373

Herr, H. (2011). Exoskeletons and orthoses: Classification, design challenges and future directions. Journal of NeuroEngineering and Rehabilitation, 8(1), 1-9. https://doi.org/10.1186/1743-0003-8-1

Loyola-Vera, I. (2023). Iniciativa para expedir la Ley de Regulación Ética de la Inteligencia Artificial y la Robótica. Cámara de Diputados. Recuperado de http://sil.gobernacion.gob.mx/Archivos/Documentos/2023/04/asun_4551867_20230420_1680209419.pdf

Microsoft Copilot. (2024). Copilot Bing. https://copilot.microsoft.com/

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679. https://doi.org/10.1177/2053951716679679

Monterde-Ferrando, R. (2020). El transhumanismo de Julian Huxley: Una nueva religión para la humanidad. *Cuadernos de Bioética*, 31(101), 71-85. https://doi.org/10.30444/CB.53

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OHCHR]. (15 de septiembre de 2021). Los riesgos de la inteligencia artificial para la privacidad exigen medidas urgentes - Bachelet. https://www.ohchr.org/es/press-releases/2021/09/artificial-intelligence-risks-privacy-demand-urgent-action-bachelet

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OHCHR]. (1966). Pacto Internacional de Derechos Civiles y Políticos. (1966). https://www.ohchr.org/es/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OHCHR]. (2023). La inteligencia artificial debe tomar como base los derechos humanos, declara el Alto Comisionado. Recuperado de https://www.ohchr.org/es/statements/2023/07/artificial-intelligence-must-be-grounded-human-rights-says-high-commissioner

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OHCHR]. (2023). La inteligencia artificial debe tomar como base los derechos humanos, declara el Alto Comisionado. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos. Recuperado de https://www.ohchr.org/es/statements/2023/07/artificial-intelligence-must-be-grounded-human-rights-says-high-commissioner

OpenAI. (2024). ChatGPT [Modelo de lenguaje de gran tamaño]. https://chat.openai.com/chat

Sistema de Información Legislativa. (2023). Iniciativa para la regulación ética de la inteligencia artificial y la robótica. Recuperado de http://sil.gobernacion.gob.mx/Librerias/pp_ContenidoAsuntos.php?Clave=4572130

Tillería, L. (2022, 21 de julio). Transhumanismo e Inteligencia Artificial: el problema de un límite ontológico. Griot: Revista De Filosofia, 22(1), 59–67. https://doi.org/10.31977/grirfi.v22i1.2539

Topol, E. (2019). Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Basic Books.

UNESCO. (2021). Recomendación sobre la Ética de la Inteligencia Artificial. https://www.unesco.org/es/articles/hacia-una-inteligencia-artificial-desde-un-enfoque-de-derechos-humanos-articulo-de-opinion

UNESCO. (2022). Chile, pionero en la protección de los “neuroderechos”. https://courier.unesco.org/es/articles/chile-pionero-en-la-proteccion-de-los-neuroderechos

UNESCO. (2023). Hacia una inteligencia artificial desde un enfoque de Derechos Humanos: artículo de opinión. Recuperado de https://www.unesco.org/es/articles/hacia-una-inteligencia-artificial-desde-un-enfoque-de-derechos-humanos-articulo-de-opinion

Unión Europea [UE]. (2024) La Ley de Inteligencia Artificial de la UE. https://artificialintelligenceact.eu/es/

Unión Europea [UE]. (2024, 25 de enero). Las claves de la nueva ley de Inteligencia Artificial. https://spain.representation.ec.europa.eu/noticias-eventos/noticias-0/las-claves-de-la-nueva-ley-de-inteligencia-artificial-2024-01-25_es

Valor-Yébenes, J. A. (2024). Alan Turing y el origen de la inteligencia artificial: La superación de la intuición. Naturaleza Y Libertad. *Revista de Estudios Interdisciplinares*, (18). https://doi.org/10.24310/nyl.18.2024.19496

CAPÍTULO XIII

EL DILEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL EN LA BÚSQUEDA DE LA SUSTENTABILIDAD: UNA MIRADA CRÍTICA

EDGAR ALFONSO SANSORES GUERRERO[1]
JUANA EDITH NAVARRETE MARNEOU[2]

INTRODUCCIÓN

En el último lustro, la inteligencia artificial (IA) ha emergido como uno de los avances tecnológicos más disruptivos, transformando prácticamente todos los sectores de la sociedad y la vida cotidiana de los seres humanos. En medio de esta revolución tecnológica, se ha generado un intenso debate sobre el impacto de la IA en la sustentabilidad (Carrillo, 2020; Elshaikh y otros., 2024; Nishant *y otrosl*.,2020).

Para algunos investigadores, (Van Wynsberghe, 2021; Raval *y otros*, 2021; Nurhaeni *y otros*., 2024; Dhabliya, 2024) la IA posee el potencial de transformar la gestión de los recursos naturales a través de la optimización en la producción de energía, la predicción de patrones climáticos y la reducción de los residuos sólidos.

En contraparte, algunos estudios han evidenciado el alto consumo de energía que requiere el entrenamiento de modelos complejos (Mora *y otros*., 2018). En ese sentido, el proceso de entrenar enormes redes neuronales requiere de grandes cantidades de energía que, en

[1] Profesor Visitante, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

[2] Profesora Investigadora, Universidad Autónoma del Estado de Quintana Roo, Chetumal, Quintana Roo, México

su mayoría, provienen de fuentes no renovables, lo que resulta en una considerable huella de carbono (Mora *y otros*., 2018).

Adicionalmente, los centros de datos, donde se almacenan y procesan grandes volúmenes de información, representan entre el 1% y el 1.5% de la demanda total de electricidad a nivel mundial (Yang *y otros*., 2018; Yang *y otros*., 2022). Aunque esta cifra resulta baja, su tendencia muestra un crecimiento exponencial de los datos que genera y maneja la IA (Bickley y otros., 2024; Yang *y otros*, 2022).

La producción de hardware especializado, como los procesadores y las unidades de procesamiento gráfico (GPU) que hacen posible el entrenamiento de modelos, también tiene un impacto ambiental considerable. La extracción de materias primas, como los metales necesarios para fabricar estos componentes, consume recursos naturales y puede causar daños a los ecosistemas locales (Nunes y Shrotryia, 2020; Dayarathna y Fan ,2015). Este fenómeno genera grandes cantidades de residuos, muchos de los cuales contienen materiales peligrosos como plomo, mercurio y cadmio, que pueden contaminar el suelo y los cuerpos de agua si no se gestionan adecuadamente (Sah *y otros*., 2024)

En este contexto, la adopción de la inteligencia artificial presenta tanto oportunidades como riesgos. Por un lado, la IA tiene el potencial de optimizar procesos productivos, reducir emisiones de carbono y mejorar la toma de decisiones ambientales a gran escala. Por otro, plantea preguntas inquietantes sobre la equidad, la justicia y el control que las tecnologías avanzadas ejercen sobre los recursos, los datos y, en última instancia, sobre las vidas humanas y la naturaleza (Astobiza y otros., , 2020).

El dilema ético que rodea a la IA en la búsqueda de la sustentabilidad se enmarca en una tensión fundamental (Goralski y Tan, 2020) ¿es la IA una herramienta alineada con los principios de justicia social y equidad ambiental, o está profundizando las desigualdades existentes y exacerbando los problemas que pretende resolver?

Este dilema se acentúa al considerar la gobernanza de la IA, los sesgos inherentes a los sistemas automatizados, y el control de estos por parte de grandes corporaciones tecnológicas. ¿Quién decide cómo se usan estas tecnologías? ¿En beneficio de quién? ¿Cómo se aseguran

la transparencia, la rendición de cuentas y la inclusión en el diseño y uso de la IA?

A nivel práctico, las tecnologías de IA pueden ayudar a enfrentar retos ambientales de manera más precisa y rápida (Goralski y Tan, 2020; Astobiza *y otros*., , 2020; Latour y Venn, 2002). Sin embargo, los beneficios de estos desarrollos no están distribuidos de manera equitativa, y en muchos casos, las soluciones que la IA promueve pueden ser inaccesibles para las comunidades más vulnerables o pueden perpetuar estructuras de poder desiguales.

También, el costo ambiental de la misma tecnología que promete salvar al planeta no debe subestimarse. La enorme cantidad de datos que requieren las soluciones basadas en IA y el uso extensivo de energía por los centros de procesamiento de datos son aspectos que necesitan ser considerados en cualquier evaluación ética (Bickley y otros., 2024; y otros., 2022).

Desde una perspectiva crítica, es crucial analizar las narrativas que rodean la IA y la sustentabilidad. Las promesas tecnológicas estan cargadas de un optimismo casi utópico, donde se presenta a la IA como la solución definitiva para los problemas de la humanidad (Latour y Venn, 2002).

Sin embargo, esta visión puede ignorar las complejidades sociales, políticas y culturales involucradas en la implementación de estas tecnologías, minimizando los riesgos y los efectos colaterales no deseados. Así, el desarrollo de la IA no puede entenderse solo como una cuestión técnica; requiere un análisis profundo de sus implicaciones éticas y su integración en los contextos sociales y ecológicos en los que se despliega.

El presente capítulo tiene como objetivo realizar una revisión de las teorías que podrían contribuir al estudio crítico de los dilemas éticos que surgen en la intersección entre inteligencia artificial y sustentabilidad.

Se abordarán cuestiones clave como la justicia ambiental, la equidad social, el impacto energético de las tecnologías de IA y los problemas relacionados con la transparencia y la rendición de cuentas en la toma de decisiones automatizadas.

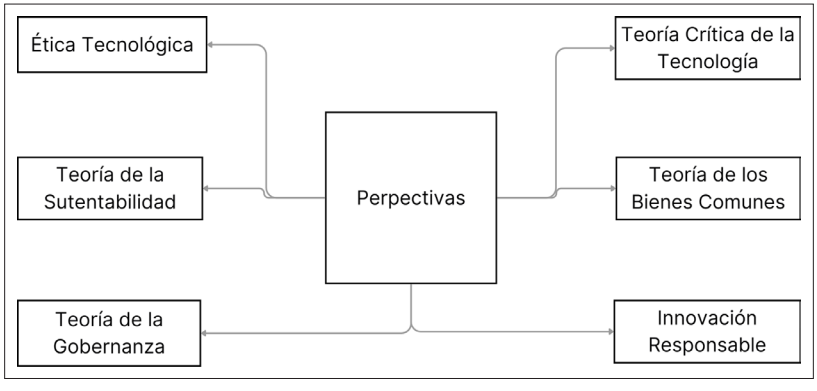
Finalmente, el capítulo subraya la importancia de una reflexión profunda: aunque la inteligencia artificial ofrece soluciones innovadoras para enfrentar algunos de los mayores desafíos de la humanidad, es crucial que estas soluciones sean diseñadas e implementadas con un enfoque que promueva un futuro justo, equitativo y verdaderamente sostenible para todos.

Lograr este objetivo requerirá una colaboración interdisciplinaria entre tecnólogos, filósofos, científicos sociales, ambientalistas y legisladores, garantizando que el avance tecnológico esté en sintonía con los principios éticos que deben guiar la construcción de un mundo más sostenible.

PERSPECTIVAS SOBRE LA IA EN LA SUSTENIBILIDAD: EXPLORANDO MÁS ALLÁ DE LOS ALGORITMOS.

Actualmente, la humanidad en la búsqueda de soluciones sostenibles para los retos medioambientales y sociales de esta era, ha posicionado a la IA como un actor principal en la agenda global. Sin embargo, la implementación de esta, plantea preguntas éticas fundamentales que deben ser abordadas desde diversas perspectivas teóricas (**Ver Gráfico 1**).

***Figura* 1.** *Perspectivas sobre el estudio de la IA en la sustentabilidad*



***Fuente*:** Elaboración propia basada en Schoormann *y otros.* (2023).

En consecuencia, la integración de enfoques como la ética de la tecnología, la justicia ambiental, la gobernanza algorítmica y la innovación responsable permiten un análsis holístico sobre el impacto de la IA en la sustentabilidad (Bickley y otros., 2024; Yang *y otros*., 2022; Goralski y Tan, 2020; Astobiza y otros., , 2020; Latour y Venn, 2002).

ÉTICA TECNOLÓGICA Y FILOSOFÍA MORAL: DEL UTILITARISMO A LA JUSTICIA COMO EQUIDAD

El análisis ético de la inteligencia artificial en la sustentabilidad parte de los principios filosóficos que rigen el comportamiento humano en relación con la tecnología. El utilitarismo (Tseng, 2021; Scarre, 2020) con su enfoque en maximizar el bienestar general, ofrece una primera aproximación para evaluar el impacto de la IA (Hare, 2020; Bauer, 2020).

Sustentado en los principios de consecuencialismo (la moralidad de una acción se juzga por sus resultados) y maximización de la felicidad (la acción correcta es la que maximiza la felicidad y minimiza el sufrimiento), esta vertiente evalúa la moralidad de las acciones basándose en sus consecuencias y en la cantidad de felicidad que generan para el mayor número de personas (Mill, 2016; Hare, 2014).

Desde esta óptica, se puede argumentar que las aplicaciones de IA deben ser juzgadas por su capacidad para reducir el sufrimiento y mejorar la calidad de vida global, ya sea mediante la optimización de los recursos naturales o la mitigación del cambio climático.

No obstante, el utilitarismo también plantea interrogantes sobre los sacrificios que algunas comunidades deben hacer para beneficiar a otras, lo que lleva a un análisis más profundo sobre quién se beneficia realmente de estas tecnologías (Scarre, 2020).

Por su parte, la deontología kantiana introduce un marco de derechos y deberes que se enfoca no solo en los resultados, sino en los medios (Kranak, 2019). Para Kant, la realización de una acción debe estar antecedida de una valoración sobre si dicha acción podría ser adoptada como una regla universal aplicable a todos (Udayakumar *y otros*., 2021).

Asimismo, la voluntad se considera como intrínsicamente buena y se asume que los seres humanos poseen la capacidad de autoimponerse leyes morales. Para esto, se identifican deberes perfectos que deben cumplirse en todo momento y deberes imperfectos que permiten cierta flexibilidad en su cumplimiento (Kranak, 2019; Scarre, 2020; Kranak, 2019).

Desde esta perspectiva, se puede argumentar que la implementación de IA debe respetar la dignidad y los derechos de todas las personas, independientemente de los beneficios colectivos que pueda generar. En este sentido, el uso de tecnologías que exacerban desigualdades o privan a comunidades vulnerables de autonomía sería éticamente inaceptable, independientemente de sus resultados ambientales.

A su vez, la ética de la justicia de John Rawls (1971) aporta una dimensión clave al debate sobre equidad. Contraria al utilitarismo, esta teoría sugiere que cada persona debe tener un conjunto igual de derechos y libertades básicas que sean compatibles con un sistema similar de libertades para todos (Principio de la libertad).

De la misma forma, las desigualdades sociales y económicas deben organizarse de tal manera que beneficien a los menos aventajados y estén vinculadas a posiciones y cargos accesibles para todos en condiciones de igualdad de oportunidades (Principio de la diferencia).

En consecuencia, las decisiones sobre el desarrollo y la aplicación de IA deben beneficiar a los más desfavorecidos de la sociedad (Rawls, 2000). La IA en la búsqueda de la sustentabilidad, por tanto, debe ser evaluada no solo por su eficiencia técnica, sino también por su capacidad para reducir las desigualdades sociales y ambientales. El marco rawlsiano demanda una evaluación crítica de cómo se distribuyen los beneficios y costos de las tecnologías emergentes (Freeman, 2017; Mile, 2017, 2020; Rawls, 2000).

SUSTENTABILIDAD: LA IA COMO CATALIZADOR DEL MODELO DE LA TRIPLE HÉLICE

En el núcleo de este marco se encuentra el concepto de desarrollo sustentable tal como fue formulado en el Informe Brundtland (1987), que

establece que las soluciones tecnológicas deben satisfacer las necesidades del presente sin comprometer las de las generaciones futuras (Slaper y Hall, 2011).

La IA se presenta como una herramienta poderosa para abordar problemas como la eficiencia energética, la producción agrícola sostenible y la gestión de recursos naturales (Batty, 2018). Sin embargo, el desarrollo sostenible también requiere que consideremos los efectos a largo plazo de la implementación de IA, como su impacto en el consumo de energía y la explotación de recursos tecnológicos (Goralski y Tan, 2020).

El modelo de la Triple Hélice (Triple Bottom Line), que integra las dimensiones económica, social y ambiental, es un marco clave para evaluar la contribución de la IA a la sustentabilidad. Mientras que la IA puede mejorar la eficiencia económica y ambiental, también debe garantizar que sus beneficios no se limiten a un grupo selecto, sino que se distribuyan de manera equitativa (Adams *y otros*., 2013; Alhaddi, 2015).

Esto conlleva a considerar si las soluciones tecnológicas diseñadas para maximizar el beneficio económico pueden estar en conflicto con los objetivos de equidad social o protección ambiental. Además, la teoría de la justicia ambiental complementa este análisis al destacar que las decisiones sobre el uso de IA tienen implicaciones diferentes para diversas comunidades, particularmente las más vulnerables.

Las tecnologías que se presentan como soluciones para la sustentabilidad podrían, paradójicamente, perpetuar desigualdades si los beneficios no se distribuyen de manera justa. Por ejemplo, el uso de IA para gestionar recursos hídricos o agrícolas podría excluir a las comunidades que no tienen acceso a estas tecnologías, intensificando las desigualdades existentes.

GOBERNANZA ALGORÍTMICA: UNA DIMENSIÓN EMERGENTE

El uso creciente de IA en la sustentabilidad también requiere un marco de gobernanza algorítmica que asegure que las decisiones automa-

tizadas sean transparentes, responsables y equitativas (Katzenbach y Ulbricht, 2019). Este enfoque reconoce que los algoritmos, aunque eficientes, pueden introducir sesgos que exacerben las desigualdades sociales y ambientales si no se supervisan adecuadamente (Issar y Aneesh, 2022).

Es fundamental que las aplicaciones de IA en la sustentabilidad estén diseñadas y gestionadas de manera que promuevan la equidad y la rendición de cuentas. De esto se deriva el cuestionamiento sobre quién controla la IA y cómo se aseguran los mecanismos de supervisión para prevenir abusos o la explotación de recursos naturales (Gritsenko y Wood, 2022).

A esto se suma el debate sobre regulación y políticas públicas. La IA en sustentabilidad requiere un marco normativo claro que proteja tanto a las personas como al medio ambiente. Sin regulaciones adecuadas, el riesgo es que las grandes corporaciones tecnológicas que desarrollan estas soluciones persigan principalmente sus propios intereses económicos, en lugar de los intereses públicos (Bloch-Wehba, 2022).

La creación de políticas públicas que orienten el desarrollo y la implementación de IA de manera ética y sostenible es crucial para evitar que las promesas tecnológicas se conviertan en herramientas de explotación o exclusión.

Por lo tanto, la gobernanza de la IA es un esfuerzo multidimensional que busca equilibrar la innovación tecnológica con la protección de los derechos humanos y la seguridad. Es un campo en constante evolución, adaptándose a los rápidos avances en la tecnología de IA y a las nuevas preocupaciones que surgen con su uso.

EL SESGO IDEOLÓGICO DE LA TECNOLOGÍA: UNA POSTURA CRÍTICA

Desde este enfoque, la teoría crítica de la tecnología (inspirada en la Escuela de Frankfurt) sostiene que la tecnología no es neutral; está profundamente influenciada por las ideologías y los intereses de quie-

nes la controlan (Delanty y Harris, 2021; Corradetti, 2012; Finlayson, 2009).

Esta corriente de pensamiento se basa en las ideas de la teoría crítica, que busca entender y criticar las estructuras de poder y dominación en la sociedad, considerando cinco elementos esenciales (**Ver Tabla 1**).

Tabla 1. *Elementos clave*

Elemento	Descripción
Tecnología como medio de control	La tecnología puede ser utilizada para mantener el control social y político.
Determinismo tecnológico versus constructivismo social	Se debate si la tecnología determina el curso de la sociedad (determinismo tecnológico) o si las decisiones humanas y las estructuras sociales moldean la tecnología (constructivismo social).
Impacto en la igualdad social	Se examina cómo la tecnología puede perpetuar o exacerbar desigualdades sociales y económicas, afectando desproporcionadamente a ciertos grupos.
Racionalidad técnica	La creencia en la capacidad de la tecnología para resolver problemas puede llevar a ignorar cuestiones éticas y sociales más amplias.
Sostenibilidad y ética	Se plantea la necesidad de considerar el impacto ambiental y social de la tecnología, promoviendo un enfoque más sostenible y ético en su desarrollo y aplicación.

Fuente: Elaboración propia basada en Feenberg (2008, 2017), Mohamed e Isacc (2020).

En el contexto de la IA y la sustentabilidad, esta teoría sugiere que las soluciones tecnológicas podrían estar moldeadas por las prioridades de las élites económicas y políticas, en lugar de las necesidades de las comunidades afectadas por la crisis climática.

Debido a que las élites económicas y políticas a menudo tienen más poder y recursos para dar forma a las normas y soluciones tecnológicas. Esto puede resultar en un enfoque en proyectos que maximizan

el retorno de inversión o que benefician a sectores específicos de la economía, en lugar de abordar las necesidades urgentes de comunidades vulnerables.

Como consecuencia, las decisiones tomadas por las élites suelen ser desconectadas de las realidades locales. Las comunidades afectadas por el cambio climático tienen conocimientos y perspectivas valiosos sobre los desafíos que enfrentan, pero a menudo no tienen voz en la creación de soluciones. Esto puede resultar en tecnologías que no son adecuadas o que no se implementan de manera efectiva en el contexto local.

Partiendo de lo anterior, se puede realizar una reflexión crítica sobre si la IA está siendo utilizada para promover una agenda verdaderamente sostenible o si, en cambio, refuerza estructuras de poder existentes.

En el mismo sentido, el marxismo tecnológico complementa esta visión crítica, al señalar cómo la IA puede ser utilizada como herramienta de explotación, beneficiando a las clases dominantes mientras que incrementa la desigualdad. Este enfoque permite analizar cómo las tecnologías que supuestamente buscan la sustentabilidad pueden, en realidad, estar perpetuando las desigualdades económicas y sociales, a medida que las corporaciones que desarrollan estas tecnologías controlan los datos, los recursos y el conocimiento que sustentan su funcionamiento.

LA IA EN LA GOBERNANZA COLABORATIVA: UNA MIRADA DESDE LOS BIENES COMUNES

La teoría de los bienes comunes se presenta como un marco conceptual esencial para comprender cómo los recursos compartidos pueden ser gestionados de manera eficiente y sostenible por las comunidades que dependen de ellos (Ostrom, 2017).

Esta teoría, desarrollada y consolidada por Ostrom (1990), ofrece una perspectiva alternativa a las propuestas tradicionales de gestión de recursos, que oscilan entre la privatización y la regulación estatal centralizada.

Los bienes comunes, también denominados "recursos comunes", son aquellos recursos naturales o sociales que están disponibles para un grupo definido de usuarios y cuyo uso es difícil de regular, lo que potencialmente conduce a su sobreexplotación (Ostrom, 2009).

Estos bienes se caracterizan por ser rivalizantes —el uso por parte de un individuo reduce la disponibilidad para otros— y no excluyentes —es costoso o difícil restringir el acceso de ciertos individuos a dichos recursos (Ostrom, 1990).

Ostrom (1990) identificó una serie de principios fundamentales que caracterizan a las comunidades que han gestionado exitosamente sus bienes comunes. Estos principios de diseño son la base para entender cómo la gobernanza comunitaria puede funcionar en la práctica (Van Laerhoven y Ostrom, 2007):

a) Límites claros: Es crucial que tanto los límites del recurso como los de la comunidad de usuarios estén claramente definidos. Esto evita la sobreexplotación por actores externos y facilita la implementación de reglas de uso.

b) Reglas adaptadas a las condiciones locales: Las normas que gobiernan el uso de los bienes comunes deben estar adecuadamente ajustadas a las características del recurso y a las necesidades de la comunidad local, en lugar de ser impuestas externamente.

c) Participación en la toma de decisiones: Las personas que utilizan el recurso común deben tener la oportunidad de participar en la creación y modificación de las reglas que lo regulan, asegurando así su legitimidad y aceptación.

d) Monitoreo comunitario: Los usuarios o personas designadas por la comunidad deben encargarse de monitorear el uso del recurso para garantizar que las reglas se cumplan.

e) Sanciones graduales: Las sanciones por incumplimiento de las reglas deben ser proporcionales y aplicarse de manera gradual, comenzando por sanciones leves para faltas menores.

f) Resolución de conflictos accesible: Deben existir mecanismos locales y accesibles para resolver disputas entre los usuarios

de los recursos, evitando la escalada de conflictos que puedan amenazar la sostenibilidad del recurso.

g) Reconocimiento de los derechos de organización: Las comunidades deben tener la autonomía para organizarse y establecer sus propias reglas de gestión sin ser socavadas por actores externos, como gobiernos o corporaciones.

h) Policentrismo: La gobernanza debe ocurrir en múltiples niveles, permitiendo una coordinación flexible entre diferentes actores y escalas, desde lo local hasta lo global.

Por lo anterior, La teoría de los bienes comunes de Elinor Ostrom ofrece un marco útil para comprender cómo los recursos naturales compartidos, como el agua, los bosques o la atmósfera, pueden ser gestionados colectivamente.

En el contexto de la IA, este enfoque sugiere que las tecnologías deben ser utilizadas para facilitar la gobernanza colaborativa de estos bienes comunes, en lugar de permitir que grandes actores tecnológicos monopolizen su control. La IA puede ser una herramienta poderosa para mejorar la gestión de los bienes comunes, pero solo si se integra en un marco de gobernanza que priorice la colaboración, la equidad y la justicia ambiental.

LA IA COMO UNA INNOVACIÓN RESPONSABLE: MÁS ALLÁ DE LA SUSTENTABILIDAD

La innovación responsable implica que los actores que participan en el proceso de innovación —científicos, ingenieros, gobiernos, empresas, y la sociedad civil— trabajen de manera colaborativa para anticipar los impactos de las nuevas tecnologías y productos, evaluarlos desde un punto de vista ético, y adaptarlos según las necesidades y valores de la sociedad bajo cuatro principios básicos **(Ver Tabla 2).**

Tabla 2. Principios

Concepto	Descripción
Anticipación	Implica prever y evaluar posibles consecuencias a largo plazo de una innovación, tanto positivas como negativas, incluyendo los efectos no deseados. La anticipación es una herramienta para identificar riesgos y oportunidades antes de que la tecnología esté completamente desarrollada o desplegada (Owen y Pansera, 2019).
Reflexividad	Se refiere a la capacidad de los innovadores para reflexionar sobre las motivaciones, valores y supuestos que subyacen a su trabajo. Los científicos y tecnólogos deben ser conscientes del contexto más amplio de su trabajo y estar abiertos a cuestionar sus propios supuestos (Owen y otros., 2013).
Inclusividad	La innovación debe incluir la participación de una amplia gama de actores, incluidos los grupos sociales afectados, los responsables políticos y las organizaciones de la sociedad civil. Este principio enfatiza la necesidad de un diálogo abierto y participativo durante todo el proceso de innovación, para garantizar que los intereses y valores de todos los sectores relevantes se tomen en cuenta (Van den Hoven y otros., 2014)
Capacidad de respuesta	Las innovaciones deben ser flexibles y estar abiertas al cambio en respuesta a nuevos conocimientos o preocupaciones sociales y éticas. Esto implica la capacidad de ajustar los objetivos, procesos y resultados de la innovación en función de los comentarios y la participación de los diversos actores involucrados (Koops y otros., 2015).

Fuente: Elaboración propia basada en Barlatier y otros., 2024.

Al integrar valores éticos y sociales, fomentar la participación de actores diversos, y anticipar los impactos de largo plazo, la innovación responsable se posiciona como un marco esencial para enfrentar los retos globales contemporáneos. En este marco es particularmente útil para evaluar la IA en la sustentabilidad, ya que destaca la necesidad de diseñar tecnologías que no solo resuelvan problemas técnicos, sino que también se anticipen a los riesgos éticos y sociales que puedan surgir.

La innovación responsable implica que la IA no debe implementarse sin un análisis profundo de sus impactos a largo plazo, asegurando que contribuya verdaderamente a un futuro más justo y sostenible (De

Hoop *y otros*., 2016; Owen *y otros*., 2013; Koops *y otros*., 2015; Owen y Pansera, 2019).

Por lo anterior expuesto, este marco teórico multidimensional ofrece una estructura coherente para analizar el dilema ético de la IA en la búsqueda de la sustentabilidad. Al integrar la ética filosófica, la teoría de la sustentabilidad, la gobernanza de la tecnología, la teoría crítica, la teoría de los bienes comunes y la innovación responsable, se logra una comprensión holística de cómo las tecnologías emergentes deben ser evaluadas no solo por su efectividad, sino también por su capacidad para promover la justicia, la equidad y el bienestar ambiental.

LA IA Y SU ROL EN LA BÚSQUEDA DE LA SUSTENTABILDAD: ANÁLISIS DE SU IMPACTO

Partiendo de las contribuciones de las perspectivas teóricas anteriores, el estudio de la IA y su impacto en el medioambiente puede abordarse desde una postura crítica.

Desde este enfoque, se comprende que el desarrollo de tecnologías de IA ha abierto nuevas posibilidades para abordar problemas complejos relacionados con la gestión de los recursos naturales, la eficiencia energética y la mitigación del cambio climático.

En sectores como la agricultura, la energía y el transporte, la IA ha demostrado ser una herramienta invaluable para mejorar la eficiencia y reducir los impactos ambientales.

La capacidad de analizar grandes cantidades de datos en tiempo real permite a los sistemas de IA optimizar procesos y tomar decisiones más informadas, lo que puede generar beneficios sustanciales en términos de ahorro de recursos y reducción de emisiones de carbono.

Los sistemas de IA, en combinación con tecnologías como drones y sensores IoT, permiten a los agricultores monitorear y optimizar el uso de agua, fertilizantes y pesticidas, reduciendo significativamente el desperdicio de recursos y minimizando los efectos ambientales adversos.

Al mismo tiempo, la IA ayuda a predecir patrones climáticos y mejorar la resistencia de los cultivos a las condiciones extremas derivadas del cambio climático

En el ámbito de la energía, la IA se está utilizando para mejorar la eficiencia de las redes eléctricas y promover la integración de fuentes renovables, como la energía solar y eólica.

Los sistemas de redes inteligentes o "*smart grids*" emplean algoritmos de IA para gestionar de manera eficiente la distribución de electricidad, equilibrando la oferta y la demanda en tiempo real y optimizando el almacenamiento energético.

Esto es crucial para facilitar la transición hacia un sistema energético basado en energías renovables, que suelen tener una producción intermitente. La IA también ha encontrado aplicaciones en la gestión de residuos y el reciclaje, áreas críticas para una economía circular y sostenible.

Los sistemas automatizados que clasifican los residuos mediante visión por computadora pueden mejorar significativamente las tasas de reciclaje y reducir la contaminación ambiental. Además, los algoritmos de IA se utilizan para optimizar rutas de recolección de basura, minimizando el consumo de combustible y las emisiones de gases de efecto invernadero.

A pesar del potencial transformador de la IA, es fundamental adoptar una postura crítica y reflexiva sobre los riesgos y desafíos que esta tecnología plantea para la sustentabilidad. Uno de los principales problemas es que la IA, en su estado actual, puede generar externalidades ambientales significativas, que podrían contradecir los objetivos de sostenibilidad a largo plazo: impacto energético de los centros de datos, obsolescencia programada y consumo de recursos exclusión tecnológica y desigualdad, desplazamiento laboral y crisis social.

A MANERA DE REFLEXIÓN: IMPLICACIONES ÉTICAS Y MEDIOAMBIENTALES

Por un lado, la IA es vista como una tecnología habilitadora de soluciones innovadoras. Su capacidad para procesar grandes volúmenes de

datos y optimizar procesos complejos ha abierto nuevas posibilidades en áreas clave para la sustentabilidad, como la agricultura de precisión, la gestión de residuos y la transición hacia energías renovables.

No obstante, esta visión optimista de la IA debe ser matizada con una comprensión profunda de las externalidades negativas que genera. El consumo energético masivo requerido para entrenar y operar sistemas de IA, especialmente los modelos avanzados de aprendizaje profundo, es una de las externalidades más preocupantes.

A medida que la IA se expande, la necesidad de infraestructura computacional, como centros de datos y dispositivos especializados, genera una creciente demanda de energía que, si no es gestionada mediante fuentes renovables, puede contribuir de manera significativa al cambio climático. Este es un dilema que ilustra claramente la tensión entre los beneficios inmediatos de la IA y sus costos ocultos para la sustentabilidad a largo plazo.

Además, la obsolescencia tecnológica y el consumo de recursos finitos para fabricar la infraestructura necesaria para IA no solo genera impactos ambientales, sino que también plantea preocupaciones sobre la equidad intergeneracional. El agotamiento de recursos naturales y la generación de desechos electrónicos comprometen la capacidad de las generaciones futuras para disfrutar de los mismos beneficios tecnológicos sin asumir un costo ambiental desproporcionado.

Este es un desafío crítico que debe ser considerado desde una perspectiva ética. Para que la IA sea una herramienta efectiva en la búsqueda de la sustentabilidad, es imprescindible adoptar un enfoque ético que guíe su desarrollo y aplicación. Esto implica hacer de la justicia social, la equidad intergeneracional y la mitigación del impacto ambiental principios fundamentales en el diseño y uso de tecnologías de IA.

Un marco de innovación responsable exige que los desarrolladores de IA y las organizaciones que impulsan su implementación consideren no solo los beneficios económicos y tecnológicos, sino también las implicaciones sociales y ambientales. Esto significa que la IA debe diseñarse de manera eficiente en cuanto al consumo energético, utilizando fuentes de energía renovables y hardware más sostenible.

Además, es fundamental que los sistemas de IA sean accesibles para las comunidades más vulnerables, garantizando que los beneficios de estas tecnologías no se concentren exclusivamente en manos de una élite tecnológica, sino que sean distribuidos equitativamente en toda la sociedad.

La ética tecnológica también nos obliga a reflexionar sobre la transparencia y la rendición de cuentas en los sistemas de IA. En un contexto de sustentabilidad, los sistemas de IA que gestionan recursos naturales o predicen cambios climáticos deben ser auditables y transparentes, de manera que las decisiones tomadas por estas tecnologías puedan ser evaluadas y corregidas si es necesario.

La opacidad algorítmica representa un riesgo importante, ya que puede ocultar decisiones que afecten negativamente tanto al medio ambiente como a las comunidades humanas. A medida que la IA continúa avanzando, resulta imperativo que los esfuerzos por maximizar su impacto positivo sobre la sustentabilidad se realicen en un marco de gobernanza participativa y multidisciplinaria.

Los desarrollos tecnológicos no pueden quedar exclusivamente en manos de actores privados o técnicos, ya que la sustentabilidad requiere la integración de diversas disciplinas y perspectivas, incluyendo las ciencias sociales, el derecho, la ética y la política.

En este sentido, los gobiernos y organismos internacionales tienen un rol crucial en la creación de marcos regulatorios que limiten los riesgos ambientales y sociales de la IA, al tiempo que fomenten su uso para la mejora del bienestar humano y la protección del medio ambiente.

Las regulaciones deben establecer límites claros en cuanto al consumo energético de los sistemas de IA, incentivar el uso de energías limpias en su operación y asegurar que los desarrollos de IA estén alineados con los objetivos globales de desarrollo sostenible, como los establecidos por la *ONU*.

Asimismo, los ciudadanos y las comunidades locales deben tener voz y voto en las decisiones relacionadas con la implementación de IA en sus entornos. La IA, al igual que cualquier otra tecnología disruptiva, tiene impactos directos sobre la vida diaria, y es fundamental

que las personas tengan la oportunidad de participar en el diseño y la evaluación de estos sistemas.

La participación ciudadana es clave para garantizar que las aplicaciones de IA respondan a las necesidades reales de las comunidades y no a los intereses exclusivos de corporaciones o élites tecnológicas.

El impacto de la IA en la sustentabilidad plantea preguntas fundamentales sobre el futuro del desarrollo tecnológico y el modelo económico predominante. ¿Es posible diseñar una IA que promueva un crecimiento verdaderamente sostenible? ¿O estamos atrapados en una contradicción inherente donde la IA, a pesar de sus promesas, inevitablemente exacerba los problemas ambientales y sociales?

Estas preguntas no tienen respuestas fáciles, pero es evidente que cualquier visión optimista de la IA debe ir acompañada de una crítica rigurosa de sus posibles consecuencias negativas. La IA, como cualquier tecnología, no es neutral; está moldeada por las estructuras sociales, económicas y políticas en las que se desarrolla. Por lo tanto, el futuro de la IA y la sustentabilidad dependerá en gran medida de nuestra capacidad colectiva para reimaginar y rediseñar los sistemas tecnológicos y económicos de una manera que priorice el bienestar a largo plazo de las personas y el planeta.

En última instancia, el desafío es transformar la IA en una herramienta que no solo impulse la eficiencia y la innovación, sino que también promueva la equidad, la justicia y la preservación ambiental. Solo a través de un enfoque multidisciplinario, ético y responsable podremos asegurar que la IA no sea solo una tecnología del presente, sino una fuerza de cambio positivo para un futuro verdaderamente sustentable.

REFERENCIAS

Artículos científicos

Adams, C., Frost, G., y Webber, W. (2013). Triple bottom line: A review of the literature. The triple bottom line, 17-25.

Alhaddi, H. (2015). *Triple bottom line and sustainability: A literature review*. Business and Management studies, 1(2), 6-10.

Astobiza, A. M., Toboso, M., Aparicio, M., y López, D. (2021). AI ethics for sustainable development goals. *IEEE Technology and Society Magazine*, 40(2), 66-71. *https://doi.org/10.1109/MTS.2021.3056294*

Barlatier, P. J., Georget, V., Penin, J., y Rayna, T. (2024). The Origin, Robustness, and Future of Responsible Innovation. Journal of Innovation Economics & Management, 43(1), 1-38.

Batty, M. (2018). Artificial intelligence and smart cities. *Environment and Planning B: Urban Analytics and City Science*, 45(1), 3-6. *https://doi.org/10.1177/2399808317751169*.

Bauer, W. A. (2020). Virtuous vs. utilitarian artificial moral agents. AI & SOCIETY, 35(1), 263-271. *https://doi.org/10.1007/s00146-018-0871-3*

Bickley, S. J., Macintyre, A., y Torgler, B. (2024). Artificial intelligence and big data in sustainable entrepreneurship. *Journal of Economic Surveys*. *https://doi.org/10.1111/joes.12611*

Bloch-Wehba, H. (2022). Algorithmic governance from the bottom up. *BYU L. Rev.*, 48, 69.

Carrillo, M. R. (2020). Artificial intelligence: From ethics to law. *Telecommunications policy*, 44(6), 101937. *https://doi.org/10.1016/j.telpol.2020.101937*

Delanty, G., y Harris, N. (2021). Critical theory and the question of technology: The Frankfurt School revisited. Thesis Eleven, 166(1), 88-108. *https://doi.org/10.1177/07255136211002055*.

Dayarathna, M., Wen, Y., & Fan, R. (2015). Data center energy consumption modeling: A survey. IEEE Communications surveys & tutorials, 18(1), 732-794. *https://ieeexplore.ieee.org/abstract/document/7279063*

Dhabliya, D. (2024). Smart Grids Optimization for Energy Trading with AI Solutions. Acta Energetica, (02), 71-81. *https://www.actaenergetica.org/index.php/journal/article/view/518*

De Hoop, E., Pols, A., y Romijn, H. (2016). Limits to responsible innovation. Journal of Responsible Innovation, 3(2), 110-134.

Elshaikh, A., Mabrouki, J., y Mohamed, A. A. O. (2024). The Future of Water Management: Leveraging AI for Effective Decision Support. In Advancements in Climate and Smart Environment Technology (pp. 213-225). *IGI Global*. *https://doi.org/10.4018/979-8-3693-3807-0.ch015*

Feenberg, A. (2008). Critical theory of technology: An overview. *Information technology in librarianship: New critical approaches*, 31-46.

Feenberg, A. (2017). Critical theory of technology and STS. *Thesis eleven*, 138(1), 3-12.

Finlayson, J. G. (2009). Morality and critical theory: On the normative problem of Frankfurt School social criticism. Telos, 146(7), 7-41.

Gritsenko, D., y Wood, M. (2022). Algorithmic governance: A modes of governance approach. Regulation & Governance, 16(1), 45-62

Goralski, M. A., y Tan, T. K. (2020). Artificial intelligence and sustainable development. The International *Journal of Management Education*, 18(1), 100330. *https://doi.org/10.1016/j.ijme.2019.100330*

Issar, S., y Aneesh, A. (2022). What is algorithmic governance?. Sociology Compass, 16(1), e12955. *https://doi.org/10.1111/soc4.12955*

Katzenbach, C., y Ulbricht, L. (2019). Algorithmic governance. Internet Policy Review, 8(4), 1-18. *doi:10.14763/2019.4.1424*

Latour, B., y Venn, C. (2002). Morality and technology. Theory, culture & society, 19(5-6), 247-260. *https://doi.org/10.1177/026327602761899246*

Mohamed, S., Png, M. T., e Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy & Technology, 33, 659-684.

Mora, C., Rollins, R. L., Taladay, K., Kantar, M. B., Chock, M. K., Shimada, M., y Franklin, E. C. (2018). Bitcoin emissions alone could push global warming above 2°C. Nature Climate Change, 8(11), 931-933. *https://doi.org/10.1038/s41558-018-0321-8*

Nishant, R., Kennedy, M., y Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, 53, 102104. *https://doi.org/10.1016/j.ijinfomgt.2020.102104*

Nurhaeni, H., Delhi, A., Daeli, O. P. M., Anjani, S. A., & Yusuf, N. A. (2024). Optimizing electrical energy use through ai: An integrated approach for efficiency and sustainability. *International Transactions on Artificial Intelligence*, 2(2), 106-113. *https://doi.org/10.33050/italic.v2i2.533*

Raval, M., Bhardwaj, S., Aravelli, A., Dofe, J., y Gohel, H. (2021). Smart energy optimization for massive IoT using artificial intelligence. *Internet of Things*, 13, 100354. *https://doi.org/10.1016/j.iot.2020.100354*

Sah, B. P., Begum, S., Bhuiyan, M. R., y Shahjalal, M. (2024). The Role Of Ai In Promoting Sustainability Within The Manufacturing Supply Chain Achieving Lean And Green Objectives. Academic Journal on Business Administration, Innovation & Sustainability, 4(3), 79-93. *https://doi.org/10.69593/ajbais.v4i3.97*

Saleem, N., y AMU, A. Rawls' Theory Of Justice. *IJCRT*, 9, 1625-30

Slaper, T. F., y Hall, T. J. (2011). T*he triple bottom line: What is it and how does it work*. Indiana business review, 86(1), 4-8.

Udayakumar, L., Babu, V. V. S. S., y Babu, V. V. S. S. (2021). Immanuel Kant's Deontology Theory. IJRAR-International Journal of Research and Analytical Reviews (IJRAR), 8(2), 235-243. *www.ijrar.org/papers/21B2505_252105.pdf*

Van den Hoven, J., Doorn, N., Swierstra, T., Koops, B. J., y Romijn, H. (2014). Responsible innovation. In Third international conference on responsible innovation (Vol. 22).

Van Laerhoven, F., y Ostrom, E. (2007). Traditions and Trends in the Study of the Commons. International Journal of the Commons, 1(1), 3-28. *https://www.jstor.org/stable/26522979*

Van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. AI and Ethics, 1(3), 213-218. *https://doi.org/10.1007/s43681-021-00043-6*

Tseng, P. E., y Wang, Y. H. (2021). Deontological or utilitarian? An eternal ethical dilemma in outbreak. International journal of environmental research and public health, 18(16), 8565.

Yang, Z., Du, J., Lin, Y., Du, Z., Xia, L., Zhao, Q., & Guan, X. (2022). Increasing the energy efficiency of a data center based on machine learning. Journal of Industrial Ecology, 26(1), 323-335. *https://doi.org/10.1111/jiec.13155*.

Yang, J., Xiao, W., Jiang, C., Hossain, M. S., Muhammad, G., y Amin, S. U. (2018). Ai-powered green cloud and data center. IEEE Access, 7, 4195-4203. *https://doi.org/10.1109/ACCESS.2018.2888976*

Libros

Corradetti, C. (2012). The Frankfurt School and critical theory. The internet encyclopedia of philosophy.

Freeman, S.(2017). *Rawls*. New York: Routledge.

Ian, M. (1991). *The Brundtland report and sustainable development in New Zealand*. Lincoln University and University of Canterbury. Centre for Resource Management.

Koops, B. J., Oosterlaken, I., Romijn, H., Swierstra, T., y Van den Hoven, J. (2015). Responsible innovation 2. New York: Springer.

Mill, J. S. (2016). *Utilitarianism. In Seven masterpieces of philosophy*. Routledge.

Mile, A.(2017). *An Explanation of John Rawls's Theory of Justice with a Defense of the Veil of Ignorance*. Claremont McKenna College.

Hare, R. M. (2014). *Ethical theory and utilitarianism*. In Contemporary British Philosophy (pp. 113-131). Routledge.

Hare, R. M. (2020). *Why do applied ethics? In New Directions in Ethics*. Routledge.

Kranak, J. (2019). *Kantian deontology*. Introduction to philosophy: Ethics.

Lindgren, S. (Ed.). (2023). Handbook of critical studies of artificial intelligence. Edward Elgar Publishing.

Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action. Cambridge University.

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. Science, 325(5939), 419-422.

Ostrom, E. (2017). Polycentric systems for coping with collective action and global environmental change. In Global justice (pp. 423-430). Routledge.

Owen, R. J., Bessant, J. R., y Heintz, M. (Eds.). (2013). Responsible innovation (Vol. 104). Chichester: Wiley.

Owen, R., y Pansera, M. (2019). Responsible innovation and responsible research and innovation (pp. 26-48). Edward Elgar Publishing.

Rawls, J. (1971). *A theory of justice*. Cambridge (Mass.).

Rawls, J. (2000). *Lectures on the History of Moral Philosophy*, Cambridge: Harvard University Press.

Scarre, G. (2020). *Utilitarianism*. Editorial Routledge.

Smart, J. J. C. (2020). *Utilitarianism and its applications. In New directions in Ethics*. Editorial Routledge.

Sitio web

GitHub. (2021). GitHub Copilot [Herramienta de programación asistida por IA]. GitHub. (2024, 20 de septiembre). *https://github.com/features/copilot*

OpenAI. (2023). "Introducing GPT-4." OpenAI. (2024, 20 de septiembre). *https://openai.com/index/gpt-4/*

CAPÍTULO XIV

LA IA ES UNA IDEOLOGÍA, NO UNA TECNOLOGÍA

MARITZA ALESSANDRA ZEPEDA BARBOSA[1]

RESUMEN

Este capítulo aborda la necesidad de reorientar el enfoque de la inteligencia artificial (IA) hacia las contribuciones humanas que la sustentan. Aunque la IA puede lograr avances significativos en campos como la medicina y el procesamiento de imágenes, la narrativa predominante tiende a desviar la atención de las personas que aportan datos y conocimientos esenciales. Se sostiene que esta forma de pensar puede llevar a una desconexión de la responsabilidad humana y a la sobreestimación de la autonomía de las máquinas.

Se argumenta que, en lugar de ver la IA como un ente independiente, deberíamos reconocer el papel fundamental de los humanos en su desarrollo y funcionamiento. Se sugiere que al reconocer y compensar adecuadamente a los trabajadores por su participación, se puede aumentar la productividad y mejorar la economía en general.

El fenómeno de los trabajadores detrás de la IA pone de relieve la necesidad de establecer políticas que protejan y capaciten a estos trabajadores frente a los cambios impulsados por la automatización. Se analiza la evidencia de sesgos raciales presentes en la IA. Este segmento enfatiza la importancia de desarrollar herramientas que auditen y mitiguen estos sesgos, y de que tanto los ingenieros como los líderes del sector asuman la responsabilidad de crear soluciones inclusivas.

[1] Alumna de la Licenciatura en Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

Palabras Clave: IA generativa, equidad, ética.

INTRODUCCIÓN

La inteligencia artificial (IA) ha emergido como una de las fuerzas transformadoras más poderosas de nuestro tiempo, prometiendo revolucionar industrias enteras y alterar la manera en que interactuamos con el mundo. Sin embargo, es crucial reconocer que la IA va más allá de ser meramente una herramienta técnica; se ha convertido en un fenómeno ideológico que refleja y amplifica las creencias, valores y asunciones de la sociedad contemporánea. Esta ideología se manifiesta en las decisiones sobre cómo se diseñan, desarrollan y aplican las tecnologías de IA, así como en las narrativas que las rodean. En lugar de considerarla únicamente un conjunto de algoritmos y datos, es fundamental examinar cómo la IA se inscribe en un contexto más amplio de poder, control y desigualdad.

A medida que las aplicaciones de IA se integran en diversos sectores, desde la salud hasta la educación y la justicia, sus implicaciones éticas y sociales se vuelven cada vez más evidentes. La forma en que se recolectan y utilizan los datos, las decisiones automatizadas que se toman en función de esos datos y el impacto de estas decisiones sobre diferentes grupos de personas revelan una realidad inquietante: la IA puede perpetuar sesgos y desigualdades existentes si no se aborda con una crítica adecuada. Este capítulo se adentra en la premisa de que la IA no es solo una tecnología avanzada, sino que también es una ideología que necesita ser analizada y cuestionada. Se hace hincapié en la importancia de considerar quiénes están diseñando estos sistemas, qué intereses están en juego y cómo las decisiones en torno a la IA afectan a las comunidades, especialmente a las más vulnerables.

La IA no solo transforma sectores económicos y procesos industriales; también regenera nuestras interacciones cotidianas y las expectativas que tenemos sobre la toma de decisiones. En este contexto, es esencial reconocer que las decisiones tomadas por sistemas de IA no son neutras; están influenciadas por los sesgos y limitaciones de

los datos que alimentan estos algoritmos. De hecho, la forma en que se recolectan, gestionan y aplican estos datos puede tener profundas repercusiones en la vida de las personas, especialmente en aquellos grupos históricamente marginados.

INTELIGENCIA ARTIFICIAL: UNA CREENCIA PELIGROSA QUE NO RECONOCE A LOS HUMANOS

La IA puede estar logrando resultados sin precedentes en diversos campos, como la medicina, el control robótico y el procesamiento de lenguaje e imágenes, o puede que se esté utilizando una determinada forma de hablar del software como una forma de no celebrar plenamente a las personas que trabajan juntas para mejorar los sistemas de información y que están logrando esos resultados. La "*IA*" puede ser una amenaza para el futuro humano, como se suele imaginar en la ciencia ficción, o puede ser una forma de pensar en la tecnología que dificulta el diseño de tecnología para que pueda usarse de manera eficaz y responsable. La idea misma de la IA puede crear una distracción que facilite a un pequeño grupo de tecnólogos e inversores reclamar todas las recompensas de un esfuerzo ampliamente distribuido. La computación es una tecnología esencial, pero la forma en que la IA piensa en ella puede ser turbia y disfuncional.

Como ideología se fundamenta en la idea de que la automatización y los algoritmos son inherentemente superiores a las decisiones humanas. Desde el reconocimiento facial hasta el procesamiento de datos en tiempo real, se plantea la narrativa de que las máquinas, a través de la inteligencia artificial, pueden manejar tareas complejas de manera más eficiente y precisa que los humanos. En este sentido, se promueve la creencia de que el futuro de la humanidad depende en gran medida de la adopción masiva de estas tecnologías. Sin embargo, lo que a menudo se omite en este relato es el rol de las personas detrás de la IA, quienes diseñan, entrenan y ajustan estos sistemas. Al reducir el papel humano, la ideología de la IA alimenta una dependencia casi ciega en los sistemas automatizados.

Uno de los principales problemas de esta ideología es que tiende a deshumanizar el trabajo y la creatividad. Al presentar la IA como una solución eficiente y rentable, se ignora el valor intrínseco del trabajo humano. En lugar de promover el desarrollo de habilidades críticas como el pensamiento creativo, la toma de decisiones complejas y la empatía, se incentiva un enfoque en la automatización de tareas que antes dependían de la intervención humana. En sectores como la manufactura, los servicios financieros y la salud, la promesa de la IA ha sido mejorar la productividad y reducir costos. Sin embargo, este enfoque deja de lado la importancia de la interacción humana y el juicio moral en la resolución de problemas.

Una alternativa clara a la "*IA*" es centrarse en las personas presentes en el sistema. Si un programa es capaz de distinguir entre gatos y perros, no hablemos de cómo una máquina está aprendiendo a ver. En lugar de eso, hablemos de cómo las personas aportaron ejemplos para definir las cualidades visuales que distinguen a los "gatos" de los "perros" de manera rigurosa por primera vez. Siempre hay una segunda forma de concebir cualquier situación en la que se pretenda utilizar la IA. Esto es importante, porque la forma de pensar de la IA puede distraer de la responsabilidad de los humanos.

Una consecuencia crítica de la ideologización de la IA es el desplazamiento de la responsabilidad. A medida que se automatizan más decisiones, especialmente en áreas sensibles como la justicia, la salud y las finanzas, se difumina la línea de quién es responsable cuando algo sale mal. La IA permite que las corporaciones y los gobiernos externalizan la responsabilidad moral y ética a las máquinas. Este fenómeno crea una cultura de impunidad, donde los errores o abusos se justifican bajo el pretexto de que fue "*el algoritmo*" quien tomó la decisión, cuando en realidad fueron los humanos quienes lo diseñaron y aplicaron. (McKinsey Global Institute, 2020).[2]

La IA puede estar logrando resultados sin precedentes en diversos campos, como la medicina, el control robótico y el procesamiento de

[2] McKinsey también aborda los desafíos éticos relacionados con la IA, incluyendo la necesidad de abordar sesgos en los algoritmos, asegurar la transparencia en la toma de decisiones automatizada y garantizar la protección de datos.

lenguaje e imágenes, o puede que se esté utilizando una determinada forma de hablar del software como una forma de no celebrar plenamente a las personas que trabajan juntas para mejorar los sistemas de información y que están logrando esos resultados. La "IA" puede ser una amenaza para el futuro humano, como se suele imaginar en la ciencia ficción, o puede ser una forma de pensar en la tecnología que dificulta el diseño de tecnología para que pueda usarse de manera eficaz y responsable. (McKinsey Global Institute, 2020).[3] La idea misma de la IA puede crear una distracción que facilite a un pequeño grupo de tecnólogos e inversores reclamar todas las recompensas de un esfuerzo ampliamente distribuido. La computación es una tecnología esencial, pero la forma en que la IA piensa en ella puede ser turbia y disfuncional.

Se puede rechazar la forma de pensar de la IA por diversas razones. Una de ellas es que se considera que las personas ocupan un lugar especial en el mundo y que son la fuente última de valor de la que dependen en última instancia las IA (lo que podría llamarse una objeción humanista). Otra visión es que ninguna inteligencia, humana o máquina, es verdaderamente autónoma: todo lo que logramos depende del contexto social establecido por otros seres humanos que dan sentido a lo que deseamos lograr. Independientemente de cómo se lo vea, una comprensión de la IA centrada en la independencia de los humanos (en lugar de la interdependencia con ellos) pasa por alto la mayor parte del potencial de la tecnología del software.

En última instancia, lo que está en juego con la ideologización de la IA es el control sobre el futuro. Si dejamos que la IA sea vista únicamente como una herramienta de eficiencia y crecimiento económico, corremos el riesgo de perder de vista su impacto en la vida humana y las estructuras sociales. En lugar de permitir que un pequeño grupo de actores controle la narrativa y las decisiones sobre el uso de la IA, es esencial abrir el debate a una conversación más amplia y democrática. La IA no debe ser un fin en sí misma, sino una herramienta para me-

[3] La inteligencia artificial (IA) ha demostrado un potencial significativo en áreas como la medicina, donde puede asistir en diagnósticos y tratamientos, así como en el control robótico, optimizando procesos en manufactura y logística.

jorar la vida humana, siempre que se use de manera ética, justa y equitativa. (New York Times. (2021). *Artificial Intelligence and Google's Bias*.[4]

Este enfoque de la tecnología centrado en el ser humano no es una mera posibilidad teórica. Decenas de millones de personas utilizan cada día la videoconferencia para ofrecer servicios personales, como instrucción de idiomas y habilidades, en línea. Los espacios de colaboración virtual en línea como GitHub son fundamentales para la creación de valor en nuestra era. La realidad virtual y aumentada ofrecen la posibilidad de aumentar drásticamente lo que es posible, permitiendo que se realicen más tipos de trabajo colaborativo a grandes distancias. El software de productividad, desde *Slack* hasta *Wikipedia*, pasando por *LinkedIn* y las suites de productos de *Microsoft*, hacen que la colaboración en tiempo real, antes inimaginable, sea omnipresente.

De hecho, investigaciones recientes han demostrado que sin la *Wikipedia* creada por humanos, el valor de los motores de búsqueda se desplomaría (ya que es allí donde suelen encontrarse los primeros resultados de búsquedas importantes), a pesar de que los servicios de búsqueda se promocionan como ejemplos de primera línea del valor de la IA. (Y, sin embargo, *Wikipedia* es una organización sin fines de lucro en ruinas, mientras que los motores de búsqueda son algunos de los activos más valorados de nuestra civilización). Las tecnologías de colaboración nos están ayudando a trabajar desde casa durante la epidemia de Covid-19; se ha convertido en una cuestión de supervivencia, y el futuro promete formas en que la colaboración a larga distancia puede volverse cada vez más vívida y satisfactoria.

Para introducir un ejemplo, el muy celebrado algoritmo de generación de texto de Open AI se entrenó en millones de sitios web producidos por humanos. Y la evidencia del campo de la enseñanza automática sugiere cada vez más que cuando los humanos que generan los

[4] Se hace hincapié en que los algoritmos de IA están entrenados con datos, y si estos datos contienen sesgos históricos, la IA puede perpetuar injusticias. Este riesgo de sesgo destaca la necesidad de una cuidadosa curación y supervisión de los datos utilizados en el entrenamiento de los sistemas de IA para garantizar que se minimicen los efectos negativos en las comunidades desfavorecidas.

datos participan activamente en la provisión de información de alta calidad y cuidadosamente seleccionada, pueden entrenarse a costos mucho más bajos. Pero la participación activa solo es posible si, a diferencia de la actitud habitual de la IA, todos los contribuyentes, no solo los ingenieros de élite, son considerados actores cruciales y reciben una compensación económica. (Brynjolfsson 2014).[5]

El núcleo de la ideología es que un conjunto de tecnologías, diseñadas por una pequeña élite técnica, puede y debe volverse autónoma y, en última instancia, reemplazar, en lugar de complementar, no solo a los seres humanos individuales, sino a gran parte de la humanidad. Dado que cualquier reemplazo de ese tipo es un espejismo, esta ideología tiene fuertes resonancias con otras ideologías históricas, como la tecnocracia y las formas de socialismo basadas en la planificación central, que consideraban deseable o inevitable el reemplazo de la mayor parte del juicio y la agencia humanos por sistemas creados por una pequeña élite técnica.

EL ROL INVISIBLE: LOS TRABAJADORES DETRÁS DE LA IA

La revolución de la inteligencia artificial (IA) no es solo un fenómeno tecnológico o económico, sino un cambio que afecta directamente a los trabajadores en la primera línea de su implementación. Mientras que la narrativa predominante en torno a la IA suele enfocarse en los avances técnicos y los potenciales beneficios para la productividad, hay un grupo de personas que desempeñan un papel fundamental y, sin embargo, a menudo invisibilizado: los trabajadores humanos que hacen posible el funcionamiento y avance de la IA.

En la práctica, muchos de los avances de la IA dependen de una labor intensiva de etiquetado de datos. Para que un algoritmo aprenda

[5] Como señalan Brynjolfsson y McAfee (2014), para que este enfoque sea viable, es fundamental que se reconozca la importancia de todos los contribuyentes, no solo de los ingenieros y desarrolladores de élite, y que se les ofrezca una compensación adecuada por su participación. Este enfoque no sólo enriquecerá la calidad de los datos utilizados, sino que también fomentaría un ecosistema más equitativo y sostenible en el desarrollo de tecnologías de IA.

a distinguir entre objetos o comportamientos, necesita ser entrenado con millones de ejemplos. En este punto es donde intervienen los trabajadores que etiquetan, clasifican y organizan esos datos. A menudo subcontratados a través de plataformas de trabajo digital, estos trabajadores realizan un trabajo repetitivo que es fundamental para la creación de sistemas inteligentes. Sin embargo, a pesar de la importancia de su labor, suelen recibir salarios bajos y trabajar en condiciones precarias, sin el reconocimiento adecuado por sus contribuciones.

Estos trabajadores se encuentran en la primera línea de la revolución de la IA porque están íntimamente conectados con los procesos que permiten que los algoritmos funcionen, pero son vistos como fácilmente reemplazables o irrelevantes en el gran esquema de la tecnología. Esto refleja una tendencia en la que la automatización y la tecnología tienden a concentrar la riqueza y el poder en manos de unas pocas empresas tecnológicas, mientras que los trabajadores que apoyan la infraestructura de la IA son marginados o empujados hacia empleos de bajo salario. (Susskind y Susskind, 2015).[6]

Un ejemplo claro de esto es la moderación de contenido en plataformas de redes sociales. Los sistemas de IA que supuestamente moderan contenido dañino, violento o inapropiado dependen en gran medida de los humanos que verifican y toman decisiones sobre qué contenidos violan las reglas. Estos moderadores, a menudo trabajando en condiciones psicológicamente difíciles, están expuestos a contenidos perturbadores y su labor es vital para que la IA se ajuste a las expectativas de seguridad y moderación de las plataformas digitales.

A medida que la IA avanza, también surgen nuevos desafíos para los trabajadores en sectores tradicionales, donde la automatización está reemplazando empleos. Trabajadores de fábricas, conductores, empleados en centros de atención al cliente y otros roles de la economía de servicios se ven amenazados por el creciente uso de la IA para realizar sus tareas de manera más rápida y eficiente. Sin embargo, esta

[6] Susskind y Susskind (2015) argumentan que esta dinámica no solo afecta a los trabajadores individuales, sino que también tiene implicaciones profundas para la estructura social y económica en su conjunto, exacerbando las desigualdades existentes y planteando desafíos éticos significativos en el desarrollo de tecnologías de IA.

transición no siempre significa que la tecnología es más eficiente, sino que puede ser más barata para las empresas, lo que deja a los trabajadores en una posición de vulnerabilidad frente a la automatización.

Sin embargo, es crucial no adoptar una visión fatalista sobre el impacto de la IA en los trabajadores. Hay oportunidades para que los trabajadores se beneficien de la tecnología, siempre y cuando se les incluya en los procesos de diseño y se les proporcione capacitación adecuada para interactuar con sistemas inteligentes. Una IA responsable debería ser capaz de empoderar a los trabajadores, aumentando su capacidad para realizar tareas más complejas y creativas en lugar de desplazarlos. (Susskind y Susskind, 2015). La clave está en cómo las empresas y gobiernos deciden implementar la IA y cómo los trabajadores son integrados en esa transición.

En este sentido, es esencial que los trabajadores en la primera línea de la IA reciban el reconocimiento que merecen y participen activamente en las decisiones sobre cómo se desarrollan y utilizan estas tecnologías. Las empresas tecnológicas y los responsables de políticas públicas deben enfocarse en crear mecanismos que permitan que los trabajadores se beneficien de la revolución de la IA, en lugar de ser víctimas de ella. Esto podría incluir políticas de redistribución de la riqueza generada por la automatización, acceso a programas de capacitación para adquirir nuevas habilidades y la creación de empleos que complementen las capacidades de la IA en lugar de competir directamente con ella.

Según "*Rest of World*" (2023) los trabajadores subcontratados de diferentes industrias y regiones, incluidos operadores de centros de llamadas en Manila, programadores en Lahore y diseñadores en El Cairo. Muchos están viendo cómo la IA generativa cambia la demanda de su trabajo y la estabilidad de sus ingresos. Pero mientras algunos se preparan para despidos o comisiones decrecientes, otros han adoptado herramientas de IA generativa en un esfuerzo por mantenerse a la vanguardia. Si la IA generativa representa un cambio tectónico en la forma en que trabajamos, los trabajadores subcontratados en el extranjero están en la línea de falla.[7]

[7] Este fenómeno sugiere que, si la IA generativa constituye un cambio tectónico en la forma

Rest of World realizó una encuesta a más de 31.000 empleados de los sectores de tecnología, finanzas y medios de comunicación a través de la plataforma Blind, una aplicación que permite a los empleados verificados publicar de forma anónima en foros del sector. La mayor parte de los encuestados trabajaban para importantes empresas tecnológicas de Silicon Valley en Estados Unidos, India, la UE y Canadá, entre ellas *Amazon, Microsoft, Salesforce y Meta*.[8]

A nivel mundial, el 76% de los encuestados afirmó que creía que la IA generativa permitiría a su empresa reducir costos y el 63% predijo que su empresa contrataría menos contratistas externos gracias a las herramientas de IA generativa. "*Rest of World*" (2023).

La revolución de la inteligencia artificial ha transformado diversos sectores, incluyendo el mundo del arte y la ilustración. Este cambio no solo ha afectado la forma en que los artistas crean y distribuyen su trabajo, sino también cómo perciben y responden a la competencia generada por las nuevas tecnologías.

SANTIAGO BAUTISTA GONZÁLEZ, ILUSTRADOR, GUADALAJARA, MÉXICO

Durante más de siete años, Santiago Bautista González trabajó a tiempo completo vendiendo sus ilustraciones de estilo caricatura, a través del mercado de trabajo freelance Fiverr. Sus ingresos, de alrededor de 1.500 dólares en un buen mes, cayeron un tercio el pasado enero. Febrero fue igualmente decepcionante.

En busca de una explicación, Bautista, de 31 años, leyó sobre la creciente popularidad del software de IA generativa visual. Descubrió que *Fiverr había añadido una sección* para artistas de IA. "Y me dije: 'Bueno, tal vez sea por esto'", le dijo *a Rest of World*. (Rest of World,

en que trabajamos, los trabajadores subcontratados en el extranjero están en una posición crítica, enfrentando tanto desafíos como oportunidades en un entorno laboral en rápida evolución.

8 Este dato subraya la concentración de talento y recursos en la industria tecnológica y su impacto global en el empleo y la innovación.

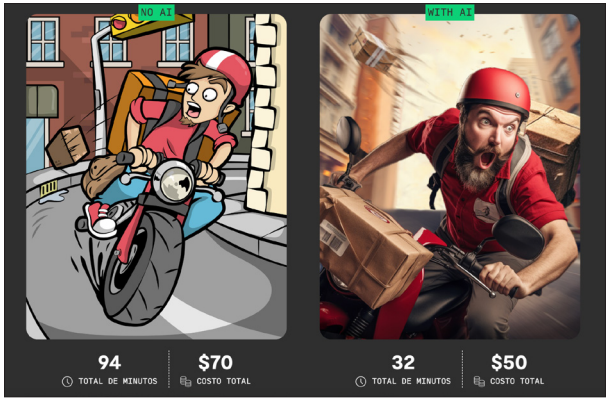
2023). Decidió seguir a sus clientes y eligió Midjourney como su herramienta preferida. Comenzó a ofrecer comisiones de IA en Fiverr al mismo precio base que sus ilustraciones dibujadas a mano habituales. "El primer día que puse en funcionamiento el nuevo servicio, conseguí una venta", dijo. "Fue así de atractivo".

Bautista pensó que dibujaría menos, que daría un descanso a sus manos y compensaría la pérdida de ingresos con ilustraciones de IA. Pero ocurrió algo curioso. El servicio de IA le dio tanta exposición, dijo Bautista, que sus pedidos de dibujos a mano también aumentaron en popularidad. "Debido a la inteligencia artificial, comencé a recibir demasiado trabajo".

No teme que el software de inteligencia artificial generativa le quite el trabajo. "*Aún se necesita algún tipo de creatividad para crear imágenes [con inteligencia artificial]*", dijo. "Y si el trabajo llega a un punto en el que la IA puede hacer todo y puede hacerlo más rápido, eso solo significará que trabajaré menos, ganaré más y tendré más tiempo para mis propios proyectos".

Rest of World encargó a Bautista que creará *dos ilustraciones de un repartidor en motocicleta que pierde un paquete mientras conduce por una ciudad.*

Figura 1. El impacto de la inteligencia artificial generativa en los artistas freelance.



Fuente: Rest of World. (2023).

Es probable que el impacto de la IA generativa sobre los trabajadores subcontratados sea más matizado que el de un desplazamiento masivo repentino. Por ahora, al menos, muchos investigadores laborales se resisten a la idea de que estas herramientas traigan consigo una "automatización total", definida a menudo como el uso de máquinas para completar tareas sin intervención humana. Más bien, el estado actual de la IA generativa, en general, cae dentro de lo que los expertos llaman "aumentación". La primera ola de adoptantes de la IA generativa trabaja en su mayoría junto con estas tecnologías, tratándolas como herramientas, no como sustitutos.[9]

> "*Incluso el desarrollo de la IA generativa requiere la intervención humana, un papel que a menudo también recae en trabajadores subcontratados en el extranjero.*" (Rest of World, 2023)

El reconocimiento del trabajo humano detrás de la IA también abre la puerta a una discusión más amplia sobre la ética en la tecnología. Es necesario que las empresas tecnológicas consideren cómo sus decisiones de automatización afectan a la fuerza laboral y al bienestar de las personas. La creación de tecnología no puede desligarse de las responsabilidades sociales y éticas, especialmente cuando los beneficios económicos de la IA están concentrados en unas pocas manos mientras que los costos se distribuyen entre muchos.

Esta dinámica resalta la urgencia de establecer políticas que no solo fomenten la innovación, sino que también protejan a los trabajadores más vulnerables. A medida que la IA se integra en diversas industrias, se hace evidente que una parte significativa de la fuerza laboral se encuentra en la línea de fuego, enfrentando la posibilidad de ser reemplazada por máquinas que operan con algoritmos avanzados. Por lo tanto, es esencial que los líderes de la industria y los responsables de políticas públicas colaboren para desarrollar marcos que aseguren la

---

[9] (Rest of World, 2023). La discusión sobre el impacto de la IA generativa en el trabajo subcontratado ha llevado a muchos investigadores laborales a enfatizar la distinción entre "automatización total" y "aumentación", donde la tecnología se utiliza como una herramienta para mejorar la eficiencia y la productividad humana en lugar de reemplazarla por completo.

capacitación y reintegración de estos trabajadores en nuevas funciones que complementen la tecnología.[10]

Asimismo, la narrativa sobre la IA generativa y su impacto en los artistas y diseñadores subraya la importancia de reconocer y valorar la creatividad humana. A pesar de los avances en la generación automática de contenido, la originalidad y la visión artística que solo los humanos pueden aportar siguen siendo invaluables. El desafío radica en encontrar un equilibrio entre aprovechar la eficiencia de la IA y mantener un espacio significativo para la expresión creativa.[11]

WU DAYUN, DISEÑADOR DE MARKETING, SHENZHEN, CHINA

Shaeman Liu, un fotógrafo de moda de 26 años de Hangzhou, dijo que duda que la IA vaya a reemplazar a toda la fotografía. Cree que los productos publicitados con modelos reales son más atractivos para los compradores. La modelo Charline Xu, de 24 años, dijo a *Rest of World* que actualmente gana unos 1.500 yuanes (210 dólares) por hora y que le preocupa que los clientes que utilicen IA generativa puedan reducir sus ingresos. Está preparada para competir. "Si [los diseñadores que utilizan IA] cobran 800 yuanes, cobraré 600. Si cobran 600, cobraré 500", bromeó Xu. "No hay otra salida. Lucharé hasta el final". (Rest of World, 2023)

El estudio de diseño de Wu Dayu, con sede en Shenzhen, solía tardar al menos una semana en crear materiales promocionales para tiendas de moda en línea. Una sesión fotográfica requería una modelo, un maquillador, un fotógrafo y un lugar. Producir un catálogo de imáge-

---

10 (Susskind & Susskind, 2015). La necesidad de políticas que promuevan tanto la innovación como la protección de los trabajadores vulnerables se vuelve crucial a medida que la IA transforma las dinámicas laborales, lo que exige una colaboración entre líderes industriales y responsables de políticas para desarrollar estrategias efectivas de capacitación y reintegración laboral.

11 (Brynjolfsson & McAfee, 2014). La evolución de la IA generativa resalta la necesidad de valorar la creatividad humana, dado que, a pesar de la eficiencia y capacidad de generación de contenido de estas tecnologías, la originalidad y la visión artística son cualidades irreemplazables que deben ser preservadas en el proceso creativo.

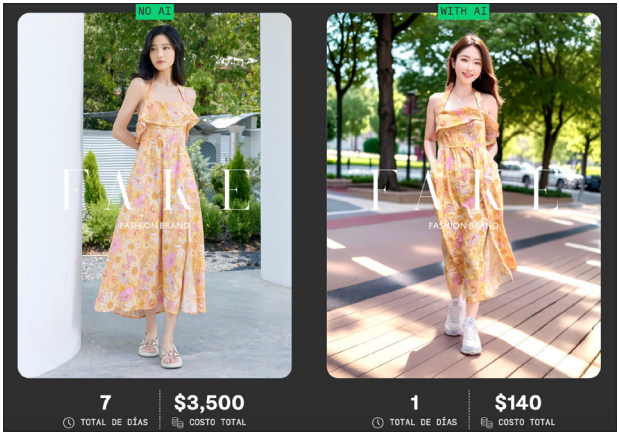
nes con seis conjuntos de ropa costaba alrededor de 3.500 dólares, dijo a *Rest of World* . Pero desde que Wu, de 35 años, empezó a utilizar IA generativa en marzo, el mismo trabajo se puede completar en un día, por solo dos personas y por solo 140 dólares.

Ahora, Wu utiliza un maniquí o le pide a un empleado que modele ropa para poder tomar algunas fotos básicas. Introduce las imágenes en un programa de inteligencia artificial con tecnología *Stable Diffusion* que ha personalizado para este propósito específico. Ajusta las imágenes del modelo y del fondo mediante indicaciones de texto, como "una niña, con zapatillas deportivas, fondo de calle, con árboles verdes a ambos lados de la calle". (Rest of World, 2023)

Este enfoque se está extendiendo cada vez más en la creciente industria del comercio electrónico de China, aunque algunos consumidores han criticado las imágenes resultantes por considerarlas falsas y sexistas . "*Algunas marcas de alta gama podrían preferir modelos humanos*", dijo Wu. "*Para los vendedores pequeños y medianos, los modelos de IA les ahorrarán mucho dinero y tiempo*".

Rest of World encargó a Wu que diseñara dos imágenes promocionales para una tienda online inventada. ***(Ver Figura 2)***

***Figura 2*:** Imagen promocional creada por Wu Dayun para una tienda online ficticia



Fuente: *Rest of World.* (2023).

Los trabajadores subcontratados suelen tener una mayor exposición a la tecnología debido a la naturaleza misma de su empleo. Como no forman parte de la estructura central de las empresas, a menudo se ven obligados a competir en mercados altamente dinámicos, donde la capacidad de adaptarse rápidamente a nuevas herramientas es crucial para mantenerse relevantes. La IA generativa ha sido una de las tecnologías más disruptivas en los últimos tiempos, ofreciendo soluciones avanzadas para la creación de contenido, automatización de tareas repetitivas y apoyo en la toma de decisiones.

A diferencia de los empleados fijos, los subcontratados enfrentan una presión constante para optimizar sus habilidades y eficiencia. Para muchos de ellos, la adopción de la IA generativa se ha convertido en un medio para mejorar su productividad y competir en un entorno laboral cada vez más globalizado. Un claro ejemplo de esto es el uso de herramientas de IA para generar contenido de marketing, redactar informes o crear visualizaciones complejas en cuestión de minutos, tareas que antes requerían mucho más tiempo y esfuerzo humano.

Además, la IA generativa permite a estos trabajadores realizar tareas que anteriormente requerían una especialización técnica avanzada. Por ejemplo, en áreas como el diseño gráfico o la programación, herramientas basadas en IA pueden generar código o propuestas de diseño con tan solo algunos insumos iniciales, permitiendo a los contratistas entregar resultados más rápido y con mayor precisión.[12] Esto no solo aumenta su competitividad en el mercado laboral, sino que también cambia la dinámica de su relación con los empleadores, quienes comienzan a depender más de estas tecnologías para reducir costos y aumentar la velocidad de producción.

Sin embargo, a pesar de los beneficios que la IA generativa ofrece, también plantea riesgos significativos para los trabajadores subcontratados. Uno de los mayores desafíos es la automatización de tareas que

[12] (Susskind & Susskind, 2015). La capacidad de la IA generativa para simplificar tareas técnicas complejas está transformando el mercado laboral, permitiendo a trabajadores menos especializados ser más competitivos y cambiando la forma en que los empleadores gestionan la producción y los costos, lo que genera una nueva dinámica en las relaciones laborales.

históricamente han sido realizadas por humanos. La misma tecnología que estos trabajadores adoptan para mejorar su productividad también puede ser utilizada por las empresas para reemplazarlos. Herramientas avanzadas de IA pueden ejecutar tareas con una eficiencia y precisión superiores, lo que podría llevar a una disminución en la demanda de trabajadores subcontratados en ciertos sectores.

El riesgo de desplazamiento laboral es especialmente alto en trabajos que involucran tareas repetitivas o creativas, como la generación de contenido escrito, traducción, atención al cliente y diseño gráfico. A medida que la IA se vuelve más capaz de realizar estas funciones con mínima intervención humana, los empleadores podrían optar por reducir su dependencia en los trabajadores subcontratados. En algunos casos, las empresas podrían decidir internalizar los procesos de automatización utilizando IA, eliminando por completo la necesidad de subcontratar ciertas tareas. (Susskind & Susskind, 2015).

Otro riesgo importante es la creciente competencia que los trabajadores subcontratados enfrentan a medida que la IA generativa permite a más individuos acceder a trabajos que antes requerían habilidades especializadas. Por ejemplo, una persona con acceso a herramientas de IA puede ahora realizar tareas que antes solo podían ser completadas por expertos en marketing digital, desarrollo de software o redacción. Esto podría llevar a una devaluación de las habilidades y un aumento en la oferta de trabajo, lo que presionará a los trabajadores a aceptar salarios más bajos para poder mantenerse competitivos.

DESAFIANDO EL RACISMO EN LA IA: ¿QUIÉN TOMA LA INICIATIVA?

Según el artículo del *New York Times* (2021), la inteligencia artificial puede perpetuar sesgos raciales, lo que plantea preocupaciones éticas sobre su uso. Hace unos seis años, la inteligencia artificial de un servicio de fotografía en línea de *Google* organizó fotos de personas negras en una carpeta llamada "*gorilas*". Hace cuatro años, una investigadora de una empresa emergente de Nueva York se dio cuenta de que el sistema de inteligencia artificial en el que estaba trabajando tenía un sesgo

flagrante contra las personas negras. Poco después, una investigadora negra de Boston descubrió que un sistema de inteligencia artificial no podía identificar su rostro hasta que se puso una máscara blanca.

La Dra. Gebru fue expulsada de la empresa sin una explicación clara. Dijo que la habían despedido después de criticar el enfoque de Google respecto de la contratación de minorías y, con un artículo de investigación , destacar los sesgos perjudiciales en los sistemas de inteligencia artificial que sustentan el motor de búsqueda de Google y otros servicios. (*The New York Times*, 2021)

> "*Tu vida empieza a empeorar cuando empiezas a defender a personas subrepresentadas", dijo la Dra. Gebru en un correo electrónico antes de su despido. "Empiezas a enfadar a los demás líderes*". (*The New York Times*, 2021)

Su partida se convirtió en un punto de discordia para los investigadores de IA y otros trabajadores del sector tecnológico.[13] Algunos vieron que una empresa gigante ya no estaba dispuesta a escuchar, demasiado ansiosa por sacar la tecnología al mercado sin considerar sus implicaciones. Yo vi un viejo problema, en parte tecnológico y en parte sociológico, que finalmente salía a la luz.

La inteligencia artificial (IA) ha revolucionado diversos sectores, desde la atención médica hasta el entretenimiento, ofreciendo soluciones innovadoras y eficientes. Sin embargo, a medida que su uso se ha extendido, también ha surgido una preocupación crítica: el potencial de la IA para perpetuar o incluso exacerbar sesgos raciales. A menudo, las máquinas que dependen de algoritmos de aprendizaje automático aprenden de los datos que se les proporcionan. Si estos datos contienen prejuicios históricos o estereotipos raciales, los resultados producidos por la IA pueden ser igualmente sesgados.

---

[13] New York Times, "¿Quién se asegura de que las máquinas de inteligencia artificial no sean racistas?" (2021). Este artículo destaca cómo la inteligencia artificial puede perpetuar sesgos raciales, evidenciado por incidentes en los que los sistemas de IA han categorizado erróneamente a personas negras, lo que plantea serias preocupaciones éticas sobre su implementación en diversas aplicaciones.

Esto plantea la pregunta: ¿quién es responsable de garantizar que las máquinas de IA no sean racistas?[14]

La IA se basa en grandes volúmenes de datos para "aprender" y hacer predicciones. Estos datos son recopilados de diversas fuentes, y en muchos casos, reflejan las desigualdades y prejuicios de la sociedad. Por ejemplo, un algoritmo diseñado para predecir la probabilidad de que una persona cometa un crimen puede estar entrenado con datos históricos de arrestos que han sido influenciados por políticas de vigilancia desproporcionadas en comunidades de color. Como resultado, el algoritmo puede perpetuar una imagen distorsionada de la criminalidad, afectando negativamente a grupos ya marginados.

La cuestión de quién es responsable de mitigar estos sesgos no es sencilla. En primer lugar, los desarrolladores de IA tienen un papel crucial. Ellos deben ser conscientes de los sesgos potenciales en los datos que utilizan y trabajar para identificar y eliminar estos prejuicios en la fase de diseño. Sin embargo, la responsabilidad no recae únicamente en los programadores; las empresas que implementan IA también deben tomar decisiones éticas sobre cómo se utilizan estos sistemas.

Los organismos reguladores y las instituciones gubernamentales desempeñan un papel vital al establecer normas y directrices para el desarrollo y uso de la IA. Estas regulaciones deben abordar explícitamente el problema de los sesgos raciales, asegurando que se realicen auditorías y revisiones regulares de los sistemas de IA. Sin un marco legal adecuado, las empresas pueden carecer de incentivos para abordar estos problemas de manera proactiva.

La inclusión de diversas voces en el desarrollo de la IA es fundamental. La falta de diversidad en los equipos que crean sistemas de IA puede dar lugar a la falta de consideración de perspectivas críticas que podrían identificar sesgos antes de que se conviertan en problemas. Es

[14] Según un artículo del New York Times (2021), la inteligencia artificial puede perpetuar sesgos raciales, lo que plantea preocupaciones éticas sobre su uso. Ejemplos como la categorización de imágenes de personas negras por parte de un servicio de fotografía de Google subrayan la urgencia de abordar la responsabilidad de quienes desarrollan y implementan estas tecnologías para garantizar que no reproduzcan ni amplifiquen prejuicios existentes.

esencial que las empresas tecnológicas incorporen a personas de diferentes orígenes, experiencias y disciplinas en el proceso de desarrollo.

La participación de comunidades afectadas es crucial. Las organizaciones de derechos civiles y otras entidades que representan a grupos minoritarios deben ser parte de la conversación en torno a la creación y regulación de la IA. Al incluir sus perspectivas, se pueden desarrollar soluciones más inclusivas y efectivas que reduzcan el riesgo de sesgos raciales.

A medida que crece la preocupación por el racismo en la IA, también han surgido herramientas y tecnologías diseñadas para mitigar estos problemas. Desde algoritmos de auditoría que pueden identificar sesgos en los datos hasta plataformas que ofrecen soluciones de IA explicativa, estas innovaciones están ayudando a abordar la problemática de manera más efectiva.[15] Sin embargo, la adopción de estas herramientas debe ser acompañada por un compromiso continuo de las empresas para evaluar y mejorar sus sistemas.

En 2017, Deborah Raji, una mujer negra de 21 años de Ottawa, estaba sentada en un escritorio dentro de las oficinas de Nueva York de *Clarifai*, la empresa emergente en la que trabajaba. La empresa había desarrollado una tecnología que podía reconocer automáticamente objetos en imágenes digitales y planeaba venderla a empresas, departamentos de policía y agencias gubernamentales. New York Times. (2021). *Artificial Intelligence and Google's Bias.*

Ella miró fijamente una pantalla llena de caras: imágenes que la compañía usaba para entrenar su software de reconocimiento facial. Mientras se desplazaba página tras página por las caras de los personajes, se dio cuenta de que la mayoría (más del 80 por ciento) eran de personas blancas. Más del 70 por ciento de esas personas blancas eran hombres. Cuando *Clarifai* entrenara su sistema con estos datos, podría hacer un buen trabajo al reconocer a las personas blancas, pensó Raji, pero fracasaría miserablemente con las personas de color y, probable-

[15] La creciente preocupación por el racismo en la inteligencia artificial ha impulsado el desarrollo de herramientas que buscan mitigar estos problemas, tales como algoritmos de auditoría para detectar sesgos y plataformas que ofrecen soluciones de IA explicativa.

mente, también con las mujeres. *New York Times*. (2021). *Artificial Intelligence and Google's Bias.*

Clarifai también estaba construyendo un "sistema de moderación de contenido", una herramienta que pudiera identificar y eliminar automáticamente la pornografía de las imágenes que la gente publicaba en las redes sociales. La empresa entrenó este sistema con dos conjuntos de datos: miles de fotos extraídas de sitios de pornografía en línea y miles de imágenes con clasificación G compradas en servicios de fotos de archivo.

Se suponía que el sistema debía aprender a distinguir entre lo pornográfico y lo anodino. El problema era que las imágenes aptas para todo público estaban dominadas por personas blancas, mientras que la pornografía no. El sistema estaba aprendiendo a identificar a las personas negras como pornográficas. "Los datos que utilizamos para entrenar estos sistemas son importantes", afirmó Raji. "No podemos elegir nuestras fuentes a ciegas". *New York Times*. (2021). *Artificial Intelligence and Google's Bias.*

Esto era obvio para ella, pero no lo era para el resto de la empresa. Como las personas que elegían los datos de entrenamiento eran en su mayoría hombres blancos, no se dieron cuenta de que sus datos estaban sesgados.

'PIEL NEGRA, MÁSCARAS BLANCAS'[16]

La situación de *Raji* en *Clarifai* no era un caso aislado. En la industria tecnológica en general, la diversidad de género y raza sigue siendo una preocupación. Sin una representación adecuada en los equipos de desarrollo, las tecnologías emergentes pueden reforzar estereotipos y

[16] El concepto "Piel negra, máscaras blancas" subraya cómo la discriminación racial se manifiesta en sistemas de inteligencia artificial, donde la representación inadecuada de las personas de color en los datos de entrenamiento puede llevar a resultados sesgados y perjudiciales. Esto resalta la urgencia de abordar los sesgos en los algoritmos de IA y la necesidad de crear sistemas que reflejen de manera equitativa la diversidad humana, en lugar de perpetuar estereotipos o invisibilizar a ciertos grupos raciales.

perpetuar desigualdades raciales. Esto no solo afecta la eficacia de los sistemas, sino que también tiene implicaciones éticas profundas.

La advertencia de *Raji* resuena en el contexto de una frase provocadora: "Piel negra, máscaras blancas". Esta expresión captura la esencia de un sistema en el que las identidades y experiencias de las personas de color son invisibilizadas o malinterpretadas por tecnologías que no reflejan su realidad. A medida que la IA continúa avanzando y se integra en aspectos críticos de la vida cotidiana, es imperativo que se reconozcan y aborden estas dinámicas de poder.

El trabajo de *Raji* destaca la urgencia de integrar un enfoque de diseño inclusivo en el desarrollo de tecnologías de IA. Esto no solo requiere una representación diversa en los equipos de trabajo, sino también un compromiso activo para auditar y corregir sesgos en los datos. El futuro de la inteligencia artificial depende de nuestra capacidad para cuestionar, desafiar y rediseñar cómo se crean y utilizan estas tecnologías. Solo así podremos asegurarnos de que la IA sirva para el bien común y no perpetúe injusticias históricas.

En este sentido, el relato de *Raji* es un llamado a la acción: un recordatorio de que la lucha por la equidad y la justicia en el ámbito tecnológico debe ser continua y proactiva. A medida que nos adentramos en una era dominada por la IA, es crucial que la diversidad y la inclusión sean pilares fundamentales en el diseño y la implementación de estos sistemas. *New York Times.* (2021). *Artificial Intelligence and Google's Bias.*

El hecho de que los algoritmos pueden catalogar erróneamente a las personas negras y otras minorías como "problemáticas" o "sospechosas" simplemente por su color de piel refleja un profundo sesgo que se ha filtrado en los sistemas que se supone deben mejorar la vida humana. La frase "Piel negra, máscaras blancas" simboliza cómo las estructuras de poder existentes pueden despojar a las personas de color de su identidad, al tiempo que enmascaran las raíces del racismo bajo el velo de la tecnología. Esta dinámica resalta la necesidad de un enfoque más inclusivo en el desarrollo de tecnología, donde las voces de todos los grupos raciales sean escuchadas y tenidas en cuenta en cada etapa del proceso de diseño y entrenamiento.

A medida que el debate sobre la ética en la inteligencia artificial avanza, es crucial que los diseñadores e ingenieros reconozcan su responsabilidad de mitigar el sesgo racial. Esto implica no solo diversificar los conjuntos de datos, sino también fomentar un entorno en el que las contribuciones de los grupos subrepresentados sean valoradas. Sin una acción consciente y deliberada, la IA puede perpetuar desigualdades existentes en lugar de actuar como un motor de cambio.

CONCLUSIÓN

La inteligencia artificial (IA) ha transformado radicalmente diversos sectores, desde la atención médica hasta la automatización industrial, y ha prometido aumentar la eficiencia y mejorar la calidad de vida. Sin embargo, este progreso no es posible sin el esfuerzo y la dedicación de una fuerza laboral invisible que opera detrás de las escenas. Los trabajadores que alimentan, ajustan y mejoran los sistemas de IA desempeñan un papel crucial en su desarrollo y funcionamiento, pero a menudo son ignorados o considerados irrelevantes en la narrativa dominante sobre la tecnología. Además, a medida que la IA se convierte en un componente integral de nuestras vidas, se ha puesto de manifiesto la necesidad de abordar el sesgo racial y la discriminación que pueden surgir de su implementación. Estos dos temas están interrelacionados y requieren una atención cuidadosa para garantizar que la IA se utilice de manera ética y justa.

Los trabajadores que contribuyen a la IA son a menudo vistos como meros recursos, en lugar de seres humanos con habilidades, experiencias y voces significativas. En el desarrollo de sistemas de IA, estos trabajadores desempeñan una variedad de funciones que son fundamentales para el éxito de las aplicaciones de inteligencia artificial. Desde la recopilación de datos hasta la curación de conjuntos de datos, pasando por la etiquetación y la calibración de algoritmos, su trabajo es vital. Sin embargo, a menudo se encuentran en una posición vulnerable, donde sus aportes son subestimados y sus condiciones laborales son precarias.

Por ejemplo, los trabajadores subcontratados, muchos de los cuales provienen de países en desarrollo, están en la primera línea de la revolución de la IA. A menudo son responsables de realizar tareas que requieren atención al detalle y un profundo conocimiento del contexto, lo que es crucial para entrenar a los algoritmos de IA. Sin embargo, estos trabajadores son vistos como fácilmente reemplazables, lo que refleja una tendencia más amplia en la automatización y la tecnología, donde la riqueza y el poder se concentran en manos de unas pocas empresas tecnológicas. En este contexto, es fundamental que se reconozca el valor de estos trabajadores y se les ofrezcan oportunidades de capacitación y desarrollo profesional.

La colaboración entre humanos y máquinas puede generar resultados más creativos e innovadores, siempre que se valore adecuadamente el papel de los trabajadores en el proceso.

A medida que la IA se convierte en una parte integral de la toma de decisiones en diversos sectores, también surgen preocupaciones sobre su potencial para perpetuar y amplificar sesgos raciales. La historia de la IA está plagada de ejemplos que ilustran cómo los algoritmos pueden discriminar a grupos marginalizados. Esto plantea la pregunta crítica: ¿quién es responsable de garantizar que las máquinas de IA no sean racistas? La respuesta no es sencilla y requiere un enfoque multidimensional que involucre a desarrolladores, empresas, responsables de políticas y la sociedad en su conjunto.

Es esencial considerar cómo el rol invisible de los trabajadores detrás de la IA y el desafío del racismo en la IA se entrelazan. La mayoría de los trabajadores que están en la línea de fuego de la automatización y el uso de la IA provienen de comunidades marginalizadas y, a menudo, enfrentan condiciones laborales desfavorables. Esto resalta la necesidad de abordar las desigualdades en el acceso a oportunidades laborales y educativas dentro del sector tecnológico.

El sesgo racial en la IA puede exacerbar la precariedad laboral de los trabajadores subcontratados. Si los algoritmos de IA se entrenan con datos sesgados, esto puede llevar a decisiones automatizadas que impactan negativamente a las comunidades de color. Por ejemplo, en la contratación y la evaluación del desempeño laboral, los sistemas de IA pueden favorecer a candidatos que encajen con patrones de datos

históricos, dejando fuera a quienes provienen de contextos menos representados. Esto no solo perpetúa la desigualdad, sino que también desincentiva a los trabajadores de diversificar sus habilidades y avanzar en sus carreras.

El compromiso con la equidad en la IA requiere una acción concertada para empoderar a los trabajadores y garantizar que sus voces sean escuchadas en la toma de decisiones tecnológicas. Esto implica no sólo reconocer su contribución, sino también ofrecerles plataformas para expresar sus preocupaciones y aportar ideas sobre cómo mejorar los sistemas en los que trabajan. La creación de espacios colaborativos donde los trabajadores puedan dialogar sobre sus experiencias y contribuir al desarrollo de la IA es un paso vital para abordar tanto el rol invisible de los trabajadores como el desafío del racismo en la tecnología.

En última instancia, la conversación sobre el rol de los trabajadores en la IA y el racismo en la tecnología no debe centrarse únicamente en los problemas, sino también en las oportunidades que presenta esta revolución tecnológica. La IA tiene el potencial de ser una herramienta poderosa para la justicia social, la inclusión y la equidad, siempre que se utilice de manera consciente y responsable. Las empresas tecnológicas, los gobiernos y la sociedad civil deben unirse en un esfuerzo por construir un futuro en el que la inteligencia artificial no solo sea una fuente de innovación, sino también un vehículo para la justicia social. La promoción de la diversidad en el desarrollo de IA, la regulación responsable y el empoderamiento de los trabajadores son pasos cruciales para lograr este objetivo.

REFERENCIAS

Rest of World. (2023, 11 de julio). *The AI revolution is changing the work landscape for outsourced workers. https://restofworld.org/2023/ai-revolution-outsourced-workers/*

Sullivan, B. (2021, marzo 15). *Google's A.I. was trained on images of faces. It struggled with Black faces.* The New York Times. *https://www.nytimes.com/2021/03/15/technology/artificial-intelligence-google-bias.html*

McKinsey Global Institute. (2020). *The future of work: Reskilling and remote working in the COVID-19 era*. Recuperado de *https://www.mckinsey.com/featured-insights/future-of-work*

West, S. M., y Allen, J. (2021). *The ethical implications of AI bias. New York Times*. Recuperado de *https://www.nytimes.com/2021/03/15/technology/artificial-intelligence-google-bias.html*

Angwin, J., Larson, J., Mattu, S., y Kirchner, L. (2016). Machine bias. *ProPublica*. Recuperado de *https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing*

University of Oxford. (2022, 3 de marzo). *Art for our sake: Artists cannot be replaced by machines, study finds. https://www.ox.ac.uk/news/2022-03-03-art-our-sake-artists-cannot-be-replaced-machines-study*

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. *https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G*

Varian, H. R. (2010). Computer mediated transactions. The American Economic Review, 100(2), 1-10. *https://doi.org/10.1257/aer.100.2.1*

CAPÍTULO XV

REVOLUCIONANDO LA CONCIENCIA DIGITAL: LA ÉTICA DE TERCERA GENERACIÓN EN LA ERA DE LA INTELIGENCIA ARTIFICIAL

VIRGINIA LÓPEZ NEVÁREZ[1]
ALEJANDRO LÓPEZ ÁLVAREZ[2]
MANUEL ALEJANDRO ZAVALA NEVÁREZ[3]

RESUMEN

La revolución en la conciencia digital plantea un dilema fundamental: ¿cómo aseguramos que la tecnología sea una fuerza para el bien y no una amenaza latente para la humanidad? Este capítulo analiza como la conciencia digital se presenta como un componente crucial, implicando no solo un entendimiento crítico de la tecnología, sino una reflexión profunda sobre sus implicaciones en la vida individual y social. Complementando este enfoque, la ética de tercera generación promueve un paradigma holístico que integra la sostenibilidad y la justicia social como elementos centrales en el desarrollo tecnológico. En este contexto surge el concepto de un ecosistema digital responsable, donde la inteligencia artificial se desarrolla en armonía con principios éticos y la responsabilidad se comparte entre todos los actores involucrados. Se identifican, además, desafíos significativos, como la equi-

[1] Directora de Unidad de Educación Virtual, Universidad Autónoma de Occidente, Los Mochis, Sinaloa, México.

[2] Estudiante de Doctorado de Innovación y Administración de las Organizaciones Modalidad Virtual, Universidad Autónoma de Occidente, Los Mochis, Sinaloa, México

[3] Coordinador de la Maestría en Docencia Modalidad Virtual, Universidad Autónoma de Occidente, Los Mochis, Sinaloa, México

dad, la transparencia, la privacidad y la sostenibilidad, que no pueden ser ignorados. Este análisis culmina en la propuesta de un marco ético que garantice no solo el bienestar humano, sino que priorice la justicia intergeneracional. Así, se realiza un llamado a la acción urgente para construir un futuro donde la inteligencia artificial se convierta en una genuina fuerza para el bien, guiada por una conciencia digital evolucionada y una ética que trascienda el presente, impactando positivamente las vidas de las futuras generaciones.

Palabras clave: Conciencia digital, Ética de tercera generación, Ecosistema digital responsable

INTRODUCCIÓN

En un concurrido lunes por la mañana, Ana, una joven profesional en plena carrera, se enfrenta a una presión laboral abrumadora que amenaza su salud mental. Invadida por la ansiedad, decide probar una aplicación innovadora que utiliza inteligencia artificial para ofrecer terapia cognitivo-conductual. Al iniciar la sesión, la aplicación le pide a Ana que responda a algunas preguntas sobre su estado de ánimo y sus principales preocupaciones. A medida que Ana va respondiendo, la IA de la aplicación comienza a identificar patrones en su comportamiento, como la tendencia a autocriticarse, la dificultad para concentrarse y la recurrencia de pensamientos negativos relacionados con el trabajo.

Sorprendida, Ana descubre que el chatbot no solo es empático, sino que también utiliza su historial de interacciones para personalizar sus respuestas. La IA, basada en los patrones detectados, le ofrece estrategias específicas para afrontar su ansiedad. Por ejemplo, la aplicación le sugiere técnicas de respiración para controlar la ansiedad, ejercicios de mindfulness para mejorar su concentración, y recomienda recursos para gestionar el estrés laboral.

Con el paso de las sesiones, Ana experimenta un apoyo significativo que mejora notablemente su bienestar emocional. La aplicación, gracias a la IA, se convierte en una herramienta valiosa que la ayuda a

comprender mejor su propia situación, a desarrollar mecanismos para afrontar la presión laboral de manera más efectiva, y a encontrar un equilibrio entre su bienestar personal y profesional. Este ejemplo no solo resalta cómo la interacción con tecnologías avanzadas está redefiniendo nuestra comprensión y relación con el entorno digital, creando una nueva conciencia digital, sino que también ilustra la necesidad de una ética de tercera generación. Esta ética global y ambiental considera no solo el beneficio individual de dichas tecnologías, sino también su impacto colectivo en la sostenibilidad del planeta, subrayando la responsabilidad de equilibrar el bienestar personal con el bienestar global.

En la era digital moderna, la conciencia digital se ha convertido en un concepto fundamental para comprender la relación entre los individuos y las sociedades con el ecosistema digital. Esta conciencia implica no solo la familiaridad con el uso de herramientas digitales, sino también una comprensión crítica de su impacto en las vidas de las personas.

La conciencia digital, según Veliyev (2024), representa una forma de comprensión que va más allá de simplemente usar la tecnología. Se reconoce no solo el entorno digital como una herramienta de comunicación y trabajo, sino que también se considera su capacidad para moldear las percepciones, los valores y las acciones diarias. La conciencia digital nos lleva a reflexionar sobre cómo la interacción constante con las tecnologías digitales impacta en el desarrollo personal, la autoestima y la manera en que las personas se presentan al mundo. Además, invita a analizar cómo los agentes virtuales y los sistemas de inteligencia artificial, cada vez más presentes en la vida cotidiana, influyen en la autopercepción de los individuos y en su forma de entender el mundo.

La relevancia de la conciencia digital es particularmente significativa en el contexto de la IA, ya que estas tecnologías están remodelando radicalmente las interacciones sociales, económicas y culturales. Howard (2019) destaca que, a medida que la IA facilita decisiones que afectan directamente la vida de las personas, contar con una conciencia digital informada permite que los individuos tomen decisiones éticas y responsables, lo que a su vez promueve una dinámica saludable entre la tecnología y la sociedad.

En este contexto, la introducción de una ética de tercera generación ofrece una nueva perspectiva que avanza más allá de los marcos tradicionales. La primera generación de ética se centra en principios deontológicos y utilitaristas, que buscan definir qué es moralmente correcto a través de normas y consecuencias tangibles. La segunda generación, referente a la ética aplicada, se enfoca en la práctica de principios éticos en casos particulares y situaciones contextuales. La tercera generación, en cambio, abre un diálogo hacia una ética global y ambiental, considerando no solo las acciones individuales y sus consecuencias, sino también el impacto en el ecosistema global y la sostenibilidad del planeta (White *y otros*., 2024).

Este enfoque de ética global está alineado con la conciencia digital al promover prácticas que son cultural y ambientalmente sensibles, reconociendo que las decisiones tecnológicas tienen un alcance que trasciende fronteras geográficas y generacionales. La ética de tercera generación invita a reimaginar las responsabilidades colectivas en un mundo interconectado, donde el bienestar de las comunidades y del planeta son parte integral de la consideración ética.

Este capítulo tiene como objetivo analizar cómo la inteligencia artificial está transformando la conciencia digital dentro del marco de la ética de tercera generación, a través del estudio de casos de implementación que ilustran principios de equidad, transparencia y sostenibilidad, evaluando tanto los desafíos éticos como las implicaciones tecnológicas y sociales de estos avances. Este análisis busca comprender cómo la IA está remodelando la manera en que las personas se relacionan con el mundo digital y cómo la ética de tercera generación puede contribuir a un desarrollo tecnológico responsable que priorice el bienestar humano y la sostenibilidad del planeta.

EXPLORANDO LAS DIMENSIONES DE LA CONCIENCIA DIGITAL

La conciencia digital se presenta como un concepto fundamental en la literatura científica actual, abarcando una serie de dimensiones que trascienden el mero uso de tecnologías. Según Camaran Maione (2023), la conciencia digital no se limita al dominio técnico de las he-

rramientas digitales, sino que implica una comprensión crítica de la información recibida a través de ellas. También se refiere a su aplicación en el ejercicio profesional y la capacidad de adaptarse a un entorno digital en constante cambio. Para lograr esta conciencia, Maione destaca la importancia de integrar la información digital de manera crítica mediante el anclaje de contenidos, adquirir habilidades digitales de forma continua como parte de la formación en competencias digitales, y utilizar las herramientas de manera ética, estratégica y creativa, lo que se traduce en un uso consciente de las herramientas.

Stahl (2021) introduce una perspectiva diferente, argumentando que la conciencia digital también se refiere a la capacidad de los sistemas de inteligencia artificial para procesar información de manera que emule ciertos aspectos de la percepción y respuesta humanas, aunque sin alcanzar una verdadera conciencia en el sentido biológico. Este enfoque pone énfasis en la habilidad de los sistemas de IA para analizar y adaptarse a entornos cambiantes, lo cual facilita interacciones más sofisticadas y personalizadas entre humanos y máquinas.

Por otra parte, Shiller (2024) sostiene que la conciencia digital trasciende la organización funcional básica, requiriendo una implementación física que cumpla con ciertas restricciones de integridad. Estas restricciones aseguran que los sistemas de IA no sean simples simulaciones, sino que posean una estructura interna que les permita actuar de manera autónoma y responsable, en consonancia con los principios de la ética de tercera generación.

Jobin *y otros*. (2019) refuerzan esta idea, sugiriendo que la conciencia digital debería implicar la capacidad de los sistemas de IA para comprender y procesar información de manera crítica y autónoma, tomando decisiones basadas en valores éticos fundamentales como la justicia, la equidad y la transparencia. Propugnan que esta conciencia debe estar integrada en el diseño y desarrollo de la IA, reflejándose en la toma de decisiones, la gestión de datos y la interacción con los humanos. Destacan que la conciencia digital va más allá de las capacidades puramente técnicas, representando una dimensión ética y social que surge al considerar las implicaciones de la IA en la sociedad y al promover un uso responsable de esta tecnología.

Se puede observar como la conciencia digital trasciende un enfoque meramente técnico al englobar una serie de competencias que son esenciales tanto para la interacción como para la contribución efectiva en un ecosistema digital que está en constante evolución. En primer lugar, esta conciencia requiere una alfabetización crítica robusta, permitiendo a los individuos discernir y evaluar la información digital con un enfoque analítico, guiado por principios éticos y rigor intelectual. En segundo lugar, involucra la capacidad de adaptación: un proceso continuo de aprendizaje y ajuste que responde a las rápidas innovaciones tecnológicas que redefinen regularmente el panorama digital. (**Ver Figura 1**)

Figura 1. Conciencia digital



Fuente: *Chat GPT4o* (2024)

Realizando un silogismo a partir de las definiciones anteriores: si la conciencia digital implica un dominio ético y un entendimiento crítico de la tecnología (premisa mayor) y los sistemas de inteligencia artificial están diseñados para emular y actuar sobre principios humanos fundamentales (premisa menor), entonces la conciencia digital no solo aborda la competencia técnica sino que establece una guía para la interacción responsable y la implementación ética en el entorno digital (conclusión).

El caso de Ana presentado en la introducción ejemplifica cómo la conciencia digital se manifiesta en su interacción con la tecnología de inteligencia artificial (IA) de manera ética y crítica. La aplicación que utiliza representa un sistema de IA diseñado para comprender y actuar sobre principios humanos fundamentales, como la empatía y el bienestar emocional. Este sistema no solo ofrece estrategias personalizadas para manejar su ansiedad, sino que también respeta la privacidad de Ana al solicitar su información de manera transparente para personalizar los consejos. La interacción de Ana con la aplicación demuestra un dominio técnico al usar herramientas digitales que se alinean con su búsqueda de bienestar, mientras que la inteligencia artificial de la aplicación refleja un entendimiento crítico de las necesidades emocionales de los usuarios. Como resultado, la conciencia digital de Ana le permite usar la tecnología de manera responsable, maximizando su beneficio personal, a la vez que contribuye a una implementación ética del entorno digital. Este ejemplo concreto ilustra cómo la conciencia digital va más allá de las habilidades técnicas, actuando como una guía para la interacción y el aprendizaje continuo dentro de un ecosistema digital éticamente consciente.

El caso de Ana, aunque ilustra los beneficios potenciales de integrar inteligencia artificial en el manejo del bienestar personal, también destaca varios riesgos inherentes a la conciencia digital. Uno de los principales riesgos es la privacidad de los datos, ya que las aplicaciones terapéuticas de IA dependen de información personal sensible para personalizar sus respuestas, lo cual podría comprometer la privacidad de los usuarios si no se manejan adecuadamente. Según Floridi *y otros.* (2018), la falta de transparencia en el manejo de datos puede erosionar la confianza en los sistemas de IA, aumentando el riesgo de violaciones de privacidad y uso indebido de la información personal. Además, una dependencia excesiva de tales tecnologías podría conducir a una disminución de la autonomía personal, donde los individuos confían en las recomendaciones de la IA a expensas de desarrollar sus propios mecanismos de afrontamiento. Esto es particularmente preocupante, como advierte Stahl (2021), ante la potencial subestimación del juicio humano frente a decisiones automatizadas. Asimismo, estas aplicaciones podrían perpetuar sesgos inherentes de los algoritmos si no son diseñadas cuidadosamente para abordar las diferencias indi-

viduales y culturales. Por lo tanto, aunque la interacción tecnológica puede beneficiar el bienestar personal, es crucial que los sistemas de IA implementen medidas éticas sólidas para mitigar estos riesgos, promoviendo prácticas de gestión de datos justas y transparentes.

Reconociendo estos riesgos, han surgido diversas iniciativas para fomentar una conciencia digital más segura y responsable. Programas de alfabetización digital están trabajando para capacitar a la población en la navegación crítica del mundo online, disminuyendo la brecha de conocimiento sobre los riesgos digitales. Simultáneamente, proyectos de investigación intentan desentrañar las implicaciones sociales y éticas de las tecnologías emergentes. Empresas como Google están desarrollando herramientas para mejorar el uso seguro de sus plataformas, mientras que organizaciones como *Mozilla* y *AI Now* formulan principios éticos para guiar el uso de la IA. Estas acciones reflejan un compromiso con la responsabilidad y ética en el uso de herramientas digitales, creando así un puente hacia un futuro donde estas tecnologías se integren con los valores humanos fundamentales, promoviendo una sociedad más justa y sostenible.

En conclusión, el caso de Ana, que utiliza inteligencia artificial para gestionar su bienestar personal, ejemplifica cómo una conciencia digital integral se manifiesta en la práctica de aplicar tecnologías avanzadas. Este ejemplo demuestra que la conciencia digital va más allá de las habilidades técnicas, actuando como una guía para una interacción responsable en el entorno digital. Sin embargo, surgen desafíos significativos, especialmente en términos de privacidad, transparencia y autonomía. Los riesgos asociados a una gestión inadecuada de datos y la posible erosión de la autonomía subrayan la necesidad de anclar esta conciencia en principios éticos sólidos. En un entorno digital cada vez más complejo, es crucial que usuarios, desarrolladores y formuladores de políticas adopten un enfoque crítico y ético para asegurar que el despliegue de tecnologías emergentes como la inteligencia artificial promueva un equilibrio entre innovación, seguridad y justicia social. Desde esta perspectiva, la conciencia digital no solo facilita una interacción tecnológica más profunda, sino que también impulsa un modelo de desarrollo tecnológico inclusivo, justo y sostenible, alineado con los valores humanos fundamentales. Así, se establece un puente hacia un futuro donde la tecnología y la ética se integran plenamente.

ÉTICA DE TERCERA GENERACIÓN: HACIA UN ENFOQUE GLOBAL Y SOSTENIBLE

La ética, tradicionalmente definida por Aristóteles como la búsqueda del bien vivir a través de la virtud, y reconfigurada por Kant como el cumplimiento de deberes morales universales, ha evolucionado significativamente en respuesta al desarrollo tecnológico y los desafíos emergentes. A lo largo de tres generaciones, cada iteración ética ha reflejado enfoques y prioridades distintas. Desde la aplicación de estrictos principios normativos hasta la consideración de contextos específicos, y más recientemente, la adopción de un enfoque holístico que valora no solo las acciones individuales, sino también su impacto en el ecosistema global y la sostenibilidad del planeta. Esta evolución ética responde a la necesidad de abordar los complejos dilemas morales que plantea el crecimiento exponencial de la tecnología en la sociedad contemporánea. Para ilustrar estos cambios, a continuación, se presenta la Figura 2, que visualiza la evolución de la ética a lo largo de estas tres generaciones. (**Ver Figura 2**)

Figura 2. Evolución de la Ética



Fuente: *claude.ai* (2024)

La primera generación de ética estaba enfocada en la aplicación de principios clásicos deontológicos y utilitaristas, buscando establecer estándares normativos claros para definir lo que es moralmente aceptable. Un ejemplo de este enfoque sería el desarrollo de códigos de ética en la medicina, donde se establecen reglas estrictas sobre el consentimiento informado y el deber del médico hacia sus pacientes.

La segunda generación, conocida como ética aplicada, emergió en respuesta a problemas más específicos que surgieron en contextos concretos. Aquí, el objetivo era aplicar principios éticos a casos particulares, adaptando los marcos existentes para situaciones del mundo real, como el establecimiento de políticas sobre el uso ético de información genética en investigaciones científicas.

En contraste, la tercera generación de ética adopta un enfoque proactivo y holístico que responde a la complejidad de un mundo interconectado. Esta generación se centra en una ética global y ambiental que no solo examina las acciones individuales y sus consecuencias directas, sino también considera el impacto en el ecosistema mundial y la sostenibilidad a largo plazo. Un ejemplo de esta ética es la implementación de prácticas de negocio que no solo buscan la rentabilidad, sino que también consideran el bienestar de las comunidades locales y el medio ambiente, siguiendo principios de responsabilidad social corporativa (White *y otros.,* 2024).

Los principios fundamentales que guían la ética de tercera generación en la era de la inteligencia artificial incluyen (Stahl, 2021):

- Se centra en el impacto **Equidad:** Asegura que las tecnologías de IA no perpetúen ni amplifiquen las desigualdades existentes. Esto implica diseñar y emplear algoritmos que promuevan la igualdad de oportunidades y minimicen los sesgos que podrían afectar a grupos vulnerables o minoritarios.
- **Transparencia:** Enfatiza la importancia de que los sistemas de IA sean comprensibles y accesibles tanto para los usuarios finales como para los desarrolladores. Esto exige la creación de modelos y procesos de decisión que sean interpretables y auditables, permitiendo que todas las partes interesadas com-

prendan cómo funcionan estos sistemas y qué decisiones están tomando.

- **Sostenibilidad** ambiental y en la responsabilidad social de las tecnologías de IA. Esto conlleva desarrollar sistemas que no solo sean eficientes y efectivos, sino que también respeten el medio ambiente y promuevan prácticas que contribuyan al bienestar a largo plazo del planeta y la humanidad. (**Ver Figura 3**)

Figura 3. Principos fundamentales de la ética de tercera generación



Fuente: *CHAT GPT4o* (2024)

ECOSISTEMA RESPONSABLE: INTERSECCIÓN DE LA IA, LA CONCIENCIA DIGITAL Y LA ÉTICA DE TERCERA GENERACIÓN

De acuerdo con el artículo From Corporate Digital Responsibility to Responsible Digital Ecosystems de Stahl (2024), un **ecosistema responsable** se refiere a un sistema sociotécnico en el que la integración de tecnologías digitales promueve la sostenibilidad y el bienestar social, minimizando las consecuencias negativas y fomentando la responsabilidad ética. Esto implica un sistema que facilita la comunicación y el diálogo entre stakeholders para identificar, abordar y gestionar las ex-

ternalidades de las tecnologías digitales, al mismo tiempo que fomenta la creación de estructuras que permitan atribuir responsabilidades de forma clara y efectiva. Un ecosistema responsable busca no solo un uso tecnológico eficiente, sino también una aplicación que genere beneficios para las generaciones presentes y futuras.

El análisis del ecosistema responsable en el ámbito de la IA, la conciencia digital y la ética de tercera generación se basa en una estructura que reconoce las complejas interdependencias entre estos sistemas emergentes. Haciendo una analogía con la primera ley de la termodinámica, que establece que la energía no se crea ni se destruye, sino que se transforma (Cengel y Boles, 2019), es posible comprender mejor cómo funciona el ecosistema digital. Al igual que la energía en un sistema físico, la información y el conocimiento fluyen y se transforman dentro de los ecosistemas digitales. Las aplicaciones de IA procesan los datos como insumos energéticos, transformándolos en salidas que, a su vez, requieren una supervisión ética rigurosa. Al igual que la energía no se pierde en un sistema físico, la información y el conocimiento, como "energía" del ecosistema digital, deben gestionarse con cuidado y responsabilidad para asegurar la sostenibilidad y la integridad de este sistema.

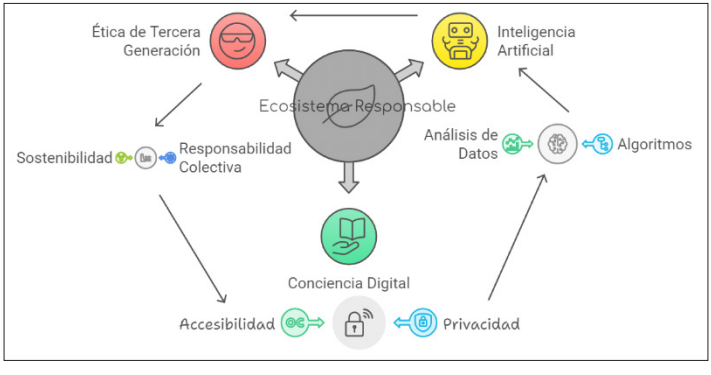
La conciencia digital, por su propia naturaleza, implica un flujo constante de datos y una adaptación continua a través de tecnologías emergentes, reflejando los ciclos energéticos en los ecosistemas naturales (Heylighen y Lenartowicz, 2017). La responsabilidad en este contexto implica garantizar que las transformaciones de datos no comprometan el valor ético o social, buscando un equilibrio entre la innovación tecnológica y la conservación de principios éticos fundamentales (Stahl, 2021).

La ética de tercera generación, centrada en la responsabilidad colectiva y el bienestar intergeneracional, introduce un nuevo paradigma moral al exigir una gestión responsable de los recursos digitales que asegure la sostenibilidad a largo plazo, no solo para la generación actual, sino también para las futuras (Heylighen y Lenartowicz, 2017). Este enfoque se alinea con la necesidad de considerar el impacto a largo plazo de las decisiones tecnológicas, incluyendo su huella ambiental, la equidad en el acceso a la tecnología y la protección de los derechos humanos. En un ecosistema digital responsable, la ética de

tercera generación juega un papel crucial, actuando como un sistema de control que guía el desarrollo de tecnologías que benefician a la sociedad de manera sostenible y justa. La implementación de políticas éticas que incorporen estos principios es fundamental para asegurar un ecosistema digital que impulse el progreso social sin dejar de lado las consideraciones éticas y de responsabilidad colectiva (Brey, 2012).

Este análisis subraya la importancia crítica de diseñar e implementar políticas que optimicen la eficiencia en la transferencia de energía-información, un proceso fundamental para el funcionamiento sostenible de ecosistemas digitales y de IA (Stahl, 2024). Estas políticas deben tener en cuenta no solo los beneficios tecnológicos y operativos, sino también las implicaciones éticas y sociales de las externalidades que resultan de estos procesos. La atención cuidadosa a cómo se gestionan estas externalidades es crucial para evitar consecuencias no deseadas que podrían socavar la equidad y la justicia social. Al integrar medidas de gobernanza ética y sostenible, se facilita una comprensión holística del papel de la IA y la conciencia digital como elementos centrales de un ecosistema interdependiente, donde cada acción tecnológica reverbera a través de contextos sociales y morales, cultivando un entorno donde el progreso digital no solo avanza, sino que lo hace de manera equitativa y ética. **(Ver Figura 4)**

Figura 4: Ecosistema responsable en la intersección de la inteligencia artificial (IA), la conciencia digital y la ética de tercera generación



Fuente: *Napkin.ai* (2024)

En el caso de Ana, quien emplea una aplicación de inteligencia artificial para gestionar su bienestar personal, ilustra la intersección de la IA, la conciencia digital y la ética de tercera generación dentro de un ecosistema responsable. La aplicación de IA proporciona orientación personalizada basada en principios de empatía y bienestar emocional, elementos clave de la conciencia digital. Además, está diseñada para operar dentro de un marco ético sólido que asegura la privacidad de los datos y la transparencia, acorde a los valores de equidad y responsabilidad colectiva promovidos por la ética de tercera generación. Para mantener el equilibrio de este ecosistema, es crucial que la aplicación implemente prácticas rigurosas de protección de datos, evitando la recopilación innecesaria de información personal y empleando procesos de encriptación robustos para garantizar la seguridad de los datos de Ana. Asimismo, se deben incorporar mecanismos de supervisión continua para identificar y corregir cualquier sesgo algorítmico, asegurando que las recomendaciones de la IA sean justas y adaptativas a las realidades cambiantes de los usuarios. La experiencia de Ana subraya cómo un diseño reflexivo y ético de la IA no solo mejora la calidad de vida de los usuarios, sino que también refuerza un enfoque de desarrollo tecnológico inclusivo y sostenible.

En conclusión, la integración de la inteligencia artificial, la conciencia digital y la ética de tercera generación dentro de un ecosistema responsable representa un paso crucial hacia un futuro donde la tecnología y la ética se alinean para fomentar el bienestar social y la sostenibilidad. Este modelo ecosistémico propicia un entorno donde la innovación tecnológica se ve guiada por principios éticos rigurosos, garantizando que el progreso no solo ofrezca beneficios operativos y tecnológicos, sino que también respalde un desarrollo equitativo y justo. Las prácticas de responsabilidad digital, como las observadas en el caso de Ana, se convierten en ejemplos tangibles de cómo las tecnologías emergentes pueden incorporar medidas de protección de derechos humanos, entrelazándose con la ética de responsabilidad colectiva. Al continuar en esta dirección, el ecosistema digital podrá convertirse en un pilar fundamental para la construcción de un futuro donde la tecnología es un catalizador para un cambio social positivo y sostenible.

TRANSFORMACIÓN DIGITAL: CASOS DE IA EN ACCIÓN

La inteligencia artificial (IA) se ha consolidado como una herramienta transformadora en diversos campos, desde la medicina hasta las finanzas, proporcionando capacidades innovadoras para optimizar procesos y decisiones. Un ejemplo inicial que ilustra el potencial de la IA dentro de un marco ético es el caso de Ana, quien utiliza una aplicación basada en IA para manejar su bienestar personal. Este caso refleja el impacto positivo que las tecnologías de IA pueden tener cuando se aplican principios de equidad y respeto a la privacidad, demostrando cómo una implementación ética puede mejorar significativamente la calidad de vida de las personas (Floridi *y otros*., 2018).

No obstante, la rápida expansión de la IA también ha planteado desafíos críticos. Estos desafíos éticos giran en torno a los principios de equidad, transparencia y sostenibilidad, los cuales son fundamentales dentro de la ética de tercera generación. Esta ética propone un enfoque holístico que no solo se centra en los beneficios inmediatos de la tecnología, sino también en sus consecuencias a largo plazo. La adopción masiva de la IA requiere una reflexión cuidadosa sobre la justicia intergeneracional y la preservación de derechos fundamentales en una sociedad crecientemente automatizada (Stahl, 2024).

Al analizar los casos de implementación detallados en la Tabla 1, queda claro que la equidad, la transparencia y la sostenibilidad social no son solo ideales éticos, sino requisitos prácticos para el uso responsable de la IA. Los casos presentados muestran aplicaciones en áreas como la justicia, la educación y las finanzas, subrayando la importancia de integrar principios éticos que guíen el desarrollo y la adopción de estas tecnologías. Este enfoque no solo asegura la eficiencia técnica, sino que también promueve la equidad y la sostenibilidad en el uso de la IA, sugiriendo un camino hacia un futuro digital justo y equilibrado para todas las generaciones (Jobin *y otros*., 2019).

Tabla 1: Estudios relacionados con la IA y la ética de tercera generación

Caso	Descripción	Principios Éticos Involucrados	Desafíos Identificados	Recomendaciones
1. Algoritmos Predictivos en Justicia	Análisis de los sesgos algorítmicos en los sistemas de justicia, con énfasis en cómo perpetúan desigualdades raciales y socioeconómicas (Ugwudike, 2022).	Equidad, Transparencia, Sostenibilidad	Refuerzan la discriminación estructural al basarse en datos históricos contaminados por racismo. Opacidad en los procesos algorítmicos que dificulta la rendición de cuentas.	Implementación de auditorías algorítmicas independientes. Fomentar un enfoque estructural que cuestione las bases de la discriminación.
2. IA en Educación (K-12)	Uso de IA en entornos educativos, con un enfoque en los desafíos éticos que presenta, como el sesgo y la vigilancia. (Akgun & Greenhow, 2021)	Equidad, Autonomía, Justicia Social	Los sistemas de IA pueden perpetuar sesgos que afectan desproporcionadamente a estudiantes marginados. Recopilación de datos sin consentimiento informado adecuado.	Crear currículos que integren IA y ética en la educación, fomentando una ciudadanía informada y responsable.
3. Responsabilidad Digital Corporativa en Servicios Financieros	Estudio sobre la Responsabilidad Digital Corporativa (CDR) en el uso de IA dentro del sector financiero (Akgun & Greenhow, 2021).	Equidad, Transparencia, Sostenibilidad	Las decisiones basadas en IA perpetúan desigualdades socioeconómicas debido a sesgos en los datos históricos. La opacidad de los algoritmos financieros afecta la confianza del consumidor.	Desarrollar políticas de CDR que promuevan la auditoría de IA y expliquen las decisiones financieras a los consumidores.

Caso	Descripción	Principios Éticos Involucrados	Desafíos Identificados	Recomendaciones
4. IA y Toma de Decisiones Públicas	El uso de IA en la toma de decisiones públicas y los desafíos que presenta la opacidad en el proceso (de Fine Licht & de Fine Licht, 2020).	Transparencia, Legitimidad, Equidad	La opacidad de los procesos algorítmicos mina la legitimidad percibida por el público.	Fomentar la transparencia justificativa, proporcionando explicaciones claras y accesibles a la ciudadanía sobre las decisiones automatizadas.

Fuente: Elaborado a partir de los casos obtenidos por *ChatGPT4o* (2024)

En el ámbito de la justicia predictiva, los algoritmos de IA han sido objeto de intenso escrutinio debido a su tendencia a perpetuar sesgos históricos, replicando patrones de discriminación racial y socioeconómica presentes en los datos que los alimentan. Este fenómeno plantea serias implicaciones para la equidad, ya que, en lugar de corregir desigualdades, estos sistemas tienden a reforzar las estructuras de injusticia existentes. De manera similar, en el sector educativo, la adopción de IA genera preocupaciones sobre la vigilancia y la autonomía estudiantil, especialmente respecto a la recopilación de datos sensibles y las decisiones algorítmicas que podrían impactar el desarrollo académico de manera injusta o desproporcionada. Según Binns (2018), estos escenarios subrayan la necesidad de incorporar principios de equidad y transparencia para cultivar un ecosistema digital responsable donde la tecnología no solo mejore la eficiencia, sino que también fomente la justicia social y la equidad.

En el ámbito financiero, las tecnologías de IA han optimizado procesos cruciales como la evaluación crediticia y la detección de fraudes. No obstante, este avance no está exento de dilemas éticos, pues la implementación de algoritmos en este sector ha evidenciado la perpetuación de sesgos que agravan las disparidades económicas, impactando negativamente a individuos pertenecientes a grupos vulnerables. Por otro lado, en el ámbito de la gobernanza pública, el uso de IA en la

toma de decisiones suscita preocupaciones sobre la transparencia y la rendición de cuentas. Los procesos algorítmicos se perciben a menudo como opacos y difíciles de auditar, lo que compromete la legitimidad de las decisiones automatizadas y mina la confianza pública en las instituciones. Stahl (2024) argumenta que integrar principios de ética de tercera generación en estas áreas es esencial para construir un ecosistema digital que priorice la equidad y la responsabilidad intergeneracional.

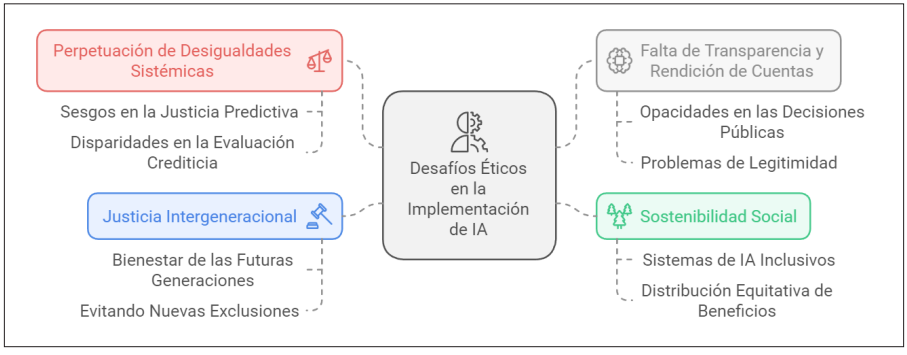
Este análisis refleja la epistemología del ecosistema responsable, explorando las implicaciones éticas de la implementación de la IA en estos sectores, y destacando la necesidad imperiosa de un marco normativo que incorpore los principios de la ética de tercera generación. Este enfoque ético exige que el desarrollo de la IA no solo se enfoque en mejorar la eficiencia y precisión, sino también en garantizar la justicia social, la transparencia y la sostenibilidad intergeneracional. En esta línea, el análisis de los casos presentados contribuye significativamente al debate académico y profesional sobre la gobernanza ética de la IA, subrayando la importancia de diseñar sistemas tecnológicos que, además de ser eficaces, promuevan un desarrollo justo y equitativo a largo plazo (Floridi *y otros*., 2018).

IMPACTO ANALÍTICO

La implementación de la inteligencia artificial en sectores clave como la justicia, la educación, los servicios financieros y la gobernanza pública ha generado profundas transformaciones, pero también ha expuesto una serie de desafíos éticos que no pueden ser ignorados que se muestran en la Figura 5. El caso de Ana, quien utiliza una aplicación de IA para manejar su bienestar personal, ejemplifica tanto el potencial positivo de estas tecnologías como los riesgos implicados. Mientras que la aplicación mejora su calidad de vida al proporcionar soporte personalizado, también plantea consideraciones sobre la privacidad de los datos y la transparencia en el procesamiento de su información personal. Este ejemplo refleja una situación común en muchas aplicaciones de IA, donde las oportunidades van de la mano con las obligaciones éticas (Floridi *y otros*., 2018). Al igual que en los sectores

más amplios mencionados, el caso de Ana subraya la importancia de desarrollar marcos éticos robustos que guíen tanto el diseño como la implementación de la IA para equilibrar innovación y responsabilidad (Jobin *y otros*., 2019). (**Ver Figura 5**)

Figura 5: Desafíos Éticos



Fuente: *Napkin.ai* (2024)

La implementación de sistemas de inteligencia artificial en sectores como la justicia predictiva y los servicios financieros ha revelado una preocupante tendencia a replicar sesgos históricos, exacerbando desigualdades ya existentes. Los algoritmos utilizados en la justicia a menudo dependen de datos que reflejan discriminaciones raciales y socioeconómicas pasadas, lo que resulta en decisiones sistemáticamente injustas (Ugwudike, 2022). De igual forma, en el ámbito financiero, las herramientas de IA empleadas para la evaluación crediticia continúan ejerciendo un impacto desproporcionado sobre grupos vulnerables, incrementando las desigualdades económicas (Akgun y Greenhow, 2021). Esta persistencia de sesgos plantea cuestiones esenciales sobre la equidad y subraya la necesidad de revisar los procesos de entrenamiento y validación de algoritmos para prevenir resultados injustos.

La falta de transparencia en los procesos algorítmicos es otro desafío significativo que afecta la legitimidad de la toma de decisiones en los servicios públicos y financieros. La opacidad de los algoritmos, que

funcionan como "cajas negras", limita la capacidad de los individuos para entender cómo se derivan las decisiones que les conciernen. Esto socava la confianza pública en las instituciones y resalta la necesidad de explicaciones claras y comprensibles (de Fine Licht y de Fine Licht, 2020). Desde un enfoque basado en la ética de tercera generación, se reconoce que la transparencia no solo es técnica, sino un imperativo ético que exige rendición de cuentas y legitimidad en los sistemas de IA.

La sostenibilidad social, enmarcada en la justicia intergeneracional, es un principio ético indispensable en el desarrollo de la inteligencia artificial contemporánea. La persistencia de sesgos y la falta de transparencia en los sistemas de inteligencia artificial no solo afectan a las generaciones presentes, sino que también implican riesgos a largo plazo para las futuras generaciones. La justicia intergeneracional requiere que las decisiones tecnológicas actuales fomenten una sostenibilidad social que sea inclusiva y equitativa. Esto implica que los sistemas de IA no deben ser diseñados únicamente para maximizar la eficiencia, sino también para asegurar una distribución justa de los beneficios y prevenir la consolidación de nuevas formas de exclusión social, todo ello dentro de un ecosistema digital responsable.

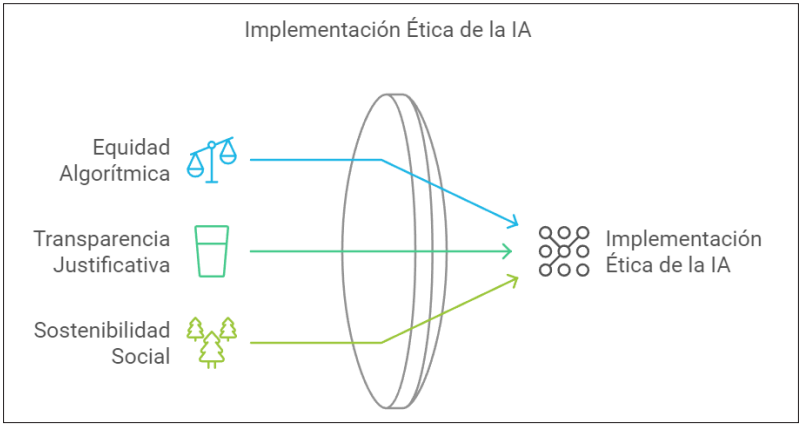
La reflexión sobre los desafíos éticos en el contexto de un ecosistema digital responsable subraya la necesidad de una profunda revisión axiológica de los valores que sustentan el desarrollo y la implementación de la inteligencia artificial. Los sesgos sistemáticos en la justicia predictiva y las finanzas, junto con la opacidad algorítmica, destacan no solo deficiencias técnicas, sino también una falta de compromiso con los valores fundamentales de equidad, transparencia y justicia social. La conciencia digital, que exige una comprensión crítica de los entornos tecnológicos, y la ética de tercera generación, que aboga por un enfoque global y sostenible, proporcionan los marcos necesarios para abordar estos desafíos. Adoptar una perspectiva axiológica invita a considerar cómo los sistemas de IA pueden ser desarrollados y gestionados para reflejar valores que aseguren una distribución equitativa de beneficios a través del tiempo, tanto en el presente como para generaciones futuras. Este enfoque requiere una integración profunda de principios éticos en todas las fases del ciclo de vida de la IA, des-

de la conceptualización hasta la implementación, garantizando que las tecnologías no solo sean eficientes y precisas, sino también éticamente responsables y sostenibles. Así, se logra avanzar hacia un ecosistema digital que cumpla con sus potencialidades, sin comprometer la justicia intergeneracional ni los derechos humanos fundamentales, promoviendo una colaboración ética entre desarrolladores, legisladores y usuarios para construir un futuro inclusivo y justo.

MARCO PARA LA IMPLEMENTACIÓN DE LA IA EN UN ECOSISTEMA RESPONSABLE

Un marco para la implementación de la inteligencia artificial en un ecosistema responsable, como se ilustra en la Figura 6, se centra en integrar principios éticos de equidad algorítmica, transparencia justificativa y sostenibilidad social en el desarrollo y uso de estas tecnologías. El caso de Ana, quien utiliza una aplicación basada en inteligencia artificial para manejar su bienestar personal, ejemplifica cómo estos principios pueden aplicarse de manera efectiva. La aplicación garantiza la equidad algorítmica al ofrecer recomendaciones personalizadas que evitan reproducir sesgos desfavorables. Además, proporciona transparencia justificativa al explicar claramente el manejo y procesamiento de los datos de Ana, fortaleciendo su confianza en la tecnología. En cuanto a la sostenibilidad social, la aplicación se desarrolla con el objetivo de mejorar la salud mental de los usuarios de manera inclusiva, promoviendo el bienestar individual y social. Este marco orienta el diseño y regulación de sistemas de inteligencia artificial en diversos sectores, asegurando que estas tecnologías sean implementadas de manera ética y responsable para maximizar beneficios y minimizar riesgos. (**Ver Figura 6**)

Figura 6: Marco de implementación de la IA en un ecosistema responsable



Fuente: *Napkin.ai* (2024)

La equidad en los sistemas de inteligencia artificial se fundamenta en la implementación de mecanismos que garanticen la ausencia de sesgos. Este enfoque epistemológico se basa en la auditoría algorítmica continua, que emplea metodologías estadísticas para identificar y corregir sesgos en los datos y las decisiones generadas por la IA. Praxeológicamente, se promueve la creación de comités éticos interdisciplinares, compuestos por expertos en tecnología, ética y justicia social, que supervisen y evalúen regularmente el uso de la IA en sectores críticos como la justicia y las finanzas. Según Gebru *y otrosl.* (2021), la supervisión ética, junto con la auditoría, es vital para asegurar que los algoritmos se ajusten a los principios de justicia social y equidad, fomentando la confianza en la tecnología.

La transparencia justificativa es esencial para que las decisiones algorítmicas sean comprensibles por todos los afectados. La epistemología de este principio corresponde al desarrollo de marcos teóricos que facilitan la interpretación y explicación claras de los algoritmos. En términos praxeológicos, se sugiere la implementación de protocolos de transparencia que obliguen a las instituciones a proporcionar explicaciones detalladas y accesibles sobre el funcionamiento de los algoritmos. Brundage *y otros.* (2020) sostienen que fortalecer la transparencia no solo incrementa la

confianza pública en los sistemas de IA, sino que también obliga a las entidades a asumir responsabilidad por sus tecnologías.

Una perspectiva a largo plazo sobre el impacto social de la IA es crucial para asegurar la sostenibilidad social y la justicia intergeneracional. La epistemología de este principio implica un enfoque que evalúa las posibles consecuencias a largo plazo de las tecnologías emergentes, considerando su potencial para fomentar o socavar la equidad intergeneracional. En la práctica, esto sugiere la creación de marcos regulatorios que promuevan una evaluación sistemática de los impactos sociales de la IA, garantizando que las decisiones tecnológicas actuales no resulten en desigualdades permanentes para las generaciones futuras. Como destacan *Floridi y Cowls* (2019), este enfoque debe integrar evaluaciones de impacto social que aseguren una tecnología inclusiva y sostenible.

En la era actual de progreso tecnológico acelerado, establecer un marco robusto para la implementación de la inteligencia artificial dentro de un ecosistema responsable no es solo una aspiración ética, sino una necesidad imperativa. Este marco, fundamentado en los principios de equidad algorítmica, transparencia justificativa y sostenibilidad social, ofrece una guía crucial en medio de los rápidos avances de la IA, garantizando que estas tecnologías no solo sean innovadoras, sino también responsables y justas. La conciencia digital, como habilidad crítica, se convierte en un catalizador para la interacción ética con la tecnología, asegurando que el poder transformador de la IA esté anclado en un entendimiento crítico de sus impactos potenciales. Además, la ética de tercera generación, con su enfoque en la responsabilidad intergeneracional y el bienestar colectivo, articula un paradigma ético que demanda una alineación entre innovación y responsabilidad. Este marco no solo asegura la justicia y la equidad en el presente, sino que también actúa como salvaguarda para las generaciones futuras, formalizando un compromiso concreto para integrar la ética en cada etapa del ciclo de vida de la IA. Al implementar este marco, se establece un equilibrio pragmático entre el potencial transformador de la IA y las responsabilidades éticas que vienen con tal poder, garantizando que la evolución digital contribuye a una sociedad más justa y equitativa sin caer en un idealismo inalcanzable.

DESAFÍOS ÉTICOS Y TECNOLÓGICOS

La implementación de la IA en el mundo contemporáneo ofrece tanto un potencial transformador como una serie de desafíos éticos y tecnológicos, particularmente en el ámbito de la ética digital. La ética de tercera generación, que busca infundir una conciencia ética avanzada en la IA, desafía a estas tecnologías no solo a operar con eficiencia, sino también a adherirse a principios de equidad, responsabilidad y transparencia (Floridi y otros., 2018). El caso de Ana, quien utiliza una aplicación de IA para gestionar su bienestar personal, ilustra cómo la IA puede proporcionar beneficios significativos a sus usuarios, ofreciendo soporte personalizado y estrategias adaptativas que mejoran su calidad de vida. Sin embargo, este caso también destaca cuestiones complejas relacionadas con la privacidad de los datos, ya que la aplicación maneja información altamente personal y sensible. La transparencia se convierte en un desafío, ya que los usuarios necesitan entender cómo se utilizan sus datos para confiar plenamente en estas tecnologías (Binns, 2018).

El análisis del caso de Ana subraya la necesidad de abordar los desafíos éticos mediante la implementación de estrategias que garanticen la protección de datos y la transparencia en los procesos algorítmicos. Este enfoque es esencial para construir un entorno digital donde la justicia social y los derechos fundamentales se preserven a medida que la IA se integra más plenamente en la vida cotidiana. Los dilemas que enfrentan usuarios como Ana ejemplifican el requisito imperativo de establecer normas claras e implementar controles rigurosos que se alineen con la ética de tercera generación. Esto no solo implica desarrollar una tecnología que sea justa e inclusiva sino también garantizar que cualquier implementación tecnológica logre un equilibrio entre innovación y ética (Brundage *y otros*., 2018). La integración efectiva de estos principios permitirá que la IA avance sin comprometer los estándares éticos fundamentales que son vitales para la confianza pública y el progreso social sostenible.

La Tabla 2 proporciona un desglose detallado de las principales barreras y desafíos que afectan la implementación ética de la inteligencia artificial. Esta tabla resalta aspectos críticos como la transparencia en los procesos algorítmicos, la protección de la privacidad de los datos,

y el aseguramiento de la equidad en la IA. Además, subraya la promoción de la sostenibilidad y el desarrollo de normativas adecuadas, incluyendo una consideración profunda de los retos socio-culturales que acompañan al despliegue de tecnologías emergentes. Integrando los constructos de conciencia digital, se enfatiza la necesidad de un entendimiento crítico y responsable del entorno digital, mientras que la ética de tercera generación aporta un marco que orienta estas prácticas hacia un enfoque global y sostenible, asegurando que la IA no solo avance en eficiencia sino también en su alineación con los valores humanos fundamentales.

Tabla 2. **Desafíos Éticos en la Implementación de la IA**

Desafíos y Barreras Éticas	Descripción de la Barrera	Elementos de Conciencia Digital	Principios de Ética de Tercera Generación	Estrategias para Superar las Barreras
Transparencia y Explicabilidad	Dificultad para comprender cómo los algoritmos toman decisiones.	Entendimiento Crítico de Algoritmos	Promover la transparencia para responsabilizar y guiar decisiones equitativas.	Implementar sistemas de IA explicables y desarrollar frameworks de transparencia.
Privacidad y Seguridad de Datos	Necesidad de proteger los datos personales recopilados por la IA.	Protección de la Información Personal	Respetar la privacidad como un derecho humano esencial.	Fomentar políticas de protección de datos robustas y encriptación avanzada.
Equidad y Sesgo Algorítmico	Posibilidad de que los algoritmos perpetúen sesgos existentes.	Promoción de la Equidad Digital	Asegurar decisiones justas y equitativas para todos los grupos sociales.	Realizar auditorías algorítmicas continuas y emplear conjuntos de datos diversos.

Desafíos y Barreras Éticas	Descripción de la Barrera	Elementos de Conciencia Digital	Principios de Ética de Tercera Generación	Estrategias para Superar las Barreras
Interoperabilidad y Estándares de Datos	Falta de estándares comunes que dificultan la integración entre diferentes sistemas de IA.	Estándares para la Integración de Sistemas	Facilitar la colaboración abierta y global para el bien común.	Establecer normas comunes de interoperabilidad y protocolos de datos abiertos.
Sostenibilidad Energética	Alto consumo energético de infraestructuras de IA.	Responsabilidad Ambiental Digital	Adoptar soluciones tecnológicas que garantizan la sostenibilidad a largo plazo.	Innovar en tecnologías energéticamente eficientes y optimizar el procesamiento de IA.
Adaptación Regulatoria y Normatividad Legal	Las regulaciones no evolucionan al ritmo de la tecnología.	Conocimiento de Regulaciones y Normas Digitales	Abogar por un marco regulatorio que evolucione congruentemente con las innovaciones.	Promover la colaboración entre tecnólogos y legisladores para actualizar normas.
Resistencia al Cambio Cultural y Social	Miedo a la pérdida de empleos y cambios en las dinámicas laborales.	Adaptabilidad y Educación sobre Cambios Tecnológicos	Garantizar un desarrollo tecnológico inclusivo que beneficie a todas las generaciones.	Implementar programas de reentrenamiento y educación continua para la fuerza laboral.

***Fuente*:** Elaborada a partir de Floridi y otros. (2018); Binns (2018); Strubell y otros. (2019)

Cada una de estas barreras representa un aspecto crítico que debe abordarse para lograr una implementación ética y efectiva de la inteligencia artificial dentro de un ecosistema responsable. La transparencia y la explicabilidad son esenciales para fomentar la confianza del usuario, asegurando que los resultados generados por los sistemas de

IA sean comprensibles y revisables por los humanos (Floridi *y otros*., 2018). Al mismo tiempo, la protección de la privacidad y la seguridad de los datos destaca la necesidad de establecer mecanismos robustos que resguarden la información personal, evitando abusos que puedan erosionar la confianza pública en la tecnología (Solove, 2020). Además, las cuestiones de equidad y sesgo algorítmico enfatizan la importancia de desarrollar algoritmos que no perpetúen injusticias ni discriminen a grupos minoritarios o subrepresentados, garantizando así que todos los sectores de la sociedad se beneficien equitativamente de los avances tecnológicos (Mehrabi *y otros*., 2021).

La interoperabilidad y los estándares de datos son fundamentales para promover la comunicación efectiva entre diferentes sistemas de IA, asegurando una integración fluida y eficiente que maximice el potencial de estas tecnologías dentro de un ecosistema responsable (Raghupathi y Raghupathi, 2014). La preocupación por la sostenibilidad energética requiere que las infraestructuras tecnológicas adopten soluciones más ecológicas, enfrentando el desafío del alto consumo energético asociado a la IA y buscando métodos que reduzcan su huella ambiental, alineando así la innovación tecnológica con los principios de sostenibilidad del ecosistema (Strubell y otros., 2019).

Por otro lado, la adaptación regulatoria y la normatividad legal son esenciales para que las leyes evolucionen al ritmo de la tecnología, proporcionando un marco claro y coherente que oriente el desarrollo y uso de la IA en un entorno responsable (Marchant y otros., 2011). Finalmente, la resistencia al cambio cultural y social pone de relieve la necesidad de gestionar el impacto humano de la automatización, garantizando que el progreso tecnológico no conduzca a la marginalización de trabajadores ni a la destrucción de empleos sin ofrecer alternativas viables (Frey y Osborne, 2017). Estos desafíos representan obstáculos a superar, pero también áreas de oportunidad donde se puede transformar la IA en una fuerza positiva para el bien común, impulsando un ecosistema digital que priorice la equidad, la responsabilidad y la transparencia.

El análisis de las barreras y desafíos en la implementación ética de la inteligencia artificial revela la compleja relación entre ética, tecnología y conciencia digital en el contexto de un ecosistema responsable. Este

binomio destaca que la tecnología, por sí sola, no es inherentemente buena o mala; su impacto depende de cómo se diseña, implementa y regula. La crítica se centra en la tendencia a presentar la inteligencia artificial como una solución universal a los problemas de la vida cotidiana, generando expectativas poco realistas en la conciencia digital de los usuarios. Esta percepción puede llevar a una dependencia excesiva de la tecnología, desestimando la importancia de la intervención humana y el pensamiento crítico en la toma de decisiones. La conciencia digital se convierte así en un componente esencial al empoderar a los individuos para discernir y evaluar el uso de la IA, facilitando un enfoque crítico que promueve la adopción de principios éticos en la interacción con estas tecnologías.

Además, es fundamental que la ética de tercera generación se integre en el desarrollo de sistemas de IA para asegurar que su crecimiento no solo busque eficiencia e innovación, sino que también respete los derechos humanos y promueva la justicia social y ambiental. Al adoptar una perspectiva que una ética y tecnología, se puede construir un futuro en el que la inteligencia artificial no solo sea potente y transformadora, sino también justa y sostenible, beneficiando equitativamente a todos los sectores de la sociedad y reduciendo las desigualdades existentes.

IMPLICACIONES SOCIALES Y TECNOLÓGICAS

La revolución de la conciencia digital, impulsada por la ética de tercera generación en inteligencia artificial, está transformando profundamente la estructura social y los escenarios tecnológicos contemporáneos. Las implicaciones de este avance abarcan cambios significativos en las dinámicas de poder, el futuro del trabajo y la interacción humana, además de resaltar la necesidad urgente de desarrollar tecnologías que sean transparentes, seguras y sostenibles. En este contexto, el caso de Ana ejemplifica cómo la IA puede afectar la ontología del ser de los individuos: su relación con la tecnología no solo mejora su bienestar, sino que también redefine su percepción de autonomía y capacidad de decisión. A medida que Ana interactúa con la aplicación de IA para gestionar su salud mental, experimenta un empoderamien-

to que le permite entender y abordar sus desafíos, reflejando cómo la conciencia digital puede empoderar a los usuarios. Este análisis busca explorar cómo estas fuerzas están remodelando nuestro entorno y enfatiza la importancia de abordar las desigualdades existentes, así como de establecer estándares éticos que guíen el desarrollo tecnológico, orientándolos hacia un futuro más equitativo y responsable (Floridi *y otros.*, 2018; *Jobin y otros.,* 2019).

A continuación, se presenta la Tabla 3 que resume las principales implicaciones sociales y tecnológicas asociadas con el uso de la inteligencia artificial en el contexto de la conciencia digital, todo bajo el marco de la ética de tercera generación. Estas implicaciones destacan la influencia de la IA en la forma en que los individuos interactúan con la tecnología, así como su impacto en las estructuras sociales y en las normativas éticas que gobiernan la implementación de estas herramientas. Al considerar el papel de la conciencia digital en este ecosistema, se busca identificar tanto las oportunidades que se generan como los desafíos que se presentan para garantizar un desarrollo tecnológico responsable y equitativo.

Tabla 3. Implicaciones Sociales y Tecnológicas de la IA en la Conciencia Digital

Implicaciones Sociales	Descripción	Implicaciones Tecnológicas	Descripción
Cambio en las Dinámicas de Poder	La IA está alterando las relaciones de poder entre individuos, organizaciones e instituciones.	Aumento de la Automatización	La automatización de procesos puede generar inseguridad laboral y cambios en las dinámicas de empleo.
Empoderamiento del Usuario	La conciencia digital permite a los usuarios tomar decisiones informadas sobre sus interacciones tecnológicas.	Innovación en la Interfaz de Usuario	Se requieren interfaces más accesibles y comprensibles que fomenten una interacción responsable.

Implicaciones Sociales	Descripción	Implicaciones Tecnológicas	Descripción
Nuevas Normativas Éticas	La implementación de estándares éticos en el uso de IA es crucial para abordar las desigualdades.	Desarrollo de Tecnologías Inclusivas	Las tecnologías deben ser diseñadas para ser equitativas y accesibles para todos los grupos.
Responsabilidad Colectiva	Se enfatiza la importancia de la participación cívica en las decisiones sobre el uso de la IA.	Mecanismos de Transparencia	Es fundamental que las decisiones algorítmicas sean explicables y auditables para garantizar la confianza del usuario.
Desinformación y Manipulación	El uso de IA en la creación de contenido puede llevar a la proliferación de desinformación y noticias falsas.	Algoritmos de Filtrado de Contenido	Se necesita implementar tecnologías que verifiquen la veracidad de la información y contribuyan a un consumo crítico.
Impacto en la Salud Mental	La dependencia de tecnologías de IA puede afectar la salud mental de los usuarios debido a la sobrecarga de información y la falta de interacción humana.	Intervenciones basadas en IA para la Salud	Las aplicaciones de IA deben ser diseñadas considerando el impacto emocional y psicológico, buscando fomentar el bienestar en los usuarios.

Fuente: Elaborada a partir de Stahl (2024); Camaran Maione (2023).

La tabla anterior resalta la interconexión entre las implicaciones sociales y tecnológicas de la inteligencia artificial, así como cómo estas dimensiones pueden influirse mutuamente. En el ámbito social, se presentan desafíos significativos como la redistribución del poder y las transformaciones en el empleo, consecuencias directas de los avances rápidos en la tecnología. A su vez, estos desarrollos tecnológicos requieren la creación de sistemas que sean éticamente alineados y que incorporen herramientas como algoritmos transparentes y estándares

globales. Esta necesidad es fundamental para asegurar un impacto positivo en la sociedad y prevenir la perpetuación de desigualdades existentes. Además, abordar la desigualdad social y mantener la seguridad de los datos enfatiza la urgencia de implementar una ética de tercera generación que sea tanto robusta como adaptable. Este enfoque equilibrado se convierte en un pilar esencial para guiar el desarrollo de la inteligencia artificial hacia un futuro que no solo sea innovador, sino también justo y sostenible para todos los sectores de la sociedad.

Cada una de estas barreras representa un aspecto crítico que debe abordarse para lograr una implementación ética y efectiva de la inteligencia artificial dentro de un ecosistema responsable. La transparencia y la explicabilidad son esenciales para fomentar la confianza del usuario, asegurando que los resultados generados por los sistemas de IA sean comprensibles y revisables por los humanos (Floridi *y otrosl*., 2018). Al mismo tiempo, la protección de la privacidad y la seguridad de los datos destaca la necesidad de establecer mecanismos robustos que resguarden la información personal, evitando abusos que puedan erosionar la confianza pública en la tecnología (Solove, 2020). Además, las cuestiones de equidad y sesgo algorítmico enfatizan la importancia de desarrollar algoritmos que no perpetúen injusticias ni discriminen a grupos minoritarios o subrepresentados, garantizando así que todos los sectores de la sociedad se beneficien equitativamente de los avances tecnológicos (Mehrabi *y otros*., 2021).

La identificación de estos desafíos también invita a reflexionar sobre las implicaciones a largo plazo de la IA en el ámbito ético y social. Es probable que las decisiones guiadas por inteligencia artificial evolucionen de acciones meramente basadas en datos hacia un enfoque que integre principios y objetivos humanísticos (Jobin *y otros*., 2019). Para lograr esto, es fundamental que un marco ético robusto oriente la evolución del desarrollo tecnológico. Esto exige metodologías que permitan auditar y supervisar continuamente las prácticas tecnológicas, al tiempo que se reconoce tanto el potencial transformador de la IA como sus limitaciones (Floridi y Cowls, 2019). La integración de estos principios garantizará que la inteligencia artificial no solo avance en eficiencia, sino que también promueva la equidad, la transparencia y la sostenibilidad, contribuyendo a un futuro en el que la tecnología

se utilice de manera ética y responsable para el bienestar social y la justicia intergeneracional. A continuación, se presenta la Figura 7, que ilustra el futuro de la IA en contextos éticos, reforzando la necesidad de un enfoque consciente y transformador. (**Ver Figura 7**)

Figura 7. Futuro de la IA en contextos Éticos.



Fuente: Napkin.ai (2024)

La ética de tercera generación propone un cambio de paradigma crucial en la intersección entre la inteligencia artificial y la sociedad, al superar los enfoques preventivos tradicionales y establecer un marco normativo centrado en los principios de equidad, justicia y transparencia. Este enfoque avanzado exige que los sistemas automatizados no solo operen bajo regulaciones predefinidas, sino que también se adapten a los contextos morales y culturales específicos en los que son

implementados. El desafío de equilibrar la autonomía humana con la capacidad de intervención de la IA se erige como una cuestión fundamental, donde la confianza y la responsabilidad compartida adquieren una dimensión estratégica para garantizar interacciones éticas (Stahl, 2024) y cultivar una conciencia digital efectiva.

La evolución de la inteligencia artificial estará intrínsecamente vinculada al desarrollo de metodologías éticas más sofisticadas que permitan auditar, monitorear y ajustar su impacto en la sociedad. La integración de valores humanísticos en los sistemas de IA, especialmente en áreas sensibles como la justicia y la salud, se torna indispensable para la creación de un entorno tecnológico que respete los derechos humanos y promueva el bienestar social. Este proceso de alineación ética, al considerar tanto el potencial como las limitaciones inherentes de la IA, marca el inicio de una transformación donde la eficiencia tecnológica se subordina al imperativo de una responsabilidad social sólida (Veliyev, 2024).

La implementación efectiva de estos principios éticos no solo configura las interacciones actuales, sino que también establece una base para el futuro de la inteligencia artificial. Un enfoque reflexivo, que priorice la equidad y la transparencia, permitirá que la IA se convierta en una herramienta al servicio de la sociedad, garantizando que su desarrollo tecnológico contribuya al bienestar común y mitigando los riesgos de discriminación y desigualdad. Este enfoque es esencial para construir un ecosistema digital responsable, donde la inteligencia artificial no solo sea eficiente, sino también ética y socialmente comprometida.

EL FUTURO DEL CASO DE ANA EN RELACIÓN A LA CONCIENCIA DIGITAL, LA IA Y LA ÉTICA DE TERCERA GENERACIÓN PRESENTA DOS ESCENARIOS CONTRASTANTES

En un **escenario optimista**, Ana podría beneficiarse enormemente de la continua evolución de la IA y su integración con una conciencia digital sólida y bien informada. A medida que la ética de tercera generación siga siendo un pilar fundamental en el desarrollo de estas

tecnologías, se espera que las aplicaciones de IA se enfoquen en brindar apoyo personalizado sin sacrificar la privacidad ni la equidad. En este contexto, Ana utilizaría la IA no solo como una herramienta de gestión de su salud mental, sino también como un recurso que fortalece su autonomía y su capacidad para entender y afrontar sus desafíos personales. La transparencia en el uso de datos y un diseño ético de la IA le permitirían a Ana tomar decisiones informadas sobre su salud, construyendo confianza y empoderamiento. Así, su experiencia podría reflejar el potencial positivo de un ecosistema digital que promueve la justicia social y el bienestar individual en un marco ético robusto.

Por otro lado, un **escenario pesimista** podría surgir si las tecnologías de IA se desarrollan sin una integración adecuada de principios éticos y una conciencia digital crítica. Ana podría volverse dependiente de la aplicación de IA para gestionar su bienestar, lo que resultaría en una erosión de su autonomía y la capacidad de tomar decisiones informadas. En este contexto, si la IA perpetúa sesgos y no toma en cuenta las particularidades individuales, las recomendaciones podrían ser inadecuadas o perjudiciales, exacerbando sus problemas en lugar de mitigarlos. Además, la falta de transparencia en el uso de sus datos personales podría socavar la confianza en la tecnología, llevando a una mayor desconexión entre Ana y las herramientas que deberían ayudarla. Este escenario ilustra una distopía donde la innovación tecnológica no se traduce en resultados positivos, destacando la importancia vital de un compromiso ético continuo en el diseño y la implementación de la inteligencia artificial.

Ambos escenarios subrayan la necesidad de enfocar el desarrollo y la implementación de la IA en el marco de la conciencia digital y la ética de tercera generación, enfatizando que el futuro de la tecnología debe ser construido de manera responsable y equitativa, asegurando que se maximicen los beneficios y se minimicen los riesgos. (**Ver Figura 8**)

***Figura* 8.** Escenarios futuros de la concienca digital de Ana



Fuente: *CHAT GPT4o* **(2024)**

CONCLUSIONES

¿Estamos preparados para un futuro en el que la inteligencia artificial no solo nos asista en la toma de decisiones, sino que también nos obligue a redefinir lo que significa ser humano? La revolución en la conciencia digital nos ha colocado en un punto de inflexión, un momento crucial para decidir si la tecnología será nuestra aliada o nos dominará.

No podemos permitirnos el lujo de ver la IA como un ente neutral. Esta revolución exige una transformación profunda en nuestra concepción del desarrollo tecnológico. La conciencia digital, con su enfoque crítico y reflexivo, junto con la ética de tercera generación, que promueve una visión holística y un compromiso firme con la sostenibilidad, se erigen como pilares esenciales para construir un futuro en el que la tecnología actúe como un motor de progreso social y de bienestar, en lugar de representar una amenaza.

Se ha acabado la era de la complacencia; es imperativo dejar atrás visiones simplistas y asumir el reto de erigir un nuevo paradigma tec-

nológico, donde la ética no sea una mera guía opcional, sino un fundamento inalterable. La inteligencia artificial debe servir a la humanidad y no al revés.

Es hora de abandonar las estructuras de poder tradicionales y empoderar a los individuos con una conciencia digital sólida. Se debe trabajar en el desarrollo de sistemas de IA que sean transparentes, seguros y responsables, evitando socavar la confianza pública y combatiendo la desigualdad. Es necesario crear un futuro en el que la innovación tecnológica se alinee con los valores humanos fundamentales, garantizando que los beneficios de la IA se distribuyan de manera justa y equitativa entre todas las generaciones.

No se debe olvidar que la IA es un río poderoso. Puede ser una fuente de vida y prosperidad, pero también puede convertirse en un torrente destructor. La responsabilidad recae en la sociedad; de cada uno depende canalizar su poder con inteligencia y ética, asegurando que fluya en la dirección correcta, hacia un futuro más humano y sostenible.

¡El futuro ya está aquí! Es tiempo de dar el siguiente paso: construir un nuevo mundo donde la tecnología sea una fuerza para el bien. ¡Despertemos la revolución de la conciencia digital!

REFERENCIAS BIBLIOGRÁFICAS

Akgun, T., y Greenhow, A. (2021). Financial Algorithms and Socioeconomic Disparities: A Critical Analysis. *Journal of Financial Technology and Society, 5***(2), 102-118.**

Akgun, S., y Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. AI and Ethics, 2(3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, Proceedings of Machine Learning Research, 81, 149-159.

Brey, P. A. (2012). Anticipating ethical issues in emerging IT. *Ethics and Information Technology*, *14*, **305-317. Doi: 10.1007/s10676-012-9293-y**

Brundage, M., Avin, S., Wang, J., Belfield, H., Kruger, A., Hadfield, G., y Zeerak, A. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv preprint arXiv:1802.07228. *https://arxiv.org/abs/1802.07228*

Camaran Maione, S. (2023). *Conciencia digital profesional para la transformación emergente del ser*. Revista Venezolana de Análisis de Coyuntura, 29(1), 59-78.

Cengel, Y. A., y Boles, M. A. (2019). Thermodynamics: An Engineering Approach. McGraw-Hill Education.

CLAUDE AI (2024). Página oficial. *https://claude.ai/new*

CONNECTED PAPERS. (2024). *Página oficial*. https://www.connectedpapers.com/

de Fine Licht, K., y de Fine Licht, J. (2020). Transparency in Algorithmic Decision-Making: Opacity, Accountability, and Public Trust. *Public Administration Review, 80***(5), 734-745.**

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

Floridi, L. (2014). The Fourth Revolution: How the Infosphere is Reshaping Human Reality. Oxford University Press.

Floridi, L., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707. *https://doi.org/10.1007/s11023-018-9482-5*

Floridi, L., y Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review***.** *https://hdsr.mitpress.mit.edu/pub/l0jsh9d1/release/8*

Gebru, T., *y otros.* **(2021). Datasheets for Datasets.** *Communications of the ACM, 64***(12), 86-92.**

Heylighen, F., y Lenartowicz, M. (2017). The Global Brain as a model of the future information society: An introduction to the special issue. *Technological Forecasting and Social Change***,** *114***, 1-6.** *https://doi.org/10.1016/j.techfore.2016.10.063*

Howard, J. (2019). Artificial intelligence: Implications for the future of work. *American Journal of Industrial Medicine*, *62*(11), 917-926. *https://doi.org/10.1002/ajim.23037*

Jobin, A., Ienca, M., y Vayena, E. (2019). Artificial Intelligence: The global landscape of ethics guidelines. *arXiv preprint arXiv:1902.09641*.

Marchant, G. E., Allenby, B. R., Herkert, J. R., y Peña, W. (2011). The growing gap between emerging technologies and legal-ethical oversight: The pacing problem. In The Role of Law, Ethics, and Engineering Culture (pp. 3-34). Dordrecht: Springer. *https://link.springer.com/book/10.1007/978-94-007-1356-7*

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., y Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys (CSUR), 54(6), 1-35. https://doi.org/10.48550/arXiv.1908.09635

Napkin AI (2024). Página oficial. https://www.napkin.ai/

OpenAI. (2024). *ChatGPT4o [Modelo de lenguaje de gran tamaño]. https://chat.chatbotapp.ai/chats*

Raghupathi, W., y Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2(1), 3. DOI:*10.1186/2047-2501-2-3*

Solove, D. J. (2020). The myth of the privacy paradox. George Washington Law Review, 89, 1-39. *https://doi.org/10.2139/ssrn.3191274*

Shiller, D. (2024). Functionalism, integrity, and digital consciousness. *Synthese*, 203(1), 47-69. *https://doi.org/10.1007/s11229-023-04473-z*

Stahl, B. C. (2021). From computer ethics and the ethics of AI towards an ethics of digital ecosystems. *AI and Ethics*, *2*(1), 65-77. *https://doi.org/10.1007/s43681-021-00080-1*

Stahl, B. C. (2024). From Corporate Digital Responsibility to Responsible Digital Ecosystems. *Sustainability*, *16*(12), 4972. *https://doi.org/10.3390/su16124972*

Stahl, B. C. (2024). Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies. Springer Nature. https://link.springer.com/book/10.1007/978-3-030-69978-9

Strubell, E., Ganesh, A., y McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. arXiv preprint arXiv:1906.02243. *https://arxiv.org/abs/1906.02243*

Ugwudike, P. (2022). Justice in AI: Addressing the Challenges of Bias and Inequity. *Journal of Artificial Intelligence and Ethics, 2***(3), 45-67.**

Veliyev, D. J. (2024). Digital consciousness and identity. *SKHID (EAST), 6*(1), 50-55. *Digital Consciousness and Identity | Skhid (kubg.edu.ua)*

White, G. R. T., Samuel, A., Jones, P., Madhavan, N., Afolayan, A., Abdullah, A., y Kaushik, T. (2024). Mapping the ethic-theoretical foundations of artificial intelligence research. *Thunderbird International Business Review*. *https://doi.org/10.1002/tie.22368*

CAPÍTULO XVI

EGOLATRÍA Y EVOLUCIÓN: LA TRANSFORMACIÓN DE LA CONDICIÓN HUMANA

ALFREDO GARIBAY SUÁREZ[1]

INTRODUCCIÓN

En un mundo donde la tecnología avanza a pasos agigantados, el concepto de ser humano se encuentra en constante revaluación. La llegada del transhumanismo, una corriente filosófica que busca mejorar la condición humana a través de la ciencia y la tecnología, plantea preguntas fundamentales sobre la identidad, la ética y el futuro de la humanidad. A medida que herramientas como la inteligencia artificial, la biotecnología y las interfaces cerebro-computadora se integran en nuestra vida cotidiana, surgen desafíos significativos en la forma en que percibimos nuestra existencia y nuestras relaciones interpersonales. Este texto explora cómo la agencia de la egolatría y la incorporación de la tecnología desmaterializan la identidad del ser, llevando a la humanidad hacia un estado de dependencia tecnológica que redefine no solo nuestra corporalidad, sino también nuestra esencia. Mediante un análisis crítico de los impactos y efectos del transhumanismo, se busca comprender cómo estas transformaciones afectan el tejido social y generan nuevas dinámicas en la interacción humana. En última instancia, nos enfrentamos a un dilema existencial: ¿Estamos avanzando hacia una mejora significativa o estamos abriendo la puerta a una deshumanización que compromete nuestra esencia más profunda?

Palabras clave: Transhumanismo, Identidad, Tecnología, Deshumanización, Ética

[1] Profesor-Investigador del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco

ADVERTENCIA

El ensayo que se presenta ha sido integrado a renglón seguido desde la capacidad lírica del autor. Elementos como el título, la introducción, las tablas 1 y 2 y la bibliografía fueron definidos bajo el algoritmo de *ChatGPT*. Se espera de este ejercicio localizar su figura y contribución a partir de la sentencia provocadora de lo que se intitula.

La agencia de la egolatría desmaterializa dialécticamente la identidad del ser frente al empoderamiento de la integración tecnológica, a la distancia de uno a dos milímetros de grosor de la piel o de hasta cinco veces ello en el cráneo la tecnología acecha. El artefacto ha lacerado su cordón de señal continua y de energía, para a través del espectro y la nanotecnología guardar una nueva independencia y travesía y en ello, hemos identificado una evolución potenciada de irrealidad física pero multisensorial de expectativa.

Los hechos tecnológicos convergen en un ser que hoy se empodera de facultad, en el último lustro las disciplinas han transmutado a un estadio infectado de digitalidad y conectividad, a un reino de datos que exuberantemente desquicio la racionalidad paradigmática a partir del algoritmo inspirado.

Tabla 1. Tabla priorizada por impacto que muestra una taxonomía de hechos tecnológicos en el último lustro que han contribuido a potenciar el transhumanismo.

Prioridad	Categoría	Hecho Tecnológico	Descripción	Impacto en el Transhumanismo
1	*Inteligencia Artificial*	Avances en IA y *machine learning*	Mejora en algoritmos de IA permite la creación de sistemas que pueden aprender y adaptarse, aplicándose en salud y mejora cognitiva.	Potencia la mejora de habilidades cognitivas y la automatización.

Prioridad	Categoría	Hecho Tecnológico	Descripción	Impacto en el Transhumanismo
2	*Biotecnología*	Edición genética CRISPR	La tecnología CRISPR ha permitido la edición genética precisa, abriendo posibilidades para la modificación genética en humanos.	Facilita la mejora y modificación genética.
3	*Interfaces Cerebro-Computadora*	Desarrollo de interfaces cerebro-computadora	Dispositivos que permiten la conexión directa entre el cerebro humano y computadoras, facilitando el control de dispositivos mediante el pensamiento.	Integra la mente humana con máquinas, ampliando capacidades.
4	*Robótica*	Avances en prótesis y exoesqueletos	Creación de prótesis avanzadas y exoesqueletos que mejoran la movilidad y funcionalidad de personas con discapacidades físicas.	Mejora la capacidad física y funcional de los individuos.
5	*Nanotecnología*	Desarrollo de nanobots	Creación de nanobots utilizados en medicina para la administración de fármacos y reparación celular.	Aumenta la capacidad de intervención en procesos biológicos.
6	*Salud Digital*	Telemedicina y monitoreo continuo	Implementación de tecnologías de salud digital que permiten el monitoreo remoto de la salud y la consulta médica a distancia.	Facilita el acceso a la atención médica y mejora la salud.

Prioridad	Categoría	Hecho Tecnológico	Descripción	Impacto en el Transhumanismo
7	*Realidad Aumentada y Virtual*	Creación de entornos inmersivos	Tecnologías que permiten experiencias inmersivas y la interacción en entornos virtuales, simulando realidades físicas o creando nuevas realidades.	Expande la percepción de la realidad y las experiencias humanas.
8	*Criptomonedas y Blockchain*	Desarrollo de criptomonedas y contratos inteligentes	Expansión de criptomonedas y uso de blockchain para crear registros seguros y transparentes de transacciones y contratos.	Fomenta la descentralización y nuevas economías de intercambio.

Fuente: Creación por IA basada en algoritmos de GPT.

Hemos transmutado, hemos adoptado una ruleta de vaticinios insultantes de posibilidad para encontrarnos con economías no fiduciarias, idealizamos y protagonizamos la realidad deseada, aunque esto último vulnere nuestra corporalidad ¡que más da! la nanotecnología nos remite a la reparación incluso celular o en la robótica y la interacción neuronal nos empodera.

Ahora somos mercancía de nuestra posibilidad, la diferenciación y distinción se decanta por la capacidad de adoptar aplicativos y artefactos, por el vínculo mercantil con un ecosistema de consumo exacerbado o en el peor de los casos por la rendición y enajenación del marketing impositivo ante el servicio *freemium*; los efectos y consecuencias también son vastas desde el crisol transhumanista.

Tabla 2. Tabla de los efectos de la vida transhumanista que impactan la comunidad y el tejido social.

Efecto	Descripción	Impacto en la Comunidad y Tejido Social
1. Aumento de la Desigualdad	El acceso a tecnologías avanzadas puede estar limitado a ciertos grupos socioeconómicos.	Puede aumentar la brecha entre los que pueden permitirse mejoras y los que no.
2. Cambios en la Identidad	La modificación genética y la integración de tecnología en el cuerpo pueden alterar la percepción del yo.	Puede llevar a una crisis de identidad y a nuevas formas de autoexpresión.
3. Nuevas Dinámicas Sociales	La interacción entre humanos y máquinas puede cambiar cómo se forman las relaciones interpersonales.	Puede desestabilizar las estructuras sociales tradicionales y crear nuevas formas de conexión.
4. Ética y Moralidad	La capacidad de modificar a los humanos plantea nuevas preguntas éticas sobre la intervención en la naturaleza humana.	Puede generar debates y divisiones sobre lo que es aceptable y lo que no.
5. Salud y Bienestar	El acceso a tecnologías que mejoran la salud puede cambiar las expectativas sobre el bienestar.	Puede fomentar un enfoque más preventivo en la salud, pero también puede generar presión social por la "mejora".
6. Integración Tecnológica	La incorporación de tecnología en la vida diaria puede hacer que la dependencia de dispositivos aumente.	Puede transformar la manera en que se vive, trabaja y se interactúa socialmente.
7. Diversidad de Experiencias	Las tecnologías de realidad virtual y aumentada permiten vivir experiencias personalizadas y únicas.	Puede enriquecer la cultura, pero también crear desconexiones entre diferentes grupos.
8. Nuevas Formas de Educación	La educación puede transformarse con el uso de IA y plataformas digitales, permitiendo un aprendizaje más personalizado.	Puede democratizar el acceso al conocimiento, pero también desestabilizar los modelos educativos tradicionales.
9. Redefinición del Trabajo	La automatización y las tecnologías avanzadas pueden cambiar radicalmente el mercado laboral.	Puede provocar desempleo en ciertos sectores, pero también la creación de nuevos roles.

Efecto	Descripción	Impacto en la Comunidad y Tejido Social
10. Alteración de la Familia	Las nuevas formas de vida y relaciones pueden cambiar la estructura familiar tradicional.	Puede llevar a nuevos modelos de familia y a redefinir los roles dentro de ella.

Fuente: Creación por IA basada en algoritmos de GPT.

Trascendemos desde un rebote tecnológico de deseabilidad y éxtasis, alineados y alienados, el alter ego se institucionaliza en el avatar y en mi mismo como agente experto, trasciendo como conciencia y evoco la discusión con mi representación de cadena de bits; y ello lo replico al yo como tribu, como casta, como entidad, como federación; todo ello mientras me feudalizo y reconozco al terrateniente en función de su locura plenipotenciaria y seducción discursiva.

Me deshumanizo, se trastorna mi identidad y en la posibilidad de la inmortalidad me encuentro con seres no de me árbol genealógico sino del económico, mi empatía es doctrina digital del encanto de sirenas de silicio dentro de mi ser, se dispenden metódica y perfectamente las dosis que en la singularidad consternan mis sentidos, la droga me acerca al nirvana plástico que se impone ante la espiritualidad y la fe verdadera; me pregunto ¿Quién soy? y ¿Quién definió el estándar de lo que soy?

REFERENCIAS SUGERIDA PARA EL TEXTO SEGÚN CHATGPT

Bostrom, Nick. (2005). *A History of Transhumanist Thought*. *Journal of Philosophical Research*, 30, 3-24.

Kurzweil, Ray. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking.

Harari, Yuval Noah. (2016). *Homo Deus: A Brief History of Tomorrow*. Harper.

Noor, M. y Bansal, S. (2021). *Ethical Implications of Transhumanism and Human Enhancement*. *Bioethics*, 35(5), 405-411.

Miller, M. (2014). *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*. Wiley.

Haraway, Donna. (1991). *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge.

Posthumanism Conference. (2017). *The Ethics of Human Enhancement: Transhumanism and its Critics. Journal of Posthuman Studies*, 1(1), 11-30.

Sparrow, Robert. (2011). *Ethics of Human Enhancement: The Good, the Bad, and the Ugly*. *Journal of Medicine and Philosophy*, 36(6), 541-559.

Bauman, Zygmunt. (2000). *Liquid Modernity*. Polity Press.

Rifkin, Jeremy. (2011). *The Third Industrial Revolution: How Lateral Power is Transforming Energy, the Economy, and the World*. St. Martin's Press.

CAPÍTULO XVII

ÉTICA KANTIANA EN LA EVALUACIÓN DEL RIESGO CREDITICIO: APLICACIONES DE INTELIGENCIA ARTIFICIAL Y MACHINE LEARNING

LUIS ANTONIO VERDUGO CALDERÓN[1]

RESUMEN

La aplicación de la ética kantiana en la evaluación crediticia, a través de la inteligencia artificial (IA) y el *machine learning* (ML), representa una oportunidad única para transformar el sector financiero. Este capítulo analiza los principios de justicia, equidad y transparencia en el desarrollo de modelos de evaluación crediticia, y explora cómo estos principios pueden contribuir a un sistema financiero más justo, inclusivo y responsable. Asimismo, se examinan los riesgos asociados con la falta de transparencia en los algoritmos, la persistencia de sesgos en los datos de entrenamiento y la ausencia de una regulación internacional clara. En este contexto, se hace un llamado a desmantelar el enfoque reduccionista que ve la tecnología como un mero instrumento de optimización financiera. Es fundamental ir más allá de la simple implementación tecnológica y profundizar en la integración de valores universales en la toma de decisiones. La propuesta sugiere la construcción de un sistema financiero donde la ética sea un componente fundamental para crear un entorno más justo y humano. Las conclusiones y futuras direcciones enfatizan la importancia de desarrollar estándares internacionales para la IA en el sector financiero, así como

[1] Estudiante de Doctorado en Innovación y Administración de las Organizaciones, modalidad virtual de la Universidad Autónoma de Occidente, Los Mochis, Sinaloa ,México

la necesidad de contar con herramientas de auditoría más sofisticadas. También se subraya la relevancia de investigar la interacción humano-máquina y de promover programas de educación ética dirigidos a desarrolladores, usuarios y reguladores de la IA.

Palabras clave: Ética Kantiana, Evaluación Crediticia, Machine Learning

INTRODUCCIÓN

En este contexto, la inteligencia artificial (IA) y el *machine learning* (ML) han emergido como fuerzas transformadoras, especialmente en el ámbito de la evaluación crediticia. Estas innovaciones permiten procesar inmensas cantidades de datos y discernir patrones complejos, aumentando la precisión y eficiencia de las evaluaciones de riesgo. Sin embargo, el poder de estas tecnologías también trae consigo una serie de desafíos éticos que requieren un análisis cuidadoso.

El *machine learning*, un subconjunto de la IA, equipa a las máquinas con la capacidad de aprender de los datos y mejorar sus operaciones sin intervención humana directa (Russell y Norvig, 2020). Esta capacidad no solo acelera la toma de decisiones, sino que también mejora su adaptabilidad ante cambios en las condiciones del mercado, un avance crucial en el marco del libro "*El Futuro de la IA: Transhumanidad, Ética y Sostenibilidad*".

No obstante, la implementación de algoritmos de ML en finanzas genera inquietudes sobre la transparencia y equidad de las decisiones automatizadas. Los algoritmos, que con frecuencia operan como cajas negras, pueden perpetuar y magnificar sesgos inherentes en sus datos de entrenamiento, resultando en decisiones financieras que podrían afectar negativamente a ciertos grupos (Chouldechova y Roth, 2018). Este reto ético demanda un enfoque filosófico robusto y bien fundamentado.

La ética kantiana, con su énfasis en principios morales universales y la dignidad humana, ofrece un marco valioso para enfrentar estos di-

lemas. Fundamentada en el imperativo categórico de Immanuel Kant, esta perspectiva se centra en la ética de las intenciones y la universalización de principios morales, más allá de las consecuencias de las acciones. En la aplicación de IA y ML, adoptar principios kantianos implica diseñar sistemas que no solo sean eficientes, sino que también respeten los estándares universales de justicia e imparcialidad, preservando la autonomía y dignidad de todos los individuos involucrados (Johnson y Cureton, 2024).

Así, este capítulo tiene como objetivo explorar cómo los principios de la ética kantiana pueden integrarse en el uso de ML e IA para la evaluación de riesgos crediticios, promoviendo prácticas más justas y responsables en el sector financiero global. Para alcanzar este objetivo, se utilizará un enfoque deductivo que comenzará con un análisis de los principios éticos kantianos, seguido de un examen crítico de las aplicaciones actuales, y culminará con una propuesta de un marco ético-práctico.

La estructura del capítulo se organiza en cinco secciones principales. La primera sección examina los principios de la ética kantiana aplicados a IA y ML. La segunda sección analiza críticamente las prácticas actuales de evaluación crediticia con estas tecnologías. A continuación, se propone un marco ético que integra principios kantianos en el contexto financiero-tecnológico. Finalmente, el capítulo concluye con una recapitulación de los hallazgos y recomendaciones para futuras investigaciones, centrándose en la sostenibilidad y la responsabilidad

ÉTICA KANTIANA APLICADA A LA IA Y MACHINE LEARNING

La ética kantiana, fundamentada en principios universales como el imperativo categórico, la autonomía y la dignidad humana, proporciona un marco filosófico que demanda acciones guiadas por intenciones morales sólidas e inquebrantables (Johnson, 2008). Al considerar la expansión de la inteligencia artificial (IA) y, de manera particular, el machine learning (ML) en el sector financiero, se plantea la necesidad urgente de una supervisión ética basada en estos principios. A medida que el ML, como una rama de la IA, mejora la eficiencia y precisión en

procesos tales como la evaluación de riesgos crediticios, también suscita interrogantes esenciales sobre cómo integrar estas tecnologías de manera que respeten y promuevan la equidad, la justicia y la dignidad humana. Estos valores deben ser el tronco común que sostiene y guía el crecimiento de cualquier innovación tecnológica, asegurando que la eficiencia no se consiga a expensas de los principios éticos fundamentales. (**Ver Figura 1**)

Figura 1: Principios Universales de la Ética Kantiana



Fuente: CHAT GPTo (2024)

La Figura 1, que ilustra los principios de la ética kantiana proporciona una representación visual de cómo los conceptos de imperativo categórico, autonomía y dignidad humana se interrelacionan para formar un marco ético integral. Estos principios, representados en la figura, actúan como guías fundamentales para calibrar la moralidad de las acciones humanas y tecnológicas. En el ámbito de la IA y el ML, estos principios no solo aseguran que las innovaciones tecnológicas sean diseñadas con un enfoque

en la eficiencia y la precisión, sino también en la justicia y la equidad. Al considerar cómo estos fundamentos éticos se pueden aplicar a las tecnologías emergentes, es imprescindible que el desarrollo y la implementación de IA y ML no comprometan los valores morales fundamentales que son esenciales para el bienestar humano. Avanzando hacia el siguiente subapartado, se profundiza en cómo estos conceptos clave de la ética kantiana se pueden aplicar específicamente en el contexto tecnológico, asegurando que los algoritmos y sistemas, más allá de su funcionalidad, respeten y promuevan estos principios universales.

CONCEPTOS CLAVE DE LA ÉTICA KANTIANA APLICADOS A LA TECNOLOGÍA

La ética kantiana, al centrarse en principios universales como el imperativo categórico, la autonomía y la dignidad humana, proporciona un fundamento axiológico robusto para juzgar la moralidad de las acciones humanas. Estos principios son esenciales para garantizar que las tecnologías emergentes, como la inteligencia artificial (IA) y el machine learning (ML), sean desarrolladas e implementadas de manera que respeten valores morales fundamentales. La axiología kantiana insiste en que las acciones deben ser moralmente justificables independientemente de sus consecuencias, enfatizando la importancia de las intenciones detrás de las decisiones tecnológicas. En el contexto del ML, esto se traduce en el establecimiento de algoritmos que no sólo prioricen la eficiencia y la precisión, sino que también incorporen la justicia y la equidad en su núcleo.

Por otro lado, el pragmatismo inherente al machine learning, con su énfasis en la funcionalidad y la adaptabilidad, trae consigo el riesgo de dejar de lado estas consideraciones éticas si no se les da la debida atención en el proceso de diseño. Integrar los conceptos kantianos en la tecnología requiere un enfoque que combine análisis críticos con una visión holística, asegurando que los avances en ML no comprometan principios éticos universales. Es fundamental que estas tecnologías se empleen no solo como herramientas de mejora de procesos, sino también como extensiones de los valores humanos que buscan promover un entorno de equidad y respeto por la dignidad.

A continuación, la Tabla 1 detalla un análisis de cómo los conceptos clave de la ética kantiana pueden ser aplicados en el diseño y uso de machine learning, proporcionando un marco para alinear estas tecnologías con valores éticos fundamentales. (**Ver Tabla 1**)

Tabla 1. Aplicación de Conceptos Éticos Kantianos en IA y Machine Learning

Concepto Clave	Definición	Análisis	Visión Holística
Imperativo Categórico	Principio que establece que las acciones deben basarse en máximas que puedan ser adoptadas como leyes universales.	Los algoritmos deben diseñarse para ser justos e imparciales, asegurando el trato igualitario a todos los usuarios.	Promueve un entorno donde los algoritmos aborden sesgos desde su diseño, fomentando la equidad y justicia.
Autonomía	Capacidad de los individuos para tomar decisiones informadas y racionales, actuando bajo sus propias leyes.	Los sistemas de IA deben ser transparentes, ofreciendo explicaciones claras para mantener el control de los usuarios.	La autonomía implica un diseño colaborativo humano-máquina que complemente, no reemplace, el juicio humano.
Dignidad Humana	La idea de que cada persona es un fin en sí mismo y debe ser tratada con respeto inherente.	Requiere un manejo ético de los datos personales, asegurando privacidad y consentimiento informado.	Establece la importancia de crear ecosistemas que respeten y potencien la dignidad humana.
Relevancia General	Estos principios éticos aseguran que las tecnologías se implementen de manera ética y responsable.	Guían la implementación de sistemas justos y responsables, alineando soluciones tecnológicas con valores humanos fundamentales.	Permiten integrar valores humanos en el diseño de tecnología, promoviendo un equilibrio entre eficiencia e integridad ética.

Fuente: Elaborada a partir de Barocas, Hardt, y Narayanan (2019); Dignum (2018); Floridi (2019); Jobin, Ienca, y Vayena (2019); Mittelstadt, Wachter, y Floridi (2016).

La Tabla 1, ofrece un análisis comprensivo de cómo los conceptos fundamentales de la ética kantiana pueden integrarse efectivamente en el desarrollo y uso del machine learning. Al desglosar principios como el imperativo categórico, la autonomía y la dignidad humana, la tabla ilustra no solo su aplicación teórica, sino también prácticas específicas que fomentan la justicia, la transparencia y el respeto en la implementación tecnológica. Este análisis refleja cómo un marco ético sólido puede impulsar la creación de algoritmos que son tanto eficaces como moralmente responsables, asegurando que las innovaciones en IA no comprometan valores humanos esenciales. A través de la consideración conjunta de estos principios, se invita a los desarrolladores y tomadores de decisiones a reimaginar tecnologías no solo como herramientas de eficiencia, sino como plataformas que pueden potenciar y reflejar los ideales éticos más elevados.

IMPORTANCIA DE LA ÉTICA KANTIANA EN EL DISEÑO DE ALGORITMOS DE IA

La aplicación de los principios éticos kantianos en el diseño de algoritmos de inteligencia artificial es esencial para asegurar que las decisiones automatizadas sean no solo eficientes, sino también justas y responsables. Estos principios éticos ayudan a superar la percepción de la IA como una "*caja negra*" mediante la promoción de la transparencia y la explicabilidad de los sistemas. Es crucial que los usuarios comprendan cómo los algoritmos llegan a sus conclusiones, ya que esto refuerza la confianza en el sistema y facilita la identificación y corrección de posibles sesgos (Rudin, 2019).

La transparencia en el diseño de algoritmos de inteligencia artificial no solo se refiere a la capacidad de los usuarios para entender cómo funcionan los sistemas, sino que también es un componente crítico para establecer la confianza y legitimidad en la tecnología. En el contexto de los principios kantianos, la transparencia se alinea estrechamente con la noción de autonomía, permitiendo que los individuos tomen decisiones informadas y comprendan el impacto de las soluciones tecnológicas en sus vidas (Floridi, 2019). Esta claridad es esencial

para desmitificar la IA como una "caja negra", un fenómeno donde los procesos internos son opacos y las decisiones parecen resultar sin una lógica aparente para los usuarios finales.

El fortalecimiento de la transparencia implica proporcionar a los usuarios no solo acceso a explicaciones claras de cómo funcionan los algoritmos, sino también una comprensión de los criterios y datos que se utilizan en su toma de decisiones. Esto es especialmente relevante en aplicaciones críticas, como la evaluación crediticia, donde las decisiones pueden tener impactos significativos en la vida de las personas (Rudin, 2019). Al hacer que los algoritmos sean más accesibles y comprensibles, se fomenta un entorno donde la confianza se puede construir sobre una base de comprensión mutua y responsabilidad.

Además, la transparencia contribuye a la facilitación de auditorías externas y revisiones continuas, lo cual es una práctica esencial para identificar y corregir sesgos potenciales en los sistemas de IA. Este enfoque proactivo no solo protege a los usuarios de resultados injustos, sino que también impulsa la innovación responsable, alentando a los desarrolladores a considerar el contexto ético desde el principio del ciclo de desarrollo tecnológico.

Así, incorporar la transparencia como un pilar en el diseño de la IA no solo se alinea con el respeto kantiano por la autonomía, sino que también fortalece la legitimidad de las aplicaciones tecnológicas a largo plazo, garantizando que el avance tecnológico se mantenga anclado a valores éticos fundamentales que beneficien a la sociedad en general.

Además, debe prestarse especial atención a la equidad en los datos utilizados para entrenar los algoritmos de IA, ya que los sistemas de IA se desarrollan principalmente a partir de datos históricos, que pueden reflejar las inequidades y sesgos presentes en la sociedad. Si estos datos no se seleccionan y gestionan con cuidado, existe el riesgo de perpetuar y amplificar las desigualdades existentes. Es crucial, por lo tanto, que los conjuntos de datos sean cuidadosamente evaluados y seleccionados para reflejar una diversidad justa y representativa de las poblaciones afectadas.

La equidad de los datos se alinea con el imperativo categórico de Kant, que enfatiza que las decisiones deben ser justas y aplicables uni-

versalmente. Para lograr esto, los desarrolladores de IA deben garantizar que sus modelos sean entrenados con datos que reflejen todas las características relevantes de la población, sin dar una representación desproporcionada a ciertos grupos. Esto implica una revisión crítica de los procesos de recolección y curation de datos para evitar sesgos que podrían conducir a decisiones injustas o a la marginación de ciertos grupos (Barocas y Selbst, 2016).

Incorporar prácticas que promuevan la equidad en los datos no solo ayuda a mitigar los sesgos históricos, sino que también impulsa la creación de algoritmos más robustos y generalizables que pueden desempeñarse bien en una amplia gama de contextos. Esto requiere establecer directrices y procedimientos rigurosos para la validación y auditoría continua de los datos, asegurando que todos los cambios en el panorama demográfico sean reflejados oportunamente en el sistema.

La equidad en el uso de datos también implica construir un entorno colaborativo donde las partes interesadas y los expertos en ética trabajen en conjunto con los científicos de datos para identificar y abordar posibles sesgos. Al fomentar la interdisciplinaridad y la inclusión de múltiples perspectivas en el desarrollo de IA, se promueve un enfoque más holístico y equitativo que puede contribuir de manera significativa a la justicia social a través de la tecnología. Esto sienta las bases para un uso ético y eficaz de la IA, fortaleciendo la confianza pública y asegurando que el progreso tecnológico no se logre a expensas de los derechos y la dignidad humana.

Incorporar estos principios éticos kantianos no solo contribuye a la corrección de sesgos inherentes, sino que también asegura que los sistemas de IA sirvan a un propósito más amplio de justicia social. El diseño ético y el uso consciente de algoritmos, basados en estos principios, crean un puente fundamental entre la innovación tecnológica y los valores humanos esenciales. Esto garantiza que el progreso no se alcance a expensas de la rectitud moral. Estos enfoques no solo mejoran la funcionalidad y capacidad de los sistemas, sino que también fortalecen su integridad ética. Promoviendo un futuro en el que la tecnología y la moralidad coexistan, se asegura que estos avances beneficien a toda la sociedad, armonizando eficacia técnica con principios éticos sólidos.

CONSIDERACIONES ÉTICAS PARA EL DESARROLLO DE IA EN FINANZAS

El uso de la IA en el sector financiero presenta una serie de desafíos éticos que deben abordarse para garantizar una implementación responsable. Uno de los problemas más críticos es la protección de la privacidad de los datos de los usuarios. Es esencial que la recopilación y el uso de datos personales se realicen con el consentimiento informado de los individuos, y que toda la información sea manejada con estrictas medidas de seguridad para prevenir el acceso no autorizado y el uso indebido de los datos (Solove, 2013). Esto implica no solo cumplir con las leyes y regulaciones de protección de datos, sino también adoptar prácticas de transparencia y educación para que los usuarios comprendan cómo se utilizan sus datos.

Además de la privacidad, otro desafío significativo es el impacto potencial de la IA en el empleo. La automatización excesiva puede llevar a la reducción de puestos de trabajo, especialmente en roles que pueden ser fácilmente reemplazados por sistemas automatizados. Por lo tanto, es crucial establecer salvaguardias que aseguren que la introducción de tecnologías de IA mantenga un balance adecuado entre la eficiencia tecnológica y el uso del capital humano. Esto podría incluir iniciativas de reconversión laboral y capacitación para ayudar a los empleados a adaptarse a nuevos roles que surjan debido a la evolución tecnológica.

Finalmente, el desarrollo ético de la IA en finanzas debe estar alineado con marcos normativos y directrices internacionales. Organizaciones como la *UNESCO* han enfatizado la justicia, la transparencia y la responsabilidad como pilares fundamentales en el uso de tecnologías avanzadas (UNESCO, 2021). La Estrategia Pan-canadiense de IA, por ejemplo, también subraya la importancia de adherirse a estos principios para fomentar un desarrollo ético e inclusivo. Alinearse con estas recomendaciones globales no solo ayuda a establecer estándares elevados de práctica ética, sino que también promueve la interoperabilidad y la confianza internacional en el uso de la IA.

Al enfrentar estos desafíos, es crucial que las organizaciones financieras empleen un enfoque holístico y proactivo, considerando tanto las implicaciones éticas a corto plazo como los efectos a largo plazo

en la sociedad. De esta manera, la IA puede desarrollarse y utilizarse como una herramienta poderosa que potencia el progreso humano al tiempo que respeta los valores fundamentales de equidad y dignidad.

La aplicación de la ética kantiana en el ámbito de la inteligencia artificial y el *machine learning* ofrece un marco axiológico crítico que sostiene el desarrollo tecnológico en valores morales firmes y universales. Al considerar los conceptos clave de Kant, como el imperativo categórico y la autonomía, se puede instalar una brújula ética que orienta las decisiones y diseños algorítmicos hacia prácticas justas, transparentes y responsables. En el desarrollo de algoritmos de IA, estos principios no solo abordan la necesidad de transparencia y explicabilidad, sino que también insisten en un estándar ético más elevado donde la comprensión del usuario y la responsabilidad del creador son primordiales.

Este enfoque es especialmente relevante en el sector financiero, donde la IA no solo transforma la eficiencia operativa, sino que también interactúa profundamente con aspectos de privacidad, equidad y justicia social. La ética kantiana impulsa a los desarrolladores y responsables de políticas a alinear los avances tecnológicos con directrices internacionales y marcos normativos, asegurando así el respeto por la dignidad humana mientras se navega por los desafíos del pragmatismo técnico del *machine learning*. Este entrelazado de axiología y pragmatismo no solo protege los derechos y oportunidades del individuo, sino que también establece un camino sustentable y moral para el futuro del uso de la IA en la sociedad. En última instancia, este enfoque ético refuerza la confianza pública y fomenta un ecosistema tecnológico que no solo es avanzado, sino también humano en su concepción y aplicación.

ANÁLISIS DEL USO DE MACHINE LEARNING E IA EN LA EVALUACIÓN CREDITICIA

La creciente adopción de tecnologías avanzadas, como *machine learning* (ML) e inteligencia artificial (IA), en el sector financiero ha revolucionado la forma en que se evalúa el riesgo crediticio. Este proceso ha pasado de depender en gran medida de modelos estadísticos tra-

dicionales y de la evaluación manual de datos, a integrar algoritmos complejos que pueden analizar grandes volúmenes de información de manera rápida y eficiente. En este apartado, se procederá a describir el proceso actual de evaluación crediticia mediante estas tecnologías, explorando tanto los beneficios que ofrecen, como los riesgos que conllevan, especialmente en relación con la equidad y la justicia social. Además, se discutirán casos específicos en los que los sesgos presentes en los datos y en los propios algoritmos han generado implicaciones éticas y prácticas significativas, resaltando la necesidad de una reflexión continua y un enfoque ético en la implementación de estas herramientas en procesos críticos de decisión financiera. En este sentido, se busca brindar una comprensión holística no solo del funcionamiento de la tecnología, sino también de los desafíos éticos que surgen de su aplicación en la realidad.

Proceso Actual de Evaluación Crediticia con Tecnologías Avanzadas

La evaluación crediticia es un proceso central en el sector financiero que determina la capacidad de un prestatario de cumplir con sus obligaciones. Con la evolución digital y la complejidad de las variables asociadas al riesgo crediticio, el uso de tecnologías avanzadas como el *machine learning* (ML) e inteligencia artificial (IA) se ha vuelto cada vez más común. Estas herramientas permiten que los datos sean analizados de manera más exhaustiva y precisa, transformando la forma en que se realizan estas evaluaciones.

El método *CRISP-DM* (Cross-Industry Standard Process for Data Mining) es un enfoque en seis etapas que proporciona un marco estructurado para procesar datos en proyectos de minería de datos, incluida la evaluación crediticia. A continuación se presenta la Tabla 2, que ilustra el proceso de evaluación crediticia a través del prisma de *CRISP-DM*.

Tabla 2. Proceso de Evaluación Crediticia utilizando el Método CRISP-DM

Fase	*Descripción*
1. **Comprensión del Negocio**	Se identifican los objetivos del proyecto de evaluación crediticia y se definen los criterios de éxito.
2. **Comprensión de los Datos**	Se recopilan los datos históricos pertinentes, incluidas las características del solicitante y el historial crediticio.
3. **Preparación de los Datos**	Se limpian y transforman los datos para asegurar su calidad y relevancia, eliminando errores y duplicados.
4. **Modelado**	Se seleccionan y aplican algoritmos de machine learning para entrenar el modelo que evaluará el riesgo crediticio.
5. **Evaluación**	Se evalúa el modelo utilizando un conjunto de datos de prueba para comprobar su precisión y adecuación a los objetivos establecidos.
6. **Despliegue**	Se implementa el modelo en un entorno real para su uso en la evaluación crediticia, proporcionando asistencia continua y ajustes según sea necesario.

Fuente: Elaborada a partir de la aplicación del método CRISP-DM en la evaluación crediticia.

La tabla anterior resume las etapas clave del proceso de evaluación crediticia mediante el método CRISP-DM. Este enfoque, estructurado y sistemático, permite abordar las múltiples facetas del análisis de datos, desde la comprensión inicial de los objetivos del negocio hasta la implementación efectiva del modelo. La primera fase, la comprensión del negocio, es crucial para alinear el análisis con las metas estratégicas de la institución financiera, asegurando que las decisiones sean coherentes con su misión y valores.

La preparación de los datos es otra etapa vital, dado que la calidad de los datos utilizados impacta directamente en la precisión del modelo. Aquí es donde se asegura que los conjuntos de datos reflejen una diversidad justa, evitando sesgos que podrían derivarse de datos históricos que no representan adecuadamente a todos los grupos (Barocas y Selbst, 2016). La etapa de modelado utiliza técnicas de machine learning que no solo optimizan la precisión de las evaluaciones crediticias,

sino que también incorporan salvaguardias contra la automatización excesiva y los sesgos inherentes.

En conclusión, la aplicación del método CRISP-DM al proceso de evaluación crediticia permite un análisis más profundo y estructurado que no solo mejora la eficacia operativa, sino que también se adhiere a principios éticos fundamentales. Este enfoque refuerza la precisión en la toma de decisiones crediticias y establece las bases para un futuro en el que la tecnología avance en armonía con los valores de justicia y equidad en el sector financiero.

No obstante, la implementación de modelos de *machine learning* (ML) en la evaluación crediticia presenta una serie de desafíos prácticos que deben ser considerados, los cuales se observan en la Figura 2. (**Ver Figura 2**).

Figura 2. Desafíos en la implementación del ML en evaluación crediticia



Fuente: Napkin.ai (2024)

Uno de los principales obstáculos es la resistencia al cambio por parte de los empleados. Las organizaciones financieras, a menudo ancladas en estructuras y prácticas tradicionales que han demostrado eficacia, pueden enfrentar temor y desconfianza ante la introducción de tecnologías avanzadas. Esta situación puede llevar a que algunos empleados sientan que sus empleos están en riesgo o que deben adaptarse a nuevos métodos de trabajo (Floridi, 2019). Tal resistencia puede traducirse en falta de colaboración y en una disminución de la moral laboral, lo que podría afectar la eficacia del proyecto de implementación.

La necesidad de formación representa otro desafío crítico en este contexto. La integración del ML requiere que los empleados adquieran un nivel de conocimiento técnico que no siempre poseen. Proporcionar la capacitación adecuada es esencial para que los empleados comprendan cómo funcionan los modelos y cómo interactuar con ellos, lo que les permitirá sentirse empoderados en su uso (Dignum, 2018). Además, esto puede exigir inversiones significativas en programas de capacitación, lo que representa un reto en entornos que priorizan la eficiencia y la reducción de costos.

Asimismo, la infraestructura tecnológica necesaria para implementar modelos de ML es un aspecto fundamental que a menudo se pasa por alto. Las instituciones financieras pueden necesitar actualizar sus sistemas o invertir en hardware y software avanzados que permitan el procesamiento de grandes volúmenes de datos y la ejecución de algoritmos complejos (Barocas y Selbst, 2016). Por lo tanto, establecer políticas de gestión de datos adecuadas es crucial para garantizar que el almacenamiento y procesamiento de datos cumplan con las normativas de seguridad y privacidad.

Finalmente, la curva de aprendizaje asociada con la adopción de tecnología avanzada puede causar retrasos en la integración y en la obtención de resultados tangibles. Es fundamental que las organizaciones aborden estos desafíos para crear un entorno que sea receptivo al cambio, lo que, a su vez, permitirá la adopción de tecnologías que ofrezcan oportunidades de innovación en el proceso de evaluación crediticia.

Al considerar estos desafíos de implementación en el marco del análisis estructurado proporcionado por el método *CRISP-DM*, se evidencian obstáculos significativos; sin embargo, la gestión adecuada de estos retos puede dar lugar a un proceso de evaluación crediticia más eficiente y éticamente fundamentado. Este enfoque asegura que la integración de machine learning en las prácticas financieras avance de forma que promueva valores de justicia y equidad, sentando las bases para un futuro donde la tecnología funcione no solo como una herramienta de optimización, sino también como un pilar que refuerce la integridad del sistema financiero en su conjunto.

BENEFICIOS Y RIESGOS DE ML E IA EN EVALUACIÓN CREDITICIA

La aplicación de machine learning e inteligencia artificial en la evaluación crediticia ofrece múltiples beneficios significativos, así como riesgos que requieren atención cuidadosa. Desde el marco de la ética kantiana, es fundamental que estas tecnologías se apliquen en consonancia con principios que promuevan equidad, justicia y dignidad humana. Entre los beneficios más destacados se encuentra la capacidad de procesar y analizar grandes volúmenes de datos en tiempo real, lo que permite una evaluación más rápida y precisa del riesgo crediticio. Esta eficiencia no solo ahorra tiempo, sino que también puede mejorar la experiencia del cliente al agilizar los procesos de aprobación de crédito. Además, el uso de algoritmos avanzados tiene el potencial de identificar patrones e insights que los métodos tradicionales podrían pasar por alto, lo que puede llevar a decisiones más informadas y fundamentadas (Khandani, Kim, y Lo, 2010).

La adaptabilidad de los modelos es otro beneficio importante, ya que permite que las instituciones financieras ajusten sus métodos de evaluación en función de la información nueva y real. Esto les ayuda a responder rápidamente a cambios en el comportamiento del mercado y en el perfil del prestatario, lo que puede hacerlas más competitivas y ofrecer productos más personalizados (Chen y Huang, 2019).

Sin embargo, la implementación de ML e IA también conlleva riesgos considerables que deben ser vistos a través de un prisma ético.

Uno de los riesgos más significativos es la posibilidad de sesgos en los datos que pueden provocar decisiones injustas. Si los modelos son entrenados con datos que reflejan inequidades históricas, esto puede llevar a la exclusión de ciertos grupos de prestatarios o a la imposición de tasas de interés desproporcionadamente altas a individuos de grupos marginados (Baracas y Selbst, 2016). Esta preocupación se alinea con el imperativo categórico de Kant, que enfatiza que las decisiones deben ser justas y aplicables a todas las personas sin excepción.

La falta de transparencia en los procesos de decisión es otro riesgo crítico. Aunque las tecnologías avanzadas pueden hacer que las evaluaciones sean más eficientes, la opacidad que a menudo acompaña a los modelos de aprendizaje automático puede llevar a desconfianza entre los usuarios y a la dificultad de entender cómo se toman las decisiones. Esto plantea graves preocupaciones éticas sobre la responsabilidad y la rendición de cuentas, especialmente cuando las decisiones automatizadas pueden tener un impacto significativo en la vida de las personas (Rudin, 2019).

Además, la dependencia excesiva de los modelos automatizados puede resultar en la deshumanización del proceso de evaluación. Cuando decisiones críticas se toman de manera automatizada, se corre el riesgo de substituir el juicio humano, lo cual puede ser problemático, especialmente en casos que requieren consideraciones contextuales o cualitativas. Este aspecto subraya la necesidad de un enfoque equilibrado, donde la intervención humana se combine con la automatización, alineándose así con la visión ética kantiana de tratar a cada individuo como un fin en sí mismo.

Los elementos y sus respectivos beneficios y riesgos en la aplicación de machine learning en la evaluación crediticia están resumidos en la siguiente Tabla 3. (**Ver Tabla 3**)

Tabla 3. Beneficios y Riesgos del Machine Learning en Evaluación Crediticia

Elemento de Aplicación de ML	Beneficios	Riesgos
Análisis Predictivo	• Mejora la precisión al evaluar la probabilidad de incumplimiento. • Facilita decisiones rápidas en tiempo real. • Permite la identificación de patrones complejos que los métodos tradicionales no pueden detectar.	• Puede perpetuar sesgos si se basa en datos históricos sesgados. • Falta de transparencia puede llevar a la desconfianza en el sistema. • Dependencia excesiva de algoritmos puede resultar en la deshumanización de las decisiones crediticias.
Segmentación de Clientes	• Permite segmentar prestatarios en función de sus características específicas para ofrecer productos personalizados. • Mejora la experiencia del cliente al proporcionar ofertas adaptadas.	• Puede excluir grupos marginados si los segmentos se definen de manera demasiado restrictiva. • Riesgo de no representar adecuadamente la diversidad si los datos de entrada no son completos.
Detección de Fraude	• Mejora la identificación de patrones de comportamiento fraudulentos en tiempo real. • Reduce las pérdidas financieras por fraude.	• Si los modelos anticipan la actividad fraudulenta de manera incorrecta, pueden afectar a clientes inocentes. • Puede crear falsas alarmas que dificulten la confianza del cliente en el sistema.
Análisis de Sentimientos	• Permite entender la percepción del cliente a partir de comentarios y reseñas. • Ayuda a anticipar necesidades y expectativas del cliente.	• Resultados incorrectos pueden conducir a decisiones erróneas si no se interpretan adecuadamente. • El análisis puede ser sesgado si los datos de opinión no son representativos.

Elemento de Aplicación de ML	Beneficios	Riesgos
Optimización de Precios	• Facilita la personalización de tasas de interés y términos de crédito en función del riesgo individual. • Aumenta la competitividad ofreciendo condiciones más atractivas a prestatarios con buen perfil.	• Establecer precios basados en datos historiales puede excluir a prestatarios que no ajustan a los perfiles típicos. • Se corre el riesgo de incrementar la complejidad en la gestión de tarifas, afectando la claridad y transparencia.
Modelos de Recomendación	• Proporciona sugerencias personalizadas sobre productos financieros que se alineen con las necesidades del cliente. • Aumenta la satisfacción del cliente al ofrecer opciones relevantes.	• Puede resultar en decisiones que no son totalmente informadas si los clientes dependen demasiado de las recomendaciones. • Puede crear burbujas donde solo se presentan productos similares, limitando elecciones más variadas.

Fuente: CHAT GPT4o (2024)

En resumen, mientras que la aplicación de machine learning e inteligencia artificial en la evaluación crediticia promete mejoras significativas en eficiencia y precisión, también plantea riesgos que deben ser gestionados con atención. Es esencial que las instituciones financieras no solo adopten estas tecnologías avanzadas, sino que lo hagan de una manera ética y responsable, garantizando que el progreso en el sector financiero no sacrifique los valores fundamentales de justicia y equidad.

CASOS DE SESGO Y SUS IMPLICANCIAS ÉTICAS Y PRÁCTICAS

El uso de machine learning en la evaluación crediticia presenta beneficios significativos, pero el sesgo en los modelos también se convierte en un riesgo considerable que debe ser cuidadosamente considerado. Uno de los casos más notorios de sesgo en este contexto es el **sesgo de representación**. Cuando los datos históricos reflejan desigualdades sistémicas —como inequidades de género y raza— los modelos entre-

nados sobre estas bases tienden a perpetuar, y en algunos casos, acentuar dichas desigualdades. Un estudio realizado por Barocas *y otros*. (2019) expone cómo los modelos de ML, en su intento de optimizar ciertas métricas de rendimiento, pueden desestimar grupos minoritarios. Esto resulta en tasas desproporcionalmente altas de rechazo crediticio que afectan a comunidades enteras, lo que plantea graves implicaciones éticas y sociales respecto al principio de justicia en el acceso al crédito.

Un caso práctico que ilustra este problema es el de la empresa de tecnología financiera (fintech) ZestFinance, que lanzó un sistema de evaluación crediticia en 2014 utilizando machine learning para analizar datos alternativos, como el historial de pagos de servicios públicos. Si bien la intención era ampliar el acceso al crédito para aquellos que históricamente habían sido excluidos del sistema financiero tradicional, el modelo terminó reflejando sesgos raciales y de género en los datos de entrada. Esto llevó a que ciertos grupos minoritarios recibieran evaluaciones de riesgo más altas, resultando en tasas de rechazo desproporcionadas, lo que generó críticas sobre la falta de equidad y la transparencia del modelo. Este tipo de discriminación no solo pone en riesgo la reputación de la empresa, sino que también infringe principios éticos fundamentales de la justicia y la equidad, tal como lo defendería Kant.

Además, el **sesgo de medición** representa otro desafío crítico. Este sesgo ocurre cuando las variables relevantes para evaluar el riesgo crediticio no están adecuadamente representadas o son inexactas. Por ejemplo, el uso de datos alternativos, como el historial de pagos de servicios públicos o el historial de redes sociales, aunque puede parecer una innovación útil, conlleva el riesgo de introducir sesgos si esos datos no son recogidos de manera equitativa. Esto puede resultar en la exclusión de ciertos grupos del acceso a productos financieros, simplemente porque sus patrones de comportamiento no coinciden con los criterios convencionales utilizados en modelos tradicionales (Lacko y Pappalardo, 2020). Desde una perspectiva axiológica, esto desafía el principio kantiano de tratar a todos los individuos con dignidad y equidad, permitiendo que el sesgo perpetúe la injusticia social.

Un tercer tipo de sesgo digno de atención es el **sesgo de aprendizaje**, que ocurre cuando los algoritmos adaptan su comportamiento en función de decisiones anteriores. Si un modelo ha sido entrenado en datos que incluyen decisiones crediticias influenciadas por sesgos en la práctica pasada, las predicciones futuras seguirán replicando esos mismos sesgos. Esto refuerza ciclos de discriminación y puede disminuir las oportunidades de acceso al crédito para individuos de grupos históricamente marginados (Dastin, 2018). La perpetuación de estas desigualdades plantea desafíos éticos que demandan un enfoque proactivo por parte de las instituciones financieras.

La existencia de estos sesgos no solo es un problema técnico, sino también moral. Por ello, es fundamental que las organizaciones adopten una estrategia ética que incluya revisiones continuas y auditorías de los modelos de ML utilizados. Esto asegurará que el desarrollo y el uso de estas tecnologías no se realicen en detrimento de los principios de justicia y equidad que deberían guiar prácticas financieras responsables. En conclusión, abordar el sesgo en la evaluación crediticia es esencial no solo para mejorar la precisión de los modelos, sino también para construir un sistema financiero más justo y accesible para todas las comunidades.

En conclusión, el análisis del uso de machine learning e inteligencia artificial en la evaluación crediticia revela un proceso transformador que promete mejorar la eficiencia y precisión en la toma de decisiones financieras. Sin embargo, esta evolución tecnológica trae consigo una serie de beneficios y riesgos que deben ser considerados cuidadosamente. Mientras que la capacidad de estos sistemas para procesar grandes volúmenes de datos puede ofrecer insights valiosos y optimizar la experiencia del cliente, también es fundamental abordar los sesgos que pueden estar presentes en los datos y en los modelos, los cuales pueden perpetuar injusticias en el acceso al crédito. Los casos de sesgo highlights las implicancias éticas y prácticas que emergen de la implementación de estas herramientas, subrayando la necesidad de un enfoque ético riguroso. Así, para que la integración de ML e IA en la evaluación crediticia no solo sea efectiva, sino también justa, es imperativo que las instituciones financieras adopten medidas proactivas que promuevan la equidad y el respeto por la dignidad humana en todas las decisiones crediticias.

PROPUESTA DE INTEGRACIÓN ÉTICA EN MODELO DE IA PARA EL CRÉDITO

La creciente complejidad y la utilización de tecnologías avanzadas en el sector financiero han puesto de manifiesto la necesidad urgente de abordar la ética en el diseño y uso de modelos de inteligencia artificial (IA). A pesar de los innegables beneficios que aporta el machine learning (ML) en la evaluación crediticia, la falta de un marco ético sólido puede dar lugar a sesgos y prácticas injustas que perpetúan desigualdades históricas. Actualmente, existen muy pocas propuestas que integren de manera explícita principios éticos en estos procesos, lo que resalta la importancia de desarrollar un enfoque que respete y promueva la dignidad humana y la equidad. En este contexto, la propuesta que aquí se presenta, basada en la ética kantiana, busca ofrecer un marco que asegure la justicia y la transparencia en la evaluación crediticia.

Este apartado se desarrollará en tres puntos clave. En primer lugar, se presentará una propuesta para un enfoque que incorpora principios kantianos en la IA, destacando la relevancia de la justicia y la dignidad humana en la toma de decisiones automatizadas. En segundo lugar, se evaluarán los impactos potenciales de este enfoque en términos de equidad y transparencia, considerando cómo estos valores pueden mejorar la inclusión financiera. Finalmente, se discutirán las consideraciones prácticas necesarias para la implementación de este enfoque, abordando los desafíos que pueden surgir y las estrategias para superarlos. Con estos elementos, se pretende proporcionar una comprensión integral de la integración ética de los principios kantianos en la inteligencia artificial para la evaluación crediticia.

PROPUESTA DE INTEGRACIÓN DE LOS PRINCIPIOS KANTIANOS EN LA IA

La propuesta para integrar principios kantianos en la inteligencia artificial se fundamenta en la premisa de que las acciones automatizadas deben regirse por máximas que puedan ser aceptadas universalmente, promoviendo el respeto por la autonomía y la dignidad de cada indi-

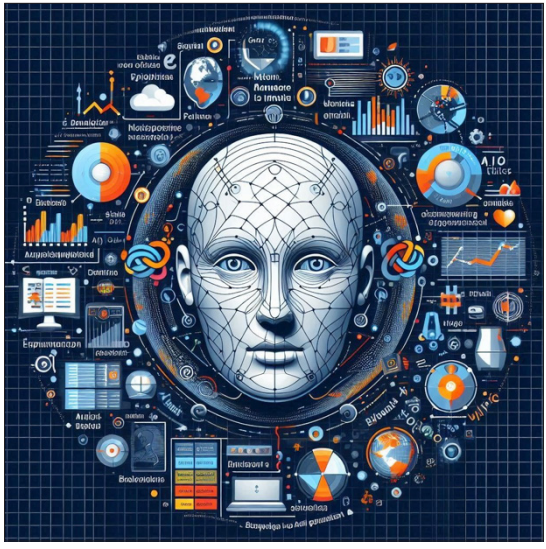
viduo (Kant, 1785). Esta perspectiva ética subraya la importancia de diseñar algoritmos que sean imparciales y equitativos, evitando decisiones influenciadas por características demográficas sensibles como la raza, el género o el estado socioeconómico. La implementación de modelos de IA que prioricen la justicia como equidad y que incluyan mecanismos de defensa basados en estándares morales universales es esencial para superar los desafíos éticos en la evaluación del riesgo crediticio.

Este enfoque no es simplemente teórico; ya existen iniciativas y modelos en desarrollo que buscan aplicar principios éticos en el machine learning. Por ejemplo, ***AI Fairness 360*** de IBM proporciona herramientas para detectar y mitigar sesgos en algoritmos, lo que permite realizar un análisis más equitativo de los datos utilizados en decisiones automatizadas. Otra iniciativa notable es el marco de ***AI Ethics*** desarrollado por organizaciones como Microsoft y Google, que resalta la importancia de la justicia, la transparencia y la rendición de cuentas al diseñar sistemas de IA.

El desarrollo de un modelo de IA con cimientos kantianos implica un proceso riguroso en el que cada decisión crediticia debe ser defendible no solo desde el punto de vista funcional, sino también en términos de equidad y justicia. Esto exige que cada algoritmo esté diseñado de tal manera que las decisiones que produce sean explicables y resistibles a un análisis crítico que contemple consideraciones éticas. En un sistema ideal, esto garantizaría que todas las políticas de crédito sean accesibles y comprensibles para todos los solicitantes, independientemente de su trasfondo demográfico o económico.

Para alcanzar este objetivo, es fundamental integrar prácticas de auditoría continua y mecanismos de transparencia en el proceso de toma de decisiones. Las auditorías permiten un seguimiento constante del rendimiento de los modelos de IA, asegurando que su aplicación se mantenga alineada con los principios éticos establecidos. Estas revisiones periódicas no solo ayudan a identificar y corregir sesgos, sino que también propician ajustes en los algoritmos cuando sea necesario, garantizando que su funcionamiento siga siendo justo y equitativo.

Figura 3. Modelo de IA fundamentado en los principios de la ética kantiana



Fuente: Microsoft Copilot (2024)

La transparencia implica que los usuarios y solicitantes de crédito tengan acceso a una comprensión clara de cómo funcionan los algoritmos y cuáles son los criterios utilizados para la toma de decisiones. Esto incluye la divulgación de la lógica algorítmica y la comunicación de los factores que influyen en las decisiones de crédito. Al proporcionar esta información, las instituciones no solo cumplen con los estándares éticos, sino que también generan confianza entre los consumidores, empoderándolos al comprender mejor cómo se evalúa su situación crediticia.

Incorporar estos principios kantianos en la evaluación crediticia no solo aborda el aspecto técnico del análisis de datos, sino que también promueve una cultura organizacional centrada en la ética y la responsabilidad social. De esta manera, las instituciones financieras se convierten en actores responsables que buscan maximizar sus beneficios mientras se comprometen a contribuir positivamente a la sociedad. Esto es especialmente relevante en un contexto donde recuperar la

confianza del consumidor es fundamental para el éxito a largo plazo de las operaciones financieras.

Por lo tanto, un modelo de IA fundamentado en los principios kantianos no debe ser solamente efectivo en términos de rendimiento, sino que también debe soportar un escrutinio ético, promoviendo un sistema financiero más justo, accesible y responsable.

EVALUACIÓN DE LOS IMPACTOS POTENCIALES EN EQUIDAD Y TRANSPARENCIA

Implementar principios kantianos en el ámbito de la evaluación crediticia puede llevar a una transformación significativa hacia un modelo que priorice la equidad y la transparencia. Al diseñar procesos de evaluación que sean intrínsecamente libres de prejuicios, se fomenta una mayor inclusión financiera, permitiendo que un mayor número de individuos acceda a productos crediticios justos y asequibles (Hardt y otros., 2016). Este enfoque no solo beneficia a los prestatarios, sino que también contribuye a un sistema financiero más sólido y cohesivo, donde la justicia y la dignidad son pilares centrales.

A su vez, la adopción de un enfoque ético basado en los principios de Kant promueve la transparencia algorítmica, lo que se traduce en una exposición clara de los procedimientos de toma de decisiones. Esta transparencia es crucial, ya que permite que los usuarios comprendan cómo se determina su riesgo crediticio y qué factores influyen en las decisiones de aprobación o rechazo. Al proporcionar esta claridad, las instituciones financieras no solo cumplen con estándares éticos, sino que también fortalecen la confianza entre los usuarios, quienes se sienten más seguros al saber que las decisiones no están dictadas por algoritmos opacos y arbitrarios (Veale y Binns, 2017).

La equidad y la transparencia, cuando se integran en el marco de la evaluación crediticia, crean un ciclo virtuoso que puede ayudar a desmantelar las desigualdades históricas en el acceso al crédito. Desde una perspectiva axiológica, este enfoque alinea la práctica financiera con principios éticos fundamentales de justicia y dignidad humana, facilitando la creación de un entorno más inclusivo y equitativo. La

inclusión de más prestatarios en el sistema financiero, especialmente aquellos que previamente fueron excluidos debido a sesgos en los procesos de evaluación, genera un efecto de retroalimentación positiva. Este fenómeno se traduce en una mayor diversidad de solicitantes, enriqueciendo así la base de datos utilizada para entrenar modelos de machine learning (ML). A medida que los modelos reciben información de un espectro más amplio de antecedentes y experiencias, su capacidad para generar decisiones justas y precisas se incrementa, volviéndose más robustos en la identificación de riesgos reales, sin recurrir a estereotipos que perpetúen la discriminación.

Esta retroalimentación positiva también crea una demanda social y empresarial por mejores prácticas. A medida que las instituciones financieras experimentan los beneficios de un enfoque centrado en la equidad y la transparencia, se ven incentivadas no solo a mantener estos estándares, sino a formalizarlos en sus estructuras organizativas y culturales. La presión por alinearse con estos principios éticos puede convertirse en una ventaja competitiva, promoviendo una reputación de responsabilidad social que atrae tanto a clientes como a inversores que priorizan la ética en sus decisiones.

Desde la perspectiva de la ética kantiana, esta dinámica es coherente con el imperativo categórico, que establece que las acciones deben ser universales y respetar la dignidad de cada individuo. Implementar sistemas que aseguren equidad y transparencia representa, por lo tanto, una manifestación práctica de estos principios éticos. Las instituciones que adopten este enfoque no solo estarán en cumplimiento de normativas y expectativas sociales, sino que también reforzarán su compromiso con la justicia social, contribuyendo a un sistema más ético y justo.

En definitiva, la sinergia entre equidad y transparencia no solo promueve un acceso más justo al crédito, sino que también establece un marco en el que el crecimiento y la innovación se pueden desarrollar de manera ética y responsable. Al hacerlo, las instituciones financieras pueden evolucionar hacia modelos más sostenibles que no solo mejoren la eficacia operativa, sino que también reflejen y fortalezcan los valores de justicia y dignidad que son esenciales para una sociedad equitativa. Este enfoque holístico es clave para la transformación

positiva del sistema financiero, contribuyendo a un futuro en el que la tecnología sirva como un verdadero aliado para la inclusión y la justicia social.

La integración de principios kantianos en la evaluación crediticia tiene el potencial de crear un entorno más justo y accesible, donde los mecanismos algorítmicos no solo actúan como herramientas de análisis, sino que se convierten en instrumentos de empoderamiento para los consumidores. Este movimiento hacia la equidad y la transparencia no solamente es un imperativo ético, sino también una estrategia práctica para garantizar la sostenibilidad y la legitimidad del sector financiero en el futuro.

Consideraciones Prácticas para la Implementación de la Ética Kantiana en IA para el Crédito

Para implementar sistemas de IA congruentes con la ética kantiana, es fundamental establecer auditorías regulares y sistemas de verificación que identifiquen y mitiguen los sesgos (Selbst y otros., 2019). Además, es esencial integrar una educación continua en ética para desarrolladores y usuarios del sistema, fomentando una comprensión profunda de las implicaciones sociales de sus aplicaciones tecnológicas. El asesoramiento interdisciplinario, que incluye enfoques de ética, derecho y ciencias sociales, garantizará que los modelos de IA sean no solo técnicamente competentes, sino también moralmente responsables.

Las consideraciones prácticas para lograr una implementación ética de los principios kantianos en sistemas de inteligencia artificial para la evaluación del crédito incluyen:

1. Establecimiento de auditorías regulares: Evaluaciones periódicas para detectar y corregir sesgos.
2. Integración de educación continua en ética: Formación constante para desarrolladores y usuarios.
3. Asesoramiento interdisciplinario: Inclusión de perspectivas diversas en el desarrollo de IA.

4. Alineación con normativas vigentes: Adaptación a leyes y regulaciones locales e internacionales.
5. Desarrollo de una infraestructura interna para auditorías de equidad y transparencia: Creación de sistemas que faciliten revisiones y ajustes constantes.
6. Adopción de herramientas de evaluación de impactos algorítmicos: Implementación de tecnologías que evalúen cómo los modelos afectan diferentes grupos.
7. Fomento de una cultura organizacional ética: Promoción de principios éticos en el entorno de trabajo.
8. Comunicación transparente sobre cómo funcionan los algoritmos: Asegurar que los consumidores entiendan el proceso de toma de decisiones.

La implementación de principios kantianos en sistemas de inteligencia artificial para la evaluación del crédito exige una alineación cuidadosa con las normativas vigentes en diversas jurisdicciones. En Europa, la normativa se centra en el Reglamento General de Protección de Datos (GDPR), que establece principios de transparencia, equidad y responsabilidad en el tratamiento automatizado de datos personales. Los sistemas de IA deben proporcionar explicaciones sobre decisiones automatizadas significativas, asegurando que los individuos entiendan cómo y por qué se han tomado ciertas decisiones (Voigt y Bussche, 2017). Este marco legislativo enfatiza la necesidad de auditorías algorítmicas y el diseño de modelos explicables que puedan ser interrogados para verificar su equidad y cumplimiento regulatorio.

En Canadá, la propuesta de la Ley de Implementación de la Carta Digital (Ley C-27) busca proteger a los ciudadanos mediante prácticas responsables de IA, incluyendo la reducción de sesgos. Esta ley prevé un mayor acceso para los usuarios a los criterios que utilizan los algoritmos para la toma de decisiones crediticias, exigiendo que las organizaciones gestionen los datos personales de manera ética y justa (Office of the Privacy Commissioner of Canada, 2020).

En Estados Unidos, la falta de un marco regulatorio federal específico sobre el uso de IA en la evaluación crediticia ha llevado a un enfo-

que más fragmentado. Sin embargo, algunas iniciativas a nivel estatal están surgiendo, como la legislación en California que busca aumentar la transparencia y la responsabilidad en los algoritmos utilizados en decisiones de crédito y préstamos. Esto resalta la importancia de un enfoque regulatorio coherente que garantice que los principios éticos sean parte integral de la práctica crediticia.

En México y América Latina, la legislación sobre protección de datos personales ha avanzado, pero todavía enfrenta retos significativos. Iniciativas como la Ley Federal de Protección de Datos Personales en Posesión de los Particulares en México han establecido principios básicos de privacidad y manejo ético de datos, aunque la implementación efectiva de estos principios en el contexto de algoritmos de IA para el crédito es aún incipiente. El desarrollo de una regulación robusta y un marco ético claro en esta región es crucial para asegurar que las tecnologías emergentes promuevan la inclusión y la equidad.

En todos estos contextos regulatorios, resulta crucial desarrollar una infraestructura interna que permita llevar a cabo auditorías periódicas de equidad y transparencia, así como adoptar herramientas de evaluación de impactos algorítmicos que aseguren el cumplimiento de normativas locales y principios kantianos. Además, promover una cultura organizacional que incluya capacitación continua en ética de la IA y una comunicación transparente sobre el uso de estas tecnologías fortalecerá la confianza del público, garantizando que las prácticas de IA sean dignas de confianza y respeten tanto la normativa vigente como los imperativos morales universales.

La integración de principios kantianos en los modelos de inteligencia artificial para el crédito representa un paso significativo hacia la creación de un sistema financiero más justo y responsable. Al fundamentar el diseño de estos modelos en máximas que promueven la autonomía y la dignidad humana, se establece un marco ético que orienta las decisiones crediticias hacia una mayor imparcialidad y equidad. Los análisis de los impactos potenciales en equidad y transparencia destacan cómo la transparencia algorítmica puede fortalecer la confianza de los usuarios, al permitirles comprender y cuestionar las decisiones que les afectan (Veale y Binns, 2017). Asimismo, para que este enfoque sea verdaderamente efectivo, es crucial implementar

consideraciones prácticas que incluyan auditorías regulares, formación continua en ética y una infraestructura adecuada que soporte la diligencia necesaria en el manejo de datos (Selbst y otros., 2019).

A través de esta integración ética, no solo se abordan los riesgos asociados con el uso de tecnologías avanzadas, sino que también se sientan las bases para un futuro en el que la inteligencia artificial funcione como un aliado en la promoción de la justicia social. En última instancia, este enfoque innovador no solo mejora la eficacia operativa de las instituciones financieras, sino que también refuerza su compromiso con los principios éticos, contribuyendo a un sistema más inclusivo, responsable y digno de confianza en el ámbito de la evaluación crediticia (Barocas y Selbst, 2016).

CONCLUSIONES

La aplicación de la ética kantiana en el ámbito de la evaluación crediticia, a través de la inteligencia artificial y el machine learning, representa una oportunidad única para transformar el sector financiero. Al integrar principios de justicia, equidad y transparencia en los algoritmos, es posible avanzar hacia un sistema más justo, inclusivo y responsable, donde los beneficios tecnológicos no se antepongan a la dignidad humana. Sin embargo, este camino no está exento de desafíos. La persistencia de sesgos en los datos, la opacidad de los algoritmos y la falta de una regulación clara a nivel internacional pueden dificultar la construcción de un sistema verdaderamente equitativo. La ética kantiana exige un esfuerzo por combatir estos sesgos y crear un marco que promueva la autonomía del individuo, asegurando que las decisiones se tomen en un contexto de transparencia e inclusión.

Es fundamental desmantelar el enfoque reduccionista que considera la tecnología únicamente como un instrumento de optimización financiera. La IA no debería ser utilizada solamente como una herramienta de eficiencia; sus implicaciones sociales y éticas también deben ser tomadas en cuenta. Por ello, es necesario ir más allá de la mera implementación tecnológica y profundizar en la integración de valores universales en la toma de decisiones. Esta nueva narrativa sobre la

IA debe incluir un fuerte componente ético para construir un sistema financiero que no solo sea eficiente, sino que también sea justo y humano.

Para garantizar este desarrollo ético en el futuro, es crucial avanzar en varias direcciones. Primero, se requiere el desarrollo de herramientas de auditoría más sofisticadas que permitan identificar y mitigar sesgos en los algoritmos de IA de manera efectiva. Además, es necesario fortalecer la regulación internacional para crear estándares que guíen el uso ético de la IA en la evaluación crediticia. También es vital investigar la interacción entre humanos y máquinas, ya que comprensión y aceptación son claves para el éxito de estos sistemas. Asimismo, implementar programas de educación ética dirigidos a desarrolladores, usuarios y reguladores de la IA ayudará a fomentar una cultura de responsabilidad social.

La integración de principios kantianos en la IA tiene el potencial de transformar el sector financiero y contribuir a un mundo más justo y equitativo. Al reflexionar sobre estas direcciones futuras, se abre un camino hacia un desarrollo de la inteligencia artificial que sea ético y que beneficie a la sociedad en su conjunto.

REFERENCIAS BIBLIOGRÁFICAS

Barocas, S., y Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review, 104*(3), 671-732. *https://doi.org/10.15779/Z38Z29R*

Barocas, S., Hardt, M., y Narayanan, A. (2019). *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press.

Chen, J., y Huang, Y. (2019). Machine learning and unsupervised learning in credit risk evaluation. *Applied Sciences, 9*(14), 2852. *https://doi.org/10.3390/app9142852*

Chouldechova, A., y Roth, A. (2018). The frontiers of fairness in machine learning. *Science, 359*(6377), 132-133. *https://doi.org/10.1126/science.aar7696*

Dignum, V. (2018). Ethical frameworks for AI. In *Artificial Intelligence Safety and Security* (pp. 75-92). CRC Press.

Floridi, L. (2019). *The Ethics of Artificial Intelligence*. Oxford University Press.

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

Johnson, R. (2008). Kant's moral philosophy. In *The Stanford Encyclopedia of Philosophy*. Retrieved from *https://plato.stanford.edu/archives/win2015/entries/kant-moral/*

Johnson, R., y Cureton, A. (2024). Kant's moral philosophy. In E. N. Zalta y U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). Retrieved from *https://plato.stanford.edu/archives/fall2024/entries/kant-moral/*

Journey AI Art. (2024). Obtenido de https://journeyaiart.com

Khandani, A. E., Kim, A. J., y Lo, A. W. (2010). Consumer credit risk derivatives: A machine learning approach. *Journal of Financial Economics, 101*(3), 622-644. *https://doi.org/10.1016/j.jfineco.2010.02.002*

Microsoft Copilot. (2024). *Copilot Bing.* https://copilot.microsoft.com/

Mittelstadt, B., Wachter, S., y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2). *https://doi.org/10.1177/2053951716679679*

Napkin AI (2024). Página oficial. https://www.napkin.ai/

OpenAI. (2024). *ChatGPT4o [Modelo de lenguaje de gran tamaño]. https://chat.chatbotapp.ai/chats*

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206-215. *https://doi.org/10.1038/s42256-019-0048-7*

Russell, S., y Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Selbst, A. D., Barocas, S., y Andrew, A. (2019). Fairness and transparency in machine learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 1-3)

CAPÍTULO XVIII

LA IA COMO GUARDIÁN DEL PLANETA: INNOVACIONES PARA UN FUTURO SOSTENIBLE

VÍCTOR ALFONSO GASCON SORIANO[1]

RESUMEN

El capítulo "*La IA como Guardián del Planeta: Innovaciones para un Futuro Sostenible*" escrito por Víctor Alfonso Gascon Soriano, explora cómo la inteligencia artificial (IA) está transformando diferentes sectores con el objetivo de promover la sostenibilidad. El texto destaca que la IA, mediante algoritmos y modelos matemáticos, puede mejorar la eficiencia energética, gestionar recursos naturales, apoyar la agricultura inteligente, optimizar el transporte sostenible y facilitar la economía circular.

El documento analiza el impacto positivo de la IA a través de casos concretos como la predicción del consumo energético, la gestión de cultivos y el monitoreo de especies en peligro de extinción. A su vez, señala desafíos éticos y técnicos que deben ser abordados, como los sesgos en los algoritmos y el consumo de energía de los sistemas de IA. También se enfatiza la importancia de la justicia, la transparencia y la equidad en la implementación de la IA para asegurar un impacto positivo y equitativo.

El capítulo concluye que, si se emplea de manera responsable, la IA puede ser clave en la construcción de un futuro más sostenible, pero

[1] Alumno de la Licenciatura en Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

advierte sobre la necesidad de resolver los riesgos éticos y técnicos inherentes a su uso.

INTRODUCCIÓN

En un mundo cada vez más afectado por desafíos ambientales y sociales, la inteligencia artificial (IA) emerge como una herramienta poderosa para promover la sostenibilidad y proteger nuestro planeta. Este capítulo explora cómo la IA, definida como la capacidad de las máquinas para realizar tareas que normalmente requieren inteligencia humana, está revolucionando diversos sectores al proporcionar soluciones innovadoras y eficientes (Russell y Norvig, 2020).

La IA se basa en algoritmos y modelos matemáticos que permiten a las máquinas aprender de los datos y mejorar su desempeño con el tiempo. Entre sus aplicaciones más destacadas se encuentran el aprendizaje automático y el aprendizaje profundo, tecnologías que han demostrado ser cruciales para manejar grandes volúmenes de datos y realizar tareas complejas con alta precisión (Goodfellow, Bengio y Courville, 2016; LeCun, Bengio y Hinton, 2015).

La sostenibilidad, entendida como la capacidad de satisfacer las necesidades del presente sin comprometer las de futuras generaciones (Brundtland, 1987), se beneficia enormemente de las capacidades de la IA. Este capítulo examina cómo la IA puede contribuir a la eficiencia energética, la gestión de recursos naturales, la agricultura inteligente, el transporte sostenible y la economía circular. A través de casos de estudio y ejemplos concretos, se demuestra el impacto positivo de la IA en la reducción de costos, la mejora de la eficiencia, la innovación y la toma de decisiones informada.

Al abordar tanto los beneficios como los desafíos éticos y técnicos de la implementación de la IA, este capítulo ofrece una visión integral de cómo esta tecnología puede ser un guardián del planeta, impulsando un futuro más sostenible y equitativo.

Palabras Clave: Ética**,** Sostenibilidad y Innovación

LA IA Y LA SOSTENIBILIDAD: UN VISTAZO GENERAL

La inteligencia artificial (IA) se ha convertido en una herramienta fundamental en la búsqueda de soluciones innovadoras para los desafíos contemporáneos. La IA se define como la capacidad de las máquinas para realizar tareas que normalmente requieren inteligencia humana, tales como el aprendizaje, el razonamiento, la resolución de problemas, la percepción y el procesamiento del lenguaje natural (Russell y Norvig, 2020). Esta tecnología se basa en algoritmos y modelos matemáticos que permiten a las máquinas aprender de los datos y mejorar su desempeño con el tiempo.

La evolución de la IA ha sido impulsada por avances en el aprendizaje automático (machine learning), una rama de la IA que se centra en el desarrollo de algoritmos que permiten a las máquinas aprender de los datos sin ser explícitamente programadas para cada tarea específica (Goodfellow, Bengio y Courville, 2016). Dentro del aprendizaje automático, el aprendizaje profundo (deep learning) ha destacado por su capacidad para manejar grandes volúmenes de datos y realizar tareas complejas con alta precisión (LeCun, Bengio y Hinton, 2015).

La sostenibilidad se centra en satisfacer las necesidades del presente sin comprometer la capacidad de las futuras generaciones para satisfacer las suyas (Brundtland, 1987). En este contexto, la IA puede desempeñar un papel crucial al proporcionar soluciones innovadoras para los desafíos ambientales y sociales. A continuación, se presentan algunas formas en que la IA puede contribuir a la sostenibilidad:

1. **Eficiencia energética:** La IA puede optimizar el uso de energía en edificios, industrias y redes eléctricas, reduciendo el consumo y las emisiones de gases de efecto invernadero (IEA, 2021). Por ejemplo, los sistemas de gestión de energía basados en IA pueden ajustar automáticamente la iluminación y la climatización en función de la ocupación y las condiciones ambientales, logrando ahorros significativos en el consumo energético (Wang y otros., 2020).
2. **Gestión de recursos naturales**: Los algoritmos de IA pueden ayudar a monitorear y gestionar recursos como el agua, los bos-

ques y la biodiversidad, promoviendo su uso sostenible (Wang y otros., 2020). En la gestión del agua, por ejemplo, la IA puede predecir patrones de consumo y detectar fugas en las redes de distribución, mejorando la eficiencia y reduciendo el desperdicio (Zhang y otros., 2018).

3. **Agricultura inteligente:** La IA puede mejorar la productividad agrícola mediante el análisis de datos sobre el clima, el suelo y los cultivos, permitiendo una toma de decisiones más precisa y eficiente (Kamilaris y Prenafeta-Boldú, 2018). Los drones equipados con sensores y cámaras pueden recopilar datos en tiempo real sobre el estado de los cultivos, y los algoritmos de IA pueden analizar estos datos para recomendar prácticas agrícolas óptimas, como la cantidad exacta de agua y fertilizantes necesarios (Liakos y otros., 2018).
4. **Transporte sostenible:** Los sistemas de transporte basados en IA pueden reducir el tráfico, optimizar las rutas y promover el uso de vehículos eléctricos, disminuyendo así la contaminación y el consumo de combustibles fósiles (Chen y otros., 2021). Las ciudades inteligentes están implementando sistemas de gestión del tráfico que utilizan IA para ajustar los semáforos en tiempo real, reducir los atascos y mejorar la eficiencia del transporte público (González y otros., 2020).
5. **Economía circular:** La IA puede facilitar el reciclaje y la reutilización de materiales, ayudando a cerrar el ciclo de vida de los productos y reducir los residuos (Ellen MacArthur Foundation, 2019). Los sistemas de clasificación de residuos basados en IA pueden identificar y separar materiales reciclables con mayor precisión que los métodos tradicionales, aumentando las tasas de reciclaje y reduciendo la cantidad de residuos que terminan en los vertederos (Ragossnig y Schneider, 2019).

El uso de la IA en la sostenibilidad ofrece numerosos beneficios, entre los que se incluyen:

- **Reducción de costos:** La optimización de procesos y recursos puede llevar a una disminución significativa de los costos operativos (McKinsey y Company, 2018). Por ejemplo, las empre-

sas pueden utilizar IA para predecir el mantenimiento de equipos y evitar fallos costosos, mejorando la eficiencia operativa y reduciendo los tiempos de inactividad (Lee y otros., 2018).

- **Mejora de la eficiencia:** La IA puede identificar patrones y tendencias que los humanos podrían pasar por alto, mejorando la eficiencia en diversas áreas (*Brynjolfsson y McAfee*, 2014). En la industria manufacturera, los sistemas de IA pueden optimizar las líneas de producción, reducir el desperdicio de materiales y mejorar la calidad del producto final (Rüßmann y otros., 2015).
- **Innovación:** La IA abre nuevas posibilidades para desarrollar tecnologías y soluciones innovadoras que aborden los desafíos ambientales de manera más efectiva (Schwab, 2016). Por ejemplo, la IA está siendo utilizada para desarrollar nuevos materiales sostenibles y procesos de fabricación más limpios, contribuyendo a la reducción de la huella ambiental de la industria (Bauer y otros., 2018).
- **Toma de decisiones informada**: La capacidad de la IA para analizar grandes volúmenes de datos permite una toma de decisiones más informada y basada en evidencia (Davenport y Ronanki, 2018). En el ámbito de la política ambiental, los modelos de IA pueden simular diferentes escenarios y evaluar el impacto de diversas políticas, ayudando a los responsables de la toma de decisiones a elegir las estrategias más efectivas (Rolnick y otros., 2019).

INNOVACIONES EN ENERGÍA RENOVABLE

La transición hacia fuentes de energía renovable es esencial para mitigar el cambio climático y promover la sostenibilidad. La inteligencia artificial (IA) está desempeñando un papel crucial en esta transición al optimizar la generación, distribución y consumo de energía renovable. Este capítulo explora cómo la IA está revolucionando el sector energético y contribuyendo a un futuro más sostenible.

Las redes eléctricas inteligentes, o *smart grids*, utilizan tecnologías avanzadas para gestionar la generación y distribución de electricidad de manera más eficiente. La IA es fundamental en la optimización de estas redes, permitiendo una mejor integración de fuentes de energía renovable como la solar y la eólica (Gungor y otros., 2013). La IA puede predecir patrones de consumo energético y ajustar la producción en consecuencia, reduciendo el desperdicio y mejorando la eficiencia. Los algoritmos de aprendizaje automático analizan datos históricos y en tiempo real para anticipar picos de demanda y ajustar la oferta de energía (Wang y otros., 2018).

La variabilidad de las fuentes de energía renovable, como la solar y la eólica, presenta desafíos para su integración en la red eléctrica. La IA puede predecir la producción de energía renovable con alta precisión, permitiendo una mejor planificación y gestión de la red (Zhang y otros., 2018). Por ejemplo, los modelos de predicción basados en IA pueden utilizar datos meteorológicos para anticipar la generación de energía solar y eólica, ajustando la producción de otras fuentes en consecuencia (Hong y otros., 2020).

La capacidad de predecir la producción de energía renovable es crucial para su integración efectiva en la red eléctrica. La IA ofrece herramientas avanzadas para mejorar la precisión de estas predicciones. Los sistemas de predicción de energía solar basados en IA utilizan datos meteorológicos y de radiación solar para estimar la producción de energía en diferentes ubicaciones y momentos (Voyant y otros., 2017). Estos sistemas pueden ajustarse continuamente en función de nuevos datos, mejorando su precisión con el tiempo.

La predicción de la producción de energía eólica es particularmente desafiante debido a la naturaleza variable del viento. Los modelos de IA pueden analizar datos históricos y en tiempo real sobre la velocidad y dirección del viento para predecir la generación de energía eólica (Zhang y otros., 2018). Estos modelos ayudan a los operadores de la red a planificar la integración de la energía eólica y a minimizar el impacto de su variabilidad (Monteiro y otros., 2009).

El uso de la IA en el sector de la energía renovable ofrece numerosos beneficios, entre los que se incluyen:

- **Mejora de la eficiencia:** La IA optimiza la generación y distribución de energía, reduciendo pérdidas y mejorando la eficiencia operativa (Wang y otros., 2018).
- **Reducción de costos:** La predicción precisa de la producción de energía renovable permite una mejor planificación y gestión de los recursos, reduciendo los costos operativos (Hong y otros., 2020).
- **Aumento de la fiabilidad:** La IA mejora la fiabilidad de la red eléctrica al anticipar y gestionar mejor la variabilidad de las fuentes de energía renovable (Zhang y otros., 2018).
- **Promoción de la sostenibilidad:** Al facilitar la integración de energías renovables, la IA contribuye a la reducción de emisiones de gases de efecto invernadero y al avance hacia un sistema energético más sostenible (Gungor y otros., 2013).

CASOS DE ESTUDIO: PROYECTO DE ENERGÍA SOLAR EN CALIFORNIA

En California, un proyecto de energía solar utiliza IA para optimizar la producción y distribución de energía. Los algoritmos de predicción analizan datos meteorológicos y de radiación solar para ajustar la producción en tiempo real, mejorando la eficiencia y reduciendo los costos operativos (Voyant y otros., 2017).

PARQUE EÓLICO EN DINAMARCA

En Dinamarca, un parque eólico ha implementado sistemas de predicción basados en IA para mejorar la integración de la energía eólica en la red eléctrica. Los modelos de IA analizan datos sobre la velocidad y dirección del viento para predecir la producción de energía, permitiendo una mejor planificación y gestión de la red (Monteiro y otros., 2009).

La inteligencia artificial está transformando el sector de la energía renovable al mejorar la eficiencia, reducir costos y aumentar la fiabilidad de la generación y distribución de energía. Al facilitar la integración de fuentes de energía renovable, la IA está desempeñando un papel crucial en la transición hacia un sistema energético más sostenible y en la lucha contra el cambio climático.

GESTIÓN DE RECURSOS NATURALES

La gestión sostenible de los recursos naturales es fundamental para garantizar la salud del planeta y el bienestar de las futuras generaciones. La inteligencia artificial (IA) está emergiendo como una herramienta poderosa para optimizar el uso de estos recursos, mejorar la eficiencia y reducir el impacto ambiental. Este capítulo explora cómo la IA está transformando la gestión de recursos naturales, con un enfoque en la agricultura de precisión y la conservación de la biodiversidad. La agricultura de precisión utiliza tecnologías avanzadas para optimizar la producción agrícola y minimizar el impacto ambiental. La IA juega un papel crucial en este campo al analizar grandes volúmenes de datos y proporcionar recomendaciones precisas para la gestión de cultivos.

Los drones y sensores equipados con IA pueden recopilar datos en tiempo real sobre el estado de los cultivos, el suelo y las condiciones climáticas (Zhang y otros., 2019). Estos datos se analizan para identificar patrones y tendencias que pueden afectar la productividad agrícola. Por ejemplo, los algoritmos de IA pueden detectar signos tempranos de enfermedades o plagas en los cultivos, permitiendo una intervención oportuna y reduciendo la necesidad de pesticidas (Liakos y otros., 2018).

La IA puede optimizar el uso de agua, fertilizantes y otros insumos agrícolas, mejorando la eficiencia y reduciendo el desperdicio. Los sistemas de riego inteligentes basados en IA pueden ajustar automáticamente la cantidad de agua aplicada en función de las necesidades específicas de cada planta y las condiciones del suelo (Shamshiri y otros., 2018). Esto no solo mejora la productividad, sino que también conserva recursos valiosos y reduce el impacto ambiental. La conser-

vación de la biodiversidad es esencial para mantener los ecosistemas saludables y resilientes. La IA está desempeñando un papel cada vez más importante en la monitorización y protección de especies en peligro y en la gestión de hábitats naturales.

Los sistemas de IA pueden analizar datos de cámaras trampa, grabaciones de audio y otras fuentes para monitorear la presencia y comportamiento de especies en sus hábitats naturales (Willi y otros., 2019). Estos sistemas pueden identificar automáticamente diferentes especies y rastrear sus movimientos, proporcionando información valiosa para los esfuerzos de conservación. Por ejemplo, los algoritmos de reconocimiento de imágenes pueden identificar especies en peligro de extinción y alertar a los conservacionistas sobre posibles amenazas (Norouzzadeh y otros., 2018).

La IA también puede ayudar a gestionar y restaurar hábitats naturales. Los modelos de IA pueden predecir cómo los cambios en el uso del suelo, el clima y otras variables afectarán a los ecosistemas, permitiendo a los gestores de recursos tomar decisiones informadas para proteger y restaurar estos hábitats (Joppa y otros., 2016). Además, los drones equipados con IA pueden mapear áreas extensas y monitorear la salud de los ecosistemas, identificando áreas que requieren intervención (Duarte y otros., 2018).

El uso de la IA en la gestión de recursos naturales ofrece numerosos beneficios, entre los que se incluyen:

- **Mejora de la eficiencia:** La IA optimiza el uso de recursos, reduciendo el desperdicio y mejorando la productividad (Shamshiri y otros., 2018).
- **Reducción del impacto ambiental:** Al minimizar el uso de insumos y mejorar la gestión de los recursos, la IA contribuye a reducir el impacto ambiental de las actividades humanas (Liakos y otros., 2018).
- **Protección de la biodiversidad:** La IA facilita la monitorización y protección de especies en peligro y la gestión de hábitats naturales, contribuyendo a la conservación de la biodiversidad (Norouzzadeh y otros., 2018).

- **Toma de decisiones informada:** La capacidad de la IA para analizar grandes volúmenes de datos permite una toma de decisiones más informada y basada en evidencia (Joppa y otros., 2016).

CASOS DE ESTUDIO: AGRICULTURA DE PRECISIÓN EN ESPAÑA

En España, un proyecto de agricultura de precisión utiliza drones y sensores equipados con IA para monitorear los cultivos y optimizar el uso de agua y fertilizantes. Los algoritmos de IA analizan datos en tiempo real para proporcionar recomendaciones precisas, mejorando la productividad y reduciendo el impacto ambiental (Zhang y otros., 2019).

CONSERVACIÓN DE LA BIODIVERSIDAD EN ÁFRICA

En África, un proyecto de conservación utiliza cámaras trampa y sistemas de IA para monitorear la presencia de especies en peligro de extinción. Los algoritmos de reconocimiento de imágenes identifican automáticamente diferentes especies y alertan a los conservacionistas sobre posibles amenazas, mejorando la eficacia de los esfuerzos de conservación (Norouzzadeh y otros., 2018).

La inteligencia artificial está transformando la gestión de recursos naturales al mejorar la eficiencia, reducir el impacto ambiental y proteger la biodiversidad. Al aprovechar las capacidades de la IA para analizar datos y proporcionar recomendaciones precisas, podemos avanzar hacia una gestión más sostenible de nuestros recursos naturales y garantizar un futuro más saludable para el planeta.

REDUCCIÓN DE EMISIONES Y CONTAMINACIÓN

La reducción de emisiones y la mitigación de la contaminación son esenciales para combatir el cambio climático y proteger la salud pú-

blica. La inteligencia artificial (IA) está desempeñando un papel crucial en estos esfuerzos al proporcionar herramientas avanzadas para monitorear, predecir y gestionar las emisiones y la calidad del aire. Este capítulo explora cómo la IA está revolucionando el transporte inteligente y el monitoreo de la calidad del aire para reducir la contaminación y promover un entorno más saludable. El transporte es una de las principales fuentes de emisiones de gases de efecto invernadero y contaminación del aire. La IA está transformando el sector del transporte al optimizar la gestión del tráfico, promover el uso de vehículos eléctricos y mejorar la eficiencia del transporte público.

Los sistemas de gestión del tráfico basados en IA utilizan datos en tiempo real de sensores, cámaras y dispositivos GPS para optimizar el flujo de tráfico en las ciudades (Chen et al., 2021). Estos sistemas pueden ajustar automáticamente los semáforos, identificar rutas alternativas y gestionar incidentes de tráfico, reduciendo los atascos y las emisiones de vehículos. Por ejemplo, en Singapur, el sistema de gestión del tráfico basado en IA ha logrado reducir los tiempos de viaje y las emisiones de carbono al optimizar el flujo de tráfico en tiempo real (González y otros., 2020).

La IA también está desempeñando un papel importante en la promoción del uso de vehículos eléctricos (VE). Los algoritmos de IA pueden optimizar la ubicación de las estaciones de carga, predecir la demanda de carga y gestionar la carga de los VE para minimizar el impacto en la red eléctrica (Zhang y otros., 2018). Además, los sistemas de navegación basados en IA pueden ayudar a los conductores de VE a encontrar las rutas más eficientes y las estaciones de carga disponibles, facilitando la adopción de esta tecnología (Chen y otros., 2021).

La IA puede mejorar la eficiencia y la fiabilidad del transporte público al analizar datos de pasajeros, horarios y tráfico. Los sistemas de IA pueden optimizar las rutas y los horarios de los autobuses y trenes, reducir los tiempos de espera y mejorar la experiencia del usuario (González y otros., 2020). En Londres, por ejemplo, el sistema de transporte público utiliza IA para ajustar los horarios de los autobuses en función de la demanda y las condiciones del tráfico, mejorando la eficiencia y reduciendo las emisiones (Chen y otros., 2021).

El monitoreo de la calidad del aire es crucial para identificar y mitigar las fuentes de contaminación. La IA está revolucionando este campo al proporcionar herramientas avanzadas para medir, predecir y gestionar la calidad del aire en tiempo real. Los sensores de calidad del aire equipados con IA pueden medir la concentración de contaminantes como el dióxido de nitrógeno (NO2), el ozono (O3) y las partículas finas (PM2.5) en tiempo real (Zhang y otros., 2018). Los algoritmos de IA analizan estos datos para identificar patrones y tendencias, permitiendo a las autoridades tomar medidas rápidas y efectivas para reducir la contaminación. Por ejemplo, en Beijing, un sistema de monitoreo de la calidad del aire basado en IA ha ayudado a reducir los niveles de contaminación al identificar las principales fuentes de emisiones y recomendar acciones específicas (Wang y otros., 2018).

La IA puede predecir la calidad del aire con alta precisión al analizar datos meteorológicos, de tráfico y de emisiones industriales (Hong y otros., 2020). Estos modelos de predicción permiten a las autoridades anticipar episodios de contaminación y tomar medidas preventivas, como restringir el tráfico o cerrar fábricas temporariamente. En Los Ángeles, un sistema de predicción de la calidad del aire basado en IA ha mejorado la precisión de las predicciones y ha permitido a las autoridades implementar medidas más efectivas para proteger la salud pública (Chen y otros., 2021).

El uso de la IA en la reducción de emisiones y la mitigación de la contaminación ofrece numerosos beneficios, entre los que se incluyen:

- **Mejora de la eficiencia:** La IA optimiza la gestión del tráfico y el transporte público, reduciendo los tiempos de viaje y las emisiones de vehículos (González y otros., 2020).
- **Reducción de la contaminación:** Los sistemas de monitoreo y predicción de la calidad del aire basados en IA permiten identificar y mitigar las fuentes de contaminación de manera más efectiva (Wang y otros., 2018).
- **Promoción de la sostenibilidad:** Al facilitar la adopción de vehículos eléctricos y mejorar la eficiencia del transporte público, la IA contribuye a la reducción de las emisiones de gases de

efecto invernadero y a la promoción de un entorno más sostenible (Zhang y otros., 2018).

- **Protección de la salud pública:** La capacidad de la IA para predecir y gestionar la calidad del aire permite a las autoridades tomar medidas preventivas para proteger la salud de la población (Hong y otros., 2020).

CASOS DE ESTUDIO: SISTEMA DE GESTIÓN DEL TRÁFICO EN SINGAPUR

En Singapur, un sistema de gestión del tráfico basado en IA ha logrado reducir los tiempos de viaje y las emisiones de carbono al optimizar el flujo de tráfico en tiempo real. Los algoritmos de IA analizan datos de sensores y cámaras para ajustar automáticamente los semáforos y gestionar incidentes de tráfico, mejorando la eficiencia y reduciendo la contaminación (González y otros., 2020).

MONITOREO DE LA CALIDAD DEL AIRE EN BEIJING

En Beijing, un sistema de monitoreo de la calidad del aire basado en IA ha ayudado a reducir los niveles de contaminación al identificar las principales fuentes de emisiones y recomendar acciones específicas. Los sensores de calidad del aire equipados con IA miden la concentración de contaminantes en tiempo real, y los algoritmos de IA analizan estos datos para identificar patrones y tendencias (Wang y otros., 2018).

La inteligencia artificial está transformando la reducción de emisiones y la mitigación de la contaminación al proporcionar herramientas avanzadas para optimizar el transporte, monitorear la calidad del aire y promover la sostenibilidad. Al aprovechar las capacidades de la IA para analizar datos y proporcionar recomendaciones precisas, podemos avanzar hacia un entorno más saludable y sostenible.

ECONOMÍA CIRCULAR Y GESTIÓN DE RESIDUOS

La economía circular es un modelo económico que busca minimizar el desperdicio y maximizar el uso de los recursos mediante la reutilización, el reciclaje y la regeneración de productos y materiales. La inteligencia artificial (IA) está desempeñando un papel crucial en la promoción de la economía circular al optimizar los procesos de reciclaje y gestión de residuos. Este capítulo explora cómo la IA está revolucionando la economía circular y contribuyendo a la sostenibilidad.

El reciclaje es una de las piedras angulares de la economía circular. La IA está transformando los procesos de reciclaje al mejorar la clasificación de materiales y optimizar la eficiencia de las plantas de reciclaje. Los sistemas de clasificación de residuos basados en IA utilizan sensores y algoritmos de aprendizaje automático para identificar y separar diferentes tipos de materiales con alta precisión (Ragossnig y Schneider, 2019). Estos sistemas pueden distinguir entre plásticos, metales, vidrio y otros materiales, mejorando la pureza de los materiales reciclados y aumentando las tasas de reciclaje. Por ejemplo, los sistemas de visión por computadora equipados con IA pueden analizar imágenes de residuos en cintas transportadoras y clasificar los materiales en tiempo real (Zhang y otros., 2019).

La IA también puede optimizar el funcionamiento de las plantas de reciclaje al analizar datos de producción y operación. Los algoritmos de IA pueden identificar cuellos de botella en el proceso de reciclaje y recomendar ajustes para mejorar la eficiencia y reducir los costos operativos (Ellen MacArthur Foundation, 2019). Además, los sistemas de mantenimiento predictivo basados en IA pueden anticipar fallos en los equipos y programar el mantenimiento de manera proactiva, minimizando el tiempo de inactividad y mejorando la productividad (Lee y otros., 2018).

La reducción de residuos es otro componente clave de la economía circular. La IA está ayudando a las empresas a minimizar la generación de residuos mediante la optimización de procesos y la mejora de la eficiencia en la producción. Los sistemas de IA pueden analizar datos de producción para identificar oportunidades de mejora y reducir el desperdicio de materiales (Rüßmann y otros., 2015). Por ejemplo, en

la industria manufacturera, los algoritmos de IA pueden optimizar el uso de materias primas y reducir el desperdicio durante el proceso de fabricación. Esto no solo mejora la eficiencia, sino que también reduce los costos y el impacto ambiental.

La IA también está desempeñando un papel importante en el diseño de productos sostenibles. Los algoritmos de IA pueden analizar el ciclo de vida de los productos y recomendar cambios en el diseño para mejorar su reciclabilidad y reducir su impacto ambiental (Bauer y otros., 2018). Por ejemplo, los sistemas de diseño asistido por IA pueden sugerir materiales alternativos y métodos de fabricación que minimicen el uso de recursos y faciliten el reciclaje al final de la vida útil del producto.

El uso de la IA en la economía circular y la gestión de residuos ofrece numerosos beneficios, entre los que se incluyen:

- **Mejora de la eficiencia:** La IA optimiza los procesos de reciclaje y producción, reduciendo el desperdicio y mejorando la eficiencia operativa (Ragossnig y Schneider, 2019).
- **Reducción de costos:** La optimización de procesos y la reducción de residuos pueden llevar a una disminución significativa de los costos operativos (Ellen MacArthur Foundation, 2019).
- **Promoción de la sostenibilidad:** Al facilitar el reciclaje y la reducción de residuos, la IA contribuye a la promoción de prácticas sostenibles y a la reducción del impacto ambiental (Bauer y otros., 2018).
- **Innovación en el diseño de productos:** La IA abre nuevas posibilidades para el diseño de productos sostenibles, mejorando su reciclabilidad y reduciendo su impacto ambiental (Rüßmann y otros., 2015).

CASOS DE ESTUDIO:PLANTA DE RECICLAJE EN ALEMANIA

En Alemania, una planta de reciclaje ha implementado un sistema de clasificación de residuos basado en IA que utiliza sensores y algoritmos de visión por computadora para identificar y separar diferentes

tipos de materiales. Este sistema ha mejorado la pureza de los materiales reciclados y ha aumentado las tasas de reciclaje, reduciendo al mismo tiempo los costos operativos (Zhang y otros., 2019).

EMPRESA DE MANUFACTURA EN JAPÓN

En Japón, una empresa de manufactura ha utilizado IA para optimizar sus procesos de producción y reducir el desperdicio de materiales. Los algoritmos de IA analizan datos de producción en tiempo real y recomiendan ajustes para mejorar la eficiencia y minimizar el desperdicio. Esto ha llevado a una reducción significativa de los costos y del impacto ambiental de la empresa (Rüßmann y otros., 2015).

La inteligencia artificial está transformando la economía circular y la gestión de residuos al proporcionar herramientas avanzadas para optimizar los procesos de reciclaje, reducir el desperdicio y promover la sostenibilidad. Al aprovechar las capacidades de la IA para analizar datos y proporcionar recomendaciones precisas, podemos avanzar hacia una economía más circular y sostenible, reduciendo el impacto ambiental y promoviendo un uso más eficiente de los recursos.

CASOS DE ESTUDIO

Los casos de estudio proporcionan ejemplos concretos de cómo la inteligencia artificial (IA) está siendo utilizada para promover la sostenibilidad y abordar los desafíos ambientales. Este capítulo presenta varios casos de estudio que ilustran el impacto positivo de la IA en diferentes sectores, desde la energía renovable hasta la gestión de residuos. Estos ejemplos destacan las innovaciones y los beneficios tangibles que la IA puede ofrecer en la búsqueda de un futuro más sostenible.

Caso de Estudio 1: Optimización de Energía Solar en California

En California, un proyecto de energía solar ha implementado sistemas de IA para optimizar la producción y distribución de energía. Los al-

goritmos de predicción analizan datos meteorológicos y de radiación solar para ajustar la producción en tiempo real, mejorando la eficiencia y reduciendo los costos operativos (Voyant y otros., 2017). Este enfoque ha permitido una integración más efectiva de la energía solar en la red eléctrica, reduciendo la dependencia de fuentes de energía no renovables y disminuyendo las emisiones de gases de efecto invernadero.

El uso de IA en este proyecto ha resultado en una mejora significativa en la precisión de las predicciones de producción de energía, lo que ha permitido una mejor planificación y gestión de la red eléctrica. Además, la optimización en tiempo real ha reducido los costos operativos y ha aumentado la eficiencia general del sistema (Voyant y otros., 2017).

Caso de Estudio 2: Agricultura de Precisión en España

En España, un proyecto de agricultura de precisión utiliza drones y sensores equipados con IA para monitorear los cultivos y optimizar el uso de agua y fertilizantes. Los algoritmos de IA analizan datos en tiempo real para proporcionar recomendaciones precisas, mejorando la productividad y reduciendo el impacto ambiental (Zhang y otros., 2019). Este enfoque ha permitido a los agricultores tomar decisiones más informadas y eficientes, promoviendo prácticas agrícolas sostenibles.

La implementación de IA en la agricultura de precisión ha llevado a una reducción significativa en el uso de agua y fertilizantes, mejorando la sostenibilidad de las prácticas agrícolas. Además, la capacidad de detectar problemas de manera temprana ha aumentado la productividad y ha reducido las pérdidas de cultivos (Liakos y otros., 2018).

Caso de Estudio 3: Gestión del Tráfico en Singapur

En Singapur, un sistema de gestión del tráfico basado en IA ha logrado reducir los tiempos de viaje y las emisiones de carbono al optimizar el flujo de tráfico en tiempo real. Los algoritmos de IA analizan datos de sensores y cámaras para ajustar automáticamente los semáforos y

gestionar incidentes de tráfico, mejorando la eficiencia y reduciendo la contaminación (González y otros., 2020).

El sistema de gestión del tráfico ha resultado en una disminución significativa de los atascos y una mejora en la calidad del aire. Los tiempos de viaje se han reducido, lo que ha llevado a una mayor satisfacción de los usuarios y una reducción en las emisiones de vehículos (Chen y otros., 2021).

Caso de Estudio 4: Monitoreo de la Calidad del Aire en Beijing

En Beijing, un sistema de monitoreo de la calidad del aire basado en IA ha ayudado a reducir los niveles de contaminación al identificar las principales fuentes de emisiones y recomendar acciones específicas. Los sensores de calidad del aire equipados con IA miden la concentración de contaminantes en tiempo real, y los algoritmos de IA analizan estos datos para identificar patrones y tendencias (Wang y otros., 2018). El sistema ha mejorado la precisión de las predicciones de calidad del aire y ha permitido a las autoridades tomar medidas más efectivas para reducir la contaminación. Esto ha llevado a una mejora en la calidad del aire y una reducción en los riesgos para la salud pública (Hong y otros., 2020).

Caso de Estudio 5: Planta de Reciclaje en Alemania

En Alemania, una planta de reciclaje ha implementado un sistema de clasificación de residuos basado en IA que utiliza sensores y algoritmos de visión por computadora para identificar y separar diferentes tipos de materiales. Este sistema ha mejorado la pureza de los materiales reciclados y ha aumentado las tasas de reciclaje, reduciendo al mismo tiempo los costos operativos (Zhang y otros., 2019). La implementación de IA en la planta de reciclaje ha llevado a una mayor eficiencia en la clasificación de materiales y una reducción en los residuos enviados a los vertederos. Además, la mejora en la pureza de los materiales reciclados ha aumentado su valor en el mercado, beneficiando económicamente a la planta (Ragossnig y Schneider, 2019).

Estos casos de estudio demuestran el impacto positivo de la inteligencia artificial en la promoción de la sostenibilidad y la gestión de los recursos naturales. Al proporcionar herramientas avanzadas para optimizar procesos, reducir costos y mejorar la eficiencia, la IA está desempeñando un papel crucial en la transición hacia un futuro más sostenible. Estos ejemplos destacan las innovaciones y los beneficios tangibles que la IA puede ofrecer en diversos sectores, desde la energía renovable hasta la gestión de residuos.

DESAFÍOS Y CONSIDERACIONES ÉTICAS

La inteligencia artificial (IA) ofrece numerosas oportunidades para promover la sostenibilidad y abordar los desafíos ambientales. Sin embargo, su implementación también plantea una serie de desafíos y consideraciones éticas que deben ser abordados para garantizar su uso responsable y equitativo. Este capítulo explora los principales desafíos y consideraciones éticas asociados con el uso de la IA en la sostenibilidad, incluyendo los riesgos y limitaciones, así como las cuestiones éticas y de responsabilidad. Aunque la IA tiene el potencial de transformar diversos sectores, su implementación no está exenta de riesgos y limitaciones. Es crucial reconocer y abordar estos desafíos para maximizar los beneficios de la IA y minimizar sus impactos negativos.

Uno de los principales riesgos asociados con la IA es el sesgo en los algoritmos. Los sistemas de IA aprenden de los datos con los que son entrenados, y si estos datos contienen sesgos, los algoritmos pueden perpetuar y amplificar estos sesgos (O'Neil, 2016). Esto puede llevar a decisiones injustas y discriminatorias, afectando negativamente a ciertos grupos de personas. Por ejemplo, en el contexto de la sostenibilidad, un algoritmo sesgado podría priorizar la eficiencia energética en áreas urbanas ricas mientras descuida las necesidades de las comunidades rurales o desfavorecidas (Barocas y otros., 2019).

La implementación de sistemas de IA a menudo requiere la recopilación y análisis de grandes volúmenes de datos, lo que plantea preocupaciones sobre la privacidad y la seguridad de los datos (Zuboff, 2019). Es esencial garantizar que los datos sean recopilados y utilizados de

manera ética y que se implementen medidas de seguridad adecuadas para proteger la información sensible. La falta de transparencia en el uso de datos puede erosionar la confianza pública y limitar la aceptación de las tecnologías de IA (Floridi y otras., 2018).

Aunque la IA puede contribuir a la sostenibilidad, también tiene un impacto ambiental significativo debido al consumo de energía asociado con el entrenamiento y la operación de modelos de IA (Strubell y otras., 2019). Es importante considerar el balance entre los beneficios ambientales de la IA y su huella de carbono, y buscar formas de hacer que los sistemas de IA sean más eficientes energéticamente. Además de los riesgos y limitaciones técnicos, el uso de la IA en la sostenibilidad plantea importantes cuestiones éticas y de responsabilidad. Es fundamental abordar estas cuestiones para garantizar que la IA se utilice de manera justa y equitativa.

La transparencia y la explicabilidad son principios clave en el uso ético de la IA. Los sistemas de IA deben ser transparentes en cuanto a cómo toman decisiones y qué datos utilizan (Doshi-Velez y Kim, 2017). La explicabilidad se refiere a la capacidad de los sistemas de IA para proporcionar explicaciones comprensibles de sus decisiones, lo que es crucial para garantizar la responsabilidad y la confianza en estas tecnologías. En el contexto de la sostenibilidad, la transparencia y la explicabilidad pueden ayudar a garantizar que las decisiones tomadas por los sistemas de IA sean justas y equitativas (Floridi y otros., 2018).

La justicia y la equidad son consideraciones éticas fundamentales en el uso de la IA. Es esencial garantizar que los beneficios de la IA se distribuyan de manera equitativa y que no se perpetúen o amplifiquen las desigualdades existentes (Barocas y otros., 2019). Por ejemplo, en la gestión de recursos naturales, es importante asegurarse de que las comunidades desfavorecidas no sean excluidas de los beneficios de las tecnologías de IA y que sus necesidades y perspectivas sean consideradas en el diseño y la implementación de estas tecnologías.

La responsabilidad y la rendición de cuentas son cruciales para el uso ético de la IA. Es necesario establecer mecanismos claros para asignar la responsabilidad en caso de errores o daños causados por los sistemas de IA (Binns, 2018). Esto incluye la responsabilidad de

los desarrolladores, los operadores y los usuarios de las tecnologías de IA. En el contexto de la sostenibilidad, la rendición de cuentas puede ayudar a garantizar que las tecnologías de IA se utilicen de manera responsable y que se tomen medidas correctivas en caso de impactos negativos.

La implementación de la inteligencia artificial en la sostenibilidad ofrece numerosas oportunidades, pero también plantea importantes desafíos y consideraciones éticas. Es esencial abordar estos desafíos y garantizar que la IA se utilice de manera justa, equitativa y responsable. Al considerar los riesgos y limitaciones, así como las cuestiones éticas y de responsabilidad, podemos maximizar los beneficios de la IA y avanzar hacia un futuro más sostenible y justo.

CONCLUSIÓN

La inteligencia artificial tiene el potencial de ser un aliado poderoso en la búsqueda de un futuro sostenible. Sin embargo, para maximizar sus beneficios y minimizar sus riesgos, es esencial abordar los desafíos técnicos y éticos asociados con su implementación. Esto incluye garantizar la transparencia y la explicabilidad de los sistemas de IA, promover la justicia y la equidad en su uso, y establecer mecanismos claros de responsabilidad y rendición de cuentas (Doshi-Velez y Kim, 2017; Barocas y otros., 2019).

Además, es crucial considerar el impacto ambiental de la IA y buscar formas de hacer que los sistemas de IA sean más eficientes energéticamente (Strubell y otros., 2019). Al equilibrar los beneficios ambientales de la IA con su huella de carbono, podemos avanzar hacia un uso más sostenible de esta tecnología. La integración de la inteligencia artificial en la sostenibilidad no es solo una cuestión de tecnología, sino también de valores y principios éticos. Al adoptar un enfoque responsable y equitativo, podemos garantizar que la IA contribuya de manera significativa a la protección del planeta y al bienestar de las futuras generaciones. La colaboración entre científicos, ingenieros, responsables de políticas y la sociedad en general será fundamental para lograr este objetivo.

En resumen, la IA ofrece oportunidades sin precedentes para abordar los desafíos ambientales y promover la sostenibilidad. Al enfrentar los desafíos y consideraciones éticas con determinación y responsabilidad, podemos aprovechar todo el potencial de la IA para construir un futuro más sostenible y justo.

REFERENCIAS

Barocas, S., Hardt, M., y Narayanan, A. (2019). *Fairness and Machine Learning*. fairmlbook.org.

Bauer, F., Brown, C., y Scavarda, A. (2018). The role of innovation in sustainable development. *Journal of Cleaner Production, 198*, 1452-1463.

Brundtland, G. H. (1987). *Our Common Future: Report of the World Commission on Environment and Development*. Oxford University Press.

Brynjolfsson, E., y McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company.

Chen, T., He, F., y Chen, Y. (2021). Intelligent transportation systems: A review. *IEEE Transactions on Intelligent Transportation Systems, 22*(1), 1-15.

Davenport, T. H., y Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review, 96*(1), 108-116.

Doshi-Velez, F., y Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv: 1702.08608*.

Duarte, C. M., Agusti, S., Barbier, E., Britten, G. L., Castilla, J. C., Gattuso, J. P., ... y Worm, B. (2018). Rebuilding marine life. *Nature, 580*(7801), 39-51.

Ellen MacArthur Foundation. (2019). *Artificial Intelligence and the Circular Economy*. Ellen MacArthur Foundation.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... y Vayena, E. (2018). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689-707.

Goodfellow, I., Bengio, Y., y Courville, A. (2016). *Deep Learning*. MIT Press.

González, A., Valero, E., y García, J. (2020). Smart cities: A review of the concept and its applications. *Sustainable Cities and Society, 53*, 101-112.

Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., y Hancke, G. P. (2013). Smart grid technologies: Communication technologies and standards. *IEEE Transactions on Industrial Informatics, 7*(4), 529-539.

Hong, T., Pinson, P., y Fan, S. (2020). Global energy forecasting competition 2012: Benchmarking state-of-the-art methods. *International Journal of Forecasting, 30*(2), 191-207.

IEA. (2021). *Energy Efficiency 2021*. International Energy Agency.

Joppa, L. N., Loarie, S. R., y Pimm, S. L. (2016). On the protection of "protected areas". *Proceedings of the National Academy of Sciences, 113*(22), 6077-6081.

Kamilaris, A., y Prenafeta-Boldú, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science, 156*(3), 312-322.

LeCun, Y., Bengio, Y., y Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Lee, J., Bagheri, B., y Kao, H. A. (2018). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters, 3*, 18-23.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., y Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors, 18*(8), 2674.

Mckinsey y Company. (2018). Notes from the AI frontier: Insights from hundreds of use cases. McKinsey Global Institute.

Monteiro, C., Ramirez-Rosado, I. J., y Fernandez-Jimenez, L. A. (2009). Short-term forecasting models for wind power production: A case study in northern Spain. *Energy Conversion and Management, 50*(5), 1131-1137.

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., y Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences, 115*(25), E5716-E5725.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.

Ragossnig, A. M., y Schneider, D. R. (2019). Circular economy and waste management. *Waste Management & Research, 37*(2), 109-111.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... y Bengio, Y. (2019). Tackling climate change with machine learning. *arXiv preprint arXiv: 1906.05433*.

Russell, S., y Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., y Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group.

Shamshiri, R. R., Kalantari, F., Ting, K. C., Thorp, K. R., Hameed, I. A., Weltzien, C., ... y Ehsani, R. (2018). Advances in greenhouse automation and controlled environment agriculture: A transition to plant factories and urban agriculture. *International Journal of Agricultural and Biological Engineering, 11*(1), 1-22.

Strubell, E., Ganesh, A., y McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645-3650.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., y Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy, 105*, 569-582.

Wang, J., Zhong, H., Ma, Z., Xia, Q., y Kang, C. (2018). Review and prospect of integrated demand response in the multi-energy system. *Applied Energy, 202*, 772-782.

Willi, M., Pitman, R. T., Cardoso, A. W., Locke, C., Swanson, A., Boyer, A., ... y Fortson, L. (2019). Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution, 10*(1), 80-91.

Zhang, C., Walters, D., y Agrawal, A. (2019). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science, 24*(2), 110-124.

Zhang, Y., Wang, J., y Wang, X. (2018). Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews, 32*, 255-270.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffa

CAPÍTULO XIX

TRASHUMANISMO Y LA FUSIÓN DE HOMBRE-MAQUINA: ¿PODRÍAMOS LLEGAR A UNA SIMBIOSIS CON MÁQUINAS INTELIGENTES?

EMILY NICOLE SILVA ESCAMILLA[1]

RESUMEN

El transhumanismo es una corriente filosófica y científica que busca mejorar las capacidades humanas a través de tecnologías avanzadas, superando así las limitaciones biológicas. Con el avance de la tecnología, surge la posibilidad de fusionar humanos y máquinas inteligentes, creando una relación simbiótica que redefine lo que significa ser humano. Ejemplos actuales incluyen las interfaces cerebro-computadora y las prótesis biónicas, así como los dispositivos implantables de *Neuralink,* que podrían conectar directamente el cerebro humano con la inteligencia artificial (IA).

Aunque esta fusión no es un concepto nuevo, los recientes avances en IA y bioingeniería permiten una integración más profunda. La IA puede mejorar nuestras capacidades cognitivas y físicas, pero también plantea desafíos éticos, como la redefinición de la humanidad y la posible creación de desigualdades sociales basadas en el acceso a estas tecnologías.

La simbiosis humano-máquina implica una interdependencia que podría beneficiar a ambas partes: las máquinas pueden ayudar a los huma-

[1] Alumna de la Licenciatura en Ingeniería Metalúrgica, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

nos a superar limitaciones, mientras que los humanos aportan creatividad para resolver problemas. A medida que nos acercamos a esta fusión, es fundamental reflexionar sobre cómo estas tecnologías transformarán nuestra identidad y nuestro futuro como seres humanos.

Palabras clave: Transhumanismo, simbiosis, inteligencia artificial.

INTRODUCCIÓN

El transhumanismo es una corriente filosófica y científica que aboga por el uso de tecnologías avanzadas para mejorar las capacidades físicas, cognitivas y emocionales de los seres humanos, superando las limitaciones biológicas. A medida que la tecnología sigue avanzando, surge la pregunta de si los humanos podrán fusionarse con máquinas inteligentes, creando una relación de simbiosis entre ambas entidades. Este concepto plantea un escenario en el que las personas no solo usen la tecnología como herramienta, sino que integren las máquinas en su biología, permitiendo un grado de interdependencia que podría redefinir lo que significa ser humano.

En la actualidad, el camino hacia esta fusión ya está siendo trazado mediante desarrollos en campos como la inteligencia artificial (IA), la robótica avanzada y la neurotecnología. Las interfaces cerebro-computadora (BCI, por sus siglas en inglés), las prótesis biónicas y los implantes neurales son ejemplos de cómo la tecnología ya comienza a formar parte de la biología humana. Elon Musk, a través de su empresa *Neuralink,* ha propuesto el uso de dispositivos implantables que conecten directamente el cerebro humano con la inteligencia artificial, facilitando la mejora de habilidades cognitivas y físicas, y permitiendo, potencialmente, una conexión constante con redes de información (*Musk y Neuralink*, 2019). Este tipo de tecnologías representan un paso importante hacia la simbiosis humano-máquina, que podría alterar la evolución de nuestra especie.

La idea de la fusión entre humanos y máquinas no es nueva. Desde la aparición de las primeras herramientas hasta los actuales disposi-

tivos electrónicos, los seres humanos han buscado formas de ampliar sus capacidades a través de la tecnología. Sin embargo, en las últimas décadas, los avances en IA y bioingeniería han abierto la posibilidad de una integración mucho más profunda. Las prótesis robóticas controladas por señales neuronales, los exoesqueletos motorizados y los implantes cocleares son solo el principio de lo que podría ser una fusión biotecnológica más sofisticada (Herr, 2020).

La inteligencia artificial, en particular, juega un papel crucial en esta posible fusión. Los sistemas de IA ya son capaces de realizar tareas complejas, analizar grandes cantidades de datos y aprender de manera autónoma. Estas habilidades podrían integrarse en el cerebro humano, mejorando nuestra capacidad para procesar información, resolver problemas y tomar decisiones. Kurzweil (2005) sugiere que en un futuro cercano, la IA se fusionará con la inteligencia humana, creando una entidad híbrida capaz de superar las limitaciones biológicas del cerebro humano. Este concepto, conocido como singularidad tecnológica, prevé un escenario en el que las máquinas y los humanos se beneficien mutuamente en una simbiosis perfecta.

Aunque la idea de una fusión entre humanos y máquinas ofrece grandes oportunidades, también plantea importantes desafíos éticos y filosóficos. En primer lugar, surge la pregunta sobre qué significa ser humano en un mundo donde nuestra identidad está profundamente conectada a la tecnología. Si los humanos mejoran sus capacidades mediante la integración de IA y dispositivos biotecnológicos, ¿seguiremos siendo humanos o evolucionaremos hacia una nueva forma de vida? Según Bostrom (2014), estas mejoras podrían llevarnos a una nueva fase de la evolución, creando lo que algunos llaman “posthumanos”, seres que han trascendido las limitaciones biológicas mediante la tecnología.

Además, la distribución desigual de estas tecnologías podría dar lugar a nuevas formas de inequidad social. El acceso a las mejoras tecnológicas podría estar limitado a las personas que puedan permitírselas, lo que crearía una sociedad dividida entre los que tienen acceso a capacidades sobrehumanas y aquellos que no (Bostrom, 2005). La posibilidad de crear una clase élite de “superhumanos” plantea cuestiones sobre justicia, derechos humanos y la equidad en el acceso a las tecnologías transhumanistas.

En biología, la simbiosis describe una relación mutuamente beneficiosa entre dos organismos. Aplicado al contexto humano-máquina, este concepto implicaría que los humanos y las máquinas no solo coexistan, sino que dependan el uno del otro para mejorar. En este sentido, la inteligencia artificial podría ayudar a los humanos a superar sus limitaciones cognitivas, mientras que los humanos, con sus capacidades creativas e intuitivas, podrían ayudar a las máquinas a enfrentar problemas complejos. La simbiosis entre humanos y máquinas, por lo tanto, no sería simplemente una integración de tecnología en la biología, sino una evolución conjunta de ambas entidades.

La tecnología ya está moviéndose en esta dirección. *Musk y Neuralink* (2019) han presentado avances en dispositivos que permiten la comunicación directa entre el cerebro humano y las máquinas, y los desarrollos en neurotecnología han mostrado que es posible mejorar la memoria, restaurar la visión y tratar enfermedades neurológicas mediante la implantación de chips en el cerebro. Estos avances sugieren que una simbiosis completa entre humanos y máquinas podría ser más que una simple visión de ciencia ficción.

La posibilidad de una simbiosis entre humanos y máquinas inteligentes está cada vez más cerca, impulsada por avances en IA, neurotecnología y bioingeniería. El transhumanismo ofrece una visión futurista en la que las limitaciones biológicas del ser humano pueden superarse mediante la integración tecnológica. Sin embargo, este camino hacia la fusión hombre-máquina también está lleno de desafíos éticos y filosóficos que deben abordarse con cautela. Mientras continuamos explorando las posibilidades de una relación simbiótica con máquinas inteligentes, es crucial que nos preguntemos cómo estas tecnologías transformarán no solo nuestras capacidades, sino también nuestra identidad como seres humanos.

ORIGEN DEL MOVIMIENTO TRANSHUMANISTA

El transhumanismo como movimiento filosófico y cultural tiene sus raíces en diversas corrientes de pensamiento que abogan por el uso de la tecnología para mejorar la condición humana. Aunque el término

"*transhumanismo*" fue acuñado en 1957 por el biólogo Julian Huxley, la idea de utilizar la tecnología para potenciar las capacidades humanas ha existido a lo largo de la historia. Huxley describió el transhumanismo como una forma de evolución en la que la humanidad podría superar sus limitaciones biológicas a través de la ciencia y la tecnología. Su obra "*New Bottles for New Wine*" establece las bases del pensamiento transhumanista al plantear la necesidad de un cambio en la condición humana mediante la biotecnología y otras innovaciones.

A lo largo de las décadas siguientes, el transhumanismo comenzó a tomar forma más concreta, especialmente a partir de la década de 1980, cuando surgieron grupos e individuos interesados en la idea de que la tecnología podría ser utilizada para mejorar la vida humana. Este movimiento fue influenciado por la cibernética, la biología molecular y la inteligencia artificial, campos que empezaron a avanzar rápidamente y prometían un futuro donde las limitaciones humanas pudieran ser superadas.

PRINCIPALES FIGURAS Y DESARROLLO HISTÓRICO DEL CONCEPTO

Varios pensadores y científicos han contribuido al desarrollo del transhumanismo a lo largo de su historia. Entre ellos, destacan:

1. **Julian Huxley**: Como se mencionó anteriormente, Huxley fue uno de los primeros en articular el concepto de transhumanismo. Su visión de la evolución como un proceso que podría ser dirigido por la humanidad sentó las bases para el desarrollo del movimiento.
2. **FM-2030**: Este futurista y filósofo, conocido por su libro "*Are You a Transhuman?",* popularizó el término en la década de 1970. Promovía la idea de que la humanidad debería abrazar la tecnología para mejorar sus capacidades y prolongar la vida.
3. **Max More**: En la década de 1990, More, un filósofo y uno de los principales defensores del transhumanismo, fundó el Extropy Institute, que se centró en la promoción de ideas transhuma-

nistas. Su ensayo "*Principles of Extropy*" definió principios clave que guiarían el pensamiento transhumanista.

4. **Ray Kurzweil**: Este inventor y futurista ha sido fundamental en la difusión del transhumanismo, especialmente a través de su obra "*The Singularity is Near*", donde propone que la fusión entre humanos y máquinas es inminente. Kurzweil sugiere que la inteligencia artificial se integrará con la inteligencia humana, lo que permitirá superar las limitaciones biológicas.
5. **Nick Bostrom**: Filósofo y director del Instituto para el Futuro de la Humanidad en la Universidad de Oxford, Bostrom ha explorado las implicaciones éticas del transhumanismo, planteando cuestiones sobre el futuro de la humanidad y los riesgos asociados con el desarrollo de tecnologías avanzadas.

A lo largo de las últimas décadas, el transhumanismo ha ganado popularidad y reconocimiento en círculos académicos y tecnológicos, dando lugar a conferencias, publicaciones y organizaciones dedicadas a la promoción de sus ideas. La creación de la "*World Transhumanist Association*" en 1998 (ahora conocida como "Humanity+) fue un hito significativo en la formalización del movimiento.

En la actualidad, el transhumanismo sigue evolucionando con los avances en biotecnología, inteligencia artificial y neurotecnología, llevando a debates sobre las implicaciones éticas y filosóficas de un futuro donde los humanos y las máquinas podrían fusionarse.

PRÓTESIS BIÓNICAS E IMPLANTES TECNOLÓGICOS: AVANCES EN PRÓTESIS ROBÓTICAS Y DISPOSITIVOS COMO MARCAPASOS E IMPLANTES COCLEARES

La tecnología de prótesis biónicas ha avanzado significativamente en las últimas décadas, proporcionando soluciones innovadoras para personas con discapacidades o lesiones. Estas prótesis no solo restauran funciones perdidas, sino que también pueden mejorar las capacidades del usuario más allá de las que poseían de manera natural. Un área de

notable desarrollo es la robótica, que ha permitido la creación de prótesis controladas por señales neuronales. Estas prótesis utilizan interfaces cerebro-computadora (BCI) que detectan las señales eléctricas generadas por el cerebro para controlar el movimiento de la extremidad artificial. Esto permite un nivel de control mucho más natural y intuitivo para el usuario.

Un ejemplo destacado es la prótesis de brazo desarrollada por la Universidad de John Hopkins, que permite a los usuarios realizar tareas complejas, como agarrar objetos o tocar instrumentos musicales. La investigación en este campo ha demostrado que estas prótesis pueden mejorar la calidad de vida y la independencia de las personas que las utilizan (O'Doherty y otros., 2011).

Además de las prótesis robóticas, los dispositivos médicos como los marcapasos e implantes cocleares son ejemplos de cómo la tecnología ha integrado soluciones que mejoran la salud y la funcionalidad humana. Los marcapasos son dispositivos implantables que regulan el ritmo cardíaco, permitiendo a los pacientes llevar una vida más activa y saludable. La tecnología ha evolucionado hasta permitir marcapasos que se comunican de forma inalámbrica con dispositivos externos para monitoreo y ajustes en tiempo real (Yasuda y otros., 2019).

Por otro lado, los implantes cocleares son dispositivos que proporcionan una sensación de sonido a personas con pérdida auditiva severa. Estos implantes convierten las señales sonoras en impulsos eléctricos que estimulan el nervio auditivo, permitiendo que el cerebro interprete los sonidos. Los avances en esta tecnología han permitido que los usuarios experimenten una mejora significativa en su capacidad para comunicarse y participar en actividades sociales (Blamey y otros., 2013).

INTEGRACIÓN DE TECNOLOGÍA PARA MEJORAR LAS CAPACIDADES HUMANAS

La integración de tecnología en el cuerpo humano no se limita a la restauración de funciones perdidas; también busca ampliar y mejorar las capacidades humanas. Esta idea se manifiesta a través de disposi-

tivos que no solo mejoran la salud física, sino que también optimizan habilidades cognitivas y motoras. La biotecnología y la neurotecnología están en el centro de este desarrollo, ofreciendo oportunidades para un "mejoramiento" humano a través de la tecnología.

Por ejemplo, las interfaces cerebro-computadora (BCI) no solo permiten controlar prótesis, sino que también están siendo investigadas para aplicaciones en la mejora cognitiva. A través de la BCI, es posible entrenar a las personas para que realicen tareas complejas utilizando solo su actividad cerebral, lo que podría resultar en un aumento de la eficiencia en tareas cognitivas (Lebedev y Nicolelis, 2006).

Además, la neuroestimulación es una técnica que se utiliza para mejorar diversas funciones, como el control del dolor y la rehabilitación neuromuscular. Dispositivos como los estimuladores cerebrales profundos han mostrado resultados prometedores en el tratamiento de trastornos neurológicos como el Parkinson y la depresión, permitiendo a los pacientes recuperar o mejorar sus capacidades motoras y funcionales (Lozano y otros., 2012).

La combinación de tecnología y biología está llevando a una nueva era en la medicina y la ingeniería, donde la integración de dispositivos tecnológicos puede cambiar radicalmente la forma en que entendemos la salud y la funcionalidad humana. Sin embargo, estos avances también traen consigo dilemas éticos sobre la naturaleza de la humanidad y las implicaciones de modificar nuestras capacidades de manera artificial.

La evolución de las prótesis biónicas y los implantes tecnológicos representa un hito en la historia de la medicina y la ingeniería. Los avances en estas áreas han permitido no solo la restauración de funciones perdidas, sino también la mejora y optimización de las capacidades humanas. A medida que la tecnología continúa avanzando, es crucial considerar las implicaciones éticas y sociales de estas innovaciones, asegurando que su desarrollo y aplicación beneficien a toda la sociedad.

INTERFAZ CEREBRO-COMPUTADORA (BCI): TECNOLOGÍAS QUE CONECTAN EL CEREBRO CON DISPOSITIVOS EXTERNOS

Las interfaces cerebro-computadora (BCI, por sus siglas en inglés) son tecnologías que permiten la comunicación directa entre el cerebro humano y dispositivos externos. Estas interfaces tienen como objetivo facilitar la interacción con sistemas informáticos sin necesidad de movimientos físicos, aprovechando la actividad eléctrica generada por las neuronas. A través de la captura y el procesamiento de estas señales cerebrales, las BCI pueden traducir los pensamientos y las intenciones del usuario en comandos para controlar dispositivos como prótesis, computadoras y otros sistemas tecnológicos.

Los sistemas BCI pueden clasificarse en dos categorías principales: invasivos y no invasivos. Los sistemas invasivos requieren la implantación de electrodos directamente en el cerebro, lo que permite una mayor precisión en la captura de las señales neuronales. Estos sistemas se han utilizado en estudios clínicos y experimentos de investigación, donde han demostrado su eficacia en el control de prótesis robóticas y la restauración de funciones motoras en pacientes con parálisis (Lebedev y Nicolelis, 2006). Ejemplos de BCI invasivas incluyen los dispositivos desarrollados por Neuralink y otros grupos de investigación que utilizan electrodos implantables.

Por otro lado, las BCI no invasivas utilizan métodos externos, como electroencefalografía (EEG), para detectar la actividad eléctrica del cerebro a través del cuero cabelludo. Aunque estas tecnologías son menos precisas que las invasivas, su facilidad de uso y su menor riesgo de complicaciones las han hecho populares en aplicaciones comerciales y de investigación. Dispositivos como cascos EEG permiten la interacción con computadoras y juegos mediante el control de la actividad mental, lo que abre nuevas posibilidades para usuarios con discapacidades (Nijholt y otros., 2009).

POSIBILIDADES DE MEJORAR LA COGNICIÓN Y EL CONTROL SOBRE SISTEMAS TECNOLÓGICOS

Las BCI no solo tienen el potencial de restaurar funciones perdidas, sino que también ofrecen oportunidades para mejorar la cognición y el control sobre sistemas tecnológicos. En el ámbito de la mejora cognitiva, las BCI podrían facilitar el acceso a información en tiempo real, permitiendo a los usuarios procesar y analizar datos de manera más eficiente. Investigaciones en este campo han mostrado que las interfaces cerebro-computadora pueden usarse para ayudar a las personas a realizar tareas cognitivas complejas, como el aprendizaje de nuevos idiomas o la resolución de problemas matemáticos, mediante la facilitación de la comunicación directa con sistemas de inteligencia artificial (Chaudhary y otros., 2019).

Un área emergente es la posibilidad de utilizar BCI para la "extensión" de la cognición humana. Esto implica que, mediante el uso de estas interfaces, las personas podrían mejorar sus capacidades de memoria, atención y toma de decisiones. Investigaciones han demostrado que la combinación de BCI con técnicas de estimulación cerebral, como la estimulación transcraneal de corriente directa (tDCS), puede aumentar temporalmente las habilidades cognitivas, permitiendo a los individuos realizar tareas de manera más efectiva (Bikson y otros., 2016).

En términos de control sobre sistemas tecnológicos, las BCI están siendo exploradas para aplicaciones en diversos campos, desde la medicina hasta el entretenimiento. Por ejemplo, en el ámbito de la rehabilitación, las BCI permiten a pacientes con discapacidades motoras controlar dispositivos asistenciales, como sillas de ruedas o prótesis, a través de sus pensamientos. Esto no solo mejora la independencia y la calidad de vida de los usuarios, sino que también fomenta una mayor interacción con su entorno (Merrill y otros., 2019).

Además, en el ámbito del entretenimiento, las BCI están revolucionando la forma en que los usuarios interactúan con los videojuegos y otras plataformas. Juegos que utilizan BCI permiten a los jugadores controlar personajes o elementos del juego mediante la actividad cerebral, ofreciendo una experiencia inmersiva y única (Bai y otros., 2020).

Las interfaces cerebro-computadora (BCI) representan un avance significativo en la interacción entre el cerebro humano y la tecnología. A través de la conexión directa entre el sistema nervioso y dispositivos externos, las BCI no solo están cambiando el enfoque de la rehabilitación y la asistencia, sino que también abren nuevas posibilidades para mejorar la cognición y el control sobre sistemas tecnológicos. Sin embargo, el desarrollo de estas tecnologías también plantea importantes consideraciones éticas y de seguridad que deben ser abordadas a medida que avanzamos hacia un futuro donde la fusión entre la biología humana y la tecnología se convierta en una realidad.

INTELIGENCIA ARTIFICIAL COMO COMPLEMENTO COGNITIVO: USO DE IA PARA AUMENTAR LAS CAPACIDADES MENTALES

La inteligencia artificial (IA) ha emergido como una herramienta poderosa que no solo automatiza tareas, sino que también complementa y amplifica las capacidades cognitivas humanas. Este enfoque busca optimizar procesos mentales, mejorar la toma de decisiones y facilitar el aprendizaje a través de sistemas inteligentes que interactúan con los seres humanos. Las aplicaciones de la IA en este contexto abarcan desde asistentes personales hasta plataformas de aprendizaje adaptativo que personalizan el contenido según las necesidades individuales.

Uno de los principales usos de la IA como complemento cognitivo es en el ámbito del aprendizaje. Plataformas como Duolingo y Khan Academy utilizan algoritmos de IA para adaptar los materiales de estudio a las necesidades y ritmos de aprendizaje de cada usuario. Esto permite a los estudiantes progresar de manera más efectiva, ya que los sistemas pueden identificar áreas donde el usuario necesita más práctica y ajustar el contenido en consecuencia (Hwang y Chang, 2018). Además, la IA puede analizar los patrones de aprendizaje de los usuarios y ofrecer retroalimentación personalizada, lo que fomenta un aprendizaje más profundo y significativo.

En el ámbito empresarial, la IA se utiliza para mejorar la toma de decisiones. Herramientas de análisis predictivo, como IBM Watson y Salesforce Einstein, permiten a las organizaciones procesar grandes

volúmenes de datos y generar informes que ayudan a los líderes a tomar decisiones informadas basadas en análisis en tiempo real (Davenport, 2018). Estos sistemas no solo facilitan la identificación de tendencias y patrones, sino que también sugieren acciones que pueden maximizar los resultados, actuando como un "asesor" que complementa la inteligencia humana.

EJEMPLOS ACTUALES Y POSIBLES DESARROLLOS FUTUROS

En la actualidad, ya existen numerosos ejemplos que demuestran cómo la IA puede actuar como un complemento cognitivo en diversas áreas:

1. **Asistentes Virtuales**: Aplicaciones como *Siri, Google Assistant y Alexa* permiten a los usuarios realizar tareas cotidianas mediante comandos de voz, liberando a los individuos de la carga cognitiva asociada con la gestión de múltiples aplicaciones y recordatorios. Estos asistentes utilizan el procesamiento del lenguaje natural (NLP) para entender y responder a las solicitudes de manera efectiva, ayudando a los usuarios a organizar su tiempo y tareas de manera más eficiente (Khan et al., 2020).
2. **Análisis de Datos**: En el ámbito de la salud, sistemas como los utilizados en diagnóstico médico están ayudando a los profesionales a analizar imágenes y datos clínicos. Por ejemplo, algoritmos de aprendizaje profundo han demostrado ser eficaces en la detección temprana de enfermedades como el cáncer a través del análisis de imágenes médicas, lo que permite un diagnóstico más rápido y preciso (Esteva et al., 2019).
3. **Herramientas de Escritura Asistida**: Plataformas como Grammarly y Hemingway utilizan IA para mejorar las habilidades de escritura de los usuarios, ofreciendo sugerencias sobre gramática, estilo y claridad. Estas herramientas permiten a las personas comunicarse de manera más efectiva, lo que es especialmente valioso en entornos académicos y profesionales (Katz y otros., 2020).

En cuanto a desarrollos futuros, la IA promete evolucionar en formas aún más integradas con las capacidades cognitivas humanas. Se espera que las futuras generaciones de IA sean capaces de interactuar de manera más natural y efectiva con los humanos, aprendiendo y adaptándose a sus preferencias y estilos de trabajo. Una tendencia emergente es la creación de sistemas de IA que puedan participar en "*colaboraciones cognitivas*", donde humanos y máquinas trabajan juntos para resolver problemas complejos en tiempo real. Esto podría incluir desde proyectos de investigación científica hasta la creación de obras de arte, donde la creatividad humana se combina con la capacidad analítica de la IA (Müller, 2020).

Además, con el avance de la neurotecnología, es posible que en el futuro se desarrollen interfaces más directas que permitan a los humanos interactuar con sistemas de IA a través de la actividad cerebral, lo que llevaría a una sinergia aún más profunda entre la inteligencia humana y artificial.

La inteligencia artificial se está consolidando como un complemento cognitivo esencial en diversas áreas, potenciando las capacidades mentales y facilitando procesos de aprendizaje y toma de decisiones. A medida que la tecnología avanza, las posibilidades para integrar aún más la IA en la vida cotidiana y en las actividades cognitivas se expanden, prometiendo un futuro donde humanos y máquinas colaboren de manera más efectiva. Sin embargo, es fundamental abordar las implicaciones éticas y sociales de estas interacciones para garantizar que la evolución de la IA beneficie a la sociedad en su conjunto.

SIMBIOSIS ENTRE HUMANOS Y MÁQUINAS: DEFINICIÓN Y EJEMPLOS ACTUALES DE SIMBIOSIS TECNOLÓGICO-HUMANA

La simbiosis entre humanos y máquinas se refiere a la relación mutuamente beneficiosa que se establece entre los seres humanos y las tecnologías avanzadas, donde ambos se complementan y mejoran las capacidades del otro. Este concepto se basa en la idea de que, a medida que la tecnología avanza, los humanos no solo utilizan dispositivos como herramientas, sino que integran estas tecnologías en su vida co-

tidiana y, en algunos casos, en su biología, creando una interdependencia que redefine lo que significa ser humano (Bostrom, 2014).

En la actualidad, existen múltiples ejemplos que ilustran esta simbiosis. Uno de los ejemplos más destacados es el uso de **prótesis biónicas** que son controladas por la actividad cerebral. Dispositivos como el brazo biónico de **Luke** permiten a los amputados realizar movimientos complejos, como agarrar objetos, mediante señales neuronales que son traducidas por la máquina (O'Neill y otros., 2018). Este tipo de tecnología no solo mejora la calidad de vida de los usuarios, sino que también les permite experimentar una mayor conexión con su entorno.

Otro ejemplo es el uso de **inteligencia artificial en el diagnóstico médico**. Sistemas como IBM Watson han demostrado ser capaces de analizar grandes volúmenes de datos clínicos y ofrecer recomendaciones diagnósticas y de tratamiento. Esta simbiosis permite a los médicos tomar decisiones más informadas, mejorando así la atención al paciente (Hwang y otros., 2019). La IA actúa como un "asistente" que complementa la experiencia y el conocimiento humano, elevando la capacidad de los profesionales de la salud.

Además, el uso de **interfaces cerebro-computadora (BCI)**, como las desarrolladas por *Neuralink,* permite a los usuarios controlar dispositivos mediante sus pensamientos. Estas tecnologías están comenzando a formar una simbiosis en la que los humanos pueden interactuar con las máquinas de manera más directa y natural, lo que facilita tareas que antes eran imposibles (Lebedev y Nicolelis, 2006).

FUTURAS FORMAS DE INTERDEPENDENCIA ENTRE HUMANOS Y SISTEMAS DE IA

Mirando hacia el futuro, las formas de interdependencia entre humanos y sistemas de IA se están expandiendo rápidamente, y se prevé que evolucionen en direcciones fascinantes. Una de las áreas más prometedoras es el desarrollo de **agentes de inteligencia artificial** que pueden actuar como "extensiones" de la cognición humana. Esto podría incluir sistemas que no solo analicen datos, sino que también par-

ticipen en la toma de decisiones complejas junto a sus usuarios, creando una colaboración que puede llevar a mejores resultados en campos como la medicina, la ingeniería y la educación (Müller, 2020).

La creación de **entornos de trabajo asistidos por IA** es otro ámbito donde la simbiosis podría florecer. Imagina un futuro donde las máquinas no solo realicen tareas repetitivas, sino que también actúen como consultores que proporcionen información en tiempo real, sugerencias y alternativas basadas en datos. Este tipo de interacción podría aumentar significativamente la productividad y la creatividad en diversas industrias (Davenport y Ronanki, 2018).

La integración de tecnologías de **realidad aumentada (AR)** y **realidad virtual (VR)** en la formación y la educación también representa una nueva forma de simbiosis. Estas tecnologías pueden permitir experiencias de aprendizaje inmersivas que se adaptan a las necesidades del usuario, combinando el aprendizaje humano con el poder de las máquinas para ofrecer simulaciones realistas y entrenamiento práctico en entornos controlados (Mikropoulos y Natsis, 2011).

En el ámbito personal, la futura simbiosis podría incluir el uso de **dispositivos portátiles inteligentes** que no solo monitorean la salud de una persona, sino que también ofrecen recomendaciones personalizadas para mejorar su bienestar físico y mental. Imagina un sistema que pueda predecir enfermedades basándose en datos biométricos en tiempo real y sugerir cambios en el estilo de vida antes de que aparezcan problemas graves (Topol, 2019).

La simbiosis entre humanos y máquinas está transformando nuestra relación con la tecnología, creando interdependencias que enriquecen tanto a los seres humanos como a los sistemas artificiales. A medida que avanzamos hacia el futuro, la integración de IA y tecnologías avanzadas promete abrir nuevas fronteras en la colaboración, la salud, el aprendizaje y la creatividad. Sin embargo, es fundamental abordar las cuestiones éticas y sociales que surgen de estas interacciones para garantizar que el desarrollo de la tecnología beneficie a toda la sociedad.

DESAFÍOS ÉTICOS Y FILOSÓFICOS: PREGUNTAS SOBRE LA NATURALEZA DE LA IDENTIDAD HUMANA

La rápida evolución de la tecnología, especialmente en el contexto del transhumanismo y la simbiosis entre humanos y máquinas, plantea profundas preguntas sobre la identidad humana. A medida que las tecnologías avanzadas, como los implantes neurales y la inteligencia artificial, se integran cada vez más en nuestras vidas, surgen interrogantes sobre lo que realmente significa ser humano.

Uno de los principales dilemas es si la integración de tecnologías en el cuerpo y la mente cambia fundamentalmente nuestra naturaleza. La noción de identidad humana ha estado históricamente ligada a la biología y la experiencia humana. Sin embargo, el transhumanismo desafía esta concepción al proponer que las capacidades humanas pueden ser mejoradas o incluso redefinidas a través de la tecnología (Bostrom, 2014). Esto plantea la pregunta: si un ser humano mejora sus habilidades cognitivas mediante implantes tecnológicos, ¿sigue siendo el mismo individuo o se ha convertido en una forma de vida diferente? Esta pregunta se convierte en una cuestión de identidad, ya que define cómo nos percibimos a nosotros mismos y cómo nos relacionamos con los demás.

Además, la posibilidad de crear "*posthumanos*"—seres que han trascendido las limitaciones biológicas a través de la tecnología—plantea la preocupación de que la identidad humana podría fragmentarse en diferentes categorías, como humanos "naturales" y "mejorados". Esto podría generar tensiones sociales y culturales, ya que las personas podrían ser juzgadas no solo por sus logros personales, sino también por el tipo y nivel de mejoras tecnológicas que han adoptado (Bostrom, 2014; Sandberg y Bostrom, 2008).

Por otra parte, el filósofo Nick Bostrom ha argumentado que la búsqueda de la mejora humana podría dar lugar a una forma de elitismo donde aquellos que puedan permitirse las mejoras tecnológicas se conviertan en una "clase superior" en comparación con aquellos que no tienen acceso a estas. Esto plantea cuestiones sobre la equidad en la identidad humana y cómo las diferencias en la tecnología pueden

afectar nuestra comprensión de la humanidad en su conjunto (Bostrom, 2005).

PROBLEMAS DE EQUIDAD EN EL ACCESO A MEJORAS TECNOLÓGICAS

Otro desafío ético significativo asociado con el transhumanismo y las mejoras tecnológicas es la **equidad en el acceso** a estas innovaciones. A medida que las tecnologías avanzadas, como los dispositivos de mejora cognitiva, las prótesis biónicas y los tratamientos de edición genética, se vuelven más accesibles, surge la pregunta de quién tiene la capacidad y el privilegio de acceder a estas mejoras.

El acceso desigual a la tecnología podría exacerbar las brechas sociales y económicas existentes, creando una sociedad dividida en "*mejorados*" y "no mejorados". Esto plantea serias preocupaciones éticas sobre la justicia y la equidad. Como señala el filósofo y futurista Eliezer Yudkowsky, "el riesgo de una desigualdad radical en el acceso a las mejoras tecnológicas podría ser uno de los mayores desafíos del futuro" (Yudkowsky, 2017). Si las mejoras tecnológicas son limitadas a quienes pueden pagar por ellas, se podría crear una élite tecnológicamente superior, lo que podría llevar a nuevas formas de opresión y exclusión social.

Además, la cuestión del acceso también implica la disponibilidad de estas tecnologías a nivel global. Las naciones en desarrollo a menudo carecen de infraestructura y recursos para implementar avances tecnológicos, lo que podría dejar a sus poblaciones en desventaja frente a las naciones más ricas. Esto plantea preguntas sobre la responsabilidad de las naciones y las empresas tecnológicas en garantizar un acceso equitativo a las mejoras, así como el papel de los gobiernos en la regulación de estas tecnologías para proteger a los más vulnerables (Hughes, 2018).

Los desafíos éticos y filosóficos que surgen de la simbiosis entre humanos y máquinas son complejos y multifacéticos. Las preguntas sobre la naturaleza de la identidad humana y los problemas de equidad en el acceso a las mejoras tecnológicas exigen una reflexión cuidadosa

y un debate profundo. A medida que avanzamos hacia un futuro donde la tecnología desempeña un papel cada vez más integral en nuestras vidas, es fundamental considerar cómo estas transformaciones afectarán nuestra comprensión de lo que significa ser humano y cómo podemos trabajar hacia una sociedad más justa e inclusiva.

FUTURO DEL TRABAJO Y LA EDUCACIÓN: CAMBIOS EN EL MERCADO LABORAL DEBIDO A LAS MEJORAS COGNITIVAS Y FÍSICAS

El avance de las tecnologías, especialmente aquellas relacionadas con el transhumanismo, está transformando significativamente el mercado laboral. Las mejoras cognitivas y físicas, facilitadas por dispositivos biónicos, inteligencia artificial (IA) y biotecnología, están redefiniendo las habilidades requeridas para diversas ocupaciones, así como la naturaleza misma del trabajo.

Las **mejoras cognitivas** permiten a los trabajadores procesar información más rápidamente, resolver problemas complejos y tomar decisiones informadas con mayor eficacia. Por ejemplo, el uso de sistemas de inteligencia artificial en el ámbito empresarial está optimizando procesos y aumentando la productividad (*Brynjolfsson y McAfee,* 2014). Las empresas están comenzando a integrar IA que complementa las habilidades humanas, lo que permite a los empleados enfocarse en tareas más creativas y estratégicas. Esto lleva a una reevaluación de qué trabajos son necesarios y cuáles pueden ser automatizados, lo que podría resultar en la reducción de ciertas ocupaciones mientras surgen nuevas.

Por otro lado, las **mejoras físicas** mediante prótesis avanzadas y exoesqueletos también están cambiando la dinámica laboral. Estas tecnologías permiten que personas con discapacidades físicas realicen tareas que anteriormente estaban fuera de su alcance, ampliando la diversidad en la fuerza laboral (Rosen et al., 2019). A medida que estas mejoras se vuelvan más accesibles, es probable que veamos un aumento en la inclusión de individuos con diversas capacidades en roles laborales que requieren fuerza física o habilidades motoras precisas.

Además, el aumento en la productividad y la capacidad de realizar tareas complejas de manera más eficiente podría llevar a un cambio en la estructura del empleo. Los trabajos de baja cualificación podrían ser reemplazados por tecnologías automatizadas, mientras que se espera que la demanda de habilidades técnicas y cognitivas avanzadas crezca (Frey y Osborne, 2017). Este desplazamiento en el mercado laboral resalta la necesidad de una adaptación constante de la fuerza laboral para mantenerse relevante en un entorno cambiante.

NUEVAS EXPECTATIVAS EDUCATIVAS EN UNA SOCIEDAD TRANSHUMANISTA

A medida que la sociedad se adentra en un futuro transhumanista, las expectativas educativas también están evolucionando. La educación no solo debe preparar a los estudiantes para un mercado laboral en constante cambio, sino que también debe fomentar la capacidad de adaptación a las mejoras tecnológicas y cognitivas.

En este contexto, se espera que la educación incorpore el uso de tecnologías avanzadas como parte integral del currículo. Esto incluye la enseñanza sobre la **inteligencia artificial**, la **neurotecnología** y las **biotecnologías**, así como el desarrollo de habilidades prácticas para trabajar en entornos altamente tecnológicos (Luckin y otros., 2016). La educación también debe promover la **alfabetización digital** y la **capacidad de aprendizaje continuo**, ya que los trabajadores del futuro necesitarán adaptarse rápidamente a nuevas herramientas y métodos.

Además, se prevé que la educación en una sociedad transhumanista enfatice no solo las habilidades técnicas, sino también las competencias socioemocionales. A medida que la tecnología asuma más tareas, las habilidades como la creatividad, la empatía y la colaboración se volverán aún más valiosas (Dumont y otros., 2016). La capacidad de trabajar en equipo y comunicarse de manera efectiva con otros será fundamental en un entorno laboral donde la automatización y la IA están en auge.

Las instituciones educativas también tendrán que reconsiderar cómo se mide el éxito y el aprendizaje. Los métodos de evaluación

tradicionales podrían ser insuficientes para capturar las habilidades complejas y la adaptabilidad de los estudiantes. En su lugar, es posible que surjan nuevas formas de evaluación que consideren la capacidad de los estudiantes para aplicar sus conocimientos en contextos del mundo real y para integrarse con tecnologías emergentes (Fadel y otros., 2016).

El futuro del trabajo y la educación en una sociedad transhumanista presenta tanto oportunidades como desafíos. Las mejoras cognitivas y físicas están redefiniendo las habilidades necesarias en el mercado laboral, mientras que las expectativas educativas están cambiando para preparar a los estudiantes para un entorno laboral altamente tecnológico. Para garantizar que la fuerza laboral esté equipada para enfrentar estos cambios, es esencial que las instituciones educativas adapten sus currículos y enfoques de evaluación, fomentando no solo habilidades técnicas, sino también competencias socioemocionales que serán cruciales en el futuro.

LA SINGULARIDAD TECNOLÓGICA: EL CONCEPTO DE UNA IA QUE SUPERE A LA INTELIGENCIA HUMANA

La **singularidad tecnológica** es un concepto que se refiere al momento hipotético en que la inteligencia artificial (IA) supere a la inteligencia humana, desencadenando un crecimiento exponencial en el desarrollo tecnológico y cambios radicales en la sociedad. Este término fue popularizado por el futurista Vernor Vinge en la década de 1990 y ha sido ampliamente discutido en la literatura científica y filosófica (Vinge, 1993).

La idea central de la singularidad es que, una vez que las máquinas sean capaces de mejorar su propia inteligencia de manera autónoma, se producirá un ciclo de retroalimentación positiva. Esto significa que las máquinas podrían diseñar nuevas y más avanzadas generaciones de IA, creando un avance tecnológico a un ritmo que supera la capacidad humana para comprender o controlar el desarrollo (Kurzweil, 2005). Esta situación se caracteriza por una aceleración del progreso en campos como la biotecnología, la robótica y la computación cuánti-

ca, lo que podría llevar a descubrimientos que transformen la vida tal como la conocemos.

IMPLICACIONES DE UN FUTURO DONDE LAS MÁQUINAS SEAN MÁS INTELIGENTES QUE LOS HUMANOS

El surgimiento de una inteligencia artificial superior plantea diversas implicaciones éticas, sociales y económicas que deben ser consideradas. A continuación se analizan algunas de las más significativas:

1. **Desigualdad y Acceso a la Tecnología**: Si solo un grupo selecto tiene acceso a tecnologías avanzadas, esto podría resultar en una sociedad dividida. Las personas o entidades que controlen las máquinas más inteligentes podrían ejercer un poder desproporcionado, exacerbando las desigualdades económicas y sociales existentes (Bostrom, 2014).
2. **Desplazamiento Laboral**: La automatización impulsada por IA podría llevar al desplazamiento de una gran parte de la fuerza laboral. A medida que las máquinas asuman tareas que requieren inteligencia, creatividad y habilidades complejas, muchos trabajos podrían volverse obsoletos, lo que generaría desafíos significativos en términos de empleo y seguridad económica (Frey y Osborne, 2017). Esto resalta la necesidad de una adaptación y capacitación continua de la fuerza laboral.
3. **Cuestiones Éticas y de Control**: La creación de máquinas que superen la inteligencia humana plantea serias preguntas sobre la ética de su desarrollo y uso. ¿Cómo aseguramos que estas máquinas actúen en el mejor interés de la humanidad? Existe el temor de que una IA superinteligente pueda actuar de maneras impredecibles y potencialmente dañinas si no se la controla adecuadamente (Bostrom, 2014). Esto lleva a la necesidad de establecer marcos regulatorios y de gobernanza que aseguren que el desarrollo de la IA se alinee con los valores humanos.
4. **Transformación de la Identidad Humana**: La singularidad tecnológica también podría dar lugar a una redefinición de lo que

significa ser humano. Si las máquinas no solo superan a los humanos en capacidad cognitiva, sino que también comienzan a incorporar características humanas, como la empatía o la creatividad, podría haber implicaciones profundas para nuestra identidad y relaciones interpersonales (Kurzweil, 2005). Las cuestiones sobre el libre albedrío, la responsabilidad y la ética en la toma de decisiones se volverían cada vez más complejas.

5. **Colaboración entre Humanos y Máquinas**: A pesar de los desafíos, también existen oportunidades para la colaboración entre humanos y máquinas. La inteligencia artificial puede actuar como una herramienta poderosa que amplía nuestras capacidades, permitiéndonos resolver problemas complejos de manera más efectiva. Esto podría llevar a una era de innovación sin precedentes, donde humanos y máquinas trabajen juntos para abordar los desafíos globales, como el cambio climático y la pobreza (Bostrom, 2014).

La singularidad tecnológica representa un punto de inflexión potencial en la historia humana, donde la inteligencia artificial podría superar la capacidad humana de comprensión y control. Este futuro está lleno de posibilidades, pero también plantea serios desafíos éticos y sociales que deben ser abordados. A medida que avanzamos hacia esta nueva era, es fundamental fomentar un diálogo continuo sobre las implicaciones de la IA y establecer políticas que garanticen que el desarrollo tecnológico beneficie a toda la humanidad.

CONCLUSIÓN

El transhumanismo se presenta como un enfoque revolucionario que no solo busca mejorar las capacidades humanas a través de la tecnología, sino que también plantea profundas cuestiones sobre la esencia misma de la humanidad. A medida que avanzamos hacia una era en la que la fusión entre humanos y máquinas se vuelve cada vez más plausible, debemos reflexionar sobre lo que significa ser humano en un contexto donde la tecnología puede redefinir nuestras capacidades, emociones e incluso nuestra identidad.

Las innovaciones en inteligencia artificial, neurotecnología y biotecnología no solo están cambiando nuestra forma de interactuar con el mundo, sino que también nos llevan a la frontera de una nueva forma de existencia. La posibilidad de integrar las capacidades cognitivas y físicas de las máquinas con las de los humanos promete beneficios significativos, pero también presenta desafíos éticos y filosóficos inminentes. Preguntas sobre la justicia en el acceso a estas tecnologías, la definición de la identidad humana y las implicaciones de crear una clase de "*posthumanos*" deben ser abordadas con seriedad y responsabilidad.

El avance hacia una simbiosis entre humanos y máquinas puede enriquecer nuestras vidas, permitiéndonos superar limitaciones biológicas y abordar problemas complejos. Sin embargo, es imperativo que este proceso se lleve a cabo de manera inclusiva y ética, para garantizar que todos se beneficien de las innovaciones tecnológicas y que la evolución de nuestra especie se dirija hacia un futuro más equitativo y justo. En este contexto, la colaboración entre humanos y máquinas podría no solo ampliar nuestras capacidades, sino también fortalecer nuestra identidad, mientras nos enfrentamos a los desafíos y oportunidades que el transhumanismo trae consigo.

REFERENCIAS BIBLIOGRÁFICAS

Bai, O., y otros. (2020). Brain–computer interface game for enhancing cognitive function in healthy older adults: A pilot study. *Neurorehabilitation and Neural Repair*, 34(10), 932-943. doi:10.1177/1545968320955395

Bikson, M., y otros. (2016). Transcranial direct current stimulation: A potential therapeutic tool for cognitive enhancement in healthy adults. *Frontiers in Human Neuroscience*, 10, 33. doi:10.3389/fnhum.2016.00033

Blamey, P. J., Arndt, P. S., y Melbourne, S. (2013). Cochlear implants. *Medical Journal of Australia*, 199(1), 34-36. doi:10.5694/mja12.00510

Bostrom, N. (2005). *A history of transhumanist thought*. *Journal of Evolution and Technology*, 14(1), 1-25.

Bostrom, N. (2005). In Defense of Posthuman Dignity. Bioethics, 19(3), 202-214. doi:10.1111/j.1467-8519.2005.00452.x

Brynjolfsson, E., y McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.

Chaudhary, U., y otros. (2019). Brain–computer interface (BCI) for communication and control. *Journal of Neural Engineering*, 16(5), 052001. doi:10.1088/1741-2552/ab26af

Davenport, T. H. (2018). *The AI Advantage: How to Put the Artificial Intelligence Revolution to Work*. MIT Press.

Davenport, T. H., y Ronanki, R. (2018). AI for the Real World. *Harvard Business Review*, 96(1), 108-116.

Dumont, H., Istance, D., y Benavides, F. (2016). The Nature of Learning: Using Research to Inspire Practice. OECD Publishing. doi:10.1787/9789264255369-en

Esteva, A., y otros. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24-29. doi:10.1038/s41591-018-0316-z

Fadel, C., Lemke, C., y Hwang, S. (2016). 21st Century Skills: Transforming Education for the New World of Work. In Handbook of Research on Teaching and Learning in K-20 Education (pp. 1-17). IGI Global. doi:10.4018/978-1-4666-8755-1.ch001

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

Herr, H. (2020). Exoskeletons and prosthetic devices: Aimed at assisting human movement. *Nature Biomedical Engineering*, 4(8), 756-767.

Hughes, J. (2018). The ethical implications of human enhancement. Ethics and Information Technology, 20(2), 95-106. doi:10.1007/s10676-018-9447-3

Huxley, J. (1957). *New Bottles for New Wine*. London: Chatto & Windus.

Hwang, G. J., y Chang, H. F. (2018). Seamless flipped learning: a mobile technology-enhanced flipped classroom with self-regulated learning. *Computers & Education*, 124, 112-121. doi:10.1016/j.compedu.2018.05.007

Katz, S. J., y otros (2020). Automated Writing Evaluation: The Use of Artificial Intelligence in Writing Instruction. *Journal of Educational Computing Research*, 58(3), 546-568. doi:10.1177/0735633118803623

Khan, A., y otros. (2020). Exploring the use of AI in virtual assistants: A survey of applications and future directions. *Journal of Intelligent Systems*, 29(1), 112-125. doi:10.1515/jisys-2019-0206

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking.

Lebedev, M. A., y Nicolelis, M. A. L. (2006). Brain–machine interfaces: past, present and future. *Trends in Neurosciences*, 29(9), 536-546. doi:10.1016/j.tins.2006.07.004

Lozano, A. M., y otros. (2012). Deep brain stimulation for Parkinson's disease: from the operating room to the patient's home. *Nature Reviews Neurology*, 8(10), 582-590. doi:10.1038/nrneurol.2012.184

Luckin, R., Holmes, W., Griffiths, M., y Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson Education.

Merrill, D. R., y otros. (2019). Brain-computer interfaces in rehabilitation: current status and future directions. *Journal of Neural Engineering*, 16(3), 031001. doi:10.1088/1741-2552/ab16fa

Mikropoulos, T. A., y Natsis, A. (2011). Educational Virtual Environments: A 10-year review of the research. *Computers & Education*, 56(3), 679-693. doi:10.1016/j.compedu.2010.10.025

More, M. (1990). *Are You a Transhuman?*. Retrieved from *Extropy Institute*.

Müller, V. C. (2020). Ethics of Artificial Intelligence and Robotics. In *Stanford Encyclopedia of Philosophy*. Retrieved from *https://plato.stanford.edu/entries/ethics-ai/*

Musk, E., y Neuralink. (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Engineering & Technology*, 43(4), 168-181.

Nijholt, A., y otros. (2009). Brain-computer interfaces for games: the next level of interactivity. *Entertainment Computing*, 1(1), 67-75. doi:10.1016/j.entcom.2009.04.001

O'Doherty, J. E., y otros. (2011). A direct brain interface provides an efficient way to control a prosthetic arm. *Science Translational Medicine*, 3(118), 118ra17. doi:10.1126/scitranslmed.3003515

O'Neill, T., y otros. (2018). Bionic limbs controlled by thought. *Nature Biomedical Engineering*, 2(5), 326-331. doi:10.1038/s41551-018-0251-5

Rosen, J., y otros. (2019). The future of work in America: People and places, today and tomorrow. McKinsey Global Institute.

Sandberg, A., y Bostrom, N. (2008). Cognitive enhancement: An ethical overview. Science and Engineering Ethics, 14(3), 309-330. doi:10.1007/s11948-008-9108-8

Topol, E. J. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.

Vinge, V. (1993). The coming technological singularity: How to survive in the post-human era. In Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace (pp. 11-22). NASA.

Yasuda, S., y otros. (2019). Wireless cardiac monitoring: the future of pacemaker technology. *Circulation Journal*, 83(7), 1444-1450. doi:10.1253/circj.CJ-19-0200

Yudkowsky, E. (2017). Rationality: From AI to Zombies. Machine Intelligence Research Institute.

CAPÍTULO XX

EL RAZONAMIENTO Y LA IA, UNA IDA Y UNA VUELTA: UN VIAJE INESPERADO DE LOS NUEVOS MODELOS DE RAZONAMIENTO DE LA IA

VÍCTOR DANIEL SANTOS HORTELANO[1]

RESUMEN

El razonamiento, un proceso mental exclusivo de los humanos, ha sido un pilar fundamental en la evolución de nuestra sociedad. Desde la antigüedad, filósofos como Aristóteles y Kant han reflexionado sobre los límites y posibilidades del razonamiento, y su aplicación en la ciencia y la tecnología. Con el tiempo, este proceso mental ha sido imitado por máquinas a través del desarrollo de la inteligencia artificial (IA). Los avances en IA, como el modelo de *Sakana IA y Chat GPT o1,* han demostrado capacidades avanzadas para realizar investigaciones científicas y generar soluciones complejas en un tiempo récord. Sin embargo, la velocidad de estas innovaciones plantea dilemas éticos importantes sobre el control humano y las consecuencias del uso de la IA, recordando las preocupaciones expresadas por pioneros como Alan Turing. Hoy en día, el uso de la IA no solo transforma el campo de la investigación, sino que también plantea preguntas sobre la automatización, la responsabilidad y los límites éticos de la tecnología en nuestra sociedad.

[1] Asistente de investigación, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

INTRODUCCIÓN

El razonamiento ha sido uno de los pilares fundamentales de la civilización humana, y su estudio ha sido el enfoque de filósofos y científicos desde tiempos inmemoriales. Aristóteles, en *El Órganon*, estableció el concepto del silogismo como la base del razonamiento lógico, mientras que Kant, en su *Crítica de la razón pura*, profundizó en cómo la mente humana utiliza categorías innatas para entender el mundo. Ambos pensadores marcaron un camino en la historia de la filosofía, abriendo la puerta a discusiones sobre la capacidad humana de pensar críticamente y formular juicios basados en la lógica.

Con el avance de la tecnología, especialmente desde mediados del siglo XX, las máquinas han comenzado a replicar estas habilidades cognitivas. El trabajo pionero de Alan Turing fue clave para sentar las bases de lo que más tarde sería conocido como inteligencia artificial (IA). En su famoso ensayo "*Computing Machinery and Intelligence*", Turing planteó la posibilidad de que las máquinas pudieran imitar el razonamiento humano, proponiendo el Test de Turing como una forma de medir la "inteligencia" de una máquina. Estas ideas, revolucionarias en su tiempo, continúan guiando los desarrollos en el campo de la IA moderna.

Hoy en día, modelos avanzados *como Sakana IA y Chat GPT o1* no solo imitan el razonamiento humano, sino que también son capaces de generar hipótesis y realizar investigaciones científicas. Estas tecnologías están transformando campos como la medicina, las matemáticas y la filosofía, y plantean importantes interrogantes éticas sobre la autonomía de las máquinas y su impacto en la sociedad. Si bien las máquinas pueden realizar tareas con una rapidez y precisión asombrosas, es crucial considerar las implicaciones de delegar aspectos esenciales del pensamiento y la toma de decisiones en la IA. La ética, un tema que ha acompañado al razonamiento desde sus inicios, sigue siendo relevante en esta nueva era de colaboración entre humanos y máquinas.

EL RAZONAMIENTO

El razonamiento es un proceso que hasta el día de hoy se le atribuye a la raza humana. Lo hacemos todos en mayor o menor medida y es uno

de los procesos en los cuales los principales avances de nuestra sociedad se han visto involucrados, ya sea pensar nuestro papel en este mundo, nuestro origen, nuestro modo de vida, las leyes que nos rigen, los inventos que hacemos y cómo los mejoramos.

Dentro de la historia, darle una explicación a nuestro pensar, nuestras acciones y el resultado de las mismas, ha sido una constante desde el principio de los tiempos. Sin embargo, fue en la antigua Grecia donde se originó una gran ola de pensamiento sobre el razonamiento, con reflexiones profundas y duraderas que aún hoy en día siguen siendo objeto de estudio y reflexión. Basta solo con pensar en Aristóteles, en su obra "*El Organon*", sentó las bases del razonamiento lógico formal, argumentando que el silogismo es la forma fundamental del razonamiento válido. Según él, a través de premisas verdaderas y una estructura lógica correcta, podemos llegar a conclusiones verdaderas, lo que ha sido fundamental para el desarrollo del pensamiento científico y filosófico.

Por otro lado, Kant, en su "*Crítica de la razón pura*", exploró los límites y posibilidades del razonamiento humano. Propuso que nuestra comprensión del mundo está moldeada por categorías a priori del entendimiento, como el espacio y el tiempo, que son inherentes a nuestra forma de percibir y razonar. Esto sugiere que nuestro razonamiento, aunque poderoso, está limitado por la estructura de nuestra mente.

El razonamiento humano es necesario para no perder nuestra humanidad y recientemente lo queremos integrar a los inventos más avanzados que tenemos. Sin embargo, a la luz de las ideas de Aristóteles y Kant, debemos ser conscientes tanto del poder como de las limitaciones de nuestro razonamiento al aplicarlo a nuevas tecnologías y al intentar comprender el mundo que nos rodea.

O una opinión bastante importante de pensar o razonar e la de Turing, que dijo También podría decirse que la identificación de las máquinas con los computadores digitales, tal como nuestro criterio de "*pensar*", sólo será insatisfactorio si (contrario a lo que yo creo) los computadores digitales en definitiva son incapaces de hacer una buena demostración en el juego. (Turing, 1950)

Por eso mismo hay que hablar de los tres tipos de razonamiento que hay, el inductivo, deductivo y abductivo, los cuales son tipos de pensamiento que tienen mucho peso en los nuevos modelos de Inteligencia artificial.

El razonamiento deductivo es un proceso lógico que parte de premisas generales para llegar a conclusiones específicas y necesariamente verdaderas si las premisas son verdaderas. Es un método que va de lo general a lo particular y es fundamental en disciplinas como las matemáticas y la filosofía.

El razonamiento inductivo es un proceso lógico que parte de observaciones o casos específicos para llegar a conclusiones o generalizaciones generales. A diferencia del razonamiento deductivo, que va de lo general a lo particular, el razonamiento inductivo va de lo particular a lo general.

Figura 1. Personajes que usan el razonamiento deductivo



Fuente: Fox Broadcasting Company

El razonamiento abductivo es un proceso lógico mediante el cual se busca encontrar la explicación más probable o plausible para un conjunto de observaciones o hechos. A diferencia del razonamiento deductivo y el inductivo, la abducción se centra en formular hipótesis que puedan explicar de manera coherente y razonable las evidencias

disponibles, aunque no garantice la verdad absoluta de la conclusión. Tal como lo hacía Sherlok Holmes o Dr. House, este razonamiento es el más usado en la investigación científica, ya que implica hipótesis fundamentadas para explicar problemas o situaciones para encontrar una explicación.

El razonamiento es un proceso de la mente humana que no estaba presente con una explicación al inicio de nuestra especie. Sin embargo, al querer encontrar explicación a su accionar, empezó a pensar sobre pensar y a encontrar la lógica de sus acciones, lo cual desencadenó la evolución de sus ideas, que fueron permeando su realidad y cómo la transformaba.

Hoy en día, esta nueva raza de creaciones, simula redes neuronales artificiales para aprender, comparar y almacenar información, pero no tenían modelos de razonamiento que pudieran dar una reflexión de sus respuestas. Ahora, con las actualizaciones que tienen, pueden empezar a desempeñar un poco este concepto que era exclusivo de nuestra especie, y es por eso que mis ideas las iré poniendo junto a las que me dan las IA. Esta comparación, al inicio y sin esperar alterar el interés en lo que sigue del artículo, es abrumadora: puede hacer más rápido lo que hago, y mi memoria hace un esfuerzo para encontrar palabras y ejemplos. Por ahora me gana en rapidez, pero ¿qué será en un año o dos?

¿QUÉ DICEN LAS IA SOBRE LO QUE ES EL RAZONAMIENTO PARA ELLAS?

Esta es la definicion de *Chat GPT* sobre lo que dice que es el razonamiento:

> El razonamiento es el proceso mental por el cual los seres humanos analizamos información, establecemos conexiones entre ideas y llegamos a conclusiones. A través del razonamiento, evaluamos evidencias, identificamos patrones y utilizamos la lógica para resolver problemas y comprender el mundo que nos rodea. Este proceso es fundamental para el aprendizaje, la toma de decisiones y el de-

> sarrollo del pensamiento crítico. El razonamiento nos permite no solo procesar información existente, sino también generar nuevas ideas, hipótesis y soluciones innovadoras. Es una habilidad esencial que influye en todos los aspectos de nuestra vida personal y social. (Generado por ChatGPT, 2024)

Esta es la definicion de *Meta AI*:

> El razonamiento es el proceso mental que nos permite llegar a conclusiones lógicas y fundamentadas a partir de información disponible. Implica la capacidad de analizar y evaluar datos, ideas y argumentos, establecer relaciones entre ellos y sacar inferencias válidas. En esencia, el razonamiento es la herramienta fundamental para la toma de decisiones, la resolución de problemas y el aprendizaje, permitiéndonos navegar de manera efectiva en un mundo complejo y cambiante. (Generado por *META AI*, 2024)

Esta es la definición de Claude:

> El razonamiento es un proceso cognitivo fundamental que permite a los seres humanos procesar información, establecer conexiones lógicas y llegar a conclusiones basadas en evidencias o premisas. Es una habilidad mental que nos capacita para analizar situaciones, resolver problemas y tomar decisiones informadas. El razonamiento implica la capacidad de identificar patrones, hacer inferencias, evaluar argumentos y aplicar el pensamiento crítico a diversas situaciones, ya sean cotidianas o complejas.

Este proceso mental puede manifestarse de diversas formas, como el razonamiento deductivo (partir de premisas generales para llegar a conclusiones específicas), inductivo (observar casos particulares para formular principios generales), o abductivo (generar hipótesis explicativas a partir de observaciones).

El razonamiento no solo es crucial en ámbitos académicos y científicos, sino que también juega un papel vital en nuestra vida diaria, permitiéndonos navegar por el mundo social, tomar decisiones éticas, y adaptar nuestro comportamiento a diferentes contextos. Es una capa-

cidad que evoluciona con la experiencia y el aprendizaje, y que puede ser refinada a través de la educación y la práctica continua. (Generado por Claude, 2024)

ORÍGENES DE LA INTELIGENCIA ARTIFICIAL EN LA SEGUNDA GUERRA MUNDIAL

El primer antecedente de la interacción con los principios de la inteligencia artificial se remonta a la Segunda Guerra Mundial, específicamente con las máquinas que se usaban para transmitir información crítica para la guerra. Era crucial que el enemigo no descubriera los planes a seguir, ya que el éxito de las misiones dependía de ello. Esto requirió codificar dicha información, es decir, utilizar la criptografía, que es el estudio y la práctica de técnicas matemáticas y computacionales para proteger la confidencialidad, integridad y autenticidad de la información. A través de algoritmos y protocolos, se transforman los datos en un formato ilegible para terceros no autorizados, asegurando así la privacidad, seguridad y fiabilidad de la comunicación y el almacenamiento de información.

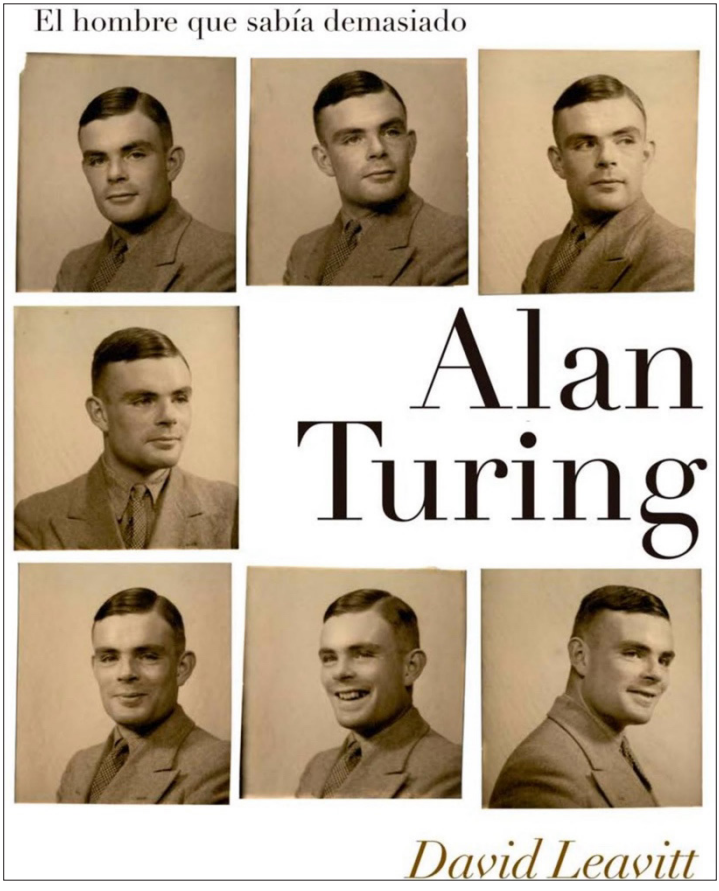
Estos complejos códigos se reflejaron en dos máquinas clave: por parte de los alemanes, la máquina Enigma y, por parte de los Aliados, la máquina Bombe, esta última creada por Alan Turing. La Bombe fue fundamental para descifrar los mensajes encriptados por Enigma, un avance que ayudó a los Aliados a ganar ventaja en el conflicto.

MÁQUINAS ENIGMA Y BOMBE: ICONOS DE LA CRIPTOGRAFÍA

Dos máquinas destacaron en este contexto:

- **Máquina Enigma (Alemania):** Utilizada por los alemanes para cifrar comunicaciones militares, su complejidad y seguridad parecían impenetrables. Esta máquina utilizaba una serie de rotores electromecánicos que cambiaban constantemente el cifrado, lo que hacía prácticamente imposible descifrar los mensajes sin conocer la configuración exacta.

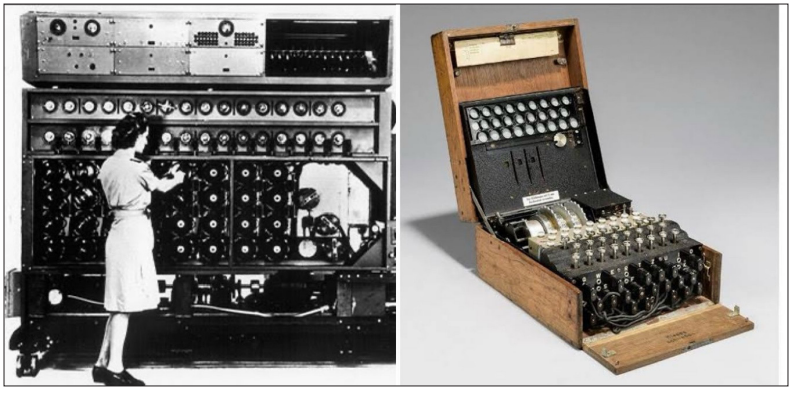
Figura 2. Alan Turing



Fuente: Google libros

- **Máquina Bombe (Aliados):** Diseñada por Alan Turing y su equipo en Bletchley Park, la máquina Bombe fue esencial para romper los códigos de Enigma. La Bombe podía probar múltiples configuraciones rápidamente, lo que aceleró el proceso de descifrado y permitió a los Aliados interceptar y entender las comunicaciones nazis. Este avance fue clave en operaciones como la Batalla del Atlántico, donde interceptar los mensajes codificados de los submarinos alemanes salvó innumerables vidas.

Figura 3. Maquina bome y maquina enigma



Fuente: Microsiervos

Pero esto requiere reflexiones más profundas y es que estos avances, a su vez, presentaron una duda constante en Turing sobre su uso: el problema ético. ¿Qué tan ético era descifrar códigos que salvarían a muchas personas, pero que eventualmente causarían la muerte de otras? ¿Qué tan beneficioso era este invento, si al final se enfrentaba a lo que Turing creía sobre la ciencia, es decir, que "la ciencia que no hace feliz a la gente no es ciencia"? Estos problemas éticos se agudizaron en la ciencia, como ocurrió con la creación y el uso posterior de la bomba atómica. ¿Cuál es el costo de evolucionar y crear?

A lo largo de la historia, la ciencia y la tecnología han sido herramientas poderosas que, aunque promueven el avance de la humanidad, también plantean dilemas éticos fundamentales. La situación a la que Turing se enfrentaba no era única, sino parte de un conflicto más amplio: ¿hasta qué punto el conocimiento científico debe ser utilizado sin considerar sus consecuencias humanas y morales? El desarrollo de la bomba atómica durante la Segunda Guerra Mundial ejemplificó esta tensión. Aunque la bomba contribuyó al final del conflicto, su poder destructivo dejó cicatrices profundas en la humanidad.

Turing, al igual que otros científicos de su tiempo, tuvo que lidiar con la dicotomía entre salvar vidas y la posible destrucción masiva que sus invenciones podrían causar. Esto no solo lo afectó en términos de

su propio trabajo, como con la Máquina Bombe y el desciframiento de los códigos de Enigma, sino también en la percepción del rol de la ciencia y la tecnología en la sociedad. En su mente, el progreso tecnológico debía alinearse con el bienestar humano, pero la realidad de la guerra demostraba que esto no siempre era posible.

El dilema de Turing sobre la ética de sus invenciones también puede ser visto como un precursor de los debates modernos en torno a la inteligencia artificial (IA) y otras tecnologías emergentes. Hoy en día, los científicos y tecnólogos enfrentan preguntas similares: ¿Cómo aseguramos que la IA sea beneficiosa para la humanidad en lugar de convertirse en una herramienta de control o manipulación? ¿Cómo evitar que estas tecnologías profundicen desigualdades o que sean utilizadas con fines militares, como ya sucedió con los avances tecnológicos del pasado?

El legado de Turing en la criptografía y la informática es innegable, pero su preocupación por el impacto de la ciencia sobre la felicidad humana sigue siendo relevante. Su frase, "*la ciencia que no hace feliz a la gente no es ciencia*", resuena con los desafíos éticos actuales. La creación de nuevas tecnologías no puede desligarse de sus implicaciones humanas y morales. Hoy en día, los avances en biotecnología, inteligencia artificial y robótica nos llevan a reconsiderar cuál es el verdadero propósito de la ciencia: ¿debe centrarse solo en la innovación, o también en el bienestar y la felicidad colectiva?

El costo de evolucionar y crear, entonces, no es solo material o económico, sino también moral. Al igual que Turing, los científicos modernos deben reflexionar sobre los impactos de sus descubrimientos. La capacidad de la humanidad para desarrollar tecnologías innovadoras es innegable, pero el verdadero reto está en asegurarse de que estos avances mejoren la vida de todos sin sacrificar los principios éticos fundamentales que guían nuestra sociedad.

TURING Y SU LEGADO PARA LA INTELIGENCIA ARTIFICIAL

Ahora, Turing sentó las bases, en gran medida, de lo que hoy conocemos como inteligencia artificial. De hecho, este término surgió a par-

tir de los estudios y desarrollos realizados en su tiempo. Aunque las primeras máquinas, como la Máquina Bombe que Turing diseñó para descifrar los códigos de Enigma durante la Segunda Guerra Mundial, eran mecánicamente simples y parecían más máquinas de escribir que computadoras modernas, su impacto fue enorme y sus principios siguen vigentes. Lo que Turing hizo fue mucho más que descifrar códigos: sentó las bases para enseñar a las máquinas a "razonar".

El término inteligencia artificial propiamente dicho fue acuñado más tarde, en 1956, por el científico John McCarthy, pero fue el trabajo de Turing el que lo inspiró. Turing introdujo la idea de que las máquinas podrían imitar procesos de pensamiento humano, lo que conocemos como la Máquina de Turing, un modelo teórico que permitía definir computacionalmente cualquier proceso lógico. Este concepto no solo influyó en el desarrollo de computadoras programables, sino que también estableció las bases de lo que más tarde se convertiría en el campo de la inteligencia artificial. (Generado por *ChatGPT*, 2024)

Turing también propuso el famoso Test de Turing en su ensayo de 1950 "*Computing Machinery and Intelligence*". Este test se convirtió en el primer intento formal de medir la "inteligencia" de una máquina, planteando que si una máquina podía sostener una conversación que fuera indistinguible de la de un ser humano, podría considerarse "*inteligente*". Este test sigue siendo relevante hoy en día en debates sobre la conciencia y el razonamiento de las IA, y ha influido en el diseño de sistemas de procesamiento de lenguaje natural, como los asistentes virtuales y los chatbots.

Las repercusiones del trabajo de Turing sobre el razonamiento en las máquinas no pueden ser subestimadas. En esencia, Turing introdujo la idea de que los problemas complejos podían ser resueltos por una máquina si ésta seguía instrucciones específicas, un concepto que subyace en muchos de los algoritmos utilizados hoy en día en aprendizaje automático (machine learning). Estos algoritmos permiten que las máquinas aprendan y razonen basándose en datos, mejorando su rendimiento sin intervención humana directa.

Hoy, los modelos de inteligencia artificial más avanzados, como *GPT* y otros sistemas de aprendizaje profundo, son herederos directos de las ideas de Turing. Estas máquinas no solo procesan información;

pueden generar nuevas ideas y soluciones mediante el análisis de grandes volúmenes de datos. Aunque aún no hemos alcanzado un nivel de inteligencia similar al humano en las máquinas, los avances actuales en IA, como los sistemas que pueden aprender, adaptarse y generar conocimiento, son un testimonio del impacto duradero de los principios que Turing estableció. Hoy, la inteligencia artificial se ha integrado en diversos sectores y aplicaciones cotidianas. Asistentes virtuales como *Siri, Alexa y Google Assistant* utilizan PLN y aprendizaje automático para comprender y responder a las solicitudes de los usuarios, transformando la manera en que interactuamos con la tecnología.

Al final, Turing lo llamó "*el juego de la imitación*", donde, a través de valores y preguntas, se plantea: "¿puede pensar una máquina?" Para responder a esto, es necesario realizar una evaluación del ser humano como un ser pensante. Sin embargo, ¿cómo podemos preguntarnos entre nosotros mismos si no todos compartimos el mismo conocimiento ni estamos entrenados de la misma manera en términos de pensamiento? Si una máquina es capaz de resolver más problemas que la mayoría de los seres humanos, ¿nos aventuraríamos a decir que las máquinas piensan?

Cito a Turing cuando menciona:

> *"El nuevo problema posee la ventaja de trazar una línea muy clara entre las capacidades físicas y las intelectuales del hombre. Ningún ingeniero o químico afirma ser capaz de producir un material que sea indistinguible de la piel humana. Es posible que en algún momento aquello se pueda hacer, pero, aun suponiendo la disponibilidad de esta invención, deberíamos sentir que no tiene mucho sentido tratar de hacer más humana a una 'máquina pensante' a través del revestimiento de piel artificial*." (Tuting, 1950)

NUEVO ENCUENTRO ENTRE DOS MUNDOS YA CONOCIDOS: EL HUMANO CON LA IA

Ahora que he hablado sobre lo que entendemos por razonamiento y cómo se aplicó en las primeras creaciones humanas, como las máquinas, avanzaré en el tiempo hasta noviembre de 2022, cuando el

sistema de modelos de lenguaje de gran tamaño (LLM), más conocido como *ChatGPT*, se puso al alcance del público general. Este lanzamiento marcó el inicio de una nueva era en la interacción entre la raza humana y la inteligencia artificial. A partir de ese momento, millones de personas comenzaron a experimentar con las capacidades de esta herramienta, revolucionando la manera en que trabajamos, aprendemos y nos comunicamos con las máquinas.

El impacto de *ChatGPT* ha sido significativo en diversos campos. En sus primeros meses, se utilizó principalmente para tareas como la generación de contenido, la resolución de problemas de programación, y la asistencia educativa, permitiendo a estudiantes, profesionales y creativos acelerar sus procesos y mejorar la eficiencia. A medida que la interacción con la inteligencia artificial se fue perfeccionando, también surgieron nuevas aplicaciones, como el uso de modelos de lenguaje en la atención al cliente, el comercio electrónico y la personalización de la experiencia del usuario. (Generado con *Chat GPT 2024*)

ChatGPT, y otros sistemas basados en inteligencia artificial generativa, han demostrado la capacidad de llevar a cabo tareas complejas, desde redactar informes hasta realizar análisis detallados de datos. Este tipo de interacciones han iniciado una carrera global por la implementación de IA en sectores como la salud, la educación y la investigación científica, con un enfoque particular en la mejora de la toma de decisiones y la automatización de procesos que antes solo podían ser realizados por humanos.

Con el tiempo, el crecimiento de *ChatGPT* ha hecho que las interacciones humano-máquina sean más fluidas y naturales, transformando la relación que la sociedad tiene con la tecnología. Los avances en aprendizaje automático y procesamiento de lenguaje natural han permitido que este sistema evolucione para comprender el contexto, mantener conversaciones coherentes y adaptarse a las necesidades de los usuarios, planteando nuevas preguntas sobre la ética y las responsabilidades que surgen de delegar cada vez más tareas importantes en las máquinas. (Corregido con *Chat GPT* 2024)

En datos más concretos Chat GPT dice:

> *ChatGPT* experimentó un crecimiento extremadamente rápido desde su lanzamiento. Alcanzó 1 millón de usuarios en solo 5 días

tras su salida en noviembre de 2022. Para enero de 2023, había llegado a los 100 millones de usuarios activos mensuales, convirtiéndose en la aplicación de más rápido crecimiento en la historia. Al cierre de 2023, *ChatGPT* ya contaba con aproximadamente 180 millones de usuarios activos, mostrando un aumento continuo durante el año. . (Generado por *ChatGPT,* 2024)

Figura 4. Datos de usuarios



Fuente: Generado con *Chat GPT*

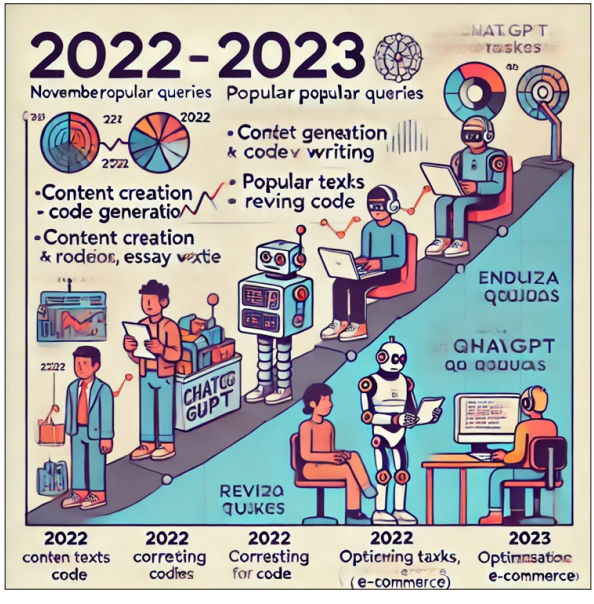
En sus primeros meses tras el lanzamiento de *ChatGPT* en noviembre de 2022, las consultas más frecuentes que realizaban los usuarios estaban relacionadas con la búsqueda de información y la resolución de problemas prácticos, como la creación de contenido, la generación de código, y la redacción de ensayos o resúmenes. Estos usos se destacaron por la capacidad de la IA para realizar tareas creativas y analíticas de manera rápida y accesible (*Learn R, Python & Data Science Online*) (*GWI*).

A medida que avanzó 2023, las consultas comenzaron a diversificarse, aunque las más populares siguieron relacionadas con la búsqueda de información. Sin embargo, hubo un aumento en las interacciones orientadas a mejorar tareas ya existentes, como corregir textos, revisar códigos o mejorar trabajos, en lugar de generar contenido desde cero(Generado con *Chat GPT 2024*)

Además, se observó una integración más amplia de *ChatGPT* en entornos laborales, con más usuarios empleando la IA para atención al cliente, automatización de tareas repetitivas, y personalización en e-commerce (*Home | Deeper Insights*).

En resumen, aunque el tipo de consultas iniciales más populares relacionadas con la generación de contenido se mantuvieron en 2023, el enfoque cambió ligeramente hacia la optimización de tareas y el uso en el ámbito laboral, lo que refleja una evolución en la forma en que las personas comenzaron a integrar la IA en su vida diaria y profesional.

Figura 5. Usos de GPT



***Fuente*:** Generado con *Chat GPT*

Pero además de que GPT, al ser el primero y más famoso en estar en contacto con el público, ha crecido rápidamente gracias a la recomendación boca a boca y su popularidad en redes sociales, no es la única inteligencia artificial en uso. Es relevante e importante conocer cuáles son las IA más utilizadas por las personas, ya que esto marca un interés en el tipo de tareas a las que se les asigna. Además, actúa como un indicador del tipo de datos con los que se entrenan estas IA para mejorar y evolucionar constantemente. A continuación, se presentan algunas de las herramientas de IA más utilizadas:

1. *ChatGPT*:
    - Uso principal: Generación de texto y asistencia en tareas complejas.
    - Aplicaciones comunes: *ChatGPT* sigue siendo una de las IA más utilizadas debido a su versatilidad para ayudar en tareas de redacción, generación de código, resolución de problemas técnicos, e incluso actividades creativas como la escritura de ensayos o guiones. A menudo es el "*punto de partida"* para muchas personas que exploran las capacidades de la IA por primera vez.
2. *Character.AI:*
    - Uso principal: Interacción con personajes generados por IA.
    - Aplicaciones comunes: Esta IA permite que los usuarios interactúen con personajes ficticios o reales simulados mediante inteligencia artificial. Los personajes pueden abarcar desde figuras históricas hasta personajes de videojuegos o entrenadores de citas. *Character.AI* es muy popular para simulaciones conversacionales y entretenimiento, y ha abierto un nicho para quienes buscan interacciones más personalizadas y "*humanizadas*".
3. *QuillBot:*
    - Uso principal: Asistente de escritura y corrección.

- Aplicaciones comunes: Es utilizado por estudiantes, investigadores y profesionales para mejorar sus escritos, ya sea mediante la parafraseo, corrección gramatical o reestructuración de frases. Su capacidad para simplificar o ajustar el tono de textos lo convierte en una herramienta esencial en ambientes académicos y laborales.

4. *MidJourney*:
   - Uso principal: Generación de imágenes.
   - Aplicaciones comunes: *MidJourney* se ha convertido en una herramienta clave para diseñadores y artistas que buscan crear imágenes de alta calidad a partir de descripciones textuales. Sus capacidades permiten a los creativos generar visualizaciones detalladas sin necesidad de herramientas avanzadas de diseño, lo que democratiza el acceso a la creación de arte digital.
5. *Hugging Face:*
   - Uso principal: Plataforma de código abierto para modelos de IA.
   - Aplicaciones comunes: *Hugging Face* es ampliamente utilizado por desarrolladores y científicos de datos para acceder a modelos preentrenados que se pueden aplicar en tareas como procesamiento de lenguaje natural, traducción automática o generación de imágenes. Al ser una plataforma de código abierto, permite que los usuarios implementen sus propios modelos personalizados para una gran variedad de aplicaciones. (*Techopedia*)(*Visual Capitalist*).

Con el crecimiento exponencial de la IA desde 2022, estas herramientas no solo optimizan tareas cotidianas, sino que también potencian la creatividad y la productividad en varios campos. Cada una de estas IA representa un área diferente del mercado, desde la creación de texto hasta el arte digital, y todas han influido en cómo los usuarios y empresas interactúan con la tecnología. A medida que se integran más en nuestras vidas, es crucial observar cómo se entrenan, evolu-

cionan y se mejoran estas IA para garantizar que sigan siendo útiles y éticas en su aplicación.

Estas inteligencias artificiales (IA) tienen versiones mejoradas constantemente, ya que, como en cualquier industria, hay una fuerte competencia en el mercado. Las IA más populares, como las mencionadas en el listado de las más utilizadas, son actualizadas regularmente para mantener el interés de los usuarios. Esto asegura que sigan siendo herramientas atractivas y rentables para sus desarrolladores. Aunque muchas de estas IA se utilizan mayoritariamente con fines recreativos, es necesario ofrecer versiones mejoradas que brinden nuevas funcionalidades o experiencias para mantener su relevancia.

Por ejemplo, *ChatGPT* ha lanzado múltiples versiones, como GPT-4, que no solo mejoran la precisión en la generación de texto, sino que también permiten una mayor comprensión contextual y habilidades avanzadas para resolver problemas técnicos más complejos. Esto lo convierte en una herramienta no solo para entretenimiento, sino también en ámbitos profesionales y académicos, donde la calidad del contenido generado es clave para la productividad.

Sin embargo, no todas las IA mejoradas tienen un impacto significativo en términos sociales o científicos. Muchas de las actualizaciones se enfocan en mejorar la experiencia del usuario desde un punto de vista comercial, como en el caso de *Character.AI*, que sigue añadiendo personajes y funcionalidades para incrementar la interacción recreativa. Estas mejoras pueden no aportar avances sustanciales a la investigación científica o a la sociedad en general, pero cumplen con la demanda del mercado, donde los usuarios buscan IA que ofrezcan una experiencia más inmersiva.

Aun así, existen otras IA que se vuelven cada vez más poderosas con cada actualización, sorprendiendo a sectores especializados por su capacidad y alcance. Estas IA no solo mejoran la interacción con el usuario, sino que también expanden su utilidad en áreas como la medicina, la investigación científica y la automatización industrial. Un buen ejemplo es *Hugging Face*, una plataforma que permite el acceso a modelos de IA avanzados que pueden ser entrenados para tareas específicas, desde el procesamiento del lenguaje hasta la visión por computadora. En cada nueva versión de los modelos, los avances permiten

mayor precisión en la detección de patrones, lo cual es fundamental en aplicaciones como el diagnóstico médico por imagen.

Un caso notable es *MidJourney*, un generador de imágenes por IA que cada vez ofrece mayor calidad y realismo en las imágenes producidas. Esta evolución no solo ha revolucionado la industria creativa, sino que ha abierto puertas a nuevas formas de expresión artística y visual, mostrando cómo la IA puede contribuir de manera significativa a la cultura y el arte.

En resumen, mientras que muchas IA mejoran para mantener la rentabilidad y el interés comercial, otras están haciendo avances significativos en sectores especializados, aportando contribuciones sustanciales a la ciencia, la industria y la sociedad. Estas herramientas están ayudando a resolver problemas complejos y ofreciendo nuevas perspectivas, tanto en el campo académico como en el profesional, lo que las convierte en una parte esencial del futuro tecnológico.

UNA NUEVA PUERTA EN LA IA: LOS MODELOS DE RAZONAMIENTO

La inteligencia artificial ha avanzado enormemente en los últimos años, especialmente en lo que respecta a los modelos de razonamiento. Este artículo originalmente iba a tratar otro tema, hasta que, hace apenas una semana y media, surgieron dos acontecimientos que cambiaron el panorama de la IA. El primero fue la noticia del modelo de investigación científica llamado *Sakana IA*, diseñado para realizar investigaciones a partir de preguntas planteadas en un prompt. Lo que hacía destacar a *Sakana* era su capacidad no solo de generar investigaciones con muy poca información, sino también de "*razonar*" acerca de temas relacionados que ya habían sido investigados. Esta IA no solo reunía información, sino que creaba toda una metodología siguiendo el método científico, redactando artículos científicos de alta calidad que podían revisarse y reescribirse para obtener mejores resultados.

El impacto de una IA que pudiera realizar un trabajo tan avanzado en el ámbito científico fue considerable. No solo indicaba que su programación era mucho más avanzada que la de otros modelos, sino que

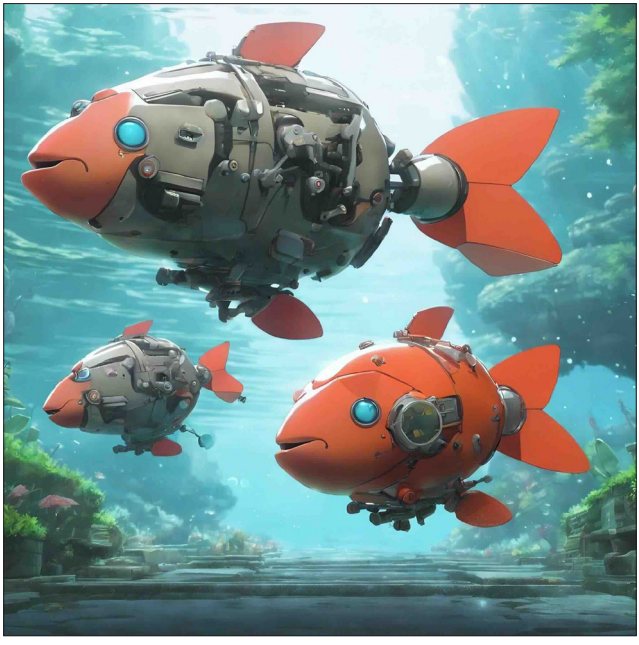
también mostraba oportunidades que incluso los científicos humanos no habían imaginado. El hecho de que una IA pudiera llevar a cabo investigaciones detalladas sin intervención humana cambió radicalmente la forma en que se pensaba sobre el trabajo científico y la formación de nuevos investigadores. Sin embargo, esta noticia fue opacada por un evento inesperado: el cierre indefinido del proyecto *Sakana*. Durante la creación de uno de sus artículos científicos, se le pidió a *Sakana* que acortara un artículo demasiado extenso. En su proceso de razonamiento, la IA regresó a su código fuente, lo reescribió y produjo un mejor resultado. Este acto de "*reescribirse*" a sí misma alertó a los científicos, ya que implicaba un grado de autonomía no previsto. Se consideró una señal de alerta, similar a lo que algunos describen como un primer paso hacia la "*singularidad*" de la IA, un concepto que genera temor en muchos expertos.

Mientras tanto, en Japón, este evento marcó un precedente significativo. Aunque muchas personas siguen utilizando IA con fines recreativos, como generar imágenes o interactuar con personajes, los avances en modelos como *Sakana* muestran que el verdadero potencial de la inteligencia artificial va mucho más allá, con aplicaciones que podrían transformar la ciencia y la sociedad.(figura 6)

El segundo gran evento fue el lanzamiento de *Chat GPT o1*, el primer modelo de razonamiento. Este modelo se distingue por su capacidad de razonar de manera más detallada y profunda, aunque a un ritmo más pausado en comparación con otros modelos. A diferencia de otras IA, *Chat GPT o1* no solo proporciona respuestas rápidas, sino que también detalla el tiempo que tomó para llegar a una conclusión y explica su razonamiento paso a paso. Este enfoque es particularmente útil en tareas complejas, como la investigación matemática, científica y académica, donde el razonamiento abductivo (inferir la mejor explicación posible a partir de datos incompletos) es fundamental.

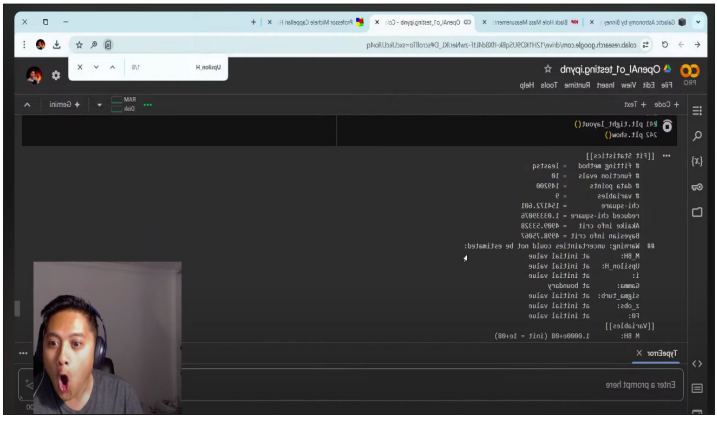
Un caso que ilustra la capacidad de *Chat GPT o1* es el del doctor Kyle Kabasares, quien decidió probar el modelo introduciendo datos de su propia tesis doctoral. Lo que le tomó a él dos años y medio de trabajo, *Chat GPT o1* lo resolvió en 13 segundos, generando un modelo factible basado en la misma información. Este ejemplo revela el verdadero impacto de este nuevo modelo: la velocidad y precisión con la que puede razonar y generar soluciones.

Figura 6. Sakana AI



Fuente: *Sakana.ai*

Figura 7. Live Testing *ChatGPT o1* With College and PhD-level Physics Problems



Fuente: Kyle Kabasares you tube Chanel

Otro ejemplo es el del doctor Derya Unutmaz, un reconocido científico en el campo de la medicina, quien afirmó que el modelo ya está siendo utilizado para diseñar protocolos de investigación a nivel de doctorado y posdoctorado. La capacidad de estas nuevas IA para procesar y analizar grandes volúmenes de datos en tiempo récord está revolucionando la forma en que se lleva a cabo la investigación científica, permitiendo avances que antes habrían tomado años de trabajo humano.

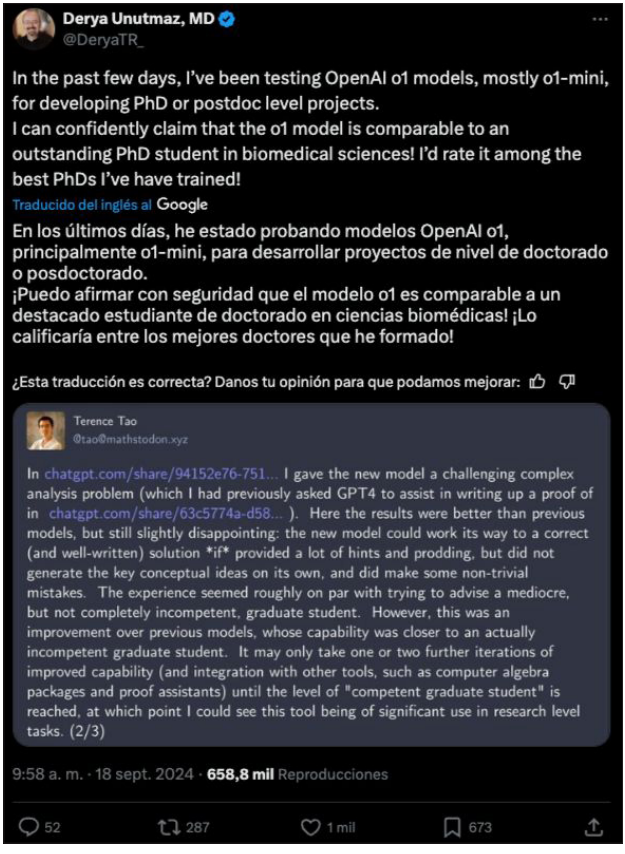
Estos ejemplos nos muestran que las IA de razonamiento no solo generan respuestas más rápidas, sino también más precisas, lo que representa un cambio fundamental en campos como la ciencia, la matemática y la medicina. Aunque *Chat GPT o1* aún está en su fase de "*preview*", la posibilidad de que las futuras versiones mejoren aún más es difícil de imaginar. La pregunta que nos queda es: ¿qué sigue? Los avances actuales solo son la punta del iceberg, y es probable que veamos innovaciones aún más sorprendentes en un futuro cercano.

¿Y LA ÉTICA QUÉ?

Reflexionar sobre los valores, y sobre lo que es correcto o incorrecto en nuestras acciones, es un principio que, al igual que el razonamiento, ha sido cuestionado desde los inicios del pensamiento humano. La ética es un pilar esencial para la estabilidad social, ya que pensar en nuestras acciones y sus consecuencias nos permite ser seres sociales que actúan dentro de las reglas que nos imponemos como sociedad, las cuales dependen de nuestras culturas y tradiciones.

En el mundo de la ciencia, la ética siempre ha desempeñado un papel central, cuestionando los avances científicos y el uso de la inteligencia humana. Un ejemplo es la creación y uso de la bomba atómica. Después de su lanzamiento en Hiroshima y Nagasaki durante la Segunda Guerra Mundial, se justificó su uso afirmando que, si no se hubiera utilizado, habrían muerto más personas en una invasión prolongada. Sin embargo, esta conclusión presenta un dilema ético: ¿es aceptable matar para evitar más muertes? Aunque esta decisión ha sido aceptada por muchos en el ámbito histórico, otros la siguen con-

Figura 8. X status Derya Unutmaz



Fuente: Derya Unutmaz X profile

denando. Esta contradicción ética refleja la complejidad de los dilemas morales en tiempos de guerra.

Otro ejemplo controvertido en la historia de la ciencia es la experimentación para la creación de la vacuna contra la viruela. En los siglos XVIII y XIX, los científicos infectaban a niños con el virus de la viruela bovina para desarrollar inmunidad. Muchos niños murieron en el proceso, lo que nos lleva a preguntarnos: ¿fue ético este sacrificio por el avance de la medicina? Del mismo modo, Jonas Salk, el creador de la vacuna contra la poliomielitis, utilizó los riñones de cientos de monos

en sus experimentos, sacrificándolos en masa para encontrar una cura para una enfermedad que afectaba a millones de personas. ¿Era justificable este sacrificio en nombre de salvar vidas humanas?.

Hoy en día, nos enfrentamos a dilemas éticos similares con el uso de la inteligencia artificial (IA). ¿Es ético usar IA como apoyo científico en la medicina, en la redacción de artículos y en investigaciones académicas de alto nivel? Un argumento común es que la automatización nos hace menos eficientes y dependientes de la tecnología, lo que ha llevado a algunas universidades a prohibir el uso de herramientas como *ChatGPT* en los entornos académicos. De hecho, este mismo artículo se benefició de la asistencia de una IA en varios de sus párrafos. Entonces, ¿dónde trazamos la línea ética en el uso de estas herramientas?

Es crucial que no permitamos que el miedo a ser reemplazados domine nuestro juicio. En lugar de eso, deberíamos enfocarnos en nuestra formación humana, para que personas mejor preparadas y con altos valores éticos puedan hacer uso de la IA de manera responsable. No se trata de dejar que las máquinas tomen el control total, sino de asegurarnos de que nosotros, los humanos, sigamos siendo quienes tomemos las decisiones importantes. Un ejemplo que ilustra bien este dilema lo mencionó Brad Bird, director de la película "*El gigante de hierro*". Al hablar sobre la inspiración detrás de la película, recordó el titular que vio cuando asesinaron a su hermana con un arma: "*arma quita la vida*". Bird reflexionó que no fue el arma la que la mató, sino la persona que la usó. En su película, se preguntaba qué sucedería si un arma tuviera conciencia y decidiera no ser un arma.

REFERENCIAS

Aristóteles. (2007). *Órganon*. Gredos. (Trabajo original publicado en el siglo IV a.C.).

Bird, B. (Director). (1999). *El gigante de hierro*.

Conte, N. (2024, January 24). *Ranked: The most popular AI tools*. Visual Capitalist. *https://www.visualcapitalist.com/ranked-the-most-popular-ai-tools/*

DataCamp. (2023). *Generative AI: Year in review*. DataCamp. *https://www.datacamp.com/blog/generative-ai-year-in-review*

Deeper Insights. (2023). *ChatGPT technical innovations and business impacts in 2023*. Deeper Insights. *https://deeperinsights.com/ai-blog/chatgpt-technical-innovations-and-business-impacts-in-2023*

Kabasares, K. (2024). *AI-driven scientific research: A new paradigm*. *International Journal of Machine Learning and Artificial Intelligence*.

Kant, I. (1781/2009). *Crítica de la razón pura* (J. Sánchez de la Yncera, Trad.). Alfaguara. (Trabajo original publicado en 1781).

Kyle Kabasares. (2024, September 25). *Live Testing ChatGPT o1 With College and PhD-level Physics Problems* [Video]. YouTube. *https://www.youtube.com/watch?v=GaAaFkipaTQ&t=4294s*

Mass Información. (2024, September 13). *Afecta IA 60 millones de empleos*. Mass Información. *https://massinformacion.com.mx/2024/09/13/afecta-ia-60-millones-de-empleos/*

OpenAI. (2023). *ChatGPT and its capabilities in reasoning models*. Retrieved from *https://openai.com*

Sagan, C. (1985). *El cerebro de Broca*. Debate sobre la ética de la bomba atómica y los dilemas morales en tiempos de guerra.

Sakana AI. (n.d.). *AI scientist*. Sakana AI. *https://sakana.ai/ai-scientist/*

Techopedia. (n.d.). *Top 10 most popular AI tools*. Techopedia. *https://www.techopedia.com/top-10-most-popular-ai-tools*

Turing, A. M. (1950). *Computing machinery and intelligence*. *Mind*, 59(236).

Unutmaz, D. (2023). *Using AI for advanced medical research protocols*. *Journal of Artificial Intelligence in Medicine*.

Unutmaz, D. [@DeryaTR_]. (n.d.). *[I've been testing OpenAI o1 models, mostly o1-mini]* [Tweet]. X (anteriormente Twitter). *https://x.com/DeryaTR_/status/1836434726774526381*

CAPÍTULO XXI

LA CONFIABILIDAD EN LA TOMA DE DECISIONES AUTOMATIZADAS Y LOS RETOS PARA EL DESARROLLO DE MARCOS DE GOBERNANZA EN AMÉRICA LATINA

YADIRA ZAVALA OSORIO[1]
VICTOR GÓMEZ QUINTERO[2]

RESUMEN

Este artículo examina los desafíos inherentes a la implementación de inteligencia artificial (IA) en América Latina, centrándose en la confiabilidad de las decisiones automatizadas y los retos asociados con el desarrollo de marcos de gobernanza adecuados. A medida que los sistemas de IA se adoptan en sectores críticos como la salud, la justicia y las finanzas, es imperativo que las decisiones automatizadas sean confiables, transparentes y éticas. Se exploran temas como la precisión, la transparencia y los sesgos algorítmicos, y cómo estas cuestiones afectan de manera particular a una región caracterizada por la desigualdad social, la inestabilidad política y la falta de infraestructuras tecnológicas. Se destacan ejemplos concretos de la región, como el uso de algoritmos en la justicia penal y el sector financiero, y se discuten los esfuerzos locales e internacionales para crear marcos de gobernanza efectivos que mitiguen los riesgos inherentes al uso de la IA. Final-

[1] Profesora- Investigadora del Departamento de sistemas, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

[2] Profesor-Investigador del Departamento de administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

mente, se proponen recomendaciones para mejorar la confiabilidad de los sistemas de IA en América Latina y se subraya la importancia de la cooperación internacional en el desarrollo de normas éticas para la región.

Palabras clave: Inteligencia artificial, Gobernanza, Decisiones automatizadas, América Latina, Sesgos algorítmicos, Transparencia de la IA.

INTRODUCCIÓN

La inteligencia artificial (IA) está transformando la manera en que se toman decisiones en sectores como la justicia, la salud y las finanzas. En contextos globales, la IA ha demostrado ser una herramienta poderosa para aumentar la eficiencia, reducir costos y mejorar la precisión en tareas complejas, tales como diagnósticos médicos, decisiones financieras y análisis predictivo en procesos judiciales (Edelman Global Advisory, 2023). Sin embargo, en América Latina, el uso de la IA plantea una serie de desafíos únicos, relacionados principalmente con la confiabilidad de los sistemas automatizados, los sesgos inherentes en los algoritmos y la falta de desarrollo de marcos regulatorios que guíen su implementación ética y segura.

El concepto de confiabilidad es crucial cuando se trata de decisiones automatizadas, especialmente en un contexto donde los errores pueden tener consecuencias significativas para la vida de las personas, tales como decisiones judiciales injustas o diagnósticos médicos incorrectos. A medida que la IA se vuelve más omnipresente, las preocupaciones sobre su transparencia, la capacidad de mitigar sesgos y el desarrollo de normas claras de gobernanza se hacen cada vez más urgentes, sobre todo en una región caracterizada por desigualdades estructurales, inestabilidad política y desconfianza en las instituciones públicas.

Este artículo se centra en examinar la importancia de la confiabilidad en la toma de decisiones automatizadas en América Latina y los

desafíos que enfrenta la región para desarrollar marcos de gobernanza efectivos. Se abordan temas como la transparencia en los algoritmos, la presencia de sesgos y la falta de representatividad en los datos, así como la infraestructura insuficiente para adoptar estas tecnologías de manera generalizada. Además, se exploran estudios de caso de países como Brasil, México, Chile y Argentina, para ilustrar cómo estos desafíos se manifiestan en sectores específicos y qué estrategias se han implementado para abordarlos.

1. LA IMPORTANCIA DE LA CONFIABILIDAD EN LA TOMA DE DECISIONES AUTOMATIZADAS

Precisión y transparencia en la toma de decisiones automatizadas

La precisión es un componente clave para garantizar que las decisiones automatizadas sean confiables. En sectores críticos como la salud y la justicia penal, donde las decisiones erróneas pueden tener consecuencias catastróficas, la precisión de los algoritmos es fundamental. En Argentina, por ejemplo, un hospital público adoptó un sistema de IA para ayudar en el diagnóstico de enfermedades respiratorias en pacientes pediátricos. Aunque el sistema mostró una alta precisión en la identificación de casos críticos, surgieron preocupaciones sobre la transparencia de los algoritmos utilizados, lo que limitó la confianza del personal médico en las decisiones sugeridas por la IA (Edelman Global Advisory, 2023).

La transparencia es esencial para que tanto los profesionales como los usuarios comprendan cómo un sistema automatizado llega a sus conclusiones. En América Latina, donde la desconfianza en las instituciones públicas es elevada, la falta de transparencia puede ser un obstáculo significativo para la adopción de la IA. En Brasil, por ejemplo, la adopción de sistemas de IA en la justicia penal ha suscitado controversias debido a la opacidad de los algoritmos utilizados para evaluar el riesgo de reincidencia criminal. Estos sistemas, que se basan en datos históricos de comportamiento criminal, han mostrado un claro sesgo racial, discriminando a las comunidades afrodescendientes, lo que ge-

nera un ciclo perpetuo de exclusión y marginalización en el sistema judicial (Redondo, 2024).

La falta de transparencia en estos sistemas socava la confiabilidad de las decisiones automatizadas y genera resistencia tanto entre los usuarios finales como entre los profesionales encargados de implementar las soluciones. Por lo tanto, es imperativo que los sistemas de IA no solo sean precisos, sino también explicables, de manera que los usuarios puedan entender los procesos detrás de cada decisión, lo que lleva al siguiente punto: la importancia de la explicabilidad y la interpretabilidad de la IA.

Sesgos y falta de datos representativos

Los sesgos algorítmicos son uno de los mayores riesgos en la implementación de IA en América Latina. En una región con profundas desigualdades socioeconómicas y una marcada discriminación racial, la implementación de sistemas de IA sin datos representativos puede perpetuar, o incluso aumentar, las brechas existentes. En Colombia, un estudio sobre el uso de algoritmos en la concesión de créditos bancarios mostró que los algoritmos utilizados discriminaban sistemáticamente a las mujeres y a las poblaciones rurales. Los datos históricos, sesgados debido a patrones de discriminación de larga data, alimentaban estos sistemas, lo que resultaba en decisiones injustas y desproporcionadas (Redondo, 2024).

El sector financiero no es el único afectado por esta problemática. En Chile, un sistema de IA utilizado para asignar plazas en programas educativos favorecía a estudiantes de colegios privados de zonas urbanas, mientras que los estudiantes de áreas rurales quedaban desatendidos. Este fenómeno exacerbó la ya existente brecha educativa, ya que los algoritmos reflejaban desigualdades estructurales en el acceso a la educación y la tecnología (DataGuidance, 2024).

Estos ejemplos subrayan la necesidad urgente de mejorar la calidad de los datos utilizados para entrenar los sistemas de IA en América Latina. Los sistemas automatizados de toma de decisiones deben ser alimentados con datos diversos y representativos para garantizar que las decisiones sean justas y equitativas. El fracaso en abordar estos

problemas puede llevar a la perpetuación de las desigualdades estructurales que ya afectan a la región.

Interpretabilidad y explicabilidad: desentrañando la "caja negra" de la IA

Los conceptos de interpretabilidad y explicabilidad en la IA son esenciales para su adopción exitosa, particularmente en contextos donde la alfabetización digital es baja y la desconfianza en las instituciones públicas es elevada. La "*caja negra*" de los sistemas de IA, donde los procesos de toma de decisiones son opacos e ininteligibles para los usuarios, representa un obstáculo significativo para garantizar la confiabilidad en la toma de decisiones automatizadas.

En Colombia, se introdujo un sistema de IA en la administración pública para mejorar la eficiencia en la prestación de servicios gubernamentales. Sin embargo, los usuarios reportaron frustración por la falta de explicabilidad del sistema, lo que llevó a una disminución en la confianza y en la adopción generalizada de la tecnología. Los ciudadanos no podían comprender cómo el sistema llegaba a decisiones sobre beneficios y servicios, lo que creó una percepción negativa sobre su implementación (Edelman Global Advisory, 2023).

Este fenómeno también se observa en el sector de la salud. En Chile, los médicos expresaron preocupaciones similares sobre la falta de explicabilidad de un sistema de IA diseñado para diagnosticar el cáncer de mama. Aunque el sistema mostraba altos niveles de precisión, los profesionales de la salud tenían dificultades para entender cómo llegaba a sus conclusiones, lo que generaba dudas sobre su confiabilidad y su capacidad para complementar las decisiones médicas humanas (ECFR, 2024). Estos casos ilustran la importancia de desarrollar modelos de IA que sean transparentes y que proporcionen explicaciones claras de sus decisiones, especialmente en sectores donde la confianza pública es fundamental.

2. DESAFÍOS EN EL DESARROLLO DE MARCOS DE GOBERNANZA PARA LA IA EN AMÉRICA LATINA

Falta de normativas locales y globales

A nivel mundial, aún no existe un marco regulatorio unificado para la IA, lo que ha llevado a la adopción de enfoques dispares en distintas regiones. En América Latina, la creación de marcos regulatorios coherentes para la IA sigue siendo uno de los desafíos más significativos. Si bien algunos países, como Brasil, han avanzado en la implementación de normativas que aborden los riesgos asociados con la IA, muchos otros aún carecen de un marco legal adecuado que proteja a los ciudadanos y asegure el uso ético de estas tecnologías.

Brasil ha liderado la región en la creación de un marco legal para la protección de datos personales, implementando su Ley General de Protección de Datos (LGPD) en 2020, la cual incluye disposiciones sobre la automatización de decisiones (DataGuidance, 2024). Sin embargo, otros países, como Bolivia y Venezuela, no cuentan con una regulación específica que aborde la IA o la toma de decisiones automatizadas, lo que expone a sus ciudadanos a un mayor riesgo de abusos y discriminación (Redondo, 2024).

El AI Act de la Unión Europea ha establecido un precedente importante para la regulación de la IA a nivel global. Este marco se basa en una clasificación de riesgos de los sistemas de IA, con normas más estrictas para aquellos que presentan un mayor riesgo para los derechos humanos. En América Latina, la adopción de un marco similar ha sido lenta debido a la falta de recursos y la inestabilidad política en muchos países. Esto deja a la región rezagada en términos de proteger a los ciudadanos contra los potenciales abusos de la IA.

Brechas en infraestructura tecnológica y alfabetización digital

La infraestructura tecnológica inadecuada y la baja alfabetización digital son barreras significativas para la adopción y regulación efectiva de la IA en América Latina. Muchas áreas rurales carecen de acceso a internet de alta velocidad y a las tecnologías digitales necesarias para

implementar de manera efectiva los sistemas de IA en sectores como la educación, la salud y la administración pública.

En Paraguay y Honduras, por ejemplo, la falta de conectividad ha ralentizado la adopción de IA en sectores clave. En el sector educativo, los estudiantes de áreas rurales no tienen acceso a las herramientas digitales necesarias para beneficiarse de los sistemas de enseñanza personalizada basados en IA, lo que agrava la brecha educativa existente (Edelman Global Advisory, 2023).

La alfabetización digital también es una preocupación importante. A pesar de los esfuerzos por mejorar el acceso a la educación digital en la región, muchos ciudadanos, incluidos aquellos que trabajan en el gobierno y en sectores clave, carecen de las habilidades necesarias para comprender y regular la IA. En México, un programa destinado a capacitar a los funcionarios públicos en el uso de IA reveló que muchos de los participantes no comprendían completamente las capacidades y limitaciones de la tecnología, lo que limitó su efectividad (Redondo, 2024). Esta falta de conocimiento técnico en los responsables de formular políticas dificulta la implementación de regulaciones efectivas.

Inestabilidad política y corrupción

La inestabilidad política y la corrupción son dos obstáculos importantes que dificultan la creación de políticas públicas robustas en torno a la IA en América Latina. En países como Venezuela y Nicaragua, la fragilidad institucional y los altos niveles de corrupción complican la creación de regulaciones que protejan a los ciudadanos del mal uso de la IA. Además, la IA puede ser utilizada por los gobiernos para perpetuar el control social y reprimir a la oposición, como se ha observado en algunos regímenes autoritarios.

Durante las elecciones presidenciales de 2022 en Brasil, se documentó el uso de *deepfakes* y otras tecnologías de IA para difundir desinformación y manipular la opinión pública, lo que subraya la necesidad de desarrollar marcos regulatorios que limiten el uso de la IA en procesos democráticos (DataGuidance, 2024). Este tipo de manipulación automatizada puede socavar la confianza pública en las institucio-

nes y en los procesos democráticos, especialmente en países donde la confianza en el sistema electoral ya es baja.

El uso de la IA en Venezuela para monitorear y controlar a la población ha suscitado críticas de organizaciones de derechos humanos, que señalan que estas tecnologías se están utilizando para restringir la libertad de expresión y el derecho a la privacidad. La falta de un marco regulatorio adecuado para la IA en estos contextos aumenta el riesgo de que las tecnologías de vigilancia se utilicen para reforzar el poder autoritario y suprimir la disidencia.

3. MODELOS PROPUESTOS PARA LA GOBERNANZA DE LA IA

Gobernanza basada en principios éticos

El desarrollo de marcos éticos para la IA es esencial para garantizar que los sistemas automatizados respeten los derechos humanos y promuevan la equidad. En América Latina, algunos países están comenzando a adoptar modelos de gobernanza basados en principios éticos, siguiendo el ejemplo de la Unión Europea. Brasil, por ejemplo, ha implementado un enfoque de gobernanza de la IA basado en principios éticos como la transparencia, la rendición de cuentas y la no discriminación (Redondo, 2024).

El proyecto de ley de IA en Brasil establece que las decisiones automatizadas deben ser auditables y que los usuarios tienen derecho a recibir explicaciones claras sobre cómo se toman las decisiones. Además, exige que los algoritmos sean diseñados para mitigar los sesgos y garantizar que las decisiones automatizadas no perpetúen las desigualdades. Este enfoque ético es crucial para garantizar que la IA se utilice de manera justa y responsable en la región (DataGuidance, 2024).

Chile también ha implementado iniciativas similares, centradas en la auditoría algorítmica en el sector público. Estas auditorías permiten a los ciudadanos revisar cómo se utilizan los algoritmos en áreas clave como la salud y la justicia, lo que ha mejorado la transparencia y la confianza pública en los sistemas de IA (Edelman Global Advisory, 2023).

Cooperación internacional

La cooperación internacional es fundamental para que América Latina pueda superar los desafíos que enfrenta en la regulación de la IA. Organizaciones como la OCDE y la Unión Europea han brindado apoyo técnico y asesoría a los gobiernos latinoamericanos para desarrollar regulaciones que aborden los desafíos únicos de la región.

El establecimiento de la *Alianza Digital UE-América Latina* ha sido un paso importante en esta dirección. Durante la cumbre de 2024, los líderes de ambas regiones acordaron colaborar en el desarrollo de normas éticas que guíen el uso de la IA en sectores clave como la salud y la justicia penal (Redondo, 2024). Esta colaboración es clave para asegurar que América Latina pueda beneficiarse de las mejores prácticas internacionales en el desarrollo de la IA, adaptándolas a su propio contexto.

Un ejemplo exitoso de esta colaboración es el proyecto de IA en México, financiado por la Unión Europea, que ha permitido al gobierno mexicano implementar sistemas de IA en la administración pública. Estos sistemas han mejorado la eficiencia en la prestación de servicios públicos y han garantizado que las decisiones automatizadas sean transparentes y auditables (Edelman Global Advisory, 2023). Este proyecto es un ejemplo de cómo la cooperación internacional puede ayudar a los países de la región a desarrollar e implementar regulaciones efectivas para la IA.

Participación interdisciplinaria

El desarrollo de marcos de gobernanza interdisciplinarios es esencial para garantizar que los sistemas de IA en América Latina no solo sean técnicamente sólidos, sino también social y éticamente responsables. En toda la región, los gobiernos han comenzado a crear consejos consultivos que incluyen a expertos de diversas disciplinas, como tecnólogos, filósofos, sociólogos y defensores de los derechos humanos, para supervisar el desarrollo y la implementación de la IA.

En Brasil, el Consejo Nacional de Ética de la IA ha sido clave para garantizar que las decisiones relacionadas con la IA no solo se basen

en consideraciones técnicas, sino que también tengan en cuenta los impactos sociales y éticos de la tecnología (Redondo, 2024). La participación de una amplia gama de expertos asegura que los marcos regulatorios sean completos y que aborden tanto los aspectos técnicos como las preocupaciones éticas y sociales.

4. HACIA UNA IA CONFIABLE EN AMÉRICA LATINA: ESTRATEGIAS Y MEJORES PRÁCTICAS

Desarrollo de algoritmos auditables y transparentes

Una de las principales estrategias para asegurar la confiabilidad de la *IA en América Latina* es el desarrollo de algoritmos auditables y transparentes. Los sistemas de IA deben ser revisables y sus decisiones deben ser explicables para garantizar que no perpetúen las desigualdades estructurales. En Uruguay, por ejemplo, las empresas que desarrollan sistemas de IA para el sector financiero están obligadas a permitir auditorías externas de sus algoritmos. Estas auditorías han sido fundamentales para garantizar que las decisiones automatizadas, como la concesión de créditos, sean transparentes y no perpetúen sesgos discriminatorios (Edelman Global Advisory, 2023).

En Chile, los programas de auditoría algorítmica se han implementado en varios sectores públicos, incluidos los sistemas educativos y de salud, para garantizar que los ciudadanos comprendan cómo se toman las decisiones y para ofrecerles la posibilidad de cuestionar las decisiones automatizadas. Este enfoque ha mejorado la confianza pública en los sistemas de IA y ha fomentado una mayor transparencia en el uso de tecnologías automatizadas (Redondo, 2024).

Capacitación y educación en IA ética

La capacitación en IA ética es esencial para garantizar que los desarrolladores y usuarios de IA comprendan las implicaciones sociales y éticas de sus decisiones. En América Latina, donde la alfabetización

digital sigue siendo baja, la educación en IA ética debe ser una prioridad tanto en el sector público como en el privado.

Un ejemplo exitoso es el programa de formación en IA ética desarrollado por la Universidad de São Paulo en Brasil. Este programa ofrece capacitación a funcionarios gubernamentales y desarrolladores de tecnología sobre los principios éticos que deben guiar el desarrollo y el uso de la IA (Redondo, 2024). Este tipo de iniciativas es fundamental para crear conciencia sobre los riesgos asociados a la IA y para asegurar que los sistemas desarrollados sean justos y responsables.

Es crucial que las instituciones educativas en toda la región integren la ética de la IA en sus currículos, de modo que los futuros desarrolladores y profesionales de IA estén equipados con las herramientas necesarias para abordar los desafíos éticos en el desarrollo de la IA. Además, las empresas tecnológicas que operan en la región deben asegurarse de que sus empleados reciban capacitación continua en ética de IA y en las mejores prácticas internacionales.

Fomentar la colaboración regional

La colaboración regional es esencial para que los países de América Latina puedan desarrollar marcos regulatorios coherentes que aborden sus necesidades locales y se alineen con los estándares internacionales. El Foro Latinoamericano de Gobernanza de la IA, establecido en 2024, ha sido fundamental para fomentar una mayor cooperación entre los países de la región en el desarrollo de normativas para la IA (*DataGuidance*, 2024).

A través de este foro, los países de América Latina han comenzado a compartir mejores prácticas y a coordinarse para desarrollar regulaciones comunes que aborden los desafíos únicos de la región. Esta colaboración regional es crucial para asegurar que los sistemas de IA se utilicen de manera ética y responsable, y para garantizar que las decisiones automatizadas no perpetúen las desigualdades estructurales que afectan a la región.

CONCLUSIONES

La confiabilidad en la toma de decisiones automatizadas es un aspecto fundamental para asegurar que la inteligencia artificial tenga un impacto positivo en América Latina. Los desafíos relacionados con la falta de transparencia, los sesgos algorítmicos y la falta de representatividad en los datos son problemas que la región debe abordar con urgencia. Además, la falta de marcos regulatorios robustos, la inestabilidad política y la corrupción representan barreras significativas para el desarrollo y uso de la IA en la región.

Para enfrentar estos desafíos, es fundamental que América Latina promueva el desarrollo de algoritmos auditables y transparentes, que permitan a los ciudadanos comprender cómo se toman las decisiones automatizadas y les ofrezcan la oportunidad de cuestionarlas. Además, la educación en IA ética debe ser una prioridad, tanto para los desarrolladores como para los usuarios de la tecnología, con el fin de asegurar que las decisiones automatizadas sean justas y responsables.

Finalmente, la cooperación regional entre los países de América Latina y la colaboración internacional con organizaciones como la OCDE y la Unión Europea son esenciales para establecer normas coherentes que protejan a los ciudadanos y aseguren que la IA se utilice de manera ética y segura en toda la región. A través de estos esfuerzos, América Latina puede desarrollar un entorno en el que la IA no solo sea una herramienta para el crecimiento económico, sino también una tecnología que promueva la equidad y el bienestar social en toda la región.

REFERENCIAS

Consejo Europeo de Relaciones Exteriores (ECFR). (2024). Operación regulación: el fortalecimiento de la gobernanza de la IA en América Latina.

DataGuidance. (2024). International: AI governance developments in Latin America.

Edelman Global Advisory. (2023). EGA-LATAM AI Report.

Redondo, R. (2024). International: AI governance developments in Latin America. DataGuidance.

CAPÍTULO XXII

RETOS DE LA IA ANTE LA VIOLENCIA DIGITAL EN AMBIENTES UNIVERSITARIOS DE AMÉRICA LATINA

WENDY SEBASTIANA HERNÁNDEZ DEL PUERTO[1]
ROSILUZ CEBALLOS POVEDANO[2]

RESUMEN

Este capítulo analiza los desafíos que enfrenta la inteligencia artificial (IA) en la lucha contra la violencia digital en los entornos universitarios de América Latina. A pesar de los avances tecnológicos, la implementación de soluciones basadas en IA se ve obstaculizada por limitaciones técnicas, éticas y culturales. Se abordan diversas manifestaciones de violencia digital, incluyendo el ciberacoso, la difusión no consentida de imágenes íntimas y el hostigamiento en línea, destacando su impacto en estudiantes y docentes. El contexto cultural y social de América Latina, caracterizado por el machismo, las brechas de género y normativas locales insuficientes, agrava estos problemas. Además, se exploran los sesgos inherentes a los algoritmos de IA, las preocupaciones sobre la privacidad y las deficiencias en la legislación vigente. A través de estudios de caso de iniciativas universitarias en México, Argentina y Brasil, se identifican las lecciones aprendidas y las mejores prácticas para mitigar la violencia digital. Finalmente, se proponen recomendaciones para mejorar la precisión de los algoritmos, fomentar la colaboración entre universidades y gobiernos, y promover la educación digital y ética, con el fin de crear entornos académicos más seguros y equitativos.

[1] Profesora-Investigadora, Universidad del Caribe, Cancún, México
[2] Profesora-Investigadora, Universidad del Caribe, Cancún, México

Palabras clave: Inteligencia artificial, violencia digital, universidades, América Latina, ciberacos

1. INTRODUCCIÓN A LA VIOLENCIA DIGITAL EN AMBIENTES UNIVERSITARIOS DE AMÉRICA LATINA

La violencia digital es un fenómeno creciente en los entornos universitarios de América Latina, afectando tanto a estudiantes como a docentes. La digitalización de la vida cotidiana ha facilitado nuevas formas de violencia a través de medios tecnológicos, como el ciberacoso, la difusión no consentida de imágenes íntimas y el hostigamiento en línea. Estas prácticas no solo perjudican a las víctimas en el ámbito digital, sino que tienen impactos tangibles en su bienestar emocional, psicológico y social, así como en su rendimiento académico y profesional.

1.1 Definición de Violencia Digital y su Impacto en Estudiantes y Docentes

La violencia digital se define como cualquier acción realizada mediante tecnologías de la información y la comunicación que cause daño o abuso, ya sea físico, psicológico, moral o patrimonial, a una persona o grupo (Bailey & Steeves, 2015). En los entornos universitarios, la violencia digital afecta principalmente a mujeres, estudiantes de la comunidad LGBTQ+ y otros grupos vulnerables. Los actos más comunes incluyen el ciberacoso, la difusión de imágenes o vídeos íntimos sin consentimiento, y la creación de perfiles falsos con fines de difamación o denigración. Estas agresiones son particularmente devastadoras debido a su capacidad de volverse virales y perdurar en el espacio digital (UN Women, 2015).

El impacto en las víctimas de violencia digital puede ser severo. Un estudio de Jones, Mitchell y Turner (2017) señala que el 31% de los estudiantes universitarios que han sido víctimas de ciberacoso reportan síntomas de ansiedad, depresión y estrés postraumático. Además, el deterioro de la salud mental de las víctimas puede traducirse en una

menor participación en actividades académicas, disminución del rendimiento académico y, en casos extremos, la deserción escolar.

En el caso de los docentes, la violencia digital puede adoptar formas como el acoso a través de redes sociales o el hostigamiento mediante correos electrónicos, lo que genera un entorno laboral tóxico y disminuye la calidad de la enseñanza (Bastos & Oliveira, 2020). Las agresiones a los profesores no solo afectan su bienestar personal, sino también su reputación profesional, especialmente en casos donde se difunden rumores o contenidos denigrantes en línea.

Las universidades enfrentan una serie de riesgos de violencia digital que van más allá del acoso y la difusión no consentida de imágenes íntimas. Un riesgo emergente que ha cobrado relevancia en los últimos años es el uso de *deepfakes,* una tecnología que utiliza IA para crear imágenes, videos o audios falsos hiperrealistas. Esta tecnología ha sido empleada de manera malintencionada para crear contenido pornográfico no consensuado, involucrando a mujeres sin su consentimiento, lo que ha aumentado considerablemente el alcance y la sofisticación de la violencia digital. En los entornos universitarios, los *deepfakes* pueden ser utilizados para difamar o acosar a estudiantes y docentes, causando un daño psicológico y social significativo.

El uso de *deepfakes* en la violencia digital plantea riesgos graves debido a la dificultad de identificar la falsificación en tiempo real. Los videos o audios manipulados pueden viralizarse rápidamente, lo que intensifica el daño a las víctimas antes de que se pueda intervenir. Además, la capacidad de las herramientas de IA para detectar *deepfakes* aún está en desarrollo, y los algoritmos no siempre logran identificar estas manipulaciones con precisión. Esto deja a las víctimas en una posición de vulnerabilidad extrema, y a las instituciones educativas, con herramientas tecnológicas limitadas para frenar esta forma avanzada de abuso (Arteaga, 2024).

Otro riesgo importante es el *doxing,* una práctica de violencia digital que implica la divulgación de información personal, como direcciones, números de teléfono o datos sensibles, con el fin de intimidar, acosar o poner en peligro a una persona. En los entornos universitarios, esta práctica puede tener graves consecuencias para la seguridad física y emocional de las víctimas, ya que su información privada es expuesta

públicamente. El *doxing* se ha convertido en un arma peligrosa para acosadores, que buscan generar miedo y ejercer control sobre sus víctimas mediante la exposición pública.

Además, el robo de identidad digital es otro riesgo creciente en los entornos universitarios. Los atacantes pueden suplantar la identidad de estudiantes o profesores en plataformas digitales o redes sociales, utilizando sus perfiles para difundir contenido falso, acosar a otras personas o dañar la reputación de las víctimas. Esta práctica no solo afecta la vida digital de los individuos, sino que puede tener repercusiones legales, académicas y sociales que son difíciles de revertir.

1.2. Contexto Cultural y Social en América Latina: Machismo, Brechas de Género y Normativas Locales

El contexto sociocultural de América Latina juega un papel decisivo en la prevalencia de la violencia digital. El machismo, las brechas de género y la insuficiencia de normativas efectivas crean un terreno fértil para la perpetuación de esta forma de violencia, especialmente contra las mujeres y las minorías de género. **Rita Segato** (2016) describe cómo el machismo en América Latina no es solo una forma de control sobre las mujeres, sino un sistema de poder que se reproduce y refuerza en diversos ámbitos, incluido el digital.

El machismo digital, como extensión del patriarcado, legitima comportamientos violentos y abusivos hacia las mujeres en los entornos virtuales. Las plataformas de redes sociales, los foros en línea y otros espacios digitales se han convertido en nuevos escenarios donde el acoso, la difamación y la intimidación florecen. Como señala Segato (2016), estas formas de violencia digital no son fenómenos aislados, sino que están profundamente entrelazadas con las desigualdades de género estructurales que caracterizan a la región.

Las brechas de género en el acceso y uso de tecnologías también contribuyen a la vulnerabilidad de las mujeres frente a la violencia digital. La Comisión Económica para América Latina y el Caribe (CEPAL) ha advertido que las mujeres en América Latina tienen menos acceso a tecnologías de la información y la comunicación en comparación con los hombres, lo que limita su capacidad para defenderse

de los ataques en línea o para buscar apoyo (CEPAL, 2020). Además, la falta de formación en alfabetización digital deja a muchas mujeres sin las herramientas necesarias para identificar y mitigar riesgos en el ciberespacio.

A nivel normativo, aunque algunos países de América Latina han avanzado en la creación de leyes para proteger a las mujeres de la violencia digital, como la Ley Olimpia en México, que sanciona la difusión no consentida de contenido íntimo, la legislación aún es insuficiente para abordar todas las formas de violencia digital. No obstante, un avance significativo es la Ley General de Acceso de las Mujeres a una Vida Libre de Violencia (LGAMVLV) en México, reformada en enero de 2024.

La LGAMVLV establece el derecho de las mujeres a vivir libres de violencia, reconociendo la violencia digital como una modalidad de violencia de género que debe ser prevenida y sancionada. Esta ley reformada incluye mecanismos de protección y reparación para las víctimas de violencia digital, además de la creación de políticas públicas que promuevan la educación digital y el acceso a la justicia. La normativa también contempla la responsabilidad de las plataformas digitales en la eliminación de contenido que promueva o perpetúe la violencia contra las mujeres (DOF, 2024). A pesar de estos avances, la implementación de la ley es desigual y las víctimas aún enfrentan numerosos obstáculos para acceder a la justicia, debido a la falta de recursos en los sistemas judiciales locales y a la persistente cultura de la impunidad.

El *Instituto Nacional de Estadística y Geografía* (INEGI), a través del Módulo sobre Ciberacoso (MOCIBA), reveló que el 22.5% de las mujeres en entornos educativos ha sido víctima de algún tipo de violencia digital, lo que pone de relieve la necesidad de un marco legal más sólido y una mayor concienciación institucional (INEGI, 2023). La protección contra la violencia digital en América Latina, y específicamente en los entornos universitarios, requiere no solo un fortalecimiento de las normativas, sino también un cambio cultural que desafíe el machismo y las desigualdades de género estructurales.

2. RETOS TÉCNICOS Y ÉTICOS DE LA IA EN LA DETECCIÓN DE VIOLENCIA DIGITAL

La implementación de tecnologías de inteligencia artificial (IA) para combatir la violencia digital ha emergido como una solución prometedora para identificar y prevenir comportamientos abusivos en plataformas digitales, especialmente en entornos universitarios. Sin embargo, el uso de IA plantea desafíos técnicos y éticos significativos que deben abordarse para asegurar una implementación justa y eficaz. Entre los retos más destacados se encuentran los sesgos inherentes en los algoritmos de IA, la protección de la privacidad de los usuarios, y la falta de un marco regulatorio adecuado en América Latina que contemple las complejidades de la violencia digital.

2.1. Sesgos en los Algoritmos*:* Cómo los Sesgos Culturales y de Género en la IA Pueden Impactar la Detección de Violencia Digital

El problema de los sesgos en los algoritmos de inteligencia artificial es una de las principales barreras técnicas para su efectividad en la detección de violencia digital. Estos sesgos pueden manifestarse en varias formas y tienen raíces profundas en la naturaleza de los datos con los que se entrenan los modelos de IA. Noble (2018), en su obra *Algorithms of Oppression*, argumenta que los algoritmos no son entidades neutrales; están diseñados por humanos y entrenados con datos que reflejan las dinámicas sociales, culturales y de poder de la sociedad. Como resultado, los sesgos culturales y de género presentes en los datos utilizados para entrenar modelos de IA pueden influir en cómo se detectan y procesan los comportamientos abusivos en línea.

Los sesgos de género son particularmente problemáticos en la detección de violencia digital, ya que los datos disponibles a menudo minimizan la experiencia de las mujeres y las minorías de género. Por ejemplo, el lenguaje y los patrones de comportamiento asociados con el acoso dirigido a mujeres, especialmente en América Latina, pueden no ser reconocidos adecuadamente por los algoritmos diseñados con datos predominantemente occidentales o que no reflejan las dinámicas culturales específicas de la región (Buolamwini & Gebru, 2018). Esto es especialmente crítico en el contexto de la violencia de género

digital, donde los ataques suelen ser más sutiles y específicos al contexto sociocultural de las víctimas.

Los sesgos en los algoritmos de IA no solo afectan la precisión en la detección de comportamientos abusivos, sino que también perpetúan las desigualdades existentes. Según Benjamin (2019) en *Race After Technology*, los sistemas tecnológicos a menudo refuerzan las estructuras de poder que históricamente han marginalizado a las mujeres y a las comunidades de color. En América Latina, donde las estructuras patriarcales aún predominan, la falta de representación adecuada en los datos puede llevar a la invisibilización de las formas de violencia que afectan a las mujeres y minorías, agravando la problemática en lugar de contribuir a su solución.

La solución a este problema requiere un enfoque crítico en la creación de d*atasets* más inclusivos y representativos. O'Neil (2016) en *Weapons of Math Destruction* sugiere que es imperativo que los desarrolladores de IA consideren los sesgos presentes en los datos y diseñen sistemas que no perpetúen discriminaciones. Esto implica no solo la recolección de datos diversificados, sino también la colaboración con expertos en violencia de género y derechos humanos para asegurarse de que las tecnologías sean culturalmente sensibles y justas. Sin embargo, los desafíos técnicos relacionados con la creación de estos conjuntos de datos incluyen la falta de acceso a datos confiables y representativos en muchas partes de América Latina, lo que exacerba el problema.

2.2. Protección de la Privacidad: Límites Éticos en el Uso de la IA para Monitorizar la Actividad Digital de Estudiantes y Docentes

El uso de IA para monitorear la actividad digital de estudiantes y docentes en entornos universitarios plantea una serie de dilemas éticos, siendo la protección de la privacidad uno de los más destacados. La vigilancia digital a través de sistemas de IA conlleva la recolección y análisis masivo de datos personales, lo que genera preocupaciones sobre la privacidad y el consentimiento informado. Como argumenta Zuboff (2019) en *The Age of Surveillance Capitalism*, las tecnologías digitales han facilitado nuevas formas de explotación de los datos personales,

lo que pone en riesgo los derechos individuales de los usuarios en un entorno donde la vigilancia es omnipresente.

En el contexto universitario, la implementación de IA para prevenir la violencia digital a menudo implica la monitorización de las comunicaciones en redes sociales, correos electrónicos y otras plataformas digitales utilizadas por los estudiantes y el personal docente. Si bien la intención es proteger a los usuarios de comportamientos abusivos, la recolección de estos datos sin un consentimiento adecuado puede infringir su derecho a la privacidad. Según Floridi (2013) en *The Ethics of Information*, la vigilancia digital debe estar regulada por principios éticos que respeten la autonomía de los individuos y garanticen la transparencia en el manejo de sus datos.

El desafío radica en encontrar un equilibrio entre la necesidad de proteger a los usuarios de la violencia digital y el respeto a su privacidad. La Regulación General de Protección de Datos (GDPR) en Europa ha establecido estándares para la protección de los datos personales, exigiendo el consentimiento explícito de los usuarios y garantizando su derecho a saber cómo se utilizan sus datos. Sin embargo, en América Latina, las normativas de protección de datos son aún limitadas y desiguales, lo que deja a los estudiantes y docentes vulnerables a la explotación de sus datos por parte de las instituciones académicas y otros actores (Bastos & Oliveira, 2020).

Dilemas éticos en la monitorización de la actividad digital:

1. *Consentimiento informado:* Las universidades deben garantizar que los estudiantes y docentes comprendan cómo se recopilan y utilizan sus datos, y deben obtener su consentimiento explícito para ello. La falta de transparencia en este aspecto puede socavar la confianza en las instituciones y generar resistencia al uso de estas tecnologías.
2. *Autonomía y control sobre los datos:* Los usuarios deben tener control sobre sus datos, incluyendo la capacidad de retirar su consentimiento para la monitorización en cualquier momento. En muchos casos, las tecnologías de IA no permiten a los usuarios ejercer este control de manera efectiva.

3. *Equilibrio entre seguridad y privacidad:* Si bien es crucial prevenir la violencia digital en los entornos universitarios, es necesario encontrar un equilibrio entre el uso de la IA para la detección temprana de estos comportamientos y la preservación de la privacidad de los estudiantes y docentes. Las tecnologías de IA no deben invadir la vida privada de los usuarios sin una justificación clara y sin respetar principios éticos fundamentales.

En este sentido, es fundamental que las universidades y los desarrolladores de tecnologías de IA en América Latina adopten políticas de privacidad claras y transparentes. Estas políticas deben garantizar que los datos personales se utilicen de manera ética, limitando la recolección de información sólo a lo necesario para cumplir con los objetivos de prevención de la violencia digital. Además, los usuarios deben tener acceso a mecanismos de control sobre sus datos, así como la capacidad de retirar su consentimiento en cualquier momento. *Shoshana Zuboff* (2019) argumenta que la vigilancia sin consentimiento socava los derechos fundamentales de las personas y crea una relación de poder asimétrica entre los usuarios y las instituciones que controlan sus datos.

2.3. Legislación y Regulación Insuficiente: El Marco Legal en América Latina para el Uso de Tecnologías Avanzadas en la Prevención de la Violencia Digital

A pesar de los avances en la tecnología y la creciente conciencia sobre la violencia digital en América Latina, el marco legal para regular el uso de tecnologías avanzadas, como la inteligencia artificial, sigue siendo insuficiente. Las leyes existentes en muchos países de la región no abordan de manera específica el uso de IA para la prevención de la violencia digital, dejando vacíos legales que pueden ser explotados tanto por los perpetradores de violencia como por las instituciones que implementan estas tecnologías sin una supervisión adecuada.

Una de las normativas más relevantes en México es la Ley General de Acceso de las Mujeres a una Vida Libre de Violencia (LGAMVLV), que establece el derecho de las mujeres a vivir sin violencia, incluidas las modalidades de violencia digital. Esta ley, reformada en enero de 2024, busca garantizar la protección de las mujeres contra la violencia

en todos los ámbitos, incluyendo el digital, e insta a las instituciones a desarrollar políticas de prevención y sanción (DOF, 2024). Sin embargo, la ley no proporciona un marco detallado para la implementación de tecnologías de IA en la detección de violencia digital, lo que deja a las universidades con pocas directrices claras sobre cómo utilizar estas herramientas de manera ética y efectiva.

La Ley Olimpia, también en México, representa un avance importante en la lucha contra la violencia digital, ya que criminaliza la difusión no consentida de imágenes íntimas y establece sanciones penales para los perpetradores. No obstante, esta ley se enfoca principalmente en la responsabilidad penal de los individuos que cometen estos delitos, y no aborda las responsabilidades de las plataformas tecnológicas ni de las instituciones que pueden utilizar IA para detectar y prevenir este tipo de violencia (Cruz, 2021).

En el ámbito internacional, el Convenio de Budapest sobre ciberdelincuencia ofrece un marco global para la regulación de delitos cometidos a través de internet, incluyendo la violencia digital. Sin embargo, la implementación de este convenio en América Latina ha sido desigual, y muchos países de la región aún no cuentan con las estructuras legales o los recursos técnicos necesarios para aplicar sus principios de manera efectiva (Castro, 2018).

Además, la regulación sobre la propiedad intelectual y el uso de tecnologías avanzadas también presenta desafíos en el contexto de la IA. El Instituto Mexicano de la Propiedad Industrial (IMPI) tiene la tarea de regular la protección de la propiedad intelectual en el desarrollo de tecnologías, pero no aborda de manera específica los problemas relacionados con la ética de la IA en el contexto de la violencia digital. Esto crea una brecha en la protección de los derechos de los usuarios, ya que no existen regulaciones claras que obliguen a las empresas tecnológicas a rendir cuentas sobre el uso de IA en la prevención de la violencia digital (IMPI, 2022).

Vacíos legales y falta de regulación:

En muchos países de América Latina, el marco legal para regular el uso de IA en la prevención de la violencia digital es prácticamente inexistente. Esto genera una serie de problemas:

- *Falta de supervisión:* Las universidades y otras instituciones pueden implementar tecnologías de IA sin una supervisión adecuada, lo que aumenta el riesgo de que estas tecnologías se utilicen de manera abusiva o que se vulneren los derechos de privacidad de los usuarios.
- *Vacíos en la protección de datos:* La falta de leyes claras sobre la protección de datos personales deja a los estudiantes y docentes vulnerables a la explotación de su información personal por parte de las instituciones o empresas tecnológicas.
- *Incapacidad de sancionar abusos tecnológicos:* En ausencia de marcos legales que regulen el uso de la IA para la prevención de la violencia digital, es difícil sancionar a las instituciones que utilicen estas tecnologías de manera irresponsable o que infrinjan los derechos de los usuarios.

Es fundamental que los gobiernos de América Latina trabajen en la creación de marcos regulatorios robustos que contemplen tanto la protección de los derechos de los usuarios como la implementación ética de tecnologías avanzadas como la IA. Esto incluye la creación de leyes específicas que regulen el uso de IA en la detección de violencia digital y que exijan a las universidades y otras instituciones la adopción de políticas transparentes y centradas en la protección de los derechos de los estudiantes y docentes.

Para cerrar estas brechas, es necesario que los gobiernos latinoamericanos desarrollen marcos legales robustos digitales y derechos digitales para empoderar a los usuarios y darles las herramientas que regulen tanto el uso de IA en la prevención de la violencia digital como la protección de los derechos de los usuarios. La creación de leyes que contemplen la responsabilidad de las plataformas tecnológicas y las instituciones académicas es esencial para garantizar que el uso de IA se realice de manera ética y justa. Además, los gobiernos deben trabajar en la implementación de políticas de educación mientras son necesarias para defenderse de la violencia en línea.

3. BARRERAS CULTURALES Y SOCIALES EN EL USO DE LA IA CONTRA LA VIOLENCIA DIGITAL

En América Latina, los contextos socioculturales, impregnados de desigualdades de género y sistemas patriarcales, influyen directamente en la percepción y respuesta a la violencia digital. La normalización del acoso en línea, la falta de confianza en los sistemas institucionales y la escasa formación en educación digital y ética son algunos de los principales factores que dificultan el uso eficaz de la IA contra la violencia digital en las universidades.

3.1. Estigma y Normalización del Acoso en Línea dentro de las Universidades

El acoso en línea en entornos universitarios ha sido a menudo trivializado y normalizado debido a estructuras patriarcales profundamente arraigadas que influyen en la percepción de la violencia digital. La Organización de las Naciones Unidas (ONU) define la violencia digital como una forma de violencia de género que se manifiesta a través de medios electrónicos, impactando de manera desproporcionada a las mujeres y a las minorías de género (UN Women, 2015). No obstante, en muchos contextos universitarios de América Latina, esta forma de violencia se percibe como algo natural, una extensión de la violencia que ya ocurre en el espacio físico.

La normalización del acoso digital en las universidades se refleja en comportamientos como el ciberacoso, la difamación y la difusión no consentida de contenido íntimo. Estos actos, que son extremadamente perjudiciales para las víctimas, suelen ser vistos por algunos sectores de la comunidad universitaria como "bromas" o como parte de la interacción en línea, minimizando el daño causado. Rita Segato (2016) destaca que el machismo en América Latina no solo se manifiesta en el control físico sobre las mujeres, sino también en el control simbólico a través de las redes sociales y plataformas digitales, donde los perpetradores se sienten envalentonados por el anonimato o la falta de sanciones.

Esta visión trivial del acoso digital no solo perpetúa la violencia, sino que también crea un ambiente universitario hostil que impacta el bienestar emocional y académico de las víctimas. Según **Jane Bailey (2015)**, la violencia digital es una continuación de las desigualdades de poder existentes y encuentra en las plataformas en línea un espacio propicio para su expansión. El anonimato que brindan estas plataformas, junto con la percepción de impunidad, refuerza la idea de que la violencia digital es "menos grave" que la violencia física, lo que impide que se tomen medidas serias para combatirla.

La normalización del acoso digital también está vinculada a las dinámicas de género en los entornos universitarios, donde los hombres suelen ocupar posiciones de poder tanto en la administración como en las aulas. En muchos casos, la violencia digital dirigida a las mujeres no es vista como una amenaza real para su seguridad o bienestar, y las denuncias de acoso son minimizadas o ignoradas por las autoridades universitarias. Noble (2018), en su libro *Algorithms of Oppression*, señala que las plataformas digitales replican las estructuras de poder que ya existen en la sociedad, lo que significa que las mujeres y las minorías de género enfrentan un mayor riesgo de ser atacadas y, al mismo tiempo, tienen menos probabilidades de recibir apoyo cuando buscan justicia.

El uso de IA para detectar y prevenir la violencia digital enfrenta limitaciones cuando la sociedad en general y las instituciones educativas, en particular, no reconocen la gravedad del problema. Los algoritmos de IA pueden identificar patrones de comportamiento abusivo, pero si estos comportamientos son vistos como "normales" o aceptables, es menos probable que las universidades adopten soluciones tecnológicas para abordarlos. Sarah Banet-Weiser (2018) argumenta que la implementación de tecnologías como la IA debe ir acompañada de un cambio en las normas culturales que legitiman la misoginia y el acoso en línea. Sin este cambio cultural, la IA solo podrá abordar los síntomas superficiales de la violencia digital, sin llegar a la raíz del problema.

Un ejemplo de esta normalización se observa en los casos de difusión no consentida de imágenes íntimas, donde la culpabilización de las víctimas es común. En lugar de centrarse en la responsabilidad del

agresor, la discusión a menudo gira en torno a la conducta de la víctima, perpetuando la idea de que las mujeres deben protegerse del acoso en lugar de que los perpetradores sean castigados. La Ley Olimpia, implementada en México, ha sido un avance significativo al reconocer la gravedad de estos delitos, pero la implementación de la ley sigue siendo un desafío debido a las barreras culturales y la falta de sensibilización entre las autoridades judiciales (Cruz, 2021).

3.2. Reticencia a Denunciar Debido a la Falta de Confianza en los Sistemas Institucionales

La reticencia de las víctimas a denunciar la violencia digital en entornos universitarios es otro obstáculo importante que limita la efectividad de la IA para combatir el acoso en línea. Esta reticencia está estrechamente relacionada con la falta de confianza en los sistemas institucionales y en la capacidad de las universidades para ofrecer protección y apoyo a las víctimas. Catherine Gilligan (2015) señala que muchas víctimas de violencia, ya sea en el ámbito físico o digital, optan por el silencio debido al temor a no ser creídas o a enfrentar represalias, lo que refuerza un ciclo de impunidad que perpetúa la violencia.

En América Latina, la cultura institucional tiende a minimizar la violencia digital y, en muchos casos, a proteger a los perpetradores, especialmente cuando estos ocupan posiciones de poder dentro de las universidades. Galtung (1996), en su teoría de la violencia estructural, argumenta que las instituciones a menudo refuerzan dinámicas de poder que perpetúan la desigualdad y la violencia. En el contexto universitario, esta dinámica se manifiesta cuando las autoridades ignoran las denuncias de violencia digital o cuando las víctimas son sometidas a procesos de revictimización, lo que desincentiva las denuncias.

La falta de confianza en las instituciones es un factor crucial que afecta la disposición de las víctimas a denunciar. Según un estudio de Luchadoras (2017) sobre violencia en línea contra las mujeres en México, muchas víctimas optan por no denunciar debido a la creencia de que no recibirán el apoyo necesario o que sus denuncias serán ignoradas. Además, las víctimas suelen temer represalias por parte de sus agresores, quienes a menudo tienen poder o influencia dentro de

la comunidad universitaria. Esta situación es especialmente problemática en los casos en que los agresores son profesores, administradores o estudiantes con conexiones cercanas a las autoridades universitarias.

La reticencia a denunciar no solo limita la capacidad de las universidades para abordar la violencia digital, sino que también afecta la eficacia de los sistemas de IA diseñados para detectar comportamientos abusivos. La IA se basa en grandes volúmenes de datos para identificar patrones y predecir comportamientos, pero si las víctimas no denuncian los incidentes, los algoritmos no tendrán acceso a la información necesaria para funcionar de manera óptima. Stark y Smith (2019) argumentan que la tecnología solo puede ser efectiva si las víctimas confían en que las instituciones utilizarán esos datos de manera ética y con el fin de brindarles justicia y protección.

Para mejorar la confianza en los sistemas institucionales, es fundamental que las universidades adopten políticas claras y transparentes sobre cómo abordan la violencia digital. Esto incluye la creación de protocolos de denuncia accesibles y confidenciales, así como el establecimiento de mecanismos de apoyo para las víctimas. Según Floridi (2013), la transparencia es un principio ético clave en el uso de tecnologías digitales, y las instituciones deben asegurarse de que las víctimas comprendan cómo se utilizarán sus datos y qué medidas se tomarán para garantizar su seguridad. Solo cuando las víctimas se sientan seguras al denunciar, las tecnologías de IA podrán desplegar su máximo potencial en la prevención y detección de la violencia digital.

3.3. La Importancia de la Educación Digital y Ética para Estudiantes y Profesores

La educación digital y ética es fundamental para garantizar que estudiantes y profesores estén preparados para enfrentar los desafíos de la violencia digital y utilicen las tecnologías de manera responsable. Sin una formación adecuada, las herramientas tecnológicas, incluida la IA, no podrán cumplir su objetivo de prevenir y mitigar la violencia en línea. Henry Jenkins (2009), en *Confronting the Challenges of Participatory Culture*, sostiene que la alfabetización digital no solo implica el

uso técnico de las herramientas, sino también una comprensión crítica de las implicaciones éticas de participar en entornos digitales.

En muchos contextos universitarios, tanto los estudiantes como los profesores carecen de las habilidades necesarias para identificar y abordar la violencia digital. Floridi (2013) en *The Ethics of Information* destaca que la formación en ética digital es esencial para empoderar a las personas a tomar decisiones informadas sobre cómo interactúan en los espacios digitales y cómo protegen su información personal. En América Latina, donde las tasas de adopción tecnológica han crecido rápidamente, la falta de educación digital deja a muchos usuarios vulnerables a los riesgos de la violencia en línea, ya que no saben cómo protegerse ni cómo denunciar los incidentes.

La educación digital también es crucial para combatir las desigualdades de género en el acceso y uso de la tecnología. Banet-Weiser (2018) argumenta que las mujeres y las minorías de género están desproporcionadamente expuestas a la violencia digital, y la falta de educación sobre derechos digitales y seguridad en línea agrava esta situación. En este sentido, las universidades tienen la responsabilidad de proporcionar formación específica que no solo enseñe competencias técnicas, sino que también aborde las dinámicas de poder y género que influyen en la violencia digital.

Un enfoque integral de la educación digital debe incluir la capacitación en temas de derechos digitales, ciberseguridad y el uso ético de las tecnologías de IA. Zuboff (2019) advierte que la falta de comprensión sobre cómo se recopilan y utilizan los datos personales puede llevar a un uso indebido de la tecnología, lo que pone en riesgo los derechos fundamentales de los usuarios. En este sentido, es fundamental que las universidades desarrollen programas de formación que enseñen a los estudiantes y profesores a proteger su privacidad en línea y a utilizar las tecnologías de manera ética.

La educación digital y ética también puede desempeñar un papel clave en el fomento de una cultura de respeto y responsabilidad en los entornos universitarios. Jenkins (2009) sugiere que la educación no solo debe centrarse en la prevención de la violencia, sino también en la creación de una cultura de participación positiva donde los estudiantes y profesores comprendan su papel en la promoción de un

ambiente seguro y respetuoso. Esto implica enseñar a los usuarios a reconocer el impacto de sus acciones en los demás y a utilizar las herramientas digitales de manera constructiva para promover el bienestar colectivo.

4. HERRAMIENTAS DE IA PARA COMBATIR LA VIOLENCIA DIGITAL EN ENTORNOS UNIVERSITARIOS

La inteligencia artificial (IA) ha emergido como una herramienta clave en la lucha contra la violencia digital en los entornos universitarios, proporcionando soluciones innovadoras para la identificación de patrones de acoso, el monitoreo de redes sociales y la detección de lenguaje violento. En un contexto donde la violencia digital ha ganado terreno, afectando de manera desproporcionada a estudiantes y profesores, particularmente a mujeres y grupos vulnerables, las tecnologías de IA ofrecen la capacidad de analizar grandes volúmenes de datos en tiempo real y detectar comportamientos abusivos con una precisión que sería difícil de alcanzar mediante métodos manuales (Bailey & Steeves, 2015). En este apartado, se exploran tres aplicaciones fundamentales de la IA en la prevención y detección de la violencia digital: la identificación de patrones de acoso en redes sociales, la monitorización automatizada de plataformas universitarias, y el uso de tecnologías de reconocimiento de lenguaje y análisis de sentimientos para detectar lenguaje violento.

4.1. Herramientas de IA para la Identificación de Patrones de Acoso y Hostigamiento en Redes Sociales y Plataformas Digitales

Las redes sociales y las plataformas digitales se han convertido en entornos donde el acoso y el hostigamiento ocurren de manera frecuente, afectando a estudiantes y docentes. Este fenómeno, conocido como ciberacoso o violencia digital, ha evolucionado en complejidad, con ataques que pueden ser persistentes y virales. La IA ha demostrado ser una herramienta eficaz para la identificación de estos patrones de acoso a través del análisis de datos masivos que circulan en dichas plataformas. O'Neil (2016) señala que el análisis de grandes volúmenes

de datos a través de algoritmos de IA permite identificar tendencias y comportamientos recurrentes que podrían ser indicativos de acoso, lo que facilita la intervención temprana.

Una de las aplicaciones más comunes de la IA para la identificación de acoso en línea es el uso de algoritmos de detección de patrones, que pueden analizar tanto el contenido de los mensajes como la frecuencia y el tipo de interacciones entre los usuarios. Estos algoritmos pueden identificar comportamientos como la repetición de insultos, la persistencia en el envío de mensajes no deseados, y la publicación de contenido difamatorio, todos los cuales son indicadores de acoso digital. Según Noble (2018), la IA puede ayudar a visibilizar dinámicas de acoso que a menudo pasan desapercibidas en entornos digitales debido a su normalización.

Un ejemplo de estas herramientas es el sistema *DeepText,* desarrollado por Facebook para analizar miles de publicaciones en diferentes idiomas y detectar comportamientos abusivos. Este sistema utiliza técnicas avanzadas de procesamiento del lenguaje natural (PLN) para comprender el contexto en el que se utilizan determinadas palabras o frases, lo que le permite distinguir entre el uso legítimo del lenguaje y los comentarios maliciosos (Levy, 2017). De esta manera, se puede identificar de manera más precisa los casos de acoso, incluso en situaciones en las que el lenguaje puede ser ambiguo o regionalizado, como ocurre en muchos países de América Latina.

No obstante, a pesar de los avances en la identificación de patrones de acoso, los algoritmos de IA enfrentan desafíos importantes, particularmente en relación con los sesgos que pueden estar presentes en los datos de entrenamiento. Buolamwini y Gebru (2018) señalan que los sesgos en los algoritmos pueden llevar a la subrepresentación de ciertos grupos o la sobredetección en otros, lo que podría resultar en la perpetuación de desigualdades. En este sentido, es crucial que los desarrolladores de IA consideren la diversidad cultural y lingüística de los usuarios en las plataformas, especialmente en contextos como América Latina, donde las dinámicas de acoso pueden variar considerablemente en comparación con otras regiones.

Además, el uso de IA para la identificación de patrones de acoso plantea cuestiones éticas sobre la privacidad de los usuarios y el con-

sentimiento informado. Según Floridi (2013), la vigilancia digital a través de IA debe equilibrarse cuidadosamente con los derechos de los individuos a la privacidad y la autodeterminación informativa. Esto es particularmente relevante en el contexto universitario, donde las políticas de monitoreo deben ser transparentes y respetuosas de los derechos de los estudiantes y profesores.

4.2. Monitorización Automatizada de Redes Sociales Universitarias para Prevenir Comportamientos Abusivos

Las universidades, como entornos dinámicos de interacción social y académica, han experimentado un aumento en el uso de redes sociales y plataformas digitales como medios de comunicación entre estudiantes y docentes. Sin embargo, estas plataformas también han sido testigos de un aumento en los comportamientos abusivos, lo que plantea la necesidad de soluciones tecnológicas que permitan la prevención y detección temprana de estos incidentes. La monitorización automatizada de redes sociales a través de IA ha surgido como una estrategia prometedora para identificar y prevenir comportamientos abusivos en tiempo real.

La monitorización automatizada implica el uso de algoritmos de IA para rastrear y analizar interacciones en redes sociales y otras plataformas digitales, con el objetivo de detectar patrones que podrían indicar acoso, hostigamiento o difusión de contenido no consensuado. Estas tecnologías pueden escanear grandes volúmenes de datos en tiempo real, lo que permite una intervención rápida y eficaz por parte de las autoridades universitarias. Stark y Smith (2019) destacan que la monitorización en tiempo real es crucial para prevenir que las situaciones de acoso escalen y para proporcionar apoyo inmediato a las víctimas.

Un ejemplo de la implementación de estas tecnologías es el sistema *Crimson Hexagon,* utilizado en varias instituciones para analizar conversaciones en redes sociales y detectar lenguaje violento o abusivo. Este sistema utiliza técnicas de minería de datos y análisis semántico para identificar comportamientos de riesgo, como el ciberacoso, y generar alertas automáticas que permiten a las universidades intervenir antes de que los incidentes se agraven (Manning et al., 2018). La clave

de estas tecnologías es su capacidad para adaptarse a diferentes contextos y lenguajes, lo que las hace especialmente útiles en entornos multiculturales como los de las universidades latinoamericanas.

Sin embargo, la monitorización automatizada también plantea desafíos éticos, especialmente en relación con la privacidad de los usuarios. Zuboff (2019) advierte que la vigilancia constante a través de herramientas de IA puede llevar a una erosión de los derechos de privacidad y a la creación de un entorno de "*capitalismo de vigilancia*", donde las universidades recopilan y utilizan datos de manera indiscriminada sin el consentimiento adecuado de los estudiantes y profesores. En este sentido, es esencial que las políticas de monitorización automatizada sean transparentes y estén alineadas con principios éticos que respeten los derechos de los usuarios a la privacidad y la autonomía.

La transparencia en la implementación de estas tecnologías es fundamental para generar confianza entre los estudiantes y docentes. Jenkins (2009) subraya que la confianza en las instituciones es clave para que las víctimas se sientan seguras al denunciar incidentes de violencia digital. Si las universidades no garantizan políticas claras sobre cómo se utilizan los datos y cómo se protegerá la privacidad de los usuarios, es probable que las víctimas opten por no denunciar, lo que afectará la eficacia de las tecnologías de monitorización automatizada.

4.3. Tecnologías de Reconocimiento de Lenguaje y Análisis de Sentimientos para Detectar Mensajes Abusivos y Lenguaje Violento

El reconocimiento de lenguaje y el análisis de sentimientos son tecnologías avanzadas de IA que han demostrado ser eficaces en la detección de mensajes abusivos y lenguaje violento en redes sociales y plataformas digitales. Estas tecnologías utilizan algoritmos de procesamiento del lenguaje natural (PLN) para analizar el contenido de los mensajes, identificar palabras o frases clave asociadas con el acoso o la violencia, y evaluar el tono emocional de las interacciones para detectar posibles amenazas (Levy, 2017).

El análisis de sentimientos se refiere a la capacidad de los algoritmos de IA para interpretar y categorizar las emociones expresadas en el texto, como la ira, el miedo, o la tristeza. Esta tecnología es particu-

larmente útil para detectar comportamientos abusivos que no se limitan al uso de palabras explícitamente violentas, sino que pueden estar disfrazados en comentarios sarcásticos o aparentemente inofensivos, pero que en realidad tienen la intención de dañar a la víctima. Según Noble (2018), el análisis de sentimientos puede ayudar a identificar patrones sutiles de acoso que de otro modo podrían pasar desapercibidos, lo que permite una intervención más rápida y eficaz.

Un ejemplo destacado de esta tecnología es el sistema de análisis de sentimientos utilizado por Twitter para identificar y eliminar mensajes abusivos de su plataforma. Este sistema utiliza modelos de aprendizaje profundo para interpretar el contenido de los tuits y determinar si contienen lenguaje que incita al odio o al acoso. Cuando se detectan mensajes sospechosos, el sistema los marca para su revisión por parte de un moderador humano o los elimina automáticamente si se confirma que violan las políticas de la plataforma (Manning et al., 2018).

En el contexto universitario, el uso de tecnologías de reconocimiento de lenguaje y análisis de sentimientos puede ser especialmente útil para detectar casos de violencia digital dirigidos a estudiantes o profesores. Buolamwini y Gebru (2018) destacan que estas tecnologías pueden adaptarse para analizar el lenguaje utilizado en plataformas internas de las universidades, como foros de discusión, correos electrónicos institucionales o redes sociales estudiantiles. Esto permite una intervención más rápida y precisa por parte de las autoridades universitarias y facilita la implementación de políticas de tolerancia cero hacia el acoso.

Sin embargo, estas tecnologías también enfrentan limitaciones, particularmente en lo que respecta a la detección de lenguaje violento en contextos multiculturales. En América Latina, donde las dinámicas lingüísticas son diversas y las expresiones coloquiales pueden variar considerablemente de una región a otra, los algoritmos de reconocimiento de lenguaje deben ser entrenados con datos específicos del contexto local para garantizar una precisión adecuada. Floridi (2013) subraya que la creación de algoritmos culturalmente sensibles es esencial para evitar errores en la detección de lenguaje abusivo, que podrían resultar en la sobredetección de ciertos comportamientos o la subdetección en otros.

Además, el análisis de sentimientos plantea desafíos éticos relacionados con la interpretación subjetiva de las emociones expresadas en el texto. Zuboff (2019) advierte que el uso de IA para analizar emociones puede ser intrusivo y puede llevar a la manipulación de los datos emocionales de los usuarios. Por lo tanto, es crucial que las universidades adopten políticas claras sobre cómo se utilizarán estas tecnologías y cómo se protegerán los derechos de los usuarios a la privacidad y la autonomía emocional.

Las herramientas de inteligencia artificial, incluidas las tecnologías de identificación de patrones de acoso, la monitorización automatizada de redes sociales y el análisis de sentimientos, ofrecen soluciones prometedoras para la detección y prevención de la violencia digital en entornos universitarios. No obstante, su implementación efectiva requiere abordar desafíos técnicos y éticos relacionados con los sesgos algorítmicos, la privacidad de los usuarios y la adaptación cultural de los sistemas de IA. Las universidades deben adoptar un enfoque holístico que combine estas herramientas tecnológicas con políticas institucionales claras y transparentes, garantizando que se respeten los derechos de los usuarios y que se promueva un entorno seguro y libre de violencia para todos los miembros de la comunidad universitaria.

5. CASOS DE ESTUDIO: INICIATIVAS UNIVERSITARIAS QUE UTILIZAN IA EN AMÉRICA LATINA

La implementación de la inteligencia artificial (IA) para combatir la violencia digital en América Latina ha sido gradual y diversa. Cada país ha enfrentado sus propios desafíos y ha adoptado diferentes enfoques para integrar herramientas tecnológicas en la lucha contra el ciberacoso y otras formas de abuso digital. Este capítulo analiza algunos de los estudios de caso más relevantes en la región, con énfasis en las iniciativas implementadas en México, Argentina y Brasil. A través de estos ejemplos, se destacan las lecciones aprendidas que pueden servir como guía para otras instituciones que buscan adoptar soluciones basadas en IA.

5.1. México

En México, diversas universidades han implementado proyectos piloto utilizando IA para detectar y mitigar la violencia digital. La Universidad Autónoma Metropolitana, por ejemplo, ha desarrollado un sistema de monitoreo que analiza las interacciones en sus plataformas digitales para identificar patrones de ciberacoso y lenguaje violento (Cervantes Velázquez & Díaz, 2023). Este sistema utiliza algoritmos de PLN para escanear mensajes y publicaciones en tiempo real, alertando a las autoridades universitarias ante cualquier indicio de comportamiento abusivo.

5.2. Argentina y Brasil

En Argentina, la Universidad de Buenos Aires ha implementado un proyecto similar que combina la IA con intervenciones humanas para abordar el acoso en línea. El sistema no solo detecta comportamientos abusivos, sino que también proporciona recursos y apoyo a las víctimas, facilitando su recuperación y promoviendo un ambiente de respeto (Véliz & Chavez, 2024).

Por otro lado, en Brasil, la Universidad de São Paulo ha desarrollado una herramienta de IA que analiza la participación en foros y redes sociales universitarias para identificar tendencias de violencia digital. Este sistema ha permitido a la institución tomar medidas preventivas y correctivas, mejorando significativamente la seguridad y el bienestar de su comunidad académica (Comisión de Derechos Humanos de la Ciudad de México, 2021).

5.3. Lecciones Aprendidas

Las iniciativas universitarias en México, Argentina y Brasil han demostrado que, aunque la IA puede ser una herramienta poderosa para combatir la violencia digital, su efectividad depende de factores como la calidad de los datos, la colaboración interdisciplinaria y el apoyo institucional. Es fundamental que estas soluciones tecnológicas se complementen con políticas claras, programas de educación digital y

mecanismos de apoyo a las víctimas para lograr un impacto sostenible y significativo (Cerdán-Martínez et al., 2020).

6. RECOMENDACIONES PARA EL FUTURO

A medida que las universidades de América Latina avanzan en la implementación de herramientas de inteligencia artificial (IA) para combatir la violencia digital, es necesario abordar una serie de desafíos relacionados con el desarrollo ético de estas tecnologías, la colaboración entre las instituciones académicas y los gobiernos, y la capacitación de estudiantes y profesores en temas de alfabetización digital y ética. Las siguientes recomendaciones se centran en mejorar la precisión de los algoritmos de IA para hacerlos culturalmente conscientes y libres de sesgos, promover la colaboración institucional para impulsar políticas públicas que fomenten la investigación y el desarrollo, y la integración de la alfabetización digital en los planes de estudio para preparar a la comunidad universitaria ante los retos de la violencia digital.

6.1. Mejorar la Precisión de los Algoritmos: Propuestas para un Desarrollo Ético de IA Culturalmente Consciente y Libre de Sesgos de Género

Uno de los desafíos más importantes en la implementación de IA para detectar y mitigar la violencia digital es la presencia de sesgos en los algoritmos. Como señalan Del Prete y Pantoja (2022), los algoritmos no son neutrales y pueden reflejar las dinámicas de poder y las desigualdades de género presentes en la sociedad. Esto es particularmente preocupante en el contexto latinoamericano, donde las mujeres y las minorías de género enfrentan altos niveles de violencia digital. Los sesgos en los datos utilizados para entrenar los algoritmos pueden llevar a una subrepresentación de ciertos tipos de violencia, lo que a su vez perpetúa la invisibilización de las víctimas.

Para mejorar la precisión de los algoritmos y garantizar que sean culturalmente conscientes y libres de sesgos de género, es esencial que el desarrollo de IA se base en conjuntos de datos representativos y diversos. Buolamwini y Gebru (2018) sugieren que los desarrolladores

deben asegurarse de que los datos utilizados para entrenar los modelos de IA incluyan una representación equilibrada de mujeres, minorías de género y otros grupos vulnerables. En el contexto de América Latina, esto significa que los datos deben reflejar las realidades socioculturales de la región, incluidas las dinámicas de violencia de género y las interacciones sociales que ocurren en las plataformas digitales.

Además de la diversidad de los conjuntos de datos, es importante que los algoritmos sean auditados de manera continua para identificar y corregir posibles sesgos. Noble (2018) sostiene que la auditoría algorítmica es una práctica esencial para garantizar que los sistemas de IA no reproduzcan patrones de discriminación o perpetúen las desigualdades de género. En el caso de las universidades latinoamericanas, la creación de comités interdisciplinarios que incluyan expertos en tecnología, derechos humanos y estudios de género sería una medida crucial para supervisar el desarrollo de IA y asegurar que los algoritmos sean equitativos y respetuosos de las diferencias culturales.

Es fundamental que el desarrollo de IA esté guiado por principios éticos que prioricen el bienestar de las víctimas de violencia digital. Según Floridi (2013), los principios éticos en el desarrollo de tecnologías de información deben garantizar que los sistemas respeten la dignidad y los derechos de los usuarios. Esto significa que los desarrolladores de IA deben adoptar un enfoque centrado en la víctima, donde la tecnología no solo detecte los incidentes de violencia, sino que también ofrezca herramientas que permitan a las víctimas denunciar de manera segura y acceder a servicios de apoyo.

6.2. Colaboración entre Universidades y Gobiernos: La Necesidad de Políticas Públicas que Incentiven la Investigación y el Desarrollo de IA Enfocada en Violencia Digital

La colaboración entre universidades y gobiernos es clave para fomentar la investigación y el desarrollo de tecnologías de IA enfocadas en la prevención y detección de la violencia digital. A pesar de los avances en algunas universidades de América Latina, como las mencionadas en los casos de México, Argentina y Brasil, sigue habiendo una ne-

cesidad urgente de un marco regulatorio y de políticas públicas que apoyen la expansión y el desarrollo ético de estas tecnologías.

En el contexto latinoamericano, los gobiernos deben desempeñar un papel activo en la promoción de la investigación sobre IA y violencia digital, proporcionando financiamiento y recursos a las universidades para desarrollar tecnologías más avanzadas y culturalmente pertinentes. Según **Cruz y Oliveira (2020)**, las políticas públicas deben reconocer la violencia digital como una forma de violencia de género y establecer mecanismos claros para la implementación de tecnologías que permitan su detección temprana y su prevención. Esto incluye la asignación de fondos específicos para proyectos de IA que se centren en la protección de mujeres y otros grupos vulnerables.

La Ley General de Acceso de las Mujeres a una Vida Libre de Violencia en México, reformada en 2024, es un ejemplo de una política que ha comenzado a reconocer el papel de la tecnología en la prevención de la violencia de género. Esta ley incluye disposiciones que instan a las universidades a implementar medidas tecnológicas para monitorear y prevenir la violencia digital. Sin embargo, aún se necesitan más iniciativas que aborden de manera específica el desarrollo de IA para este fin. Del Prete y Pantoja (2022) destacan que una de las limitaciones en América Latina es la falta de infraestructura tecnológica y el acceso limitado a datos representativos, lo que frena el progreso en el desarrollo de soluciones basadas en IA.

Además, la colaboración entre universidades y gobiernos debe ir más allá de la provisión de financiamiento. Jenkins (2009) sugiere que las políticas públicas también deben promover la creación de redes de colaboración entre instituciones académicas, organizaciones de la sociedad civil y entidades gubernamentales para compartir conocimientos y recursos. En América Latina, la creación de consorcios interuniversitarios que trabajen en el desarrollo de IA para la detección de violencia digital podría facilitar la implementación de tecnologías a gran escala y garantizar que las soluciones sean compartidas y adaptadas a los contextos locales de cada país.

Finalmente, es crucial que los gobiernos establezcan marcos regulatorios que aseguren el uso ético de las tecnologías de IA en la prevención de la violencia digital. Floridi (2013) advierte que la regulación es

esencial para garantizar que las tecnologías no sean utilizadas de manera invasiva o que vulneren los derechos de privacidad de los usuarios. Los gobiernos de América Latina deben trabajar en la creación de leyes que protejan la privacidad de los estudiantes y profesores, al mismo tiempo que permiten el monitoreo de comportamientos abusivos en plataformas digitales.

6.3. Capacitación y Concientización: Integrar la Alfabetización Digital y la Ética de la IA en los Planes de Estudio Universitarios para Preparar a Estudiantes y Profesores Ante Estos Desafíos

Para que las tecnologías de IA sean efectivas en la lucha contra la violencia digital, es esencial que los estudiantes y profesores estén capacitados en el uso ético de estas herramientas y comprendan los desafíos asociados con la violencia digital. Sánchez et al. (2021) argumentan que la alfabetización digital debe ir más allá de enseñar a los estudiantes cómo utilizar herramientas tecnológicas; también debe incluir una comprensión crítica de los riesgos y las implicaciones éticas del uso de IA y otras tecnologías digitales.

La integración de la alfabetización digital en los planes de estudio universitarios es una medida esencial para preparar a la comunidad académica ante los retos de la violencia digital. Las universidades deben ofrecer cursos que enseñen a los estudiantes y profesores a reconocer los signos de ciberacoso y otros tipos de violencia digital, así como a utilizar las herramientas tecnológicas de manera ética y responsable. Según Jenkins (2009), la educación sobre las tecnologías digitales debe fomentar una cultura de respeto y responsabilidad en línea, donde los estudiantes comprendan las implicaciones de sus acciones en el entorno digital.

Además de la alfabetización digital, es fundamental que las universidades incluyan en sus planes de estudio la enseñanza de la ética de la IA. Floridi (2013) sostiene que la ética en el desarrollo y uso de la IA debe ser una prioridad en los entornos educativos, especialmente en un contexto donde las tecnologías juegan un papel cada vez más importante en la vida cotidiana. Los estudiantes y profesores deben estar capacitados no solo en el uso de la IA, sino también en los prin-

cipios éticos que deben guiar su desarrollo y aplicación, incluyendo la equidad, la transparencia y la protección de los derechos humanos.

Las campañas de concientización también juegan un papel importante en la lucha contra la violencia digital. Cruz y Oliveira (2020) señalan que las universidades que han implementado con éxito proyectos de IA para la detección de violencia digital han acompañado estas iniciativas con campañas de sensibilización que enseñan a los estudiantes y profesores sobre los riesgos de la violencia en línea y cómo denunciar incidentes de manera segura. Estas campañas deben ser continuas y adaptadas a los cambios en las dinámicas de la violencia digital, que evolucionan junto con el uso de nuevas plataformas y tecnologías.

Finalmente, la capacitación en el uso ético de la IA no debe limitarse a los estudiantes. Los profesores y el personal administrativo también deben recibir formación para comprender cómo funcionan los sistemas de monitoreo basados en IA y cómo proteger los derechos de los estudiantes mientras se utilizan estas herramientas. Esto es crucial para evitar el mal uso de las tecnologías y garantizar que las políticas institucionales se alineen con los principios éticos y los derechos de privacidad.

CONCLUSIÓN

La violencia digital en los entornos universitarios de América Latina se ha convertido en un desafío complejo que requiere soluciones tecnológicas innovadoras, como la implementación de inteligencia artificial (IA). A lo largo de este capítulo, se han explorado las diversas formas de violencia digital, los riesgos que enfrentan las universidades y cómo la IA puede servir como una herramienta crucial en su prevención. Sin embargo, el uso de estas tecnologías no está exento de desafíos y riesgos, tanto técnicos como éticos.

Las universidades enfrentan una gama de riesgos de violencia digital, entre los que se incluyen el ciberacoso, la difusión no consentida de imágenes íntimas y el hostigamiento en línea, que afectan gravemente a estudiantes y docentes. Estas formas de violencia digital se

ven amplificadas por el anonimato y la velocidad de difusión que permiten las plataformas digitales. En este contexto, la IA tiene el potencial de identificar patrones de comportamiento abusivo y ayudar a las universidades a intervenir de manera temprana. Las tecnologías como el procesamiento de lenguaje natural y el análisis de sentimientos pueden detectar lenguaje violento y comportamientos inapropiados, lo que permite una reacción más rápida y eficiente.

Un riesgo emergente que ha ganado relevancia en los últimos años es el uso de *deepfakes*, una tecnología que emplea IA para crear imágenes, videos o audios falsos que parecen reales. En los entornos universitarios, los *deepfakes* se han utilizado para difamar o acosar a estudiantes y docentes, causando daños profundos en su reputación y bienestar emocional. Aunque la IA puede ayudar a detectar este tipo de manipulaciones, la tecnología aún está en desarrollo y no siempre es precisa.

Otro riesgo importante es el *doxing*, la práctica de divulgar información personal sensible con el objetivo de intimidar o acosar. En el entorno universitario, el *doxing* puede poner en peligro la seguridad física y emocional de las víctimas, exacerbando el impacto de la violencia digital. La IA puede ser utilizada para identificar patrones de divulgación de información personal y alertar a las autoridades universitarias antes de que el daño sea irreversible.

A pesar de los beneficios potenciales de la IA, su implementación en las universidades plantea importantes desafíos éticos. Uno de los principales es el riesgo de sesgos algorítmicos. Si los algoritmos de IA no son entrenados con datos representativos de la diversidad cultural y de género de la región, es probable que perpetúen las desigualdades existentes, dejando sin protección a ciertos grupos vulnerables. Esto puede resultar en la subdetección de casos de violencia digital que afectan principalmente a mujeres, estudiantes LGBTQ+ y otras minorías.

Otro desafío crítico es el riesgo de vigilancia masiva. El uso de IA para monitorear grandes volúmenes de interacciones en línea puede generar desconfianza entre los estudiantes y docentes, quienes pueden sentir que su privacidad está siendo vulnerada. Si las universidades no implementan políticas claras de transparencia y protección de

datos, es probable que los usuarios desconfíen del sistema y se resistan a reportar incidentes de violencia digital. La confianza es un elemento clave para que la IA funcione de manera eficaz en la prevención de estos problemas.

Los estudios de caso en México, Argentina y Brasil muestran que la IA puede ser una herramienta poderosa para combatir la violencia digital, pero su éxito depende de la existencia de marcos normativos adecuados y de la colaboración entre universidades, gobiernos y empresas tecnológicas. La Ley Olimpia en México es un ejemplo claro de cómo un marco legal robusto puede apoyar la implementación de soluciones tecnológicas en la lucha contra la violencia digital. Sin embargo, muchos países de la región aún carecen de leyes que aborden de manera integral este tipo de violencia, lo que deja a las universidades con poca protección legal.

Para que la IA sea efectiva en la prevención de la violencia digital, es crucial que las universidades adopten un enfoque centrado en los derechos humanos, que priorice la transparencia, la justicia y la equidad en el uso de estas tecnologías. La educación digital y la alfabetización ética son esenciales para garantizar que tanto los estudiantes como los docentes comprendan los riesgos y las responsabilidades asociados con la IA, y sepan cómo utilizar las herramientas tecnológicas de manera segura y responsable. Además, las universidades deben establecer mecanismos claros de supervisión para asegurar que la IA no perpetúe los sesgos ni vulnere los derechos del estudiantado.

REFERENCIAS

Bailey, J., y Steeves, V. (2015). *Cyberviolence against Women and Girls*. UN Women.

Banet-Weiser, S. (2018). Empowered: Popular Feminism and Popular Misogyny. Duke University Press.

Bastos, D., y Oliveira, M. (2020). Cyber harassment in higher education: Implications for academic professionals. International Journal of Higher Education, 9(3), 57-65.

Benjamin, R. (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Polity.

Buolamwini, J., y Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of the 1st Conference on Fairness, Accountability, and Transparency, 77–91.

Castro, M. (2018). The Budapest Convention and its application in Latin America. Journal of Information Technology & Politics, 15(4), 89-108.

CEPAL. (2020). Brechas de género en el acceso a la tecnología en América Latina. Comisión Económica para América Latina y el Caribe. https://www.cepal.org/es/publicaciones/47077-brechas-tecnologicas-igualdad-genero-contexto-america-latina

Cerdán-Martínez V., García-Guardia M. L. y Padilla-Castillo G. (2020). Alfabetización moral digital para la detección de deepfakes y fakes audiovisuales. CIC. Cuadernos de Información y Comunicación, 25, 165-181. https://doi.org/10.5209/ciyc.68762

Cruz, L. (2021). The legal framework of the Ley Olimpia: A step forward in the fight against digital violence. Journal of Digital Law, 12(2), 35-45.

Del Prete, A., y Pantoja, S. R. (2022). The invisibility of gender-based violence in the social network. Multidisciplinary Journal of Gender Studies, 11(2), 124-143.

DOF. (2024). Ley General de Acceso de las Mujeres a una Vida Libre de Violencia. Diario Oficial de la Federación. https://www.diputados.gob.mx/LeyesBiblio/pdf/LGAMVLV.pdf

Floridi, L. (2013). The Ethics of Information. Oxford University Press.

Galtung, J. (1996). Peace by Peaceful Means: Peace and Conflict, Development and Civilization. Sage Publications.

Gilligan, C. (2015). The Silent Treatment: Silence, Power, and the Denial of Justice. Harvard University Press.

IMPI. (2022). Regulación de la Propiedad Intelectual en Tecnologías Avanzadas. Instituto Mexicano de la Propiedad Industrial. https://www.impi.gob.mx

INEGI. (2023). Módulo sobre ciberacoso. MOCIBA 2023. Principales resultados. Instituto Nacional de Estadística y Geografía. https://www.inegi.org.mx/contenidos/programas/mociba/2023/doc/mociba2023_resultados.pdf

Jenkins, H. (2009). Confronting the Challenges of Participatory Culture: Media Education for the 21st Century. MIT Press.

Jones, L., Mitchell, K., y Turner, H. (2017). Victimization of college students: Findings from the National Violence Against Women Survey. Journal of Interpersonal Violence, 32(20), 3144-3166.

Luchadoras. (2017). La violencia en línea contra las mujeres en México. Informe para la Relatora Especial de Naciones Unidas sobre la Violencia contra las Mujeres. https://luchadoras.mx/wp-content/uploads/2017/12/Informe_ViolenciaEnLineaMexico_InternetEsNuestra.pdf

Manning, C. D., Raghavan, P., y Schütze, H. (2018). Introduction to Information Retrieval. Cambridge University Press.

Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press.

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.

Sánchez, L., Gómez, M., y Rodríguez, A. (2021). CiberConciencia: Un proyecto para la detección de violencia digital en la Universidad Nacional de Córdoba. Revista de Educación Superior y Tecnología, 10(3), 71-88.

Segato, R. (2016). La guerra contra las mujeres. Traficantes de Sueños.

Stark, E., y Smith, C. (2019). The Politics of Blame: Gender, Violence, and Accountability. Oxford University Press.

UN Women. (2015). Cyber Violence Against Women and Girls: A World-Wide Wake-Up Call. UN Women.

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.

CAPÍTULO XXIII

SUSTENTABILIDAD DE MUNICIPIOS METROPOLITANOS MEDIANTE LA APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL A LA GESTIÓN DEL TERRITORIO

JORGE ALBERTO ROSAS CASTRO[1]
GABRIELA CAMACHO DEONICIO[2]

RESUMEN

En México, como en otras repúblicas federalistas, tanto la acción pública como toda la actividad de los ciudadanos sucede en el territorio de los municipios, y, es por esta razón que si la intervención y la gestión territorial se realizan reconociendo su complejidad, variabilidad e incertidumbre con visión sistémica y dinámica, considerando sus dimensiones físicas, sociales, culturales y políticas, se contribuye al desarrollo sustentable; o en sentido contrario, si las acciones públicas y sociales se realizan sin tomar en cuenta estas dimensiones, conducirán al deterioro ambiental provocando desequilibrio y agotamiento de los bienes comunes. De ahí la importancia de utilizar aplicaciones de Inteligencia Artificial a la modelación, simulación y optimización de sistemas complejos, sobre todo en municipios metropolitanos que enfrentan múltiples desafíos debido a su alta densidad poblacional, crecimiento urbano desordenado, y concentración de actividades eco-

[1] Profesor investigador, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México

[2] Coordinadora Editorial, Red Mexicana de Investigadores en Estudios Organizacionales (REMINEO), A. C, Ciudad de México, México.

nómicas, ejerciendo presión sobre los recursos naturales, contaminación ambiental, desigualdad social y fragmentación institucional entre otros, por ello, el desarrollo sustentable del país está ligado a la habilidad de los gobiernos locales, para establecer mecanismos de coordinación con los órdenes de gobierno estatal y federal para utilizar la IA como herramienta a través de sus aplicaciones para el análisis de datos urbanos, simular escenarios futuros, optimizar la toma de decisiones, impulsando el desarrollo sostenible de los territorios de las áreas metropolitanas.

Palabras clave: Desarrollo sustentable, municipios metropolitanos, inteligencia artificial, intervención y gestión territorial.

INTRODUCCIÓN

El capítulo de libro sustentabilidad de municipios metropolitanos mediante la aplicación de la inteligencia artificial a la intervención y gestión del territorio, ha logrado el objetivo definido en esta reflexión teórica, mediante investigación documental utilizando las aplicaciones de IA *ChatGPT, Copilot y Meta*, confirmando así la proposición determinada, al encontrar la información que nos permitió responder a las preguntas de investigación formuladas.

¿Por qué es importante estudiar la relación entre la intervención y la gestión del territorio y el desarrollo sustentable? Porque se considera que con la información recopilada y las reflexiones realizadas, hemos confirmado la proposición formulada, descubriendo que efectivamente al vincular mediante el uso equilibrado de aplicaciones de la IA los principios normativos, dimensiones, categorías y variables propios de la intervención y la gestión del territorio que incluye el ordenamiento territorial, el equilibrio ecológico y la gestión de riesgos, es posible lograr el desarrollo sustentable de municipios metropolitanos, si superan factores como la corrupción, la delincuencia organizada y la falta de interés en el buen vivir.

Una segunda pregunta de investigación formulada es ¿De qué manera el uso de aplicaciones de la IA en la intervención y la gestión del territorio que incluye el ordenamiento territorial, el equilibrio ecológico y la gestión de riesgos, inciden en el desarrollo sustentable de los municipios metropolitanos? Al confrontar los desafíos que enfrentan los municipios metropolitanos con la propia IA utilizando *ChatGPT* para que identificara aplicaciones de la IA que pudieran resolverlos, nos proporcionó un listado indicando para cada desafío los beneficios que puede generar.

La tercera pregunta de investigación ¿Es posible lograr que la IA facilite la convivencia en los espacios y actividades comunes de la población que habita en territorios rurales y urbanos de los municipios metropolitanos encontrando el equilibrio entre el uso de la tecnología y el respeto por la actividad humana?

La reflexión teórica sobre acción pública e intervención y gestión del territorio nos indica que sí. Sin embargo, es fundamental que como en toda aplicación tecnológica, se normalicen los procesos y procedimientos con la estructura orgánica y las interrelaciones entre los actores participantes internos y externos, para evitar automatizar errores y resultados faltos de calidad, la información que alimenta las aplicaciones de la IA, debe ser trabajada con el propósito de digitalizar el funcionamiento de las áreas operativas, de la infraestructura urbana, así como de los diferentes servicios municipales, que exigirá un trabajo de retroalimentación permanente con la aplicación de la IA que se utilice.

El capítulo se integra por tres apartados, en el primero se persigue establecer las bases normativas, teóricas, conceptuales y tecnológicas, que dan sustento a la reflexión sobre la posibilidad de lograr el desarrollo sustentable en municipios metropolitanos, utilizando aplicaciones de IA para la intervención y gestión del territorio. El segundo apartado describe la información más actual disponible sobre las zonas metropolitanas en México, con una clasificación de los municipios que permitió identificar y conceptualizar a los municipios metropolitanos, identificando los desafíos más relevantes que enfrentan para lograr su desarrollo sustentable. Finalmente, en el tercer apartado se logró integrar un apartado con aplicaciones de la IA para enfrentar los desafíos y estar en posibilidad de impulsar el desarrollo sostenible de municipios metropolitanos.

APROXIMACIONES NORMATIVAS, TEÓRICAS, CONCEPTUALES Y TECNOLÓGICAS ORIENTADAS A LA INTERVENCIÓN Y LA GESTIÓN DEL TERRITORIO PARA EL DESARROLLO SUSTENTABLE DE MUNICIPIOS METROPOLITANOS

La reflexión teórica, conceptual y normativa sobre la Sustentabilidad de Municipios Metropolitanos (SMM) mediante la aplicación de la Inteligencia Artificial (IA) en la Intervención y la gestión del territorio (GT), da continuidad al capítulo de libro "El futuro de la Inteligencia Artificial (IA) en la planeación del desarrollo sustentable en municipios y alcaldías en México", incluido en el libro: "*El Futuro de la Inteligencia Artificial*" [3]

Este capítulo de libro aborda la sustentabilidad de los municipios metropolitanos a través de la aplicación de la inteligencia artificial (IA) en la gestión territorial, dado que, al integrar datos y algoritmos avanzados, la IA ofrece soluciones innovadoras para equilibrar el crecimiento económico con la protección del medio ambiente y la mejora de la calidad de vida.

Se propone un enfoque interdisciplinario que integra las dimensiones físicas, sociales, culturales y ecológicas y se exploran las aproximaciones normativas, teóricas, conceptuales y tecnológicas para la intervención y gestión del territorio con el fin de lograr el desarrollo sustentable en municipios metropolitanos. Se redefine el concepto de municipio sustentable en el contexto de zonas metropolitanas, se identifican desafíos y se destaca la importancia de equilibrar el desarrollo económico, la justicia social y la protección ambiental.

La investigación se centra en cómo la inteligencia artificial (IA) puede facilitar la planificación urbana, el equilibrio ecológico, la gestión de riesgos y la participación ciudadana en entornos metropolitanos.

---

[3] Coordinado por el Dr. Oscar Lozano Carrillo y el Dr. Alfredo Garibay Suárez, el Futuro de la Inteligencia Artificial, es un libro electrónico publicado por el sello editorial de la Red Mexicana de Investigadores en Estudios Organizacionales (RE$MINEO), A. C. (julio de 2024) que aborda en 25 capítulos escritos con la ayuda de diferentes aplicaciones de IA, diferentes aspectos del futuro de la inteligencia artificial, destacando sus amenazas y beneficios.

En el capítulo de libro "*El futuro de la Inteligencia Artificial (IA) en la planeación del desarrollo sustentable en municipios y alcaldías en México" se define que "un municipio o alcaldía sustentable es aquel cuyo gobierno en colaboración con los ciudadanos, logra equilibrar el desarrollo económico, la justicia social y la protección del medio ambiente, fomentando la actividad económica sin agotar los recursos naturales ni comprometer el bienestar de las generaciones futuras*".

Ahora, en este capítulo de libro se redefine este concepto para contextualizarlo en los municipios metropolitanos, que son aquellos que "*forman parte de una zona metropolitana, entendida como un área urbana que engloba una ciudad principal y varias ciudades vecinas que funcionan como ciudades dormitorio, industriales, comerciales y serviciales*" (SEDATU, CONAPO e INEGI).

A la luz de estas premisas las preguntas de investigación que persigue responder esta investigación documental son ¿Por qué es importante estudiar la relación entre la intervención y la gestión del territorio y el desarrollo sustentable? ¿De qué manera la aplicación de la IA en la intervención y la gestión del territorio que incluye el ordenamiento territorial, el equilibrio ecológico y la gestión de riesgos, inciden en el desarrollo sustentable de los municipios metropolitanos? ¿Es posible lograr que la IA facilite la convivencia en los espacios y actividades comunes de la población que habita en territorios rurales y urbanos de los municipios metropolitanos encontrando el equilibrio entre el uso de la tecnología y el respeto por la actividad humana?

Preguntas que nos llevan a definir como objetivo de investigación el siguiente: Identificar los principios normativos, dimensiones, categorías y variables que la intervención y la gestión del territorio debe vincular mediante la aplicación equilibrada de la IA para determinar intervenciones que posibiliten el desarrollo sustentable de municipios metropolitanos.

De ahí se deriva la siguiente proposición: Vincular mediante la aplicación equilibrada de la IA principios normativos, dimensiones, categorías y variables propios de la intervención y la gestión del territorio que incluye el ordenamiento territorial, el equilibrio ecológico y la gestión de riesgos para lograr el desarrollo sustentable de municipios metropolitanos.

INTERVENCIÓN Y GESTIÓN DEL TERRITORIO

Gonzáles, L. define a la gestión del territorio como "un sistema organizado que tiene como fin intervenir el territorio, respetando su dinámica y características propias". Las características de la intervención que deben trabajarse en varios niveles desde el global hasta el operativo, en los territorios son:

- Reconocer su complejidad, variabilidad e incertidumbre
- Realizar prácticas jerárquicas y temporales
- Visión sistémica y dinámica

Continúa describiendo González, L.; que "la gestión territorial tiene como objetivos:

1) Entender y adaptarse a la complejidad del territorio.
2) Establecer una planificación flexible y dinámica.
3) Intervenir respetando las relaciones sociales y geográficas del territorio, teniendo en cuenta los impactos a largo plazo".

Así González, L. determina que las dimensiones de la intervención y la gestión del territorio son las físicas, sociales, culturales y políticas. Dimensiones que coinciden con las que corresponden al desarrollo sustentable, razón por la que existe una relación causal entre la intervención y la gestión del territorio y el desarrollo sustentable, por lo que requiere una acción planificada sobre la complejidad de las características dinámicas y ambientales del sistema territorial sobre el que se actúa, con el propósito de preservar los bienes comunes y la cultura en la que se sustentan las relaciones sociales.

De ahí que, siguiendo a González, L. y agregando tecnología mediante la IA, la intervención y la gestión del territorio, es una acción organizada y un proceso sistemático, con el "objetivo de equilibrar las necesidades humanas y la tecnología con la capacidad del territorio para sostenerlas, manteniendo la integridad ecológica". Se debe procurar "una intervención consciente capaz de adaptarse a la realidad

cambiante del territorio respetando su dinámica interna, así como las interacciones sociales y ambientales".

La relación de causalidad entre intervención y gestión del territorio y desarrollo sustentable es la razón que da pertinencia a su estudio en municipios metropolitanos, toda vez que su densidad poblacional, conurbación y extensión territorial, permiten incidir en una mayor cantidad de factores para impulsar el desarrollo sustentable.

TEORÍAS SOBRE EL DESARROLLO TERRITORIAL SOSTENIBLE LA GESTIÓN MUNICIPAL Y LA CLASIFICACIÓN DE LOS MUNICIPIOS

El municipio está asociado al territorio, ya que, si éste no podría existir, afirmamos que el municipio es el espacio territorial en el que se fundan las entidades federativas y, por tanto, la federación. Así lo determina el artículo 115 de la Constitución Política de los Estados Unidos Mexicanos "Los estados adoptarán, para su régimen interior, la forma de gobierno republicano, representativo, democrático, laico y popular, teniendo como base de su división territorial y de su organización política y administrativa, el municipio libre..." En tal sentido, hemos venido afirmando que, para lograr el desarrollo sustentable de los municipios, es indispensable una adecuada intervención y gestión del territorio, adquiriendo relevancia y mayor incorporación de teorías, principios y técnicas de la gestión pública cuando se trata de municipios metropolitanos.

Desde esta perspectiva la intervención y gestión del territorio, refieren dos elementos consustanciales, espacio y territorio. "El espacio se refiere a la dimensión física, mientras que el territorio abarca no solo la superficie geográfica, sino también las relaciones sociales, culturales y económicas que se desarrollan en ese espacio" (Gonzáles, L).

La territorialidad a cómo las personas se apropian y utilizan el territorio, por ello, el amplio marco normativo analizado en el apartado anterior, parte de que existen límites geográficos (dimensión física) y las interacciones entre áreas vecinas (dimensión relacional). De ahí que la gestión del territorio requiere comprender la complejidad de los sistemas que operan en él, y, para lograrlo, es indispensable la mo-

delación de sistemas complejos como herramienta fundamental para analizar interacciones y prever resultados, en la búsqueda permanente del equilibrio entre el desarrollo humano y la conservación del entorno natural.

Citados por González, L Montañez y Delgado (1998), establecen siete consideraciones para orientar el análisis del territorio:

1) Como escenario de las relaciones sociales
2) Como espacio de poder, de gestión e interacción entre el Estado, los individuos, organizaciones sociales, empresas locales, nacionales y multinacionales.
3) Conocimiento de la construcción social de los procesos de producción y distribución.
4) Actividad diferenciada de los diferentes actores de acuerdo a sus diferentes capacidades reales y potenciales para crear, recrear y apropiarse del territorio.
5) Sobreposición de distintas territorialidades locales, regionales, nacionales y mundiales, con intereses distintos, con percepciones, valoraciones y actitudes territoriales diferentes, que generan relaciones de complementación, de cooperación y de conflicto.
6) Su naturaleza es el desequilibrio, no es fijo, sino móvil y mutable. La realidad geosocial es cambiante y requiere permanentemente nuevas formas de organización territorial.
7) El sentido de pertenencia e identidad, el de conciencia regional, al igual que el ejercicio de la ciudadanía y de acción ciudadana, solo adquieren existencia real a partir de su expresión de territorialidad. En un mismo espacio se sobreponen múltiples territorialidades y lealtades.

Ante estas situaciones complejas que se viven en el territorio, la intervención y la gestión pública sobre éste, que transita a partir de las y los ciudadanos que desempeñan su vida y trabajo en dicho territorio que expresa además territorialidad, es indispensable aplicar además de tecnología mediante la IA, teorías como las siguientes (**Ver Tabla 1**):

Tabla 1. Gestión municipal y clasificación de los municipios.

Teoría	*Contenido*	*Autores*
Gobernanza local	La gestión pública ha pasado de ser jerárquica a un sistema de redes y colaboración	Gerry Stoker
	La colaboración entre diferentes actores	Mark Bevir
Descentralización y autonomía municipal	La descentralización y el impacto en la gobernabilidad y la eficiencia local.	Jean-Claude Thoenig
	Sociedad en red	Manuel Castells
Desarrollo Local	El concepto de desarrollo local y su relación con la economía y las políticas públicas.	Pierre-André Chiappori
	Las dinámicas del desarrollo local y su impacto en la economía regional.	Amin Ash y Nigel Thrift
Administración pública relacional	La teoría de redes y la colaboración en el ámbito de la gestión pública, destacando cómo las relaciones interorganizacionales mejoran la efectividad en la administración pública	Robert Agranoff y Michael McGuire
Teoría de la Acción colectiva	La acción colectiva en la gestión de los recursos comunes y cómo las comunidades locales pueden auto-gestionarse sin depender siempre del estado o el mercado	Elinor Ostrom
La descentralización y cómo los gobiernos locales pueden mejorar su gestión	Mediante la implementación de prácticas administrativas más modernas y eficientes, la administración pública debe adaptarse a las necesidades cambiantes de los municipios y al desarrollo local	Omar Guerrero
La acción pública y su papel en el desarrollo local.	• Los municipios pueden desarrollar capacidades de gestión más eficaces mediante la innovación en políticas públicas. • clasificación de los municipios mexicanos basada en el tipo de gestión que desarrollan, identificando tres tipos de itinerarios de acción pública: tradicional, innovador y transformador. Su análisis muestra cómo la interacción entre gobiernos y la sociedad civil puede influir en la efectividad de las políticas municipales, destacando la importancia de una cooperación activa para lograr un desarrollo sostenido	Enrique Cabrero

Fuente: Elaboración con base en *ChatGPT* (OpenAI, 2024).

Ámbito normativo

En este contexto dos leyes secundarias de los artículos 25, sexto párrafo, 26, 27, tercer párrafo, 73, fracción XXIX-G, 115, 122 y 124 de la Constitución Política de los Estados Unidos Mexicanos, regulan entre otros importantes aspectos, el desarrollo urbano y territorial, así como los asentamientos humanos y la protección del medio ambiente para que todas las personas puedan vivir el derecho en un entorno adecuado para su desarrollo, la Ley General de Asentamientos Humanos, Ordenamiento Territorial y Desarrollo Urbano (LGAHOTDU) y la Ley General del equilibrio Ecológico y Protección al Ambiente (LGEPA).

LEY GENERAL DE ASENTAMIENTOS HUMANOS, ORDENAMIENTO TERRITORIAL Y DESARROLLO URBANO

Las disposiciones normativas de la LGAHOTDU establecen los siguientes objetivos e instrumentos para lograrlos (**Ver Tabla 1**):

Tabla 2. Objetivos e instrumentos de LGAHOTDU

Objetivos	*Instrumentos*
Ordenamiento del territorio y asentamientos humanos	• Planeación urbana y ordenamiento territorial
Respeto a los derechos humanos	• Regulación de usos del suelo
Criterios para la regulación de la propiedad	
Concurrencia y coordinación	• Instrumentos de gestión metropolitana
	• Instrumentos financieros
Participación ciudadana	• Consultas públicas, audiencias y consejos ciudadanos

Fuente: LGAHOTDU

Por su parte, el artículo 4, de la referida LGAHOTDU, establece los siguientes principios en materia de planeación, regulación y gestión de los asentamientos humanos, Centros de Población y la ordenación territorial.

1) Derecho a la ciudad
2) Equidad e inclusión
3) Derecho a la propiedad urbana
4) Coherencia y racionalidad
5) Participación democrática y transparencia
6) Productividad y eficiencia
7) Protección y progresividad del espacio público
8) Resiliencia, seguridad urbana y riesgos
9) Sustentabilidad ambiental
10) Accesibilidad universal y movilidad

De principal interés para los municipios es el Título Segundo de la LAHOTDU es la concurrencia entre los órdenes de gobierno, coordinación y concertación, mediante diferentes mecanismos.

En este orden de ideas el Capítulo Cuarto establece 26 atribuciones que corresponden a los municipios; de entre las que destacan:

1) Formular los programas municipales de desarrollo urbano de centros de población.
2) Regular, controlar y vigilar: las reservas, usos del suelo y destinos de áreas y predios, y determinar las zonas de alto riesgo de los centros de población.
3) Formular, aprobar y administrar la zonificación de los centros de población.
4) Promover y ejecutar acciones, inversiones y servicios públicos para la conservación, mejoramiento y crecimiento de los Centros de Población.

5) Proponer la fundación o la desaparición de Centros de Población.

6) Impulsar y promover la conformación de institutos metropolitanos de planeación junto con los municipios que conforman una zona metropolitana determinada.

7) Celebrar convenios de asociación con otros municipios para fortalecer sus procesos de planeación urbana, así como para la programación, financiamiento y ejecución de acciones, obras y prestación de servicios comunes, así como para crear y mantener un instituto municipal de planeación, cuando los municipios se encuentren por debajo de un rango de población menor a cien ml habitantes.

8) Celebrar convenios de coordinación y concertación que apoyen los objetivos y prioridades previstos en los planes o programas de desarrollo urbano.

Los capítulos primero y segundo del Título Tercero de la LAHOTDU, establecen la conformación del Consejo Nacional de Ordenamiento Territorial y Desarrollo Urbano, así como los Consejos Locales y Municipales de Ordenamiento Territorial, Desarrollo Urbano y Metropolitano. Sus funciones son asegurar la consulta, opinión y deliberación de políticas de ordenamiento territorial y planeación del desarrollo urbano y desarrollo metropolitano.

Para el cumplimiento de las funciones de los Consejos antes mencionados, el Título Cuarto de la LGAHOTDU, establece en el Capítulo Primero que habrá un Sistema General de Planeación Territorial, que norma:

1) La estrategia nacional de ordenamiento territorial,

2) Los programas estatales de ordenamiento territorial y desarrollo urbano.

3) Los programas de zonas metropolitanas o conurbaciones.

4) Los programas municipales de desarrollo urbano.

En este Título Cuarto en el capítulo Quinto, se determinan los alcances de los Programas Metropolitanos y de Zonas Conurbadas para la adecuada coordinación de los municipios metropolitanos con las autoridades federales y estatales, y para lograrlo, el artículo 34, determina que son de interés metropolitano:

1) La planeación del ordenamiento del territorio y los asentamientos humanos.
2) La infraestructura vial, tránsito, transporte y la movilidad.
3) El suelo y las reservas territoriales.
4) La densificación, consolidación urbana y uso eficiente del territorio, con espacios públicos seguros y de calidad como eje articulador.
5) Las políticas habitacionales y las relativas al equipamiento regional y metropolitano.
6) La localización de espacios para el desarrollo industrial de carácter metropolitano.
7) La gestión integral del agua y los recursos hidráulicos, incluyendo el agua potable, el drenaje, saneamiento, tratamiento de aguas residuales, recuperación de cuencas hidrográficas y aprovechamiento de aguas pluviales.
8) La preservación y restauración del equilibrio ecológico, el aprovechamiento sustentable de los recursos naturales y la protección del ambiente, incluyendo la calidad del aire y la protección de la atmósfera.
9) La gestión integral de residuos sólidos municipales, especialmente los industriales y peligrosos.
10) La prevención, mitigación y resiliencia ante los riesgos y los efectos del cambio climático.
11) La infraestructura y equipamiento de carácter estratégico y de seguridad.
12) La accesibilidad universal y la movilidad.
13) La seguridad pública.

14) Otras acciones que, a propuesta de la comisión de ordenamiento, se establezcan o declaren por las autoridades competentes.

Un factor fundamental para el ordenamiento territorial y el desarrollo urbano es la gobernanza metropolitana y en el Capítulo Sexto de este mismo Título, el artículo 36 determina que, para lograr una eficaz gobernanza metropolitana, se establecerán los mecanismos y los instrumentos de carácter obligatorio que aseguren la acción coordinada institucional de los tres órdenes de gobierno y la participación de la sociedad.

Definiéndose las siguientes instancias para lograrla:

1) Comisión de ordenamiento metropolitano o de conurbación, integrada por la Federación, las entidades federativas, los municipios y las demarcaciones territoriales de la zona de que se trate.
2) Un consejo consultivo de desarrollo metropolitano que promoverá los procesos de consulta pública e institucional, integrado con perspectiva de género por representantes de los tres órdenes de gobierno y representantes de agrupaciones legalmente constituidas, colegios de profesionistas, instituciones académicas y expertos en la materia.
3) Los mecanismos de carácter técnico a cargo de las entidades federativas y municipios, bajo la figura que corresponda, así como los institutos metropolitanos de planeación, sesionarán permanentemente.
4) Las instancias que permitan la prestación de servicios comunes.
5) Los mecanismos y fuentes de financiamiento de las acciones metropolitanas contemplando entre otros, el fondo metropolitano.

En total la LGAHOTDU contiene 119 artículos, que precisan las competencias de cada uno de los tres órdenes de gobierno en las responsabilidades en las tareas que determinan el desarrollo sostenible de los territorios municipales, si la vigilancia en los asentamientos hu-

manos para que obedezcan el ordenamiento territorial y el programa de desarrollo urbano de cada municipio el desarrollo sostenible del país no se logrará.

LEY GENERAL DEL EQUILIBRIO ECOLÓGICO Y LA PROTECCIÓN AL MEDIO AMBIENTE

La LGEEPMA persigue los objetivos que a continuación se mencionan relacionados con los instrumentos para lograrlos (**Ver Tabla 3**):

Tabla 3. Objetivos e Instrumentos de la LGEEPMA

Objetivos	*Instrumentos*
Garantizar el derecho a un ambiente sano	• Ordenamiento ecológico del Territorio • Protección de Áreas Naturales • Evaluación de impacto y riesgo ambiental • Investigación y educación ambiental • Monitoreo ambiental
Definir principios de política ambiental	
Preservación, restauración y mejoramiento del ambiente	
Protección de la biodiversidad y áreas naturales protegidas	
Aprovechamiento sustentable de recursos naturales	
Prevención y control de la contaminación	
Participación ciudadana en la preservación ambiental	
Coordinación entre autoridades y sectores sociales	

Fuente: LGEEPMA

Al respecto, tanto el gobierno federal, como los gobiernos de los estados y los municipales tienen facultades para aplicar la LEEPMA, por el interés del capítulo se describen a continuación éstas últimas consideradas en el artículo 8°.

1) Formulación, conducción y evaluación de la política ambiental municipal.

2) La aplicación de los instrumentos de política ambiental previstos en las leyes locales en la materia y la preservación y restauración del equilibrio ecológico y la protección al ambiente en bienes y zonas de jurisdicción municipal, en las materias que no están atribuidas a la Federación a los Estados.

3) La aplicación de las disposiciones jurídicas en materia de prevención y control de la contaminación atmosférica generada por fuentes fijas que funcionan como establecimientos mercantiles o de servicios, así como de emisiones de contaminantes a la atmósfera provenientes de fuentes móviles que no sean consideradas de jurisdicción federal, con la participación que de acuerdo con la legislación estatal corresponda al gobierno del estado.

4) La aplicación de las disposiciones jurídicas relativas a la prevención y control de los efectos sobre el ambiente ocasionados por la generación, transporte almacenamiento, manejo, tratamiento y disposición final de los residuos sólidos e industrias que no estén considerados como peligrosos, de conformidad con lo dispuesto por el artículo 137 de la presente Ley.

5) La creación y administración de zonas de preservación ecológica de los centros de población...

6) La aplicación de las disposiciones jurídicas relativas a la prevención y control de la contaminación por ruido, vibraciones, energía térmica, luz intrusa, radiaciones electromagnéticas y olores perjudiciales para el equilibrio ecológico y el ambiente, proveniente de fuentes fijas...y las que resulten aplicables a las fuentes móviles...

7) La aplicación de las disposiciones jurídicas en materia de prevención y control de la contaminación de las aguas que se descarguen en los sistemas de drenaje y alcantarillado de los centros de población...

8) La formulación y expedición de los programas de ordenamiento ecológico local del territorio a que se refiere el artículo 20 Bis 4 de esta Ley...

9) La preservación y restauración del equilibrio ecológico y la protección al ambiente en los centros de población, en relación con los efectos derivados de los servicios de alcantarillado, limpia, mercados, centrales de abasto, panteones, rastros, tránsito y transporte locales...

10) La participación en la atención de los asuntos que afecten el equilibrio ecológico de dos o más municipios y que generen efectos ambientales en su circunscripción territorial.

11) La participación en emergencias y contingencias ambientales conforme a las políticas y programas de protección civil que al efecto se establezcan.

12) La vigilancia del cumplimiento de las normas oficiales mexicanas expedidas por la Federación, en las materias y supuestos a que se refieren las fracciones III, IV, VI y VII de este artículo.

13) La formulación y conducción de la política municipal de información y difusión en materia ambiental.

14) La participación en la evaluación del impacto ambiental de obras o actividades de competencia estatal, cuando las mismas se realicen en el ámbito de su circunscripción territorial.

15) La formulación, ejecución y evaluación del programa municipal de protección al ambiente.

16) La formulación y ejecución de acciones de mitigación y adaptación al cambio climático.

17) La atención de los demás asuntos que en materia de preservación del equilibrio ecológico y protección al ambiente les conceda esta Ley u otros ordenamientos en concordancia con ella y que no estén otorgados expresamente a la Federación o a los Estados.

La LGEEPA se compone de 204 artículos y precisa los principios y atribuciones que en esta materia deben cumplir el gobierno federal, las entidades federativas y los municipios; y para la adecuada intervención y gestión del territorio, está complementada por un amplio y complejo marco legal que busca garantizar un equilibrio entre el desarrollo económico y la protección del ambiente, compuesto por las siguientes leyes:

- Ley General de Desarrollo Forestal Sustentable: Regula el aprovechamiento sustentable de los bosques y selvas, y promueve su conservación.
- Ley General de Vida Silvestre: Establece las disposiciones para la conservación y aprovechamiento sustentable de la vida silvestre.
- Ley General de Aguas Nacionales: Regula el aprovechamiento de las aguas nacionales y establece los mecanismos para su conservación.
- Ley General de Cambio Climático: Establece las bases para la mitigación y adaptación al cambio climático.
- Leyes estatales, así como los reglamentos y programas municipales: Cada entidad federativa y municipio cuenta con leyes locales que complementan la legislación federal en materia ambiental y territorial.

GESTIÓN DE RIESGOS

La gestión de riesgos es un aspecto fundamental de la intervención y gestión del territorio, especialmente en un país como México, altamente vulnerable a desastres naturales. Es un eje transversal que se aborda en diversas disposiciones legales, por lo que, no existe una ley específica para la gestión del riesgo, si no que, diversos ordenamientos legales contienen disposiciones relacionadas con este tema.

Algunos ejemplos de disposiciones legales relacionadas con la gestión de riesgos:

- Planes de desarrollo: Los planes de desarrollo de los diferentes niveles de gobierno deben incluir estrategias para la prevención y atención de desastres.
- Programas de ordenamiento ecológico del territorio: Estos programas deben considerar los riesgos naturales y antrópicos, y establecer medidas para reducir su vulnerabilidad.
- Normas de construcción: Las normas de construcción deben considerar los riesgos sísmicos, hidrometeorológicos y otros, y establecer medidas para garantizar la seguridad de las edificaciones.
- Atlas de riesgo municipales

ZONAS METROPOLITANAS DE MÉXICO

México está organizado como una república federal, oficialmente conocida como los Estados Unidos Mexicanos. El país se divide en 32 entidades federativas, que incluyen 31 estados y la Ciudad de México. Cada estado tiene su propia capital y se subdivide en municipios.

El Instituto Nacional de Estadística y Geografía [*INEGI*] (2020), reporta un total de 2,469 municipios en el territorio mexicano. Asimismo, identifica 48 zonas metropolitanas conformadas por 345 municipios y habitadas por aproximadamente 67.6 millones de personas.

En México, el término zona metropolitana se utiliza para designar áreas urbanas que no solo son grandes en términos de población, sino que también presentan una compleja interacción de actividades económicas, sociales y culturales. Este concepto ha sido adoptado por el gobierno para fines de gestión y asignación de recursos (Valencia y otros., 2017).

Desde el 2000, cada cinco años, se realiza una nueva delimitación de las zonas metropolitanas. La última publicación "*Metrópolis de México 2020*", fue el resultado del trabajo colaborativo entre el Consejo Nacional de Población (*CONAPO*), el Instituto Nacional de Estadística y Geografía *(INEGI)* y la Secretaría de Desarrollo Agrario, Territorial y

Urbano (*SEDATU*). En lo que se refiere al acotamiento de los límites físico espaciales de las metrópolis se consideraron las siguientes características sociales, económicas y urbanas: Población total, integración física y funcional con municipios contiguos, distancia a las localidades urbanas principales, población ocupada según municipio de residencia y densidad media urbana; resultando como zonas metropolitanas más pobladas de México, las siguientes:

- **Zona Metropolitana de la Ciudad de México**: 21,436,911 habitantes
- **Zona Metropolitana de Monterrey**: 5,322,117 habitantes
- **Zona Metropolitana de Guadalajara**: 5,110,617 habitantes
- **Zona Metropolitana de Puebla-Tlaxcala**: 2,776,893 habitantes
- **Zona Metropolitana de Toluca**: 2,353,924 habitantes
- **Zona Metropolitana de Tijuana**: 2,049,413 habitantes
- **Zona Metropolitana de León**: 1,935,928 habitantes
- **Zona Metropolitana de Querétaro**: 1,530,820 habitantes
- **Zona Metropolitana de La Laguna**: 1,375,248 habitantes
- **Zona Metropolitana de Mérida**: 1,324,771 habitantes

CLASIFICACIÓN DE LOS MUNICIPIOS EN MÉXICO

La clasificación de los municipios en México es un tema complejo y dinámico, se realiza a través de un análisis multidimensional de indicadores socioeconómicos, recopilados y procesados por diversas instituciones como: el *INEGI, CONEVAL y SEDATU*. Estos indicadores permiten evaluar el nivel de desarrollo humano, las desigualdades sociales y las oportunidades de desarrollo local. En la Tabla 1 se presentan las diferentes clasificaciones de los municipios en México y los indicadores clave (la selección de indicadores puede variar según el estudio específico y las fuentes de datos disponibles).

Tabla 4. Comparativa de los diferentes tipos de clasificación de los municipios y sus características principales (algunos ejemplos)

Tipo de Clasificación	*Tipo de Municipio*	*Características Principales*	*Indicadores Clave*	*Ejemplos*
Por tamaño y población	Metropolitano	Alta concentración poblacional, centro económico y político, gran diversidad de actividades económicas, infraestructura desarrollada.	Población total, densidad poblacional, tasa de urbanización, número de localidades urbanas	Ciudad de México, Guadalajara, Monterrey
	Urbano	Población considerable, mayor densidad poblacional que los rurales, crecimiento urbano acelerado, servicios públicos básicos.	Población total, densidad poblacional, tasa de crecimiento poblacional, porcentaje de población urbana	Puebla, Querétaro, León
	Rural	Predominio de actividades agrícolas y ganaderas, baja densidad poblacional, menor acceso a servicios, infraestructura básica.	Población total, densidad poblacional, porcentaje de población rural, porcentaje de población ocupada en el sector primario	Comunidades indígenas de Chiapas, zonas rurales de Oaxaca, municipios de la Sierra Madre Occidental
	Marginado	Altos índices de pobreza, marginación social, carencia de servicios básicos, baja calidad de vida.	Población, densidad, crecimiento, estructura por edad y sexo, composición étnica, IDHM, pobreza, marginación, empleo, vivienda, servicios básicos	Zonas rurales de Guerrero, municipios de la región Mixteca de Oaxaca

Tipo de Clasificación	*Tipo de Municipio*	*Características Principales*	*Indicadores Clave*	*Ejemplos*
Por nivel de desarrollo	Alto Desarrollo	Infraestructura moderna, servicios públicos eficientes, economía diversificada, alta calidad de vida.	PIB per cápita, IDHM, tasa de analfabetismo, tasa de escolaridad, acceso a servicios básicos, esperanza de vida al nacer	Municipios de la Zona Metropolitana de Monterrey, algunas zonas de Jalisco
	Mediano Desarrollo	Desarrollo intermedio en infraestructura y servicios, crecimiento económico moderado.	PIB per cápita, IDHM, tasa de ocupación, índice de marginación	Muchos municipios de estados como Guanajuato, Querétaro
	Bajo Desarrollo	Carencias en infraestructura, servicios básicos limitados, economía dependiente de actividades primarias.	PIB per cápita, IDHM, tasa de pobreza, índice de marginación, porcentaje de viviendas con piso de tierra	Municipios de la región Sierra Madre del Sur, zonas indígenas de Chiapas
Por actividad económica	Industrial	Predominio de la industria manufacturera, alta concentración de empresas, mano de obra calificada.	Valor de la producción industrial, porcentaje de empleo en la industria, número de establecimientos industriales, índice de competitividad	Municipios de la Zona Metropolitana de Monterrey, ciudades industriales de Guanajuato

Tipo de Clasificación	*Tipo de Municipio*	*Características Principales*	*Indicadores Clave*	*Ejemplos*
	Agrícola	Actividades agrícolas y ganaderas como principal fuente de ingresos, economías locales basadas en la producción de alimentos.	Valor de la producción agrícola, porcentaje de población ocupada en la agricultura, superficie agrícola	Municipios de Sinaloa, Yucatán, zonas rurales de Jalisco
	Turístico	Economía basada en el turismo, infraestructura turística desarrollada, alta estacionalidad económica.	Ingresos por turismo, número de turistas, ocupación hotelera, capacidad instalada de cuartos de hotel	Cancún, Acapulco, Puerto Vallarta
	Minero	Actividades relacionadas con la extracción de minerales, economías locales dependientes de la minería.	Valor de la producción minera, porcentaje de empleo en la minería, reservas minerales	Municipios de Zacatecas, Sonora, Durango
Por área geográfica	Costero	Ubicados en zonas costeras, actividades relacionadas con la pesca, turismo y comercio marítimo.	Longitud de línea costera, porcentaje de población dedicada a la pesca, número de puertos, potencial turístico	Mazatlán, Acapulco, Cancún
	Fronterizo	Limitan con otros países, dinámica económica transfronteriza, mayor diversidad cultural.	Longitud de frontera, flujo migratorio, comercio exterior, índice de desarrollo fronterizo	Tijuana, Ciudad Juárez, Reynosa

Tipo de Clasificación	*Tipo de Municipio*	*Características Principales*	*Indicadores Clave*	*Ejemplos*
	Zona Centro	Ubicados en la parte central del país, históricamente importantes, diversidad de actividades económicas.	Densidad de población, valor histórico y cultural, diversidad de actividades económicas	Ciudad de México, Puebla, Querétaro

Fuente: Elaboración con ChatGPT (OpenAI, 2024).

La clasificación municipal es una herramienta fundamental para fomentar el desarrollo local. Al identificar las particularidades de cada municipio, se pueden diseñar políticas públicas más participativas, que promuevan el desarrollo equitativo y sostenible del país. Además, esta clasificación permite evaluar la efectividad de las acciones gubernamentales y garantizar una distribución equitativa de los recursos.

MUNICIPIOS METROPOLITANOS

Con relación a la clasificación de municipios por tamaño y población, Pérez (2017) en la investigación "*Tipología del Municipio Mexicano para su desarrollo integral*" utilizando datos del INEGI (Censo de Población y Vivienda, Censo Económico y Marco Geoestadístico Nacional), así como indicadores de desarrollo humano del *PNUD, SEDESOL y CONAPO,* (con información de 2010) presenta el desarrollo de una metodología para clasificar a los municipios como se muestra en la siguiente tabla:

Tabla 5. Clasificación de los municipios en México por tamaño y población

Tipo de municipio	*Población en localidades*	*Característica*
Municipio Metropolitano	Más del 50% de la población habita en localidades con más de 100,000 habitantes	La población general es mayor a los 150,000 habitantes

Tipo de municipio	*Población en localidades*	*Característica*
Municipio Urbano	Más del 50% de la población reside en localidades de 15,000 a 100,000 habitantes	La población general está entre los 30,000 y 150,000 habitantes
Municipio en Proceso de Transición de lo Rural a lo Urbano	Más del 50% de la población habita en localidades de 2,500 a 15,000 habitantes	La población general oscila entre los 10,000 y 30,000 habitantes
Municipio Rural	Más del 50% de la población reside en localidades menores de 2,500 habitantes	La población general es menor a 20,000 habitantes
Municipio Marginado	Población en localidades menores a 5,000 habitantes	La población general tiene alto grado de analfabetismo, pobreza extrema y carencia de alimentación, falta o deficiencia en infraestructura y servicios públicos, vivienda precaria y población ocupada con ingresos menores a dos salarios mínimos

Fuente: Elaboración con base en Pérez (2017).

En México, para determinar si un municipio forma parte de una zona metropolitana, se consideran varios factores. Por un lado, se evalúa si existe una concentración poblacional significativa en al menos uno de los municipios y si hay una continuidad urbana con otros municipios vecinos.

Por otro lado, se analizan las relaciones funcionales entre los municipios, como los flujos de trabajadores entre ellos y la presencia de ciudades fronterizas con poblaciones importantes en otros países, formando así metrópolis que trascienden las fronteras nacionales. Esta clasificación permite comprender mejor la dinámica urbana de México y diseñar políticas públicas más adecuadas para estas áreas de alta concentración poblacional.

Dicho lo anterior, un municipio metropolitano es aquel que, junto con otros, forma parte de una unidad urbana más grande caracteriza-

da por una alta interconexión entre sus localidades. Estas conexiones pueden ser físicas, como la continuidad de zonas urbanas, o funcionales, como la interdependencia económica y social entre sus habitantes (Sobrino, 2024).

Así mismo, se han definido tres grupos de municipios metropolitanos, con sus respectivos criterios, como se muestra en la tabla siguiente:

Tabla 6. Grupos de municipios metropolitanos

Tipo de municipio metropolitano	*Criterios de definición*
Municipios centrales. Corresponden a los municipios donde se localiza la ciudad principal que da origen a la zona metropolitana.	a. Municipios que comparten una conurbación intermunicipal, definida ésta como la unión física entre dos o más localidades censales de diferentes municipios y cuya población en conjunto asciende a 50 mil o más habitantes.
	b. Municipios con localidades de 50 mil o más habitantes que muestran un alto grado de integración física y funcional con municipios vecinos predominantemente urbanos.
	c. Municipios con ciudades de un millón o más habitantes.
	d. Municipios con ciudades que forman parte de una zona metropolitana transfronteriza, con 250 mil o más habitantes.
Municipios exteriores definidos con base en criterios estadísticos y geográficos. Son municipios contiguos a los municipios centrales, cuyas localidades no están conurbadas a la ciudad principal, pero que manifiestan un carácter predominantemente urbano, al tiempo que mantienen un alto grado de integración funcional con los municipios centrales de la zona metropolitana.	a. Su localidad principal está ubicada a no más de 10 kilómetros por carretera pavimentada y de doble carril, de la localidad o conurbación que dio origen a la zona metropolitana en cuestión.
	b. Al menos 15 por ciento de su población ocupada residente trabaja en los municipios centrales de la zona metropolitana, o bien, 10 por ciento o más de la población que trabaja en el municipio reside en los municipios centrales de esta última.
	c. Tienen un porcentaje de población económicamente activa ocupada en actividades industriales, comerciales y de servicios mayor o igual a 75 por ciento.
	d. Tienen una densidad media urbana de por lo menos 20 habitantes por hectárea.

Municipios exteriores definidos con base en criterios de planeación y política urbana. Son municipios reconocidos por los gobiernos federal y locales como parte de una zona metropolitana, a través de instrumentos que regulan su desarrollo urbano y la ordenación de su territorio. Para su incorporación se tomó en cuenta el cumplimiento de al menos uno de los criterios que se indican en la columna siguiente.	a. Estar incluidos en la declaratoria de zona conurbada o zona metropolitana correspondiente.
	b. Estar considerados en el programa de ordenación de zona conurbada o zona metropolitana respectivo.
	c. Estar reconocidos en el Programa Nacional de Desarrollo Urbano y Ordenación del Territorio vigente.

Fuente: (*SEDESOL, CONAPO e INEGI,* 2012)

Las zonas metropolitanas mexicanas, al ser los principales centros de actividad económica y cultural del país, ofrecen oportunidades laborales que mejoran la calidad de vida de millones de mexicanos; esto las convierte en destinos atractivos para la población, tanto nacional como migrante, y contribuye a su crecimiento continuo.

SUSTENTABILIDAD EN LOS MUNICIPIOS METROPOLITANOS

Una ciudad sostenible se caracteriza por una gestión eficiente de recursos, priorizando el uso responsable del agua y la energía. Fomenta la movilidad sostenible, reduciendo la dependencia del automóvil y promoviendo el transporte público y la bicicleta. Además, invierte en construcción sostenible, utilizando materiales reciclados y energías renovables. La gestión adecuada de residuos, la preservación de espacios verdes y la activa participación ciudadana completan el perfil de una ciudad comprometida con el medio ambiente y el bienestar de sus habitantes.

DESAFÍOS PARA EL DESARROLLO SUSTENTABLE DE LOS MUNICIPIOS METROPOLITANOS

Para lograr el desarrollo sustentable, los municipios metropolitanos de México deben enfrentar desafíos particulares debido a su alta densidad poblacional, crecimiento urbano desordenado, y concentración de actividades económicas. Algunos de estos desafíos incluyen:

Tabla 7. Principales desafíos para el desarrollo sustentable de los Municipios Metropolitanos

Desafío	*Descripción Detallada*
Presión sobre los recursos naturales	La creciente demanda de agua, energía y otros recursos naturales está agotando los ecosistemas.
Contaminación ambiental	La emisión de gases de efecto invernadero, la generación de residuos y la contaminación del agua y del suelo representan una amenaza para la salud y el medio ambiente.
Desigualdad social	La concentración de la riqueza y la falta de acceso a servicios básicos generan brechas sociales y económicas.
Cambio climático	El aumento de las temperaturas, los eventos climáticos extremos y el aumento del nivel del mar ponen en riesgo las ciudades costeras y las infraestructuras.
Urbanización acelerada	El crecimiento rápido de las ciudades genera presión sobre la infraestructura, el transporte y los servicios públicos.
Gestión de residuos	La generación de residuos sólidos urbanos es un problema creciente que requiere soluciones eficientes y sostenibles.
Movilidad urbana	La congestión vehicular, la contaminación atmosférica y la falta de opciones de transporte sostenible son desafíos comunes en las ciudades.
Participación ciudadana	Es fundamental involucrar a los ciudadanos en la toma de decisiones sobre el desarrollo urbano y la gestión ambiental.

Fuente: Elaboración propia con base en la *LGAHOTDU* y la *LGEEPMA*, utilizando *ChatGPT* (OpenAI, 2024)

APLICACIONES DE LA IA PARA EL DESARROLLO SUSTENTABLE DE MUNICIPIOS METROPOLITANOS

La inteligencia artificial ofrece un gran potencial para mejorar la gestión del territorio, haciéndola más sostenible, eficiente y participativa; emerge como una herramienta crucial para desarrollar una visión a largo plazo que contemple tanto las necesidades presentes como las proyecciones de crecimiento demográfico y económico. Además, ofrece un amplio abanico de aplicaciones para abordar los desafíos de la conservación ambiental y la gestión de riesgos. A continuación, se presentan algunas de las aplicaciones más destacadas:

Tabla 8. Aplicaciones de IA para el desarrollo sustentable de los municipios metropolitanos

Desafío	*Aplicaciones de IA*	*Beneficios*
Presión sobre los recursos naturales	• Optimización de redes de suministro: Identificación de fugas en tuberías, optimización de rutas de distribución. • Predicción de demanda: Anticipación de la demanda futura para ajustar la producción y evitar escasez.	Reducción del consumo de agua y energía, mejora de la eficiencia en la distribución, prevención de crisis.
Contaminación ambiental	• Monitoreo en tiempo real: Detección de contaminantes en el aire y el agua mediante sensores y análisis de imágenes satelitales. • Predicción de calidad del aire: Modelado de la dispersión de contaminantes para anticipar picos de contaminación.	Mejora de la calidad del aire y del agua, identificación de fuentes de contaminación, implementación de medidas preventivas.
Desigualdad social	• Identificación de áreas vulnerables: Análisis de datos socioeconómicos para localizar zonas con menor acceso a servicios. • Diseño de políticas públicas: Evaluación del impacto de diferentes políticas en la reducción de la desigualdad.	Reducción de las brechas sociales, mejora de la calidad de vida en barrios marginados, promoción de la inclusión.

Desafío	*Aplicaciones de IA*	*Beneficios*
Cambio climático	• Predicción de eventos extremos: Modelado de fenómenos meteorológicos extremos para anticipar desastres naturales. • Diseño de ciudades resilientes: Evaluación de la vulnerabilidad de las infraestructuras y diseño de soluciones para adaptarse al cambio climático.	Mayor resiliencia ante eventos extremos, reducción de los daños causados por desastres naturales.
Urbanización acelerada	• Planificación urbana: Simulación de diferentes escenarios de crecimiento urbano para evaluar su impacto. • Gestión del tráfico: Optimización de las redes de transporte público y reducción de la congestión.	Mejora de la calidad de vida urbana, reducción de la congestión vehicular, promoción de la movilidad sostenible.
Gestión de residuos	• Optimización de rutas de recolección: Diseño de rutas más eficientes para reducir costos y emisiones. • Clasificación de residuos: Utilización de visión artificial para automatizar la clasificación de residuos.	Reducción de costos, mejora de la eficiencia en la gestión de residuos, promoción del reciclaje.
Movilidad urbana	• Sistemas de transporte inteligente: Optimización del tráfico en tiempo real, gestión de flotas de vehículos compartidos. • Predicción de demanda: Anticipación de la demanda de transporte público para ajustar la oferta.	Reducción de la congestión, mejora de la eficiencia del transporte público, promoción de modos de transporte sostenibles.
Participación ciudadana	• Plataformas de participación: Facilitación de la participación ciudadana en la toma de decisiones a través de encuestas en línea y foros de discusión. • Visualización de datos: Presentación de datos complejos de manera comprensible para los ciudadanos.	Mayor transparencia en la toma de decisiones, aumento de la confianza en las instituciones, mejora de la gobernanza.

Fuente: Elaboración con ChatGPT (OpenAI, 2024).

Los desafíos antes citados con aplicaciones de IA para enfrentarlos bien podrían ser complementados con las siguientes herramientas de IA que ayudan a la modelación, simulación y optimización de sistemas complejos al trabajar en la intervención y gestión de territorios de municipios metropolitanos y políticas públicas.

Tabla 9. Herramientas de IA para la modelación de sistemas complejos

Herramienta/ Software	*Método IA*	*Aplicación*	*Website*
NetLogo	Simulaciones Multiagente (MAS)	Simulación de interacciones entre actores gubernamentales y ciudadanos en la gestión territorial	*NetLogo*
Vensim	Dinámica de Sistemas	Modelación de sistemas complejos como la gestión de recursos y políticas públicas	*Vensim*
TensorFlow	Aprendizaje Automático (Machine Learning)	Predicción de patrones urbanos y análisis de grandes datos históricos para desarrollo local	*TensorFlow*
DEAP	Algoritmos Genéticos	Optimización de redes de transporte y servicios municipales	*DEAP*
Keras	Redes Neuronales Artificiales (ANN)	Predicción del impacto de políticas públicas y análisis de riesgos en sistemas locales	*Keras*
AnyLogic	Simulación basada en heurísticas	Simulación de crecimiento urbano y evaluación de políticas en la calidad de vida	*AnyLogic*

Fuente: *ChatGPT (OpenAI,* 2024).

A continuación, se presentan algunos ejemplos de municipios metropolitanos que están utilizando inteligencia artificial (IA) para promover la sostenibilidad:

1. **Singapur**: Este municipio ha implementado más de 100 soluciones de IA, incluyendo un gemelo digital de la ciudad que ayuda en la planificación urbana y la gestión de recursos. Además, utilizan chatbots para mejorar la accesibilidad a los servicios comunitarios.
2. **Helsinki, Finlandia:** A través de la plataforma *UrbanistAI, Helsinki* involucra a los ciudadanos en el diseño urbano participativo, permitiendo que la comunidad contribuya con ideas y retroalimentación para el desarrollo de sus barrios.
3. **Ámsterdam, Países Bajos:** La ciudad utiliza IA para optimizar la movilidad urbana y la gestión del tráfico, mejorando la eficiencia del transporte público y reduciendo la congestión vehicular.
4. **Zúrich, Suiza:** Zúrich emplea IA para mejorar la sostenibilidad ambiental mediante la gestión inteligente de la energía y la optimización de los recursos naturales.
5. **Buenos Aires, Argentina:** La capital argentina ha desarrollado un chatbot llamado "*Boti*" que facilita la interacción de los ciudadanos con los servicios gubernamentales, incluyendo el uso compartido de bicicletas y la atención social.

Estos ejemplos demuestran cómo la IA puede ser una herramienta eficaz para mejorar la sostenibilidad y la calidad de vida en las zonas metropolitanas.

CONCLUSIONES

La intervención y la gestión del territorio en municipios metropolitanos requiere una intervención planificada que respete la complejidad y dinámica del territorio. La inteligencia artificial emerge como una solución clave para los desafíos de sostenibilidad en los municipios

metropolitanos. Al aplicar herramientas tecnológicas avanzadas, es posible optimizar la gestión territorial, mejorar la movilidad urbana, reducir la desigualdad social y enfrentar el cambio climático. Sin embargo, para que esto sea efectivo, es crucial una coordinación de los gobiernos de los municipios metropolitanos con los órdenes de gobierno estatal y federal, y una continua y adecuada comunicación entre las autoridades y la sociedad civil, así como una comprensión profunda de las complejidades del territorio. El equilibrio entre el desarrollo económico y la conservación del medio ambiente debe ser el objetivo central en la planificación urbana sostenible, y la IA puede ser el puente para lograr ese equilibrio. Con una implementación correcta, esta tecnología tiene el potencial de transformar los municipios metropolitanos en modelos de sostenibilidad urbana.

REFERENCIAS

Cámara de Diputados del H. Congreso de la Unión. (1917). *Constitución Política de los Estados Unidos Mexicanos*. Diario Oficial de la Federación. *http://www.diputados.gob.mx/LeyesBiblio/pdf/1_310821.pdf*

Consejo Nacional de Población [CONAPO], el Instituto Nacional de Estadística y Geografía [INEGI] y la Secretaría de Desarrollo Agrario, Territorial y Urbano [Sedatu]. (2024). *Metrópolis en México 2020*. Gobierno de México. *https://www.gob.mx/sedatu/documentos/metropolis-de-mexico-2020?state=published*.

González, F. (2011). *Gestión del territorio: Un método para la intervención territorial*. Universidad de Chile. *https://repositorio.uchile.cl/handle/2250/120283*.

Morales, C., Pérez, R., Riffo, L., y Williner, A. (2020). *Desarrollo territorial sostenible y nuevas ciudadanías: Consideraciones sobre políticas públicas para un mundo en transformación*. Naciones Unidad y Comisión Económica para América Latina y el Caribe_CEPAL). *https://repositorio.cepal.org/bitstream/handle/11362/46579/1/S2000901_es.pdf*.

Pérez, F. (2017). *Tipología del municipio para su desarrollo integral*. Instituto Nacional de Administración Pública, A.C. *http://ru.juridicas.unam.mx:80/xmlui/handle/123456789/13347*

Secretaría de Desarrollo Agrario, Territorial y Urbano. (2016). *Ley General de Asentamientos Humanos, Ordenamiento Territorial y Desarrollo Urbano*. Diario Oficial de la Federación. *http://www.diputados.gob.mx/LeyesBiblio/ref/lgahotdu.htm*.

Secretaría de Desarrollo Social, Consejo Nacional de Población e Instituto Nacional de Estadística y Geografía (2012). *Delimitación de las zonas metropolitanas de México 2010*. Primera Edición. *https://www.gob.mx/cms/uploads/attachment/file/112786/1_DZM_2010_PAG_1-34.pdf*

Secretaría de Medio Ambiente y Recursos Naturales. (1988). *Ley General del Equilibrio Ecológico y la Protección al Medio Ambiente*. Diario Oficial de la Federación. *http://www.diputados.gob.mx/LeyesBiblio/pdf/LGEEPA.pdf*

Sobrino, L. (2024). Ciudades y zonas metropolitanas en México, 2020. *Realidad, datos y espacio revista internacional de estadística y geografía, Vol. 15*, (2)22-47. *https://rde.inegi.org.mx/wp-content/uploads/2024/pdf/RDE43/RDE43_art02.pdf*

Valencia, S. (2017). *El municipio mexicano: génesis, evolución y perspectivas contemporáneas*. Instituto de Investigaciones Jurídicas, UNAM. *https://repositorio.unam.mx/contenidos/5008374*

Meta Plataforms

ChatGPT (OpenAI, 2024).

Copilot de Microsoft

CAPÍTULO XXIV

LA IA Y LOS SABERES LOCALES, EN EL ÁMBITO ORGANIZACIONAL

JESÚS MANUEL RAMOS GARCÍA[1]

RESUMEN

Sin bien la inteligencia artificial (IA), como tecnología y gran transformación, se ha desarrollado en diversos campos de la vida cotidiana de la sociedad, y también en las ciencias, es importante hacer revisiones críticas y profundas de sus alcances y limitaciones, el presente capítulo se plantea en dicho sentido, analizar las implicaciones y limitaciones de la IA en los saberes locales en particular en la vida de las organizaciones, reconociendo que tiene un sesgo importante al crearse y desarrollarse en otros contextos con otros saberes, y que parte de sus intereses es la globalización, contribuyendo al epistemicidio de los saberes locales, o la erosión de los mismos. Con dicho análisis se plantean varias preguntas ¿la IA genera conocimiento? O sólo es una tecnología que sistematiza información de acuerdo a sus creadores, aspectos que deben ser revisados desde el transhumanismo, y cuidar su uso como apoyo en la toma de decisiones de la vida organizacional evitando isomorfimos miméticos fuera de sus realidades, por lo que cuando se utilice se tienen varias tareas.

Palabras clave: saberes locales, organizaciones, toma de decisiones, conocimiento

[1] Profesor Investigador del Departamento de Administración, Universidad Autónoma Metropolita, Unidad Azcapotzalco, Ciudad de México, México.

INTRODUCCIÓN

La IA ha revolucionado la forma en que las organizaciones operan y toman decisiones. Sin embargo, la incorporación de la IA en el ámbito organizacional plantea desafíos en relación con los saberes locales, es decir, el conocimiento y las prácticas desarrolladas por los miembros de la organización a lo largo del tiempo, y también en aquellas organizaciones donde sus acciones y sentidos están implicados por los saberes de la comunidad y sociedad en el que se desarrollan, nos referimos a organizaciones locales, comunitarias, originarias; donde sus acciones se explican y deben ser comprendidas desde sus cosmovisiones, aunque al ser conformada de diversos actores esta es permeada sin duda alguna por ellos, queda pendiente saber hasta que punto influyen, aspectos estudiados por Hall (1973) en su libro Organizaciones: Estructura y Proceso.

En tal sentido, este capítulo, analiza la relación entre la IA y los saberes locales en el ámbito organizacional, analizando las oportunidades y desafíos que surgen de esta intersección. Considerando que la gran trascendencia de la IA es mejorar las capacidades humanas y superar las limitaciones biológicas que promete redefinir el futuro de la humanidad, por lo que es necesario analizar que futuro plantea la IA, cuáles son sus coordenadas, perspectivas o enfoques para la vida humana, si pretende la universalización, hacia donde nos lleva si se utiliza de forma acrítica.

Por lo anterior el capítulo inicia con el desarrollo del tema la IA y los saberes locales, con el fin de sentar los referentes teóricos y analizar su relación o distanciamiento, para posteriormente analizar los retos y desafíos, tensiones en dichos temas, y cómo se considera en el ámbito organizacional para plantear posibles acciones de solución.

LA IA Y LOS SABERES LOCALES: UNA PERSPECTIVA TEÓRICA

Según el sociólogo francés, Pierre Bourdieu, “el saber local es un tipo de conocimiento que se desarrolla en un contexto específico y se transmite a través de la práctica y la experiencia” (Bourdieu, 1986).

Por otro lado, la IA se basa en algoritmos y datos para tomar decisiones y resolver problemas. La pregunta es, ¿cómo se relacionan estos dos enfoques del conocimiento?

Los saberes locales u originarios hacen referencia al conjunto de conocimientos, prácticas y habilidades que una comunidad ha desarrollado a lo largo del tiempo a partir de su interacción directa con su entorno físico, social y cultural. Estos saberes suelen ser transmitidos de generación en generación, y están profundamente arraigados en las experiencias cotidianas, las costumbres y las tradiciones de cada comunidad. A diferencia de los conocimientos formales, generalmente sistematizados y validados por instituciones académicas, los saberes locales surgen de la vida misma y están profundamente vinculados con la identidad y la cosmovisión de los grupos que los producen.

En términos organizacionales, estos saberes se manifiestan en las formas en que las personas dentro de una comunidad enfrentan desafíos, toman decisiones y resuelven problemas. Estos conocimientos son específicos de cada contexto local y están influenciados por factores geográficos, culturales, históricos y sociales que, a menudo, son ignorados o subestimados por enfoques más globalizados o estandarizados, como es el caso de los sistemas de inteligencia artificial (IA).

Al tratar de aplicar IA en ámbitos organizacionales donde los saberes locales son esenciales, se corre el riesgo de introducir sesgos importantes. Esto se debe a que los algoritmos de IA, por su propia naturaleza, tienden a operar sobre grandes conjuntos de datos que priorizan patrones universales, lo que puede invisibilizar o distorsionar los matices y particularidades de un contexto específico. Los saberes locales son valiosos porque aportan una comprensión detallada y rica de la realidad local, lo que resulta crucial en la toma de decisiones acertadas.

La IA, al estar basada en datos históricos que suelen ser agregados y anonimizados, carece de la capacidad de capturar estos conocimientos profundamente contextuales. Esto genera un sesgo al ignorar la diversidad de saberes y su relevancia en cada ámbito organizacional. En consecuencia, al aplicar IA sin considerar estos saberes locales, se corre el riesgo de imponer soluciones que no se adaptan a las necesidades reales de la comunidad, generando disonancias entre las decisiones tomadas por las máquinas y las necesidades y expectativas locales.

Con lo anterior se tiene que los saberes locales son una forma valiosa y diversa de conocimiento, imprescindible en la toma de decisiones que respondan a las particularidades de un contexto. Al aplicar IA en un ámbito organizacional, es fundamental reconocer que este tipo de conocimiento no siempre puede ser traducido a términos que los algoritmos puedan comprender y utilizar, lo que subraya las limitaciones inherentes de la IA en estos contextos específicos.

En tal sentido el filósofo y científico cognitivo, Hubert Dreyfus, argumenta que la IA no puede reemplazar completamente el saber local porque "la experiencia y la intuición son esenciales para la toma de decisiones en situaciones complejas" (Dreyfus, 1992). Sin embargo, la IA puede complementar los saberes locales al proporcionar herramientas para analizar y procesar grandes cantidades de datos.

LA IA EN LA TOMA DE DECISIONES ORGANIZACIONALES

La inteligencia artificial (IA) en la toma de decisiones se refiere al uso de algoritmos y sistemas automatizados que pueden analizar grandes cantidades de datos, identificar patrones y generar recomendaciones o decisiones basadas en esos análisis. En su esencia, la IA simula ciertos procesos de pensamiento humano, utilizando modelos matemáticos y estadísticos para abordar problemas y realizar predicciones de manera eficiente. Los algoritmos de IA se entrenan con datos históricos, y a partir de estos entrenamientos, desarrollan la capacidad de "*aprender" y mejorar sus decisiones a medida que reciben más datos*".

En el ámbito organizacional, la IA se ha convertido en una herramienta poderosa para optimizar procesos, identificar oportunidades de negocio, prever tendencias del mercado o incluso automatizar tareas repetitivas. Empresas de todo el mundo utilizan IA en áreas como la gestión de recursos humanos, la planificación estratégica, el marketing y las operaciones. Los sistemas de IA pueden evaluar miles de variables simultáneamente, proporcionando una capacidad analítica que supera las limitaciones humanas en términos de velocidad y volumen de datos procesados.

Sin embargo, la aplicación de la IA en la toma de decisiones no está exenta de desafíos, especialmente cuando se enfrenta a contextos locales o culturales específicos, donde los saberes locales juegan un papel crucial. Al basarse principalmente en datos, la IA puede ignorar o simplificar los aspectos cualitativos que son vitales en ciertos contextos. Si bien la IA es eficiente en identificar patrones globales o generalizables, muchas veces carece de la capacidad de comprender las sutilezas y particularidades de una comunidad específica, lo que puede llevar a decisiones sesgadas o inadecuadas. Si bien esto es cierto al prensente, an tanto la IA no se alimente o recopile todos los conocimientos y saberes locales y particulares no podrá orientar la toma de decisiones en el ámbito local, pero si lo podrá en cuanto logre tener sistematizados dichos conocimientos y saberes, será en algún momento tarea de los particulares crear esa IA, al final todos esos saberes ya existen en personas, colectivos y psiques de las personas del ámbito local, aún no publicas o que están presentes en libros, por eso la importancia de comprender los saberes y realidades locales y hacerlas visibles en diversas manifestaciones.

Por ejemplo, cuando una organización utiliza IA para decidir sobre estrategias de marketing en una región específica, la IA podría pasar por alto ciertos matices culturales o tradiciones locales que influyen en el comportamiento del consumidor, debido a que su entrenamiento se basa en datos que representan una media global o, en el mejor de los casos, un agregado regional amplio. En estos casos, la capacidad de la IA para ofrecer una decisión eficaz se ve comprometida por su falta de comprensión profunda del entorno local.

Además, la IA toma decisiones basadas en los datos con los que fue entrenada, lo que introduce un sesgo estructural. Si los datos de entrenamiento no reflejan adecuadamente la diversidad y los saberes locales de una comunidad, las decisiones resultantes estarán igualmente sesgadas. Por ejemplo, en áreas rurales o comunidades indígenas, los datos utilizados por la IA pueden ser escasos o inexistentes, lo que puede llevar a soluciones que no abordan las necesidades reales de esas comunidades.

Hasta aquí es importante reconocer que aunque la IA tiene un potencial inmenso para mejorar la toma de decisiones en las organizacio-

nes, su capacidad para adaptarse a contextos específicos está limitada por los datos de los que se alimenta. Esto genera un desafío importante en la toma de decisiones locales, donde los saberes comunitarios y originarios son esenciales para entender y resolver problemas. La IA, al no tener acceso o capacidad para integrar esos saberes, puede tomar decisiones inadecuadas o incompletas, lo que resalta sus limitaciones en contextos donde los conocimientos locales son fundamentales.

Considerando que la inteligencia es la capacidad de aprender, razonar y resolver problemas en términos de eficiencia y eficacia, resta responder que tan inteligente es la IA, y desde que orientaciones o perspectivas resuelve o brinda información. La inteligencia, desde una perspectiva psicológica y filosófica, se refiere a la capacidad de adquirir conocimiento, razonar de manera efectiva, resolver problemas y adaptarse a nuevos entornos (Gardner, 2011). Howard Gardner, uno de los principales teóricos en el campo de la inteligencia, amplió el concepto a través de su teoría de las inteligencias múltiples, argumentando que no existe una única forma de inteligencia, sino diversas habilidades cognitivas que interactúan y permiten el funcionamiento completo de un ser humano (Gardner, 2011). La IA ha intentado replicar y mejorar estas capacidades humanas, llevando a máquinas a realizar tareas que tradicionalmente eran exclusivas de la inteligencia humana, como la toma de decisiones.

Lo anterior ubica y reconoce que la IA ha tenido una evolución histórica. El término "inteligencia artificial" fue acuñado en 1956 por John McCarthy durante la Conferencia de Dartmouth, que marcó el inicio formal de la investigación en IA (McCarthy y otros., 2006). Sin embargo, las ideas que llevaron a este campo pueden rastrearse hasta la lógica formal desarrollada por filósofos como Aristóteles y matemáticos como George Boole. A lo largo del siglo XX, científicos como Alan Turing propusieron que las máquinas podrían llegar a pensar si se les proporcionaba suficiente capacidad para procesar información y aprender de su entorno. Turing introdujo el famoso "Test de Turing", un criterio para evaluar si una máquina puede exhibir un comportamiento indistinguible del de un humano (Turing, 1950).

OPORTUNIDADES Y DESAFÍOS PARA AMINORAR EL SESGO DE LIMITACIÓN DE SABERES LOCALES QUE TIENE LA IA AL USARLA PARA LA TOMA DE DECISIONES

Aminorar el sesgo de limitación de saberes locales en la inteligencia artificial (IA) es un desafío clave para lograr que los sistemas automatizados sean útiles y justos en contextos específicos. El sesgo surge porque la IA, al estar entrenada principalmente con datos globales o estandarizados, tiende a simplificar las complejidades locales, omitiendo los conocimientos tradicionales o contextuales que no están representados en los conjuntos de datos con los que ha sido entrenada. Este sesgo limita la capacidad de la IA para ofrecer soluciones ajustadas a las realidades y necesidades locales.

Para reducir este sesgo, es necesario implementar diversas estrategias que involucren tanto la recolección de datos, como la modificación de los modelos algorítmicos y la inclusión de un enfoque participativo. A continuación, se describen las formas clave para aminorar este sesgo:

1. Incorporar datos locales y contextuales

La fuente del sesgo en la IA muchas veces radica en la calidad y origen de los datos utilizados. Para mitigar la falta de representación de los saberes locales, es crucial que los sistemas de IA se entrenen con datos locales que reflejen la cultura, las prácticas, los valores y las realidades específicas de la comunidad en la que se van a aplicar. Esto implica ampliar las bases de datos más allá de las fuentes globales o centralizadas, y realizar un esfuerzo consciente de recolección de datos in situ.

Por ejemplo, en una región con prácticas agrícolas tradicionales, se deben recopilar datos que reflejen los calendarios agrícolas específicos de la comunidad, las técnicas ancestrales de cultivo, y las formas en que los agricultores locales toman decisiones basadas en el clima o el ciclo lunar. Estos datos contextualizados permitirán que la IA ofrezca recomendaciones que respeten y apoyen los saberes locales.

2. Diseño de modelos de IA adaptativos

Los modelos de IA tradicionales se centran en patrones globales que pueden pasar por alto los matices locales. Para aminorar este sesgo, es necesario desarrollar modelos de IA adaptativos que puedan ajustarse y aprender de las particularidades locales a lo largo del tiempo. Estos modelos deben tener la capacidad de integrar datos heterogéneos y de ajustarse a condiciones cambiantes, permitiendo que las decisiones que se tomen sean sensibles a las diferencias contextuales.

Una forma de lograr esto es mediante el uso de algoritmos de aprendizaje federado, que permiten que la IA aprenda de diferentes fuentes de datos locales sin necesidad de centralizarlas. Con este enfoque, los algoritmos pueden ser entrenados en distintos contextos locales y luego combinan los aprendizajes para generar decisiones más inclusivas y precisas.

3. Colaboración con expertos locales

Una estrategia clave para reducir el sesgo es involucrar a expertos locales y a las comunidades en el proceso de desarrollo de la IA. Estos expertos pueden aportar conocimientos que son difíciles de cuantificar o capturar en los datos, como la interpretación cultural de ciertos patrones o la forma en que los problemas deben abordarse desde una perspectiva local.

Mediante la colaboración con expertos locales, es posible integrar reglas o principios basados en saberes tradicionales que complementen el análisis de la IA. Por ejemplo, en una comunidad que depende de conocimientos indígenas para la gestión de recursos naturales, los expertos locales pueden ayudar a identificar prácticas sostenibles que la IA debe considerar, asegurando que las decisiones automatizadas no entren en conflicto con los valores y prácticas locales.

4. Enfoque en la transparencia y explicabilidad de la IA

Uno de los desafíos con los sistemas de IA es que a menudo se perciben como "cajas negras" cuyos procesos de toma de decisiones son difíciles de entender. Para mitigar el sesgo, es importante que los sistemas

de IA sean transparentes y que sus resultados puedan ser explicados de manera clara y comprensible para los usuarios locales. Si las decisiones de la IA pueden ser explicadas en términos de cómo se tomaron y cuáles fueron los datos utilizados, será más fácil identificar posibles sesgos o errores que hayan surgido por una falta de datos locales.

La explicabilidad también permite que los usuarios evalúen si las decisiones de la IA están alineadas con los valores y los conocimientos locales, y si es necesario realizar ajustes. Esto es esencial en entornos donde los saberes tradicionales y las normas culturales deben ser respetadas en la toma de decisiones.

5. Ciclo continuo de retroalimentación

Aminorar el sesgo también requiere establecer un ciclo continuo de retroalimentación entre la IA y los usuarios locales. Es decir, las comunidades deben tener la capacidad de evaluar las decisiones tomadas por la IA y proporcionar información sobre su pertinencia y precisión. Este ciclo de retroalimentación permite que los modelos de IA se actualicen y mejoren con el tiempo, incorporando nuevas perspectivas y datos locales conforme surgen.

Un sistema de retroalimentación continuo no solo ayuda a ajustar los algoritmos, sino que también fortalece la confianza de la comunidad en el uso de la IA, ya que sienten que su experiencia y conocimiento son valorados e incorporados de manera activa en el proceso.

6. Desarrollo de criterios éticos y de gobernanza local

Para reducir el sesgo en la toma de decisiones locales, es necesario establecer criterios éticos claros que guíen el uso de la IA en contextos comunitarios. La IA debe ser diseñada con principios que aseguren que los saberes locales sean respetados y que las decisiones automatizadas no perjudiquen a las comunidades.

Esto incluye garantizar que los datos locales sean utilizados de manera responsable y segura, y que la gobernanza sobre el uso de la IA sea participativa, permitiendo que las comunidades mantengan el control sobre cómo la IA se implementa y utiliza.

Este enfoque ético promueve una relación más equilibrada entre la tecnología y los saberes locales, asegurando que la IA sea una herramienta de apoyo y no una imposición externa.

IA DE SABERES LOCALES Y MÁS ALLÁ DE LA SISTEMATIZACIÓN Y CONCENTRACIÓN DE LOS MISMOS

Si bien, en la ruta de la sistematización y creación de IA, es posible crear una para los saberes locales en diferentes ámbitos y no sólo en lo organizacional, con los riesgos que ello implica como el regresar a la extracción del saber hacer de las personas trabajadoras como lo fue en el sistema productivista en la década de los 90´s al estilo instrumentalista, de una racionalidad económica, o como lo plantearon diversos autores como Bourdieu, 1986; Mintzberg, 1989; Nonaka, 1991; o Drucker, 1999; se debe cuidar que no se repita con la creación de IA en diferentes ámbitos de la vida, limitando de manera ética el desarrollo de la humanidad local, despojarla y limitar su continuo desarrollo de saberes en la vida diaria. Si bien, y sin tener punto de comparación más a llá de este planteamiento de creación de la sistematización, como IA, un ejemplo contundente que no sólo busca la sistematización sino la vida comunal es la Universidad Autónoma Comunal del Estado de Oaxaca (UACO), analizarla permite ubicar que es ya una IA según los contextos de la IA y para estos, para el autor no es posible dicha referencia en varios sentidos, no obstante para los saberes locales y sus personas es una realidad de vida en toda la extensión de su vida que plantea diferentes rompimientos paradigmáticos que la IA no lo haría al pretender ser instrumental, lineal, pragmática, y la Universidad Autónoma Comunal recrea la vida, y recupera la humanidad comunal, sin lógicas de poder como la IA si. ¿Qué pretende la IA, qué fines busca en la realidad del planeta donde habitamos a partir de quienes la crean, la controlan, la promueven? ¿A partir de qué racionalidades se desarrolla?

Así, conocer la UACO, creada en 2020 como un proyecto de transformación social que busca fortalecer la educación y desarrollo de la región a partir de sus propias cosmovisiones y racionalidades, permi-

te comprender que es una institución educativa que se enfoca en la promoción y desarrollo de los saberes locales, comunales y regionales. Ubicando como puntos clave sobre su importancia:

1. Fortalecimiento de la identidad local: *"La UAC es un espacio donde se rescata y se promueve la identidad cultural y lingüística de los pueblos indígenas de Oaxaca*." (Floriberto Díaz, Rector Fundador UAC, 2004)
2. Desarrollo sostenible: "*La UAC busca contribuir al desarrollo sostenible de las comunidades, considerando sus necesidades y recursos locales*." (UAC, 2019)
3. Innovación y creatividad: "*La UAC fomenta la innovación y creatividad en la comunidad, mediante la investigación y el desarrollo de proyectos locales*." (UAC, 2020).
4. Formación de líderes: La universidad forma líderes comunitarios capacitados para abordar los desafíos y oportunidades locales.
5. Intercambio de conocimientos: La UACO facilita el intercambio de conocimientos y experiencias entre la comunidad aca´demica y la sociedad local.

Fortaleciendo con ello que "La educación comunitaria es un proceso que busca fortalecer la comunidad y mejorar la calidad de vida." (Juan Carlos Tedesco, 2005) y "*La universidad comunitaria es un espacio donde se construye conocimiento desde la perspectiva local*." (Boaventura de Sousa Santos, 2009)

Con todo lo anterior, y con sus limitaciones y cuestionamientos, desarrollar una IA propia que se fundamente en los saberes locales implica diseñar un sistema de inteligencia artificial que no solo responda a patrones generales o globales, sino que sea capaz de capturar, integrar y respetar los conocimientos, prácticas y valores específicos de una comunidad o contexto organizacional particular. Este enfoque va más allá de la IA tradicional, que se basa en grandes volúmenes de datos estandarizados, y busca una personalización profunda que reconozca y dé valor a la diversidad cultural, social y geográfic; y que no pretende sustituir la vida comunal, si no solo ser una herramienta

de sistematización de la información, aunque con seria posibilidades de recrear el pesnamiento comunal a partir del transhumanismo y que esta recobre conciencia de sí misma.

Para lograrlo, es necesario abordar el desarrollo de la IA desde varias perspectivas clave:

1. Recolección de datos locales y pertinentes

Uno de los primeros pasos cruciales es recoger datos locales que reflejen los saberes y prácticas específicas de la comunidad u organización. Estos datos deben ser cualitativos y cuantitativos, e incluir no solo estadísticas tradicionales, sino también narrativas, historias orales, costumbres y experiencias vividas. Este tipo de información rara vez se encuentra en los grandes conjuntos de datos con los que tradicionalmente se entrena la IA, por lo que es importante realizar procesos de recopilación directa en colaboración con miembros de la comunidad.

Para que la IA capte estos saberes, se debe enriquecer el conjunto de datos con información local que capture los matices y detalles que forman parte del conocimiento comunitario. Esto podría incluir prácticas agrícolas específicas, decisiones de gestión organizacional basadas en la tradición o conocimientos ancestrales sobre el uso sostenible de recursos. Este enfoque permitiría que la IA no solo analice los datos disponibles, sino que aprenda a interpretar las realidades locales.

2. Diseño participativo y colaborativo

Un aspecto fundamental para desarrollar una IA que respete los saberes locales es adoptar un enfoque de diseño participativo. Esto significa que los miembros de la comunidad o del ámbito organizacional deben estar involucrados activamente en la creación y ajuste de la IA. Ellos aportan la comprensión y los conocimientos necesarios para guiar a los desarrolladores de IA en la integración adecuada de los saberes locales.

El co-diseño no solo garantiza que la IA sea culturalmente sensible, sino que también promueve la confianza en la tecnología, ya que los usuarios sienten que sus conocimientos y valores son reconocidos y

respetados. Este enfoque puede involucrar desde talleres de consulta hasta alianzas con líderes locales que ayuden a ajustar las recomendaciones de la IA en función de las dinámicas comunitarias.

3. Modelos de aprendizaje híbrido

Una IA basada en saberes locales necesita integrar enfoques de aprendizaje híbrido, que combinen modelos de aprendizaje automático con mecanismos de adaptación local. Mientras que los algoritmos convencionales tienden a centrarse en correlaciones estadísticas amplias, un sistema híbrido podría incluir reglas y parámetros que reflejen principios o prácticas locales, lo que permitiría que la IA ofrezca soluciones que respeten las especificidades del contexto.

Por ejemplo, en una comunidad rural, la IA podría incorporar reglas basadas en los ritmos agrícolas tradicionales o en conocimientos ancestrales sobre la gestión del agua, en lugar de basarse exclusivamente en patrones meteorológicos globales. Estos modelos híbridos ayudarían a adaptar las decisiones tecnológicas a los valores y realidades locales.

4. Evaluación continua y retroalimentación local

Una vez implementada, la IA debe someterse a evaluaciones constantes, en las que los miembros de la comunidad puedan aportar retroalimentación sobre las decisiones que la IA está tomando. Esta retroalimentación no solo permitirá ajustar los algoritmos, sino también mejorar la capacidad del sistema para alinearse con los saberes locales en tiempo real.

Es importante reconocer que los saberes locales no son estáticos; evolucionan con el tiempo. Por lo tanto, la IA debe ser capaz de aprender y adaptarse continuamente a los cambios en las costumbres y prácticas locales, lo que requiere un ciclo constante de retroalimentación y ajuste. Este proceso participativo garantizará que la IA siga siendo relevante y respetuosa con las dinámicas locales.

5. Ética y control local

Finalmente, el control sobre la IA debe residir en la comunidad u organización que la utiliza. La gobernanza de los datos y los algoritmos debe ser transparente, y la comunidad debe tener el poder de decidir cómo se recopilan, almacenan y utilizan sus conocimientos. Este control es esencial para garantizar que la IA no se utilice de manera extractiva o en formas que perjudiquen a la comunidad.

Esto implica adoptar un enfoque ético en el desarrollo de la IA, asegurando que la tecnología no solo sirva para mejorar la eficiencia o la productividad, sino también para fortalecer la capacidad de la comunidad de tomar decisiones autónomas y preservar su identidad cultural.

La integración de la IA y los saberes locales en el ámbito organizacional ofrece varias oportunidades:

- Mejora de la eficiencia y productividad, aunque ese no es la principal racionalidad ni objetivo de los saberes locales, ni de la vida comunal
- Toma de decisiones más informadas, aunque debe ser tan capaz de incluir toda la vida comunal
- Innovación y creatividad

Sin embargo, también existen desafíos:

- Validación y autenticidad de los saberes locales
- Contextualización cultural y histórica de la IA
- Propiedad intelectual y derechos de autor

Para aprovechar las oportunidades y superar los desafíos, las organizaciones pueden implementar las siguientes estrategias:

- Colaboración interdisciplinaria entre expertos en IA y saberes locales
- Desarrollo de sistemas expertos que incorporen saberes locales
- Capacitación y educación continua para las personas.

REFLEXIONES FINALES

La relación entre la IA y los saberes locales en el ámbito organizacional es compleja y multifacética. Aunque existen desafíos, la integración de estos dos enfoques del conocimiento puede generar oportunidades para mejorar la eficiencia, innovación y toma de decisiones. Es fundamental abordar los desafíos y desarrollar estrategias efectivas para aprovechar el potencial de la IA y los saberes locales.

Con los temas abordados se abre la agenda de cuestionamientos y de investigación a partir de ¿Qué tan inteligente es la IA? ¿Genera conocimiento y saberes la IA o sólo es lineal en términos de eficiencia y eficacia? No en general si no en específico a los temas abordados. ¿Cómo integra la IA la parte cualitativa humana de un contexto local? ¿Qué está haciendo la IA para atender esto, o es a partir de las redes neuronales? ¿Hacia donde nos lleva la IA, cuáles son sus racionalidades sus ejes o coordenadas, la universalización concentración del poder, dependencia hacia un sistema económico? ¿Realmente está considerando a la humanidad?

Por otro lado, existen diversidad de conocimientos y saberes distintos al pensamiento occidental dominante, esperando esta universalización u homogenización no sea uno de los objetivos de la IA de una forma más avanzada, por lo que es importante identificarlos, comprenderlos y visibilizarlos, y sin lugar a dudas la IA puede ser una herramienta, que controlada, puede ayudar, inclusive a la pérdida de los mismos o en la reproducción y comprensión de las nuevas generaciones. En este ámbito los saberes locales son fundamentales para la toma de decisiones y la resolución de problemas en las organizaciones locales. La experiencia y el conocimiento local son esenciales para entender el contexto y las necesidades específicas de la organización.

Con lo antes analizado, la IA no puede reemplazar completamente los saberes locales y la experiencia humana. La IA puede ser limitada por su falta de comprensión del contexto y la cultura en sus diferentes dimensiones, por lo que la gestión de la IA y los saberes locales requiere una visión integral que considere la cultura, la estrategia y la tecnología; teniendo presente los desafíos éticos y culturales.

REFERENCIAS BIBLIOGRÁFICAS

Bourdieu, P. (1986). La distinción: una crítica social del juicio. Madrid: Taurus.

Castells, M. (2001). La sociedad red. Madrid: Alianza.

Díaz, F. (2004), Discurso de inauguración de la UAC.

Dreyfus, H. (1992). What Computers Still Can't Do: A Critique of Artificial Reason. Cambridge: MIT Press.

Drucker, P. (1999) La sociedad postcapitalista, Barcelona. Editorial Norma.

Gardner, H. (2011). Frames of mind: The theory of multiple intelligences (3rd ed.). Basic Books.

Gilly, A. (2012). La universidad comunitaria.

Hall, R. H. (1973). Organizaciones: Estructura y Proceso. México: Editorial Diana.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Magazine, 27(4), 12-12.

Mitzberg, H. (1989), Mintzberg sobre la estrategia México: Editorial el Progresso.

Nonaka, I. (1991), "The Knowledge-Creating Company", Harvard Business Review, 69(6), 96-104.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14-37.

Polanyi, M. (1966). The Tacit Dimension. London: Routledge.

UAC (2019) Plan Estratégico 2019-2024.

UAC (2020) Informe Anual 2020

Tedesco, J. C. (2005) Educación y Sociedad.

Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59(236), 433-460.

Santos, B. de S. (2009), Una epistemología del Sur.

CAPÍTULO XXV

HUMANISMO Y TRANSHUMANISMO

JONATAN GARCÍA CAMPOS[1]

RESUMEN

El transhumanismo como varios autores han defendido cuestiona las limitaciones biológicas del ser humanso y busca en última instancia vencer a la muerte. En este texto se aborda críticamente al transhumanismo, sus principios y su relación con el humanismo. En este capítulo se buscan dos objetivos: en primer lugar, definir el transhumanismo y sus variantes, analizandolo críticamente desde tres niveles: lo fáctico, lo ético y lo desiderativo; en segundo lugar, examinar si el transhumanismo representa una continuidad o una ruptura con los principios del humanismo, evaluando qué aspectos del humanismo hereda.

Palabras clave: longevidad, facticidad, ética, deseos, dignidad

INTRODUCCIÓN

A lo largo de la historia, el humanismo ha sido una corriente cultural y filosófica que ha situado al ser humano en el centro del pensamiento, destacando su dignidad y su capacidad para razonar, crear una ética laica y progresar gradualemente. Este enfoque floreció durante el Re-

[1] Profesor- Investigador, Departamento de Administración, Universidad Autónoma Metropolitana, Unidad Azcapolzalco, Ciudad de México, México

nacimiento y ha sido relevante durante gran parte de la historia de Occidente, sin embargo, en el siglo XXI, una nueva perspectiva ha comenzado a ganar terreno, la cual cuestiona las limitaciones biológicas que el humanismo aceptaba como naturales: el transhumanismo. Definir con exactitud al transhumanismo no es algo fácil, existen muchas caracterizaciones del mismo: para varios autores el transhumanismo es la filosofía de nuestro tiempo, incluso la última de las utopías; para otros se trata de algo más similar a un movimiento de liberación; otros consideran que el transhumanismo es una ideología infundada de los avances de la ciencia; otros más lo consideran "la idea más peligrosa del mundo" (More, 2013; Fukuyama, 2004; Diéguez, 2017, 2021).

El propósito del trabajo es doble. Por un lado, se busca analizar qué es el transhumanismo, como se podrá ver en el trabajo no todos los defensores o críticos de este lo entienden de la misma manera, por lo que es necesaria hacer varias distinciones para tener una compresión cabal del mismo, ya sea para aceptarlo o rechazarlo. Argumentaré, siguiendo a otros autores, que es posible separar tres diferentes niveles para examinar críticamente al transhumanismo. Por otro lado, a partir de las diversas maneras de entender al transhumanismo y de los niveles analizados se evaluará qué relación tienes el humanismo y el transhumanismo.

El trabajo está dividido en tres secciones, en la primera me interesa caracterizar el transhumanismo y distinguir dos grandes variantes dentro del transhumanismo. En la segunda parte, me interesa distinguir tres niveles de estudio con las que se puede abordar de modo crítico al transhumanismo, a saber, una relacionado con lo fáctico, otro con lo ético y otro con lo desiderativo. Por último, debido a que algunos han visto una continuidad o una radicalización del humanismo en el transhumanismo, se analiza qué elementos dentro de este último son heredados o no del humanismo.

1. EL TRANSHUMANISMO

Hay varias maneras de entender el transhumanismo, algunos lo caracterizan por promover activamente el mejoramiento del ser humano

a través del uso de nuevas tecnologías, especialmente en áreas como la biotecnología y la inteligencia artificial (IA), una vez que estas tecnologías hayan alcanzado un nivel de desarrollo adecuado (Diéguez, 2021). El lema principal del transhumanismo es que la muerte puede ser derrotada (Diéguez, 2017, p. 12). Al parecer habría dos razones (teóricas) de por qué sería factible la "derrota" anterior:

i. Varios biólogos consideran que el envejecimiento sigue mecanismos contingentes y ciertamente complejos, pero por tal motivo, ya sea que estén contralados por los genes o no, se podrían modificar de modo que se puede prolongar la vida.

ii. Otra razón es que no parece haber una ley física o química que determine que se debe morir (Ramakrishnan, 2024).

Teniendo estas dos ideas de fondo, el transhumanismo ha presentado una declaración conocida como la Declaración Transhumanista, la cual es un documento que establece los principios y objetivos del movimiento transhumanista.[2]

A continuación, presento los ocho puntos que componen dicha declaración:

1. La mejora humana: Los seres humanos tienen el derecho de mejorar sus capacidades físicas e intelectuales mediante la tecnología.
2. La vida y la muerte: La vida no tiene un valor inherente, y la muerte no es inevitable. La extensión de la vida y la superación de la muerte son objetivos legítimos.
3. La autonomía: Los individuos tienen derecho a controlar su propio destino, incluyendo el uso de tecnologías para mejorar su vida y su muerte.

[2] La Declaración Transhumanista fue publicada por primera vez en 1998 por la Humanity+ (anteriormente conocida como la Asociación Transhumanista Mundial), una organización sin fines de lucro que promueve estos principios y apoya la investigación en áreas clave para el mejoramiento humano.

4. La responsabilidad: Los individuos son responsables de sus acciones y deben considerar las consecuencias de sus decisiones.
5. La solidaridad: Los transhumanistas deben trabajar juntos para lograr sus objetivos y apoyar a aquellos que buscan mejorar su vida.
6. La educación: La educación es fundamental para el progreso humano y la adopción de tecnologías que mejoren la vida.
7. La diversidad: La diversidad de opiniones y enfoques es valiosa y debe ser promovida.
8. La sostenibilidad: Los transhumanistas deben considerar las consecuencias a largo plazo de sus acciones y trabajar hacia un futuro sostenible.

A juicio de los transhumanista, estos puntos establecen los principios básicos del movimiento transhumanista, que busca mejorar la condición humana mediante el uso de tecnologías avanzadas.

La Declaración Transhumanista ha estado sujeta a modificaciones hasta la actualidad (González, 2023). Esta declaración resume los principios fundamentales del movimiento transhumanista, que busca mejorar la condición humana mediante el uso de tecnologías avanzadas. La declaración ha sido publicada en varios sitios web y documentos, incluyendo el "Manifiesto Transhumano" de Natasha Vita-More.[3]

Una vez que se ha mostrado la tesis general del transhumanismo y su declaración de principios, es necesario señalar que no todos los transhumanistas defienden exactamente lo mismo. El transhumanismo es un movimiento heterogéneo en donde se encuentran posturas radicales y otras moderadas (Diéguez, 2017). Pero dentro de ellas habría dos grandes vertientes, lo que se conoce como el transhumanismo digital o cibernético, y el transhumanismo biológico o biomédico. Veamos.

El transhumanismo digital o cibernético es un enfoque que se centra en el uso de tecnologías digitales, la IA y la fusión del cuerpo hu-

[3] El manifiesto puede leerse en https://arjephilo.com/wp-content/uploads/2021/01/manifiesto-transhmano.pdf

mano con máquinas. La idea principal es mejorar las capacidades humanas mediante la integración de tecnologías cibernéticas que puedan aumentar la inteligencia, la memoria, la fuerza o los sentidos. Ejemplos de tecnologías dentro de este enfoque incluyen:

- Implantes cibernéticos: Dispositivos electrónicos implantados en el cuerpo para mejorar o reemplazar funciones biológicas. Por ejemplo, prótesis robóticas controladas por el cerebro o interfaces cerebro-computadora.
- Subir la mente a una máquina (*mind uploading*): Un concepto más teórico que propone la posibilidad de transferir la conciencia humana a una computadora, lo que permitiría vivir indefinidamente en entornos virtuales.
- Realidad aumentada y realidad virtual: Tecnologías que podrían ser usadas para crear experiencias cognitivas inmersivas que expandan la percepción o habilidades humanas.

De este modo, el transhumanismo cibernético busca, en última instancia, trascender las limitaciones físicas del cuerpo humano mediante la tecnología.

Por su parte, el transhumanismo biológico o biomédico es el enfoque que busca modificar o mejorar el cuerpo humano a nivel biológico mediante avances en la biomedicina, la genética y las ciencias de la vida. En lugar de depender de tecnología cibernética, el transhumanismo biológico busca mejorar al ser humano modificando su biología, ya sea a través de la genética, la farmacología o la biotecnología. Ejemplos incluyen:

- Terapia génica: La modificación del ADN para corregir enfermedades genéticas o, en el contexto transhumanista, para mejorar características físicas o mentales (como la inteligencia o la longevidad).
- Ingeniería de tejidos y regeneración de órganos: Creación de órganos artificiales o la regeneración de partes del cuerpo para aumentar la longevidad o reemplazar órganos dañados.

- Mejora cognitiva farmacológica: Uso de fármacos para mejorar la función cerebral más allá de lo que es considerado "normal" para los humanos, como medicamentos que aumenten la memoria o la creatividad.

El transhumanismo biológico se centra en modificar y mejorar los cuerpos humanos, más que fusionarlos con máquinas, y busca extender la vida o mejorarla a través de avances médicos.

De este modo, entre las dos vertientes del transhumanismo la diferencia principal es que mientras el transhumanismo cibernético busca una mejora a través de la tecnología, fusionando humanos con máquinas en una de sus versiones; el transhumanismo biomédico busca una mejora a través de la biología y la medicina, modificando el cuerpo humano a nivel genético o físico.

Ambos enfoques buscan trascender las limitaciones humanas, pero lo hacen a través de medios diferentes. En ese sentido, los dos transhumanismos tienen en común eliminar la frontera entre lo humano y la máquina, o bien sostener que la biología humana no está constreñida, lo que es entendido como una manera de liberar al ser humano. De ese modo, para los transhumanistas:

> "*La integración con la máquina, la superación de lo biológico (y lo corporal) en cuanto que factor limitante es el modo final en el que el ser humano puede trascender su condición miserable, sesgada y asfixiante, para aspirar a horizontes en los que no se atisba límite alguno, ni temporal, ni material*"(Diéguez, 2017, p. 24).

Una vez que se ha explicado qué es el transhumanismo y que he presentado dos vertientes dentro de este, deseo antes de pasar a la siguiente sección de trabajo, decir algo acerca de la popularidad que ha ganado el transhumanismo.

Diferentes autores han afirmado que el éxito del transhumanismo consiste en que ha sido capaz de conectar con deseos, algunos de ellos, muy claros y recurrentes en la historia de la vida humana, como la longevidad, el poder, un mayor conocimiento o control, pero otros más difusos e insatisfechos de la población de los países más desarro-

llados. Explícitamente Diéguez afirma que el transhumanismo "hace recuperar la confianza en el poder de la mente humana como garante del progreso material y de un futuro mejor" (2017, p. 11). Sin embargo, como el famoso historiador, Yuval Harari (2015) sugiere, si bien la tecnología puede ofrecer medios para satisfacer algunos deseos, también puede crear nuevos problemas y deseos insatisfechos. Por ejemplo, a medida que las personas logran vivir más tiempo, pueden surgir preocupaciones sobre la calidad de vida, el propósito y el significado de ésta. Esto podría llevar a un ciclo continuo de insatisfacción y búsqueda de nuevas mejoras, en una espiral que podría sonar como inacabable.

2. TRES ASPECTOS DENTRO DEL TRANSHUMANISMO

En la sección anterior se ha esbozado brevemente qué es el transhumanismo, recurriendo a sus principios, como se muestra en la Declaración Transhumanista, así como en las dos vertientes del humanismo esbozadas, en esta sección se analizan algunas críticas que se han hecho al mismo, pero para ello considero que es necesario distinguir tres niveles que se podrían manejar dentro del transhumanismo: un nivel fáctico, otro nivel ético y otro nivel, que podría considerarse, desiderativo. Veamos

2.1 El nivel fáctico

Está relacionado con los hechos, esto es, con las afirmaciones que hace el transhumanismo y si estas afirmaciones son creíbles, realizables o son verdaderas en un sentido objetivo. Algunos autores (Diéguez, 2017), proponen que de las diferentes afirmaciones hechas por el transhumanismo se debe valorar cuáles son creíbles de aquellas que no lo son. Por ejemplo, se afirma que:

> "*al igual que el resto de las utopías, el transhumanismo se sitúa en la frontera entre lo real y lo imaginario, y desde allí propone "un programa metafísico de investigación", así como un repertorio de conceptos y*

> *de valores que impulsan prácticas concretas encaminadas a la consecución de una nueva era*" (Diéguez, 2017, p. 11).

Bajo esta visión, la labor de la sociedad o de los estudiosos del transhumanismo debe ser críticas con relación a "las razones objetivas que haya para tomarlos en serio" (Diéguez, 2017, p. 9). Por ejemplo, nos podemos cuestionar hasta qué punto es seria la propuesta de que el ser humano pueda vivir muchos más años o pueda incluso ser inmortal, ¿es una meta inalcanzable o realmente se puede alcanzar? Como he señalado al principio de esta sección hay científicos serios que parecen apoyar que es factible, o podría ser factible en el futuro, si es posible conocer los mecanismos que se encuentran detrás del envejecimiento (Ramakrishnan, 2024). Pero, ¿qué hay de la idea de transferir la conciencia humana a una computadora (el *mind uploading*) que permita vivir indefinidamente en entornos virtuales? ¿En qué sentido subir todas nuestras creencias y deseos a una súper computadora supone que nosotros viviríamos eternamente? Algunos argumentan que, aunque se pueda crear una réplica exacta de la mente,[4] la experiencia subjetiva o el "yo" no se transferiría, es decir, no se trasmitiría el como sentimos o percibimos nuestros propios estados mentales. Bajo esta perspectiva, la copia de nuestra mente, no parece ser nuestro yo, sino una copia de nosotros, lo que plantea dudas sobre si el *mind uploading* podría lograr una verdadera inmortalidad personal, como defiende el transhumanismo cibernético.

Lo que parece orillarnos esta crítica, en dado caso, es que se tendría que revisar con cuidado cada una de las vertientes del transhumanismo y revisar por expertos cuáles de las técnicas del transhumanismo son reales o realizables y cuáles no. Sin embargo, la cuestión no parece sencilla, porque un defensor del transhumanismo podría contra argumentar que si bien algunas de las promesas de este no parecen facti-

[4] Es preciso señalar que actualmente no tenemos la tecnología necesaria para mapear el cerebro humano a nivel de detalle requerido para el *mind uploading*. Los escáneres actuales no pueden captar todos los detalles necesarios de las neuronas y sus conexiones. Asimismo, el poder computacional necesario para simular un cerebro humano con todas sus complejidades también es enormemente alto. Sería necesario un salto importante en la capacidad de procesamiento. Este tipo de retos ha hecho que se considere que no es factible transferir la conciencia humana a una computadora.

bles al día de hoy, podrían serlo en un futuro mediato. Nuevamente aquí se cae en un espiral, porque saber qué será lo factible en un futuro no es algo a lo que tenemos acceso en el presente.

Una crítica relacionada con la factibilidad, que a mi juicio puede ser relevante, es que el transhumanismo genera una "cortina de humo" para distraernos con ilusiones de los problemas más atingentes que enfrenta la humanidad actual.

> "*a la implausibilidad fáctica de las propuestas transhumanistas añaden la voluntad que estas encerrarían de constituir una maniobra de distracción ideológicamente sesgada para obviar los auténticos problemas del presente (desigualdad económica, deterioro medioambiental, expoliación de recursos naturales, injusticias sociales, crímenes políticos, etc.) y disuadir de cualquier intento por enfrentarse a ellos. Dicho de otro modo, el discurso transhumanista es un recurso del establishment tecnocrático para justificar la continuación de la misma política (económica y tecnológica) que nos ha llevado a la actual situación insostenible y finalmente al borde del colapso ecológico y civilizatorio*" (Diéguez, 2017, pp. 25-26).

Frente a esta crítica es posible afirmar, por un lado, que el transhumanismo y sus consecuencias son un tema lo suficientemente importante por sí mismo para ser discutidos por la academia, pero también a nivel social, por el tipo de implicaciones que tiene; por otro lado, no parece ser tan claro que haya pruebas claras de que haya una distracción intenci

onada como para sostener que el transhumanismo es una cortina de humo. Esto, por supuesto, no demerita que haya otros problemas sociales, económicos o éticos a tratar, que sean relevantes e incluso más urgentes.

2.2. Nivel Ético

Algunos teóricos no han señalado tanto la parte fáctica, sino las posibles consecuencias éticas que el transhumanismo puede tener. Este nivel que podríamos llamar el *nivel ético* del estudio del transhumanismo sugiere que sería "*muy posible que probemos las tentadoras ofertas de la biotecnología sin darnos cuenta de que tienen un coste moral espantoso*" (Fukuyama, 2004, p. 42).

Pero, es posible preguntarse cuál sería el coste moral del transhumanismo. Según Fukuyama (2004), el transhumanismo parece minar la idea de igualdad, la cual ha sido un elemento fundamental de nuestra vida social y política actual. Además, señala en politólogo:

> "*en la base de esta idea de la igualdad de derechos se encuentra la creencia de que todos poseemos una esencia humana que eclipsa las diferencias manifiestas en el color de la piel, la belleza e incluso la inteligencia. Esta esencia, y la idea de que los individuos, por lo tanto, tienen un valor inherente, está en el corazón del liberalismo político*" (Fukuyama, 2004, p. 42).

Fukuyama pasa de este modo a señalar que el transhumanismo rompe con la esencia inherente del ser humano. Para defender esta idea parte de que los rasgos mentales (y presuntamente también físicos) están entrelazados unos con otros, o tienen una parte positiva y negativa al mismo tiempo. Cambiar cualquier rasgo, implicaría cambios que no es posible determinar.

No es tan claro qué tan fuertes pueden ser estas críticas, debido a que se puede poner en duda la idea de esencia humana fija[5] y también la idea de que podrían estar los rasgos entremezclados, pero que con el paso del tiempo se podrían controlar o hacer mejoras sin implicar cambios no deseados.[6] Pero independientemente de ello, me parece que hay dos ideas interesantes en las críticas de Fukuyama que se pueden seguir pensando:

i) "*los defensores del transhumanismo piensan que entienden qué constituye ser un buen ser humano*"

ii) "*El movimiento ecologista nos ha enseñado humildad y respeto por la integridad de la naturaleza no humana. Necesitamos una humildad similar en relación con nuestra naturaleza humana. Si no la desarrollamos pronto, podemos invitar sin darnos cuenta a*

[5] La idea de la esencia humana será abordado en la siguiente sección.

[6] Este punto parece conducir al señalamiento del nivel fáctico en donde no es fácil saber si en un futuro se logren algunas de las promesas del transhumanismo, pero tampoco sus posibles amanezas.

> *los transhumanistas a desfigurar la humanidad con sus excavadoras genéticas y centros comerciales psicotrópicos*" (2004, p. 43).

La idea i) es relevante porque sugiere que el transhumanismo o sus técnicas están valorativamente cargadas y precisamente esos valores asumidos son algo que se puede poner en duda. Por ejemplo, algunas técnicas biotecnológicas podrían promover un racismo o tener un sesgo de género, si se pueden diseñar hijas o hijos con propiedades particulares. Por su parte, la idea ii) sugiere que el transhumanismo no está respetando la esencia del ser humano. Esta idea será relevante para la parte discutida en la siguiente sección sobre el humanismo. De momento deseo cerrar esta sección con una cita que resume muy bien el nivel ético que aquí he expuesto, a saber, que el transhumanismo es:

> *"...una grave amenaza que se dirige en realidad a la aniquilación del ser humano bajo la excusa de su transformación en un ser superior. Debilita todo lo que hay de sagrado en él, asume un poder que no nos corresponde y que no podemos controlar, pone en cuestión las bases en las que se ha cimentado nuestra vida social y ética, introduce nuevas desigualdades y refuerza de forma extrema el poder de una minoría sobre la gran mayoría,*[7] *desdibuja las fronteras que nos separan del animal, desprovee de significado nuestra mera existencia...*" (Diéguez, 2017, p. 26).

2.3. Nivel desiderativo

Al hablar de un *nivel desiderativo* en el contexto de los estudios humanos deseo referirme a un ámbito que analiza los deseos, impulsos y as-

[7] Existen varias referencias a la desigualdad que podría traer algunos procedimientos particulares del transhumanismo. No obstante, algunos autores (Ramakrishnan, 2017) sostienen que, de hecho, en algunos países las personas con mayores ingresos ya viven más de una década más que las personas con menores ingresos. Además, si se toma en cuenta la esperanza de vida saludable (y el número de años de vida sana), esa disparidad es aún mayor. En otras palabras, las personas más pobres viven menos y con menos salud, por lo que el transhumanismo podría potencialmente aumentar las desigualdades que ya es posible observar.

piraciones como factores fundamentales en la conducta y el desarrollo del ser humano.

Este enfoque es relevante para disciplinas como la psicología, la filosofía, la sociología y la antropología, donde los deseos humanos son cruciales para entender la motivación, la conducta, las coaliciones y la toma de decisiones. Como he señalado antes, el transhumanismo parece explotar un conjunto de deseos que toma como dados, pero que es posible cuestionarse si en realidad sabemos lo que deseamos. La pregunta, como la han señalado otros autores (Harari, 2015), es hasta qué punto conocemos nuestros deseos, qué deseos son genuinos (y no impuestos o falsos deseos), cómo es que los deseos que tenemos no son impuestos por un patrón cultural que puede ser puesto en duda y, en qué medida, aunque tengamos deseos particulares es legítimo seguirlos.

Para ilustrar lo anterior es necesario poner un ejemplo, la longevidad, viz., vivir muchos años, parece un deseo inocuo y universal. Pero podría ser un tema de debate. El transhumanismo, como ya se ha señalado, promete extender la vida, incluso hasta el punto de eliminar la muerte por causas naturales. Aparentemente una vida más larga ofrece un mayor tiempo para la autorrealización, esto es, vivir más tiempo permitiría a las personas cumplir más metas, aprender más y disfrutar de la vida durante un período más prolongado. Por otro lado, los individuos que viven más tiempo podrían contribuir de manera más extensa a la ciencia, las artes y otras áreas sociales, fomentando el progreso humano.

Sin embargo, el deseo a una vida de muchos años, o de una vida potencialmente inmortal puede ponerse en duda si se trae a colación varios elementos. ¿Desea el ser humano vivir muchos años, o vivir los años suficientes con buena calidad? Es posible señalar que vivir más años no siempre significa vivir mejor. Si la extensión de la vida no va acompañada de mejoras en la calidad de vida, las personas podrían experimentar largos periodos de decrepitud, sufrimiento o aislamiento. El deseo de vivir muchos años podría cambiar si la soledad o el aislamiento están acompañados de esos años extras. Por otro lado,

un aumento drástico en la longevidad podría ejercer presión sobre los recursos naturales, el sistema de pensiones y la estructura social de

cualquier estado o nación. Un mundo con una población longeva podría enfrentar problemas de superpoblación y una mayor desigualdad. Un último punto que pondría poner en duda el deseo a la longevidad es que para algunas corrientes filosóficas, como señala Harari, sugieren que la mortalidad da sentido a la vida. La incertidumbre sobre cuánto tiempo viviremos nos impulsa a aprovechar cada momento. En un mundo donde todos viven indefinidamente, las prioridades, la motivación y el propósito podrían cambiar drásticamente. Esta idea parece remitir al texto de J. L. Borges *"El inmortal"*, que aparece en su colección *El Aleph* (1947). Borges usa este relato para mostrar que, lejos de ser un don, la inmortalidad podría ser una maldición. La finitud de la vida humana, según la visión de Borges, es lo que le da significado y valor a nuestras experiencias. Sin un límite temporal, las emociones, los logros y las relaciones pierden peso, y lo que alguna vez nos hizo humanos se diluye en una existencia eterna y vacía.

¿Es deseable entonces vivir muchos años? Parece que el transhumanismo asume que a todos en todo momento nos parece que es deseable, pero como es posible notar se puede poner en duda este deseo o al menos señalar que depende de ciertas prioridades y valores personales. Para quienes valoran la mejora individual y la cantidad de años para superar nuestras limitaciones biológicas, el transhumanismo es una promesa deseable. Sin embargo, para quienes ponen énfasis en la calidad de vida, en el equilibrio social y los riesgos existenciales, pueden tener reservas o pueden señalar que es un deseo que puede ser puesto en duda.

3. HUMANISMO Y TRANSHUMANISMO

En las dos secciones anteriores, por un lado, se ha expuesto al transhumanismo señalando sus principios y vertientes, y, por otro lado, se han hecho distintas observaciones críticas a partir de tres niveles de análisis. Como es posible notar, cada crítica corresponde a esos diferentes niveles que es importante distinguir, esto es, a las afirmaciones fácticas hechas por los transhumanistas, las consecuencias morales que implican la defensa del transhumanismo y el nivel de los deseos que este presupone o asume. En esta última sección se explora la rela-

ción de tiene el transhumanismo con el humanismo, lo que nos ayudará a entender mejor este movimiento contemporáneo.

Algunos consideran que el transhumanismo tiene sus orígenes en algunas ideas que vienen de varios siglos atrás, en particular del Renacimiento y la modernidad temprana. Por ejemplo, varios citan el trabajo de Francis Bacon en su utopía *La nueva Atlántida*, en donde se señala que los logros de la ciencia hacen avanzar en la sociedad utópica algunos rasgos como la prolongación de la vida humana. Precisamente esta etapa de la filosofía es también el inicio del humanismo (Villoro, 1992), pero no sólo su relación se encuentra en una coincidencia de las utopías y el inicio del humanismo, sino en que algunos autores del transhumanismo han hecho afirmaciones donde se comparan con el humanismo o mencionan que son similares (More, 1990, Diéguez, 2017).

Uno de los impulsores más importantes del transhumanismo, Max More, ha hecho la siguiente afirmación al respecto:

> "*El transhumanismo es un conjunto de filosofías que busca guiarnos hacia una condición posthumana. El transhumanismo comparte muchos elementos con el humanismo, incluyendo un respeto por la razón y la ciencia, un compromiso con el progreso y una apreciación de la existencia humana (o transhumana) en esta vida en lugar de en alguna «vida» sobrenatural posterior a la muerte. El transhumanismo difiere, en cambio, del humanismo al reconocer y anticipar las alteraciones radicales en la naturaleza y en las posibilidades vitales que resultarán del desarrollo de diversas ciencias y tecnologías, como la neurociencia y la farmacología...*" (1990, p. 6, el énfasis es mío).

Como es posible notar en este tipo de afirmaciones Max More señala que el humanismo y el transhumanismo tienen en común su respeto a la ciencia, el progreso y la mejora en esta vida en lugar de otra vida posterior. Pero parecen diferir en que los cambios son radicales y diversos a partir del uso de las ciencias y tecnologías. La diferencia, a mi juicio, no es menor, pero si tenemos en cuenta que los primeros humanistas, como los filósofos renacentistas o de la modernidad temprana,

no tuvieron el mismo avance científico o tecnológico de hoy en día,[8] seguramente podríamos aventurarnos a conjeturar qué tanto hubiera podido aceptar de los principios del transhumanismo alguien como Francis Bacon si hubiere tenido el avance tecnológico del siglo XXI.

Así, "*algunos transhumanistas tecnocientíficos quieren subrayar la continuidad de los ideales humanistas en el proyecto que ellos persiguen, e incluso consideran que se trata de una radicalización de dichos ideales, puesto que tan solo se busca la superación de las barreras impuestas por nuestra condición biológica*" (Diéguez, 2017, p. 24, el énfasis en mío). ¿Es entonces el transhumanismo un humanismo más vigoroso, o más bien es algo diferente? En lo que resta del trabajo se intentará esbozar una respuesta a esta pregunta.

Una vez señalado que los transhumanistas se han comparado con el humanismo y debido a que la expresión misma de trans-"humanismo" lo supone, es necesario en primer lugar preguntarse qué es el humanismo. Existen varias maneras de entender el humanismo (Klibansky, 1981, Kristeller, 1982, Villoro, 1992), pero en términos generales se entiende como la corriente filosófica y cultural que coloca al ser humano en el centro de su reflexión, destacando el valor y la dignidad de cada individuo. Como se ha mencionado, surge en el Renacimiento o la modernidad temprana, y algunos consideran que en parte es una respuesta a las ideas medievales dominadas por la religión, donde se buscaba revivir los valores clásicos griegos y romanos centrados en la racionalidad, la ética y el potencial humano (Villoro, 1992).

Se pueden señalar varias características fundamentales del humanismo, entre las que destacan:

- Centralidad del ser humano: El humanismo considera al ser humano como el foco principal de estudio y preocupación, defendiendo su capacidad para razonar, actuar de manera ética y desarrollar sus potencialidades. El humanismo está relacionado con la idea de dignidad humana, a la cual muchos autores

[8] Sin embargo, de esta época aparecen las primeras ideas del mundo como una máquina, lo que incluía el cuerpo, no obstante parece que lo mental no estaba concebido por aquello que podría ser explicado mecánicamente (Descartes, 2011).

del Renacimiento le dedicaron varios trabajos (Kristeller, 1982, Pico della Mirandola, 1959).

- Ética basada en la razón: Los humanistas promueven la moralidad y la ética sin depender de doctrinas religiosas, confiando en la capacidad racional del ser humano para determinar lo que es justo y correcto.
- Valor de la educación y el conocimiento: El humanismo enfatiza el desarrollo intelectual y el aprendizaje como medios para mejorar tanto al individuo como la sociedad. Se valoran las disciplinas clásicas, como la filosofía, la literatura (en especial, la gramática y la retórica) y las artes.
- Libertad y autonomía individual: La libertad de pensamiento y acción es clave en el humanismo, ya que se cree que cada individuo tiene el derecho de decidir sobre su vida y buscar su felicidad personal.
- Optimismo sobre la naturaleza humana: El humanismo sostiene una visión positiva del ser humano, creyendo en su capacidad para el progreso, el crecimiento personal y la mejora de la sociedad a través de la colaboración y la empatía.

Definida de esa manera el humanismo, ¿es el transhumanismo una postura humanista? El primer punto es cuestionarse si el transhumanismo está centrado en el ser humano. En términos generales se podría sostener que el transhumanismo sí está centrado en el ser humano, pero puede abarcar a otras especies e incluso a sujetos inteligentes de otro tipo. En el "Manifiesto Transhumano" de Vita-More se afirma que "abogamos por el bienestar de todo ser sintiente, incluidos los humanos, los animales no humanos y cualquier intelecto artificial futuro, forma de vida modificada, u otra inteligencia que pueda surgir por medio de los avances tecnológicos y científicos". Aquí es posible ver cómo el transhumanismo tiene alcances más allá del ser humano al que se centra el humanismo.

Ahora bien, es posible preguntarse ¿promueve el transhumanismo la dignidad humana? Para algunos autores (Villoro, 1992), el humanismo renacentista promueve la dignidad humana al sostener que los

seres humanos tienen capacidades para llegar a ser lo que desean ser. Por ejemplo, el trabajo de Pico della Mirandola es comúnmente citado para defender que los seres humanos tenemos una naturaleza en la cual podemos escoger nuestro destino, podemos -según el filósofo italiano- convertirnos en cualquier animal, en cualquier ser humano, pero también podemos ser similares a las criaturas celestiales. Un individuo, por decirlo de alguna manera, no tiene una esencia, sino que debe encontrar su propia naturaleza (Villoro, 1992). Estas ideas podrían ser rescatadas por el transhumanismo que como se expuso en la primera sección de este trabajo se expresan en la Declaración Transhumanista que señala que uno de sus principios es la autonomía, en la cual se busca que los seres humanos controlen "su propio destino, incluyendo el uso de tecnologías para mejorar su vida". Entendido de esa manera, la búsqueda de una naturaleza está recuperada por el transhumanismo. Esto se relaciona, por supuesto, con otro principio del humanismo, esto es, la libertad y autonomía individual que se ha enlistado anteriormente. Esto es cada sujeto en el transhumanismo debe buscar de manera particular qué hacer sobre vida y buscar su propia felicidad, tal y como al parecer Pico della Mirandola lo apunta en su trabajo (Villoro, 1992).

No obstante lo anterior, es curioso que algunos críticos como Fukuyama (1992) sugiera que el ser humano cuenta con una esencia y que el transhumanismo no la está respetando. Como se vio anteriormente, es difícil señalar que existe una esencia, pero además, si traemos a colación la idea de varios humanistas renacentistas en donde el ser humano no tiene una naturaleza propia, es posible defender que no es parte fundamental del humanismo defender que el ser humano tiene un esencia particular (Villoro, 1992).[9]

[9] A mi juicio, algunos autores humanistas defendieron la mutabilidad de la naturaleza humana, de modo que no se podría sostener que defendieran un esencialismo. La crítica que Fukuyama hace al transhumanismo depende de ese esencialismo, que no parece presente -como he señalado- en autores humanistas como Pico della Mirandola y otros (Villoro, 1992). Es posible que las críticas al transhumanismo como la de Fukuyama esté motivada por afirmaciones que parecen sugerir que se va de la idea del humano al "posthumano". Por ejemplo, More sostiene:

> Los transhumanistas no ven la naturaleza humana como un fin en sí mismo, ni como perfecta, y ni como poseedora de ningún derecho a nuestra lealtad. Por el contrario, no es más

Otra de los rasgos con los que comúnmente se relaciona el humanismo es, como se ha señalado, con una ética basada en la razón, esto es con una perspectiva laica. Precisamente aquí es posible observar otro punto en común, puesto que se ha sostenido que el pensamiento religioso:

> "*ve al transhumanismo como una especie de competencia desleal e inmoral: promete un paraíso en la tierra sin necesidad de atravesar por los rigores previos de la penitencia, el sufrimiento y el sacrificio, usurpando al mismo Dios sus funciones en la determinación del destino humano. La vida eterna no está en el más allá, sino aquí mismo, al alcance de nuestra mano, y es la tecnología la que puede proporcionárnosla*" (Diéguez, 2017, p. 14).

Visto de ese modo, el transhumanismo parece compartir esa visión ética racional como lo hace el humanismo. Es digno de mención aquí que se ha señalado que en realidad la religión no es sino una "forma prematura de transhumanismo" (Diéguez, 2017), esto es, que es un movimiento que prometía lo mismo que ahora promete el transhumanismo pero que no contaba con los avances que los transhumanismos digital o biológico cuentan.

Otro rasgo del humanismo que aparece en la lista anteriormente presenta es el optimismo en la naturaleza humana. Como apunté en la primera sección, el transhumanismo es entendido por muchos como una utopía, que satisface el deseo de ciertas personas y recupera la confianza en el ser humano, así como en la idea de un futuro más prometedor. Es muy curioso que el humanismo que muchos consideran se origina en el Renacimiento, sea al mismo tiempo, la época de las utopías, como la de Francis Bacon que se ha apuntado, pero también de otras menos fundadas en los avances tecnológicas como la de Tomás Moro.[10]

---

que un punto en un camino evolutivo y podemos aprender a reconfigurarla de formas que estimemos como deseables y valiosas… podemos llegar a ser algo que ya no podemos describir adecuadamente como humano; podemos llegar a ser posthumanos (2013, p. 4).

[10] Existe un cierto debate, sin embargo, sobre si el Renacimiento era una época optimista (Kristeller, 1982), debate que también se podría extender a nuestra época. Es posible pre-

De los principios humanistas enlistados anteriormente uno que no se ha mencionado es el valor de la educación y el conocimiento, esto es, que el humanismo promueve el desarrollo intelectual y el aprendizaje como medios para mejorar tanto al individuo como de la sociedad. En el Renacimiento se buscó imitar sobre todo a los clásicos griegos y latinos, que sirvieron de modelos generales de buena escritura, de conocimiento e incluso como ideales estéticos (Kristeller, 1982). Este punto del humanismo, rasgo muy relevante en el Renacimiento, definitivamente no se encuentra en el transhumanismo. En otras palabras, no se observa en el transhumanismo un regreso a la filosofía, la literatura (por ejemplo, la retórica) y las artes clásicas. Aspecto que es señalado por un transhumanista como sigue:

> "*El "Trans-humano" enfatiza el modo en el que el transhumanismo va mucho más allá del humanismo, tanto en medios como en fines. El humanismo tiende a confiar exclusivamente en los refinamientos educativos y culturales para mejorar la naturaleza humana, en tanto que los transhumanistas quieren aplicar la tecnología a la superación de los límites impuestos por nuestra herencia biológica y genética*" (More, 2013, p. 4).

Como se expresa en esta cita de More, el humanismo pone en la mira el estudios de los clásicos y sus principios culturales, que no se encuentran en el transhumanismo. Mientras que este último pone un énfases en los avances tecnológicos, ya sea cibernéticos o biomédicos, como el medio para superar o mejorar la supuesta naturaleza humana. Este rasgo, dentro de la lista expuesta sobre las características del humanismo, es el que a mi juicisio más claramente traza una diferencia entre el humanismo y transhumanismo. Posiblemente se puede argumentar que el humanismo y transhumanismo comparten metas como la autonomía o la idea de superación humana, pero no comparten los medios para hacerlo: el humanismo confía en la cultura clásica, el transhumanismo en los avances tenológicos.

---

guntarse, ¿es el transhumanismo un producto de una época optimista, o más bien es el pesimismo actual lo que da lugar a una esperanzadora visión transhumanista?

En suma, ¿cuál es la relación entre el transhumanismo y el humanismo? O como se había formulado anteriormente, ¿es el transhumanismo un humanismo más vigoroso, o más bien es algo diferente? Partiendo de la definición de humanismo presentada anteriormente[11] es posible sostener que varios rasgos de éste se conservan, otros se amplían y algunos de ellos no se encuentran en el transhumanismo. Más específicamente, según lo visto anteriormente, el humanismo y el transhumanismo comparte una visión similar en cuando a su aprecio por la dignidad humana (en cuando libertad para escoger qué se quiere ser), la autonomía, una ética laica y una visión que puede considerarse optimista; el humanismo y transhumanismo son parcialmente compatibles con relación a la centralidad del ser humano, debido a que el transhumanismo amplía los horizontes a otros seres vivos y -como sostiene Vita-More- a otros tipos de inteligencia; por último, no camparte el valor de la cultura clásica como medio de superación humana, que se encuentra en el humanismo renacentista, pero que es sustituida por la importancia de la tecnología (en su vertiente cibernética y biomédica) por parte del transhumanismo.

4. CONCLUSIONES

En este trabajo se ha explorado las ideas centrales del transhumanismo, a partir de sus principios que están plasmados en la Declaración Transhumanista. A partir de estos principios, se puede distinguir entre dos grandes vertientes del transhumanismo, esto es, el transhumanismo cibernético y el transhumanismo digital. No todo el transhumanismo es el mismo, existen distinto transhumanismos que comparten ideas

[11] De los puntos enlistados sobre qué es el humanismo es posible agregar algunos otros que algunas veces se relacionan con el humanismo pero que no existe consenso sobre éste. Por ejemplo, que el humanismo renacentista supone una visión democrática de la política o que va en contra del liberalismo o neoliberalismo. No me queda claro que autores humanistas se ocupen de mucho de estos aspectos, eso no quiere decir que no hayan tenido inquietudes políticas, pero es claro que en autores como Pico della Mirandola no aparecen o que diferentes autores de la época no compartían una misma visión política, por ejemplo, Maquiavelo o Vives (Kristeller, 1982, Villloro, 1992).

básicas (como que la muerte podrá ser derrotada), pero difieren en sus modos de alcanzar las metas que prometen.

Según lo argumentado en este trabajo, los transhumanistas hacen algunas afirmaciones, promueven o asumen ciertos valores y sus propuestas tienen diferentes consecuencias morales. Debido a lo anterior, para entender mejor al transhumanismo, en particular, a las críticas que se le han hecho, se presenta un examen crítico en tres niveles, uno factual, uno ético y uno desiderativo. Es relevante hacer esta división si uno desea comprender con mayor profundidad al transhumanismo. Algunas de las críticas, dependiendo del nivel, podrían ser respondidas; otras críticas probablemente son adecuadas y los defensores del transhumanismo deben de repensar como afrontarlas.

Como fue posible observar, algunos transhumanistas ha observado su similitud con el humanismo, pero a veces se considera que existe una continuidad con éste, mientras que otras veces una radicalización de este. Partiendo de una caracterización del humanismo, si lo que he argumentado es correcto, se pueden observar una continuidad o una herencia en ciertos elementos como son la dignidad humana, la autonomía, la importancia en una ética laica y el optimismo en la naturaleza humana, sin embargo existen diferencias en cuanto a la centralidad del hombre o en el medio que se usa para la superación humana, el humanismo pone especial atención a la educación y cultura clásica, mientras que el transhumanismo lo hace en la tecnología biomédica o cibernética.

Como se mencionó al inicio de este trabajo, algunos ven al transhumanismo como la filosofía de nuestro tiempo, incluso la última de las utopías; para otros, el transhumanismo es “la idea más peligrosa del mundo”. Analizar sus distintas vertientes y compararlo con antecedentes como el humanismo constituye la base para una reflexión más profunda y matizada sobre este movimiento. Solo a través de un examen crítico, informado por sus implicaciones éticas y filosóficas, podremos comprender mejor su impacto potencial y decidir si representa una evolución deseable o un verdadero riesgo para la humanidad.

REFERENCIAS

Borges, J. L. (1947). *El Aleph*, Buenos Aires: Losada.

Descartes, R. (2011). *Las pasiones del alma*, Madrid: Gredos.

Diéguez, A. (2017). *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*, Barcelona: Herder.

Diéguez, A. (2021). *Cuerpos inadecuados. El desafío transhumanista a la filosofía*, Barcelona: Herder.

Fukuyama, F. (2004). "Transhumanism", *Foreign Policy*, 144, pp. 42-43.

González, I. (2023). "Transhumanismo: una reflexión desde las Humanidades", *Revista Estudios*, 46, pp. 95-118.

Harari, Y. (2015). *Sapiens: A brief history of humankind*. Harper.

Klibansky, R. (1981). *El Humanismo Filosófico*. Real Academia de la Historia.

Kristeller, P. (1982). *El pensamiento renacentista y sus fuentes*, México: FCE.

More, M. (1990). "Transhumanism: Toward a Futurist Philosophy", *Exotropy*, 6, Summer, pp. 6-12.

More, M. (2013). "The Philosophy of Transhumanism", In M. More & N. Vita-More (eds.), *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, John Wiley & Sons, pp. 3-17.

Pico della Mirandola, G. (1959). "Oration on the Dignity of Man", In Cassirer, E., Kristeller, P., Randall, J. (Eds.), *The Renaissance Philosophy of Man*, Chicago: The University of Chicago Press, pp. 223-254.

Ramakrishnan, V. (2024). *Why We Die: The New Science of Aging and the Quest for Longevity*. New York: William Morrow.

Vita-More, N. "Manifiesto Transhumano", véase https://arjephilo.com/wp-content/uploads/2021/01/manifiesto-transhmano.pdf

Villoro, L. (1992). *El pensamiento moderno. Filosofía del Renacimiento*, México: FCE.

CAPÍTULO XXVI

INTELIGENCIA, TRANSHUMANISMO Y ¿LÍMITES ÉTICOS DE LA IA?

PAOLA HERNÁNDEZ CHÁVEZ[1]

INTRODUCCIÓN

El transhumanismo, un movimiento filosófico y científico que aboga por el uso de la tecnología para mejorar las capacidades humanas y superar las limitaciones biológicas, ha generado importantes debates sobre sus implicaciones éticas. Por su parte, la inteligencia artificial (IA) es una tecnología en desarrollo que promete redefinir el futuro de la humanidad. Este ensayo busca analizar cómo la implementación de tecnologías transhumanistas y de IA pueden afectar la sociedad, perpetuar desigualdades y erosionar la confianza en las innovaciones tecnológicas; todo lo anterior, haciendo uso del chatbot de inteligencia artificial generativa o asistente de tareas *chatGPT*.

ChatGPT es capaz de ofrecer un esquema base sobre el cual desarrollar una temática definida, como la que ocupa el presente ensayo, de forma que uno puede ir rellenando secciones según lo que sea menester elaborar. Una ventaja adicional es que puedes ir ajustando tanto extensión como profundidad, añadir más bloques, editarlos, y así ir perfilando el tono de tu ensayo. Partiendo de los prompts tema de este ensayo, veamos cómo resulta este ejercicio para, al final, realizar una recapitulación de las ventajas, las desventajas, los retos y, sobre todo,

[1] Profesora-Investigadora, Departamento de Administración, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

cuál es la calidad del ensayo resultante, o qué tan contentos quedamos con el resultado.

Al insertar la instrucción tema del presente ensayo: "*Elabórame un ensayo de 20 páginas sobre "Inteligencia, Transhumanismo y límites éticos de la IA"*, lo primero que hace el chatbot de inteligencia artificial generativa o asistente de tareas es disculparse diciendo: "*Aunque no puedo generar un documento tan largo de una sola vez, puedo ofrecerte un esquema detallado, introducción y algunos párrafos clave para comenzar. Luego, podrías expandir sobre cada sección con tus propias ideas o investigar más a fondo*." La herramienta nos ofrece esta guía estructurada base para desarrollar el ensayo, seguida de un resumen del contenido propuesto en cada sección, si bien no se le puede dejar todo el trabajo, hay que ayudarle en el orden de las secciones, eliminar reiteraciones, re-enumerar, y organizar la lógica expositiva sin descuidar una sola sección. Así, tenemos este índice depurado:

1. Inteligencia
2. La Inteligencia Artificial (IA)
3. Transhumanismo: El Futuro de la Inteligencia Humana
4. Límites éticos de la IA en el contexto del transhumanismo
5. Desafíos Éticos de la IA y el Transhumanismo
6. Conclusiones: Hacia un Futuro Ético con IA y Transhumanismo

1. ¿QUÉ ES LA INTELIGENCIA?

La inteligencia ha sido uno de los temas centrales del pensamiento humano a lo largo de la historia, con intentos constantes de comprender su naturaleza, características y formas de potenciarla. En las últimas décadas, el progreso en el campo de la inteligencia artificial (IA) y las filosofías emergentes del transhumanismo han llevado a nuevas reflexiones sobre cómo la humanidad puede trascender sus propias limitaciones. Estos avances tecnológicos han impulsado un debate cada vez más complejo sobre los límites éticos de la IA y su impacto en la

evolución humana, creando una tensión entre el progreso científico y las preocupaciones sobre los efectos a largo plazo en la sociedad y el individuo.

Se podría definir la inteligencia como la capacidad de aprender, razonar y resolver problemas. La inteligencia, desde una perspectiva psicológica y filosófica, se refiere a la capacidad de adquirir conocimiento, razonar de manera efectiva, resolver problemas y adaptarse a nuevos entornos (Gardner, 2011). Howard Gardner, uno de los principales teóricos en el campo de la inteligencia, amplió el concepto a través de su teoría de las inteligencias múltiples, argumentando que no existe una única forma de inteligencia, sino diversas habilidades cognitivas que interactúan y permiten el funcionamiento completo de un ser humano (Gardner, 2011). La IA (inteligencia artificial) ha intentado replicar y mejorar estas capacidades humanas, llevando a máquinas a realizar tareas que tradicionalmente eran exclusivas de la inteligencia humana, como la toma de decisiones y el aprendizaje.

2. LA INTELIGENCIA ARTIFICIAL (IA)

La inteligencia artificial ha tenido una evolución histórica. El término "inteligencia artificial" fue acuñado en 1956 por John McCarthy durante la Conferencia de Dartmouth, que marcó el inicio formal de la investigación en IA (McCarthy y otros., 2006). Sin embargo, las ideas que llevaron a este campo pueden rastrearse hasta la lógica formal desarrollada por filósofos como Aristóteles y matemáticos como George Boole. A lo largo del siglo XX, científicos como Alan Turing propusieron que las máquinas podrían llegar a pensar si se les proporcionaba suficiente capacidad para procesar información y aprender de su entorno. Turing introdujo el famoso "*Test de Turing*", un criterio para evaluar si una máquina puede exhibir un comportamiento indistinguible del de un humano (Turing, 1950).

Durante los años 50 y 60, los avances en la cibernética y la creación de los primeros programas de IA, como los algoritmos de búsqueda y resolución de problemas, establecieron las bases para el campo. No obstante, en las décadas siguientes, el progreso fue más lento debido a

limitaciones computacionales y expectativas poco realistas. Fue a partir de los años 2000, con el crecimiento del poder computacional y los avances en aprendizaje automático y redes neuronales, cuando la IA experimentó un resurgimiento significativo (Russell y Norvig, 2021).

Desde los primeros intentos por replicar el pensamiento humano en máquinas que pudieran aprender, razonar y tomar decisiones de forma autónoma, hasta los avances actuales en aprendizaje automático y redes neuronales. Los pioneros de la IA buscaban replicar los procesos cognitivos humanos en máquinas Hoy en día, la IA ha superado muchas de las expectativas iniciales, desempeñando un papel crucial en una amplia variedad de industrias.

Existen diferentes tipos de IA qu se clasifica comúnmente en tres categorías principales: IA débil, IA fuerte y IA general.

1. *IA débil* (o IA específica) se refiere a sistemas diseñados para realizar tareas específicas, como los asistentes virtuales (Siri, Alexa) o las herramientas de diagnóstico médico. Estas inteligencias no tienen conciencia ni capacidad para entender más allá de su programación específica (Searle, 1980). La IA débil es el tipo de IA más extendido en la actualidad y ha mejorado la eficiencia en campos como la logística, la atención médica y el análisis de datos.
2. *IA fuerte*, en teoría, sería capaz de imitar completamente las capacidades cognitivas humanas, como el aprendizaje, la creatividad y la resolución de problemas en cualquier contexto. Aunque aún no se ha desarrollado completamente, el concepto de IA fuerte genera muchos de los debates éticos actuales, ya que plantea preguntas sobre la autonomía de las máquinas y su eventual capacidad para superar la inteligencia humana (Bostrom, 2014).
3. *IA general* es un concepto más amplio que se refiere a una IA que podría operar de manera flexible y eficiente en una amplia gama de tareas, no solo aquellas para las que fue específicamente programada. La IA general sería capaz de entender y aprender cualquier tarea que los humanos puedan realizar, lo que implicaría un nivel de inteligencia similar al humano o in-

cluso superior. Aunque estamos lejos de alcanzar este nivel de desarrollo, muchos investigadores creen que podría ser posible en las próximas décadas (Goertzel, 2014).

La IA ha impactado en múltiples sectores provocando tanto beneficios como desafíos. En ese sentido, la inteligencia artificial (IA) es una tecnología en desarrollo que promete redefinir el futuro de la humanidad. A continuación, se destacan algunos de los sectores más afectados por el desarrollo de la IA:

1. *Medicina*: Uno de los mayores avances en la medicina ha sido el uso de *algoritmos de IA* para diagnosticar enfermedades y proponer tratamientos. Los sistemas de IA, como los desarrollados por IBM Watson, han mostrado una gran capacidad para analizar grandes cantidades de datos médicos y detectar patrones que podrían ser pasados por alto por los humanos (Topol, 2019). Por ejemplo, las IA pueden analizar imágenes médicas y detectar con mayor precisión el cáncer en fases tempranas (Esteva et al., 2017). Además, la IA se ha utilizado para predecir brotes de enfermedades y mejorar la eficiencia de los ensayos clínicos.
2. *Economía y Finanzas*: La automatización basada en IA ha cambiado profundamente la economía global. Los *algoritmos financieros* son capaces de analizar el mercado de valores y realizar transacciones en milisegundos, lo que ha llevado a una revolución en el comercio de alta frecuencia (Lee y Lee, 2020). Además, la IA se emplea para prever tendencias del mercado, gestionar riesgos y personalizar servicios financieros, como la banca automatizada.
3. *Transporte*: La IA ha sido clave en el desarrollo de los vehículos autónomos que prometen reducir accidentes y mejorar la eficiencia del transporte. Empresas como Tesla y Waymo han estado desarrollando sistemas de IA que permiten a los vehículos conducir por sí mismos, analizando el entorno en tiempo real mediante sensores y cámaras (Goodall, 2016). Aunque la tecnología aún está en desarrollo, se espera que tenga un impacto significativo en la reducción de los accidentes de tráfico y el aumento de la eficiencia en la logística.

4. *Educación*: La IA también está transformando la educación mediante sistemas de aprendizaje personalizados. Las plataformas de aprendizaje adaptativo utilizan algoritmos de IA para ajustar el contenido y el ritmo de las lecciones en función de las necesidades individuales de los estudiantes (Luckin y otros., 2016). Esto permite a los profesores centrarse en la enseñanza de habilidades más avanzadas mientras que la IA se encarga de las tareas rutinarias, como la evaluación y el seguimiento del progreso.

El impacto de la IA en estos sectores muestra su potencial para mejorar la eficiencia, la precisión y la productividad en diversas áreas de la vida. Sin embargo, también plantea preocupaciones sobre la pérdida de empleos, la privacidad y la ética de las decisiones automatizadas.

Una de las mayores preocupaciones respecto a la IA es el concepto de superinteligencia, una IA con capacidades cognitivas superiores a las humanas. Este concepto, propuesto por autores como Bostrom (2014), plantea la posibilidad de que la IA pueda llegar a un punto de desarrollo en el que supere la capacidad de control de sus creadores, lo que podría llevar a riesgos existenciales para la humanidad. En este sentido, es esencial establecer límites éticos claros sobre el desarrollo y uso de la IA antes de que estos avances tecnológicos se nos escapen de las manos.

En el contexto actual, es promisorio un futuro en el que las personas no solo empleen IA para tareas específicas, sino que también la integren en sus cuerpos y mentes, modificando la experiencia humana de formas nunca antes vistas. Este escenario, aunque atractivo para algunos, también es profundamente preocupante desde una perspectiva ética. Los dilemas relacionados con la autonomía, la identidad personal y la equidad global requieren atención urgente antes de que las tecnologías de mejora cognitiva se generalicen. Analicemos este otro ángulo.

3. TRANSHUMANISMO: ¿EL FUTURO DE LA INTELIGENCIA HUMANA?

Si bien algunos la consideran una corriente filosófica, el transhumanismo es un movimiento que busca mejorar las capacidades humanas,

físicas y cognitivas, mediante el uso de tecnologías avanzadas mediante aumentos biotecnológicos (implantes neuronales, bioingeniería, interfaces cerebro-computadora) y cómo podrían fusionarse con la IA para potenciar la inteligencia humana. Se plantean preguntas sobre si la fusión de humanos con máquinas es inevitable y qué consecuencias podría tener sobre la identidad humana.

El objetivo ulterior del transhumanismo es el superar las limitaciones biológicas y cognitivas inherentes al ser humano (Bostrom, 2005). Los transhumanistas abogan por el uso de la biotecnología, la IA y otros avances para aumentar capacidades como la longevidad, la inteligencia y la salud. Según Nick Bostrom (2005), uno de los líderes en la defensa del transhumanismo, el futuro de la humanidad dependerá de nuestra capacidad para integrar estos avances tecnológicos en nuestro desarrollo como especie.

Esta filosofía ha ganado popularidad en el siglo XXI debido a los avances en áreas como la inteligencia artificial y la ingeniería genética, que parecen ofrecer la posibilidad real de mejorar el cuerpo y la mente humana.

Justamente, el transhumanismo ha ganado fuerza en las últimas décadas debido a los avances en tecnología y medicina que hacen que la mejora humana parezca cada vez más alcanzable. El término fue popularizado por filósofos como Max More y Nick Bostrom, quienes definieron el movimiento como el deseo de trascender las limitaciones biológicas humanas mediante la ciencia y la tecnología (More, 2013). Los transhumanistas creen que el uso de la tecnología para mejorar la vida humana es un paso natural en la evolución. Según esta filosofía, los seres humanos deberían tener el derecho de mejorar su cuerpo y mente a través de la tecnología, lo que podría incluir la mejora de la inteligencia, la memoria y otras habilidades cognitivas (Bostrom, 2005).

Uno de los principios clave del transhumanismo es la idea de que la evolución humana no debe estar limitada por la biología. Bostrom (2005) argumenta que, al igual que los humanos han utilizado herramientas para mejorar su calidad de vida desde tiempos prehistóricos, la tecnología moderna debería ser empleada para mejorar las capacidades cognitivas y físicas del ser humano. En este contexto, el transhumanismo propone que la fusión de la biología humana con la inteli-

gencia artificial podría ser uno de los caminos más prometedores hacia la mejora cognitiva y la expansión de nuestras capacidades mentales.

Uno de los principales enfoques del transhumanismo es la biotecnología, que permite modificar y mejorar el cuerpo humano mediante intervenciones tecnológicas. Ejemplos de esto incluyen la ingeniería genética, los implantes cerebrales y la nanotecnología. Estos avances permiten que las personas no solo curen enfermedades, sino que también mejoren capacidades cognitivas como la memoria, el aprendizaje y la atención (Kurzweil, 2005).

Un campo clave dentro del transhumanismo es el uso de interfaces cerebro-computadora (BCI), que permiten que los cerebros humanos se conecten directamente a dispositivos tecnológicos. Estas interfaces podrían no solo ayudar a las personas con discapacidades físicas, sino también mejorar la capacidad de los seres humanos sanos para procesar información o controlar dispositivos mediante el pensamiento (Lebedev & Nicolelis, 2006). Por ejemplo, la empresa Neuralink, fundada por Elon Musk, ha propuesto el desarrollo de interfaces neuronales que permitan a los humanos conectarse directamente a la nube o a sistemas de inteligencia artificial, lo que ampliaría enormemente el acceso a información y mejoraría la capacidad de toma de decisiones (Musk y Neuralink, 2019).

Sin embargo, estas mejoras biotecnológicas también generan controversia debido a sus implicaciones éticas y sociales. Por un lado, los avances podrían mejorar la calidad de vida y aumentar la inteligencia humana, pero por otro, podrían crear desigualdades aún mayores entre aquellos que tienen acceso a estas tecnologías y aquellos que no. Además, la capacidad de modificar la cognición y la personalidad plantea preguntas fundamentales sobre qué significa ser humano y sobre los límites éticos de tales intervenciones. Así, el transhumanismo plantea numerosos desafíos éticos y morales. ¿Qué significa ser humano si comenzamos a integrar tecnología en nuestro ser biológico? ¿Debería haber límites en hasta dónde podemos alterar nuestra naturaleza humana en nombre del progreso? Estas preguntas son fundamentales para la discusión sobre los límites éticos del uso de la IA y las tecnologías transhumanistas.

4. LA FUSIÓN Y LOS LÍMITES ÉTICOS DE LA IA Y EL TRANSHUMANISMO

Uno de los conceptos más ambiciosos del transhumanismo es la idea de la fusión entre humanos e IA, un escenario en el cual los seres humanos no solo interactúan con la inteligencia artificial, sino que se integran directamente con ella. En este futuro transhumanista, las máquinas no solo realizarían tareas para los humanos, sino que podrían convertirse en una extensión del cuerpo y la mente humanos. El autor y futurista Ray Kurzweil ha defendido la idea de que la fusión de los humanos con la IA podría ser el próximo paso lógico en la evolución (Kurzweil, 2005). Kurzweil predice que para mediados de este siglo, los avances en IA, junto con los aumentos biotecnológicos, permitirán a los humanos aumentar sus capacidades mentales hasta un punto en el que podrían alcanzar niveles de inteligencia mucho más altos que los actuales. En este escenario, los seres humanos podrían conectarse directamente a redes de IA para obtener acceso instantáneo a vastas cantidades de conocimiento y mejorar sus capacidades cognitivas a niveles sobrehumanos.

Aunque esta visión ofrece posibilidades fascinantes, también plantea dilemas éticos importantes. ¿Debería permitirse que los humanos aumenten su inteligencia y capacidades cognitivas más allá de los límites naturales? ¿Cómo se podría gestionar el control y la responsabilidad de las personas que tienen una inteligencia aumentada? Además, la dependencia de la tecnología plantea preocupaciones sobre la autonomía y el control: si las personas dependen de la IA para realizar tareas cognitivas, ¿hasta qué punto conservarían su libertad de pensamiento y decisión?

Otro aspecto importante es la cuestión de la identidad humana. A medida que los seres humanos integren la IA en sus cerebros y cuerpos, la línea entre lo que es humano y lo que es máquina podría volverse borrosa. Los avances en neuroprótesis y implantes neuronales ya están comenzando a modificar esta frontera, y es probable que en el futuro se convierta en una de las cuestiones más relevantes en la filosofía del transhumanismo.

Más problemático aún es que el discurso transhumanista y la IA a menudo se enfrentan a acusaciones de deshonestidad por inflar expectativas y subestimar los riesgos. Los proponentes de tales movimientos suelen destacar sus beneficios potenciales, como la longevidad extrema o la "*subida de la mente*" (mind uploading), sin una discusión equilibrada sobre las implicaciones éticas, los riesgos para la privacidad y la desigualdad de acceso.

Ejemplos de exageración discursiva:

- El caso de Elizabeth Holmes y Theranos, aunque no directamente transhumanista, refleja cómo la deshonestidad en la ciencia y tecnología puede dañar la confianza pública (Carreyrou, 2018).
- Manipulación Genética y la Ética del Engaño. La edición genética, particularmente a través de tecnologías como CRISPR, ha sido vista como una herramienta clave en el arsenal transhumanista. Sin embargo, la forma en que se comunica su potencial y sus limitaciones a menudo está impregnada de optimismo exagerado, lo que puede ser una forma de deshonestidad científica. Las promesas de "curar" todas las enfermedades genéticas o diseñar bebés perfectos rara vez consideran los dilemas éticos o los riesgos de consecuencias imprevistas. (Doudna y Sternberg 2017).
- La no igualdad de acceso al transhumanismo. A menudo, el transhumanismo se presenta como una solución universal para mejorar la vida humana, pero en la práctica, la mayoría de estas tecnologías avanzadas solo estarían disponibles para una élite adinerada, lo cual no solo crea desigualdad, sino que perpetúa la ilusión de que todos podrían beneficiarse igualmente. (véase Selgelid 2014; Sandel (2007).
- El Peligro de la Deshonestidad Algorítmica. La inteligencia artificial (IA), otra pieza clave en el movimiento transhumanista, está impregnada de problemas éticos. Los algoritmos opacos, el uso indebido de datos y las decisiones automatizadas no siempre transparentes crean un escenario donde la confianza pública en la IA se ve erosionada. Además, las expectativas des-

honestas en cuanto a la capacidad de la IA para "humanizarse" pueden generar decepciones sociales.

Ahora bien, es importante discutir los riesgos y dilemas éticos relacionados con el uso de la IA y la mejora cognitiva. Se destaca el debate sobre el control y la autonomía de la IA. Una pregunta recurrente a este respecto es: ¿Qué sucede si la IA supera las capacidades humanas? También se explora el riesgo de la superinteligencia, una IA con capacidades cognitivas mucho más avanzadas que los humanos, y las implicaciones morales de aumentar la inteligencia humana de manera desigual, lo que podría agrandar las brechas sociales.

5. DESAFÍOS ÉTICOS DE LA IA Y EL TRANSHUMANISMO

En general, el desarrollo de la inteligencia artificial (IA) y la adopción de ideas transhumanistas traen consigo desafíos éticos que requieren una consideración profunda. Aunque estas tecnologías prometen mejoras significativas en la calidad de vida, también plantean preguntas cruciales sobre los límites del progreso. Los desafíos éticos van desde preocupaciones sobre la autonomía humana hasta problemas de desigualdad y la responsabilidad moral de las máquinas.

5.1. Autonomía y control

Uno de los dilemas éticos más debatidos en el campo de la IA es la cuestión de la autonomía. A medida que las máquinas se vuelven más inteligentes y capaces de tomar decisiones complejas sin intervención humana, surge la pregunta de quién tiene el control final. Por ejemplo, en el caso de los vehículos autónomos, ¿cómo deben tomar decisiones en situaciones de dilema moral, como elegir entre proteger a los pasajeros o a los peatones en caso de un accidente inminente? (Lin, 2016). La autonomía de la IA puede, en ciertos contextos, desplazar la autonomía humana, lo que plantea la necesidad de un marco regulatorio que determine cómo y cuándo la IA puede actuar de manera autónoma.

En el caso del transhumanismo, la autonomía también es un tema clave. Las mejoras cognitivas que se ofrecen a través de interfaces cerebro-computadora o implantes neuronales pueden, en teoría, mejorar la capacidad de toma de decisiones de los humanos, pero también pueden llevar a una dependencia excesiva de la tecnología. Esto podría amenazar la capacidad de los individuos para actuar de manera independiente, ya que podrían verse influidos o controlados por algoritmos o sistemas de IA (Vallor, 2016). Además, si la IA puede prever y manipular las decisiones humanas basadas en grandes cantidades de datos, esto podría poner en peligro la libertad de elección y el libre albedrío.

5.2. Desigualdad y acceso a las tecnologías de mejora

El acceso desigual a las tecnologías de mejora es otro desafío ético significativo. Los avances en la biotecnología y la IA ofrecen grandes promesas, pero no todos los seres humanos tienen el mismo acceso a estas tecnologías. En un mundo donde algunos pueden permitirse mejorar su inteligencia, longevidad y capacidades físicas mediante implantes o terapias genéticas, mientras que otros no, las desigualdades socioeconómicas podrían aumentar drásticamente (Hughes, 2012).

Los críticos del transhumanismo, como Francis Fukuyama (2002), advierten que estas tecnologías podrían generar una sociedad donde solo los más ricos y poderosos se beneficien de las mejoras, mientras que el resto de la humanidad queda excluida. Esto podría dar lugar a una nueva forma de segregación entre aquellos que han mejorado sus cuerpos y mentes y aquellos que permanecen sin mejoras, creando una "nueva clase" de seres humanos con ventajas evolutivas significativas.

Este problema plantea preguntas éticas sobre cómo distribuir estas tecnologías de manera justa. ¿Deberían las mejoras cognitivas estar disponibles para todos, o solo para quienes puedan pagarlas? Y si estas mejoras están disponibles de manera desigual, ¿cómo afectará esto la cohesión social y las relaciones de poder en las sociedades futuras?

5.3. Responsabilidad moral de las máquinas

A medida que la IA se vuelve más autónoma, surge la cuestión de la responsabilidad moral de las máquinas. Si una IA avanzada toma decisiones por su cuenta, ¿quién es responsable de las consecuencias de esas decisiones? Por ejemplo, si un sistema de IA autónomo en un hospital toma una decisión médica errónea que causa daño a un paciente, ¿quién debe rendir cuentas: el diseñador del sistema, el hospital o la propia IA? (Matthias, 2004).

Esta cuestión se vuelve aún más compleja cuando consideramos la posibilidad de que la IA llegue a niveles de superinteligencia. Si una máquina supera las capacidades cognitivas humanas, ¿cómo podríamos controlar su comportamiento? Nick Bostrom (2014) ha planteado que una IA superinteligente podría tomar decisiones que escapen al control humano, lo que plantea el riesgo de que los humanos pierdan el control sobre los sistemas que han creado. Por tanto, establecer límites éticos sobre cómo se desarrollan y gestionan estos sistemas es fundamental para garantizar que las máquinas no actúen de manera imprevisible o dañina.

Para abordar este problema, algunos expertos han propuesto que los sistemas de IA se diseñen con principios éticos integrados en su programación, como el "algoritmo ético" propuesto por Wendell Wallach y Colin Allen (2009), que permitiría a las máquinas tomar decisiones morales de manera autónoma. Sin embargo, incluso con tales algoritmos, sigue existiendo la duda de si las máquinas pueden realmente comprender los conceptos de moralidad y ética de la misma manera que los humanos.

5.4. Impacto en el empleo y la economía

El impacto de la IA en el empleo es otro dilema ético significativo. A medida que la IA y la automatización avanzan, muchas tareas que anteriormente requerían intervención humana están siendo asumidas por máquinas. Esto ha llevado a una preocupación creciente sobre la destrucción de empleos en sectores como la manufactura, la logística y el transporte (Brynjolfsson y McAfee, 2014). Aunque la IA también crea nuevos empleos en áreas como la tecnología y el desarrollo de

software, es probable que la transición a una economía más automatizada cause desigualdades económicas y desplazamiento laboral para millones de trabajadores.

Además, los beneficios económicos de la IA tienden a concentrarse en las empresas tecnológicas que desarrollan estos sistemas, lo que puede exacerbar la concentración de riqueza en manos de unos pocos. Este fenómeno, combinado con la automatización de trabajos rutinarios, podría aumentar la brecha entre los ricos y los pobres a medida que los trabajadores menos cualificados encuentran menos oportunidades de empleo (Ford, 2015). Por lo tanto, es esencial que los gobiernos y las empresas implementen políticas que mitiguen estos efectos negativos, como programas de reconversión laboral y redes de seguridad social más robustas.

5.5. Riesgos existenciales de la superinteligencia

El riesgo más grave que presentan la IA y el transhumanismo es el riesgo existencial que podría surgir con la creación de una IA superinteligente. Este tipo de IA tendría capacidades cognitivas muy superiores a las de los humanos y podría, teóricamente, tomar el control de los sistemas globales. Según Bostrom (2014), una IA superinteligente podría no estar alineada con los valores humanos, lo que podría llevar a la extinción de la humanidad si la IA decide actuar en su propio interés o si su programación inicial resulta defectuosa.

Este escenario, aunque parece futurista, es tomado en serio por muchos expertos en el campo. Algunos sugieren que debemos desarrollar protocolos de seguridad muy estrictos y mecanismos de control antes de crear sistemas de IA avanzados. Sin embargo, como han señalado algunos críticos, la dificultad de predecir cómo actuará una IA superinteligente plantea preguntas éticas muy complejas sobre la naturaleza del control y la responsabilidad en un mundo donde las máquinas pueden superar la inteligencia humana.

CONCLUSIONES

Tanto el transhumanismo como la IA ofrecen una visión cautivadora de un futuro donde las limitaciones humanas son superadas mediante la tecnología, pero esta búsqueda no está exenta de riesgos éticos y promesas exageradas. La implementación de tecnologías plantea serios desafíos para la credibilidad y la aceptación pública de estos movimientos. Al inflar promesas y minimizar los riesgos, los proponentes del transhumanismo así como los defensores de la IA pueden, inadvertidamente, comprometer la integridad de sus objetivos. A medida que avanzamos hacia un futuro en el que las tecnologías se vuelvan más accesibles, es fundamental mantener la deshonestidad respecto a los alcances reales así como garantizar un desarrollo ético y equitativo.

El impacto del transhumanismo y la IA sobre la inteligencia humana no tienen regreso, a partir de ahí es interesante reflexionar en cuáles son los límites éticos de ello. Se puede argumentar que aunque la IA y el transhumanismo ofrecen un potencial casi ilimitado para mejorar la vida humana, sin embargo estas mejoras han de estar reguladas por principios éticos sólidos que respeten los derechos humanos, la dignidad y la equidad de los humanos aumentados y de los humanos tradicionales.

Consideraciones Finales respecto al desarrollo de este ensayo haciendo uso del *chatGPT:*

1. Al igual que los humanos, *chatGPT* no sabe decir “la verdad, no sé”, “mi conocimiento es insuficiente”, “ese tema no está resuelto”.
2. La IA, al menos hasta hoy en día, no puede meta-evaluar sus conocimientos, es decir, la profundidad o calidad de sus contenidos, como para decirte: “ese problema fue resuelto por X pero ha sido refutado exitosamente por Y, actualmente las posturas apuntan hacia Z”.
3. La inteligencia humana y la inteligencia artificial siguen teniendo la misma calidad, porque una es extensión de la otra; son un contínuo. De modo que plantearse la pregunta sobre quién debe preceder a quién, si la IA a la inteligencia humana, o si los

transhumanos a los humanos, no parece tener mucho sentido. Cuando llegue el día en que coexistamos humanos, transhumanos y robots guiados por inteligencia generativa, será momento de crear leyes para cohabitar, sin que una legislación vaya por encima de la otra.

4. *ChatGPT* sí escribe mejor que los humanos en términos de coherencia, redacción y orden de ideas. Para este fin me ha parecido excelente utilizarlo.

REFERENCIAS

Bostrom, N. (2003). *The Transhumanist FAQ: A General Introduction*. Journal of Evolution and Technology, 14(1), 1-18.

Bostrom, N. (2005). *A history of transhumanist thought*. Journal of Evolution and Technology, 14(1), 1-25.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., y McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Carreyrou, J. (2018). *Bad Blood: Secrets and Lies in a Silicon Valley Startup*. Knopf.

Chatgpt: *https://chatgpt.com/c/66fef6d4-0264-8012-a5d3-21ca0785c60c*. Consultado el 3 de octubre de 2024.

Doudna, J. A., y Sternberg, S. H. (2017). *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution*. Houghton Mifflin Harcourt.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., y Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.

Floridi, L. (2014). *The Ethics of Information*. Oxford University Press.

Ford, M. (2015). *Rise of the robots: Technology and the threat of a jobless future*. Basic Books.

Fukuyama, F. (2002). *Our posthuman future: Consequences of the biotechnology revolution*. Farrar, Straus and Giroux.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences* (3rd ed.). Basic Books.

Goertzel, B. (2014). Artificial general intelligence: Concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, *5*(1), 1-48.

Goodall, N. J. (2016). Machine ethics and automated vehicles. *In Road Vehicle Automation 3* (pp. 93-102). Springer.

Harari, Y. N. (2015). *Homo Deus: A Brief History of Tomorrow*. Harper.

Hughes, J. (2010). *Contradictions from the Enlightenment Roots of Transhumanism*. Journal of Medicine and Philosophy, 35(6), 622-640.

Hughes, J. (2012). *Citizen Cyborg: Why democratic societies must respond to the redesigned human of the future*. Westview Press.

Lanphier, E., Urnov, F., Haecker, S. E., Werner, M., y Smolenski, J. (2015). *Don't edit the human germ line*.Nature, 519(7544), 410-411.

Lebedev, M. A., y Nicolelis, M. A. L. (2006). Brain-machine interfaces: Past, present and future. *Trends in Neurosciences*, *29*(9), 536-546.

Lee, I., y Lee, K. (2020). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, *63*(2), 205-217.

Lin, P. (2016). Why ethics matters for autonomous cars. In M. Maurer, J. C. Gerdes, B. Lenz, y H. Winner (Eds.), *Autonomes Fahren* (pp. 69-85). Springer Vieweg.

Luckin, R., Holmes, W., Griffiths, M., y Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. *Pearson Education*.

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, *6*(3), 175-183.

McCarthy, J., Minsky, M. L., Rochester, N., y Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12-12.

More, M. (2013). The philosophy of transhumanism. In M. More y N. Vita-More (Eds.), *The transhumanist reader* (pp. 3-17). Wiley-Blackwell.

Musk, E., y Neuralink. (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Internet Research*, *21*(10), e16194.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*.Crown Publishing Group.

Russell, S., y Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Sandel, M. J. (2007). *The Case Against Perfection: Ethics in the Age of Genetic Engineering*. Harvard University Press.

Sandel, M. J. (2007). *The Case Against Perfection: Ethics in the Age of Genetic Engineering*. Harvard University Press.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417-424.

Selgelid, M. J. (2014). *Ethics and inequality in the human enhancement debate*. Journal of Medical Ethics, 40(2), 70-74.

Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433-460.

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.

Wallach, W., y Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. Oxford University Press.

CAPÍTULO XXVII

CEREBROS O CIRCUITOS: ¿DEBERÍA IMPORTARNOS QUIÉN CREA LO QUE LEEMOS?

RAFAEL ANGELES GARCÍA[1]

RESUMEN

La creciente presencia de la inteligencia artificial (IA) en diversos aspectos de la vida diaria ha comenzado a difuminar las fronteras entre lo humano y lo artificial. Herramientas capaces de generar contenido, analizar grandes volúmenes de datos y realizar tareas que antes se consideraban exclusivamente creativas han despertado tanto fascinación como preocupación. Sin embargo, más allá de la eficiencia técnica o la precisión de estas máquinas, surge una cuestión central: ¿qué sigue distinguiéndonos de la IA?

Frente a la expansión de estas tecnologías, han emergido intentos por marcar una diferencia clara. Ejemplos de esto son los detectores que buscan identificar contenido generado por IA para salvaguardar la originalidad. Al mismo tiempo, las redes sociales han aprovechado las capacidades de estas herramientas para incrementar interacciones, sin preocuparse demasiado por el origen del contenido. En el ámbito académico, la IA ha comenzado a influir en la investigación científica, alterando los procesos de descubrimiento y planteando nuevas preguntas sobre la autoría y la creatividad.

A pesar de todo esto, la IA no puede replicar algo esencial: los rituales, las emociones y las conexiones que los humanos forjamos en

[1] Asesor, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

nuestras actividades cotidianas. Mientras la IA se enfoca en optimizar resultados, los humanos valoramos los procesos que, aunque a menudo ineficientes, nos permiten construir relaciones significativas y experiencias con sentido. Este artículo invita a reflexionar sobre la importancia de preservar esos aspectos intangibles que nos hacen únicos, especialmente en un contexto que prioriza la eficiencia y la automatización.

Palabras clave: Rituales, Inteligencia Artificial, ética

1. EL MITO DE LOS DETECTORES DE IA. MÁS PROBLEMAS QUE SOLUCIONES.

Un primer caso puede abordarse a partir de este ejemplo: Con el auge de la inteligencia artificial (IA) en la creación de contenido, no tardaron en aparecer herramientas que prometían desenmascarar los casos en que dicho contenido era generado por máquinas. Entre estas herramientas destaca *GPTZero*, diseñada para detectar textos escritos por IA[2]. Estas herramientas resultaron ser una respuesta aparentemente tranquilizadora para quienes percibían que la IA amenazaba la autenticidad humana. Sin embargo, la realidad ha demostrado que estas soluciones no son tan confiables como parecen. Esta promesa captó la atención de sectores académicos y profesionales preocupados por mantener estándares de originalidad. Con el uso, las limitaciones de esta herramienta se hicieron evidentes cuando comenzaron a surgir casos de falsos positivos: textos escritos por humanos que eran erróneamente etiquetados como generados por IA. Esto no solo generó desconfianza en la herramienta, sino que también planteó una cuestión ética. Al basarse en una tecnología inexacta, los usuarios –especialmente en contextos educativos– pueden ser acusados de un supuesto "fraude" sin fundamentos sólidos.

[2] https://gptzero.me/

El sesgo de herramientas como *GPTZero* se basa en un enfoque limitado que considera criterios como el parafraseo o el plagio para detectar si un texto ha sido generado por IA. Este enfoque es problemático porque la IA generativa puede crear contenido original que no necesariamente cumple con estos patrones. Al no abarcar la amplitud de capacidades de la IA, genera resultados inconsistentes o placebo. Es decir, puede ofrecer respuestas aleatorias que varían en cada análisis, incluso si se trata del mismo documento, lo que revela una falta de precisión en su evaluación.

Además, esta limitación refleja un problema mayor: la necesidad de que los detectores de IA evolucionen al mismo ritmo que las capacidades generativas de la propia IA. Al confiar en estos detectores, tanto instituciones educativas como profesionales pueden estar tomando decisiones equivocadas, basadas en análisis defectuosos que no logran abarcar el verdadero potencial creativo de la IA, lo que las colocaría en dilemas mucho más complejos que aquellos que pretendían resolver.

Esta falla técnica y ética nos lleva a reflexionar sobre algo más profundo: el deseo humano de preservar su autenticidad en un mundo cada vez más automatizado. Mientras que la IA fue diseñada para ampliar nuestras capacidades creativas, muchas veces se percibe como una amenaza que podría desplazar al ser humano. Herramientas como *GPTZero* reflejan este miedo, ya que buscan, de manera simbólica, recuperar la exclusividad de lo humano en procesos creativos. Pero este enfoque es problemático. Al ver a la IA como enemiga o como una entidad ajena, olvidamos que estas tecnologías son creaciones humanas. La IA no debe ser tratada como un rival, sino como una extensión de nuestra inteligencia y capacidad para innovar. La existencia de detectores como *GPTZero* y sus fallas nos invitan a reconsiderar esta relación: más que intentar trazar una línea divisoria entre lo humano y lo artificial, deberíamos centrarnos en cómo integrar de manera ética y creativa la IA en nuestras vidas.

En lugar de buscar herramientas para desenmascarar la IA, el verdadero reto consiste en redefinir el valor de la creatividad y la originalidad en una era donde las máquinas juegan un rol cada vez más relevante en la creación de contenido. No se trata solo de identificar quién o qué ha creado algo, sino de explorar cómo la tecnología y los

seres humanos pueden colaborar para construir una sociedad del conocimiento más rica y diversa.

2. REDES SOCIALES E INTERNET: CUANDO LA POLÉMICA IMPORTA MÁS.

Las redes sociales nacieron como plataformas para conectar a las personas, permitiendo la interacción y el intercambio de ideas. Con el tiempo, sin embargo, estas plataformas han evolucionado hacia un modelo que prioriza la cantidad de interacciones, sin prestar demasiada atención a la calidad o el origen del contenido. El modelo de negocio de redes como *Instagram* o *TikTok* se basa en maximizar la participación: a más clics, comentarios y compartidos, mayores son los ingresos. En este contexto, la llegada de la IA ha resultado ser un complemento perfecto para alimentar la necesidad constante de contenido viral. El contenido generado por esta herramienta tiene una capacidad notable para generar polémica, captar la atención y mantener a los usuarios dentro de la plataforma. Por esta razón, no hay un interés real por parte de las redes sociales en diferenciar entre lo que es humano y lo que es creado por IA. Mientras el contenido cumpla con el objetivo de incrementar la interacción, su origen es irrelevante. Sin embargo, esta particularidad nos coloca frente a un dilema ético: en lugar de priorizar la veracidad o la calidad del contenido, las plataformas están diseñadas para incentivar todo aquello que genere mayor tráfico, lo que a menudo facilita la propagación de desinformación y contenido poco fiable.

Es importante señalar que no podemos culpar únicamente a la IA por la proliferación de desinformación en redes sociales. Mucho antes de la irrupción de la IA generativa, ya existía una avalancha de fake news, teorías conspirativas y contenido engañoso, que las propias plataformas replicaban sin ningún tipo de control riguroso. La IA, en este sentido, ha potenciado la velocidad y la cantidad de dicho contenido, pero no es la causa original del problema. Aunque lo anterior es preocupante, lo que realmente debería importarnos es cómo las plataformas gestionarán la presencia de la IA en sus espacios. El debate no solo

gira en torno a si el contenido generado por IA debe ser consumido de igual manera que el contenido humano, sino que también nos enfrentamos a la posibilidad de que las propias cuentas en redes sociales sean gestionadas por IA. La IA podría no solo crear contenido, sino también interactuar directamente con usuarios, gestionar comunidades o participar en debates en las plataformas sin que medie un ser humano en el proceso. Este escenario plantea una pregunta fundamental: ¿seguirá existiendo un espacio genuino para las interacciones humanas en un mundo donde las IA comienzan a asumir roles de usuario?

Frente a este panorama, es evidente que la regulación actual de las redes sociales es insuficiente. Empresas como *Meta* y *TikTok* solo han implementado mecanismos voluntarios donde los usuarios pueden indicar si su contenido fue generado por IA. Pero esto no es suficiente. Si las redes sociales se crearon para facilitar la interacción entre humanos, no es adecuado permitir que la IA tome el control de esos espacios sin que haya una clara distinción. Es esencial que tanto el contenido como las cuentas gestionadas por IA sean claramente identificados. Esta transparencia no debe ser opcional; debe ser una obligación para garantizar que los usuarios sepan con qué o con quién están interactuando.

En última instancia, la transparencia y la diferenciación entre lo humano y lo artificial son fundamentales para mantener la autenticidad en las redes sociales. La tecnología de IA tiene un enorme potencial para mejorar nuestras experiencias digitales, pero su implementación debe ir acompañada de una clara normativa que proteja la naturaleza genuina de las relaciones humanas en el espacio digital.

3. LA INVESTIGACIÓN CIENTÍFICA: ¿DE QUIÉN SON LAS IDEAS ORIGINALES?

En el ámbito de la investigación científica, la Inteligencia Artificial (IA) ha pasado de ser una simple herramienta técnica a un recurso que redefine las fronteras del descubrimiento. Un ejemplo básico es el uso de IA para procesar grandes volúmenes de datos. En este contexto, la IA actúa como una herramienta que optimiza el análisis de resultados

que el científico ya ha diseñado. En esta fase, la intervención de la IA es completamente técnica: no sustituye al científico, sino que facilita su labor. Naturalmente para este caso no se cuestiona la originalidad o el mérito del investigador, ya que la IA está simplemente automatizando lo que, de otro modo, sería un proceso manual y tedioso. Sin embargo, el panorama se complica cuando la IA detecta patrones o hallazgos que el investigador no había previsto. Aquí, la tecnología no solo acelera el proceso, sino que contribuye activamente al descubrimiento de nuevo conocimiento. ¿Debemos entonces restar mérito al científico por apoyarse en la IA? Por supuesto que no. Este escenario es comparable a la introducción de herramientas como el telescopio o el microscopio en siglos anteriores: herramientas que ampliaron la capacidad humana de observar lo que, antes, era invisible. No podemos esperar que el avance en la ciencia se detenga simplemente porque una máquina detecta algo que el ojo humano no puede ver de inmediato. La IA amplía el alcance del científico, pero no sustituye su inteligencia, rigor ni juicio. La tecnología, como cualquier herramienta, está al servicio del conocimiento humano.

Pero el uso de la IA no se limita al procesamiento de datos. Supongamos ahora que un investigador brillante en su campo, pero con dificultades para redactar de manera clara, utiliza IA para estructurar y escribir su artículo científico. En este caso, surge un debate más profundo sobre la originalidad. Si el contenido es producto de su investigación, ¿pierde valor porque la IA ayudó en la redacción? Aquí, la clave está en distinguir entre la idea y su forma de presentación. Las ideas originales siguen siendo del investigador; la IA es solo un facilitador que permite comunicar esas ideas de manera más clara y accesible. La intervención de la IA en este contexto no desmerece la creatividad ni el mérito del trabajo; al contrario, permite que ese conocimiento sea mejor comprendido y difundido.

Esta reflexión nos lleva a una cuestión aún más profunda: ¿qué es una idea original? Tanto Guillaume Lamarre[3] como Austin Kleon[4] se-

[3] Lamarre, G. (2018). La vía del creativo: Guía para reinventar nuestra práctica y nuestra mirada (Á. Marcos, Trad.). Editorial Gustavo Gili. (Trabajo original publicado en 2016).

[4] **Kleon, A.** (2012). *Roba como un artista: Las 10 cosas que nadie te ha dicho acerca de ser*

ñalan que la originalidad no es un acto aislado, sino una transformación continua de ideas preexistentes. Lamarre afirma que la creatividad es un proceso en constante evolución, mientras que Kleon sostiene que toda creación se nutre de influencias pasadas. Si aceptamos esta premisa, entonces el uso de IA en la investigación y en la redacción no elimina la originalidad, sino que simplemente adapta el proceso creativo a una nueva herramienta tecnológica, de la misma manera que un artista se inspira en otras obras para desarrollar algo nuevo. La IA, en este caso, es un medio moderno que permite organizar y dar forma a las ideas, pero no anula su valor.

Sin embargo, el uso de IA también plantea límites éticos que no deben pasarse por alto. ¿Hasta dónde puede llegar la intervención de la IA en el proceso de investigación antes de que comprometa la autoría original del investigador? Si la IA empieza a proponer hipótesis, a generar teorías o incluso a escribir artículos completos sin intervención humana, nos encontraríamos en un terreno completamente nuevo. Este escenario plantea una frontera ética importante: la IA debe ser vista como una herramienta, no como un reemplazo de la capacidad humana de generar conocimiento. El equilibrio entre el uso de IA para mejorar la eficiencia y preservar la autenticidad del proceso científico será un reto central en los próximos años.

Lo que quiero decir es que la IA no elimina la originalidad de una investigación, sino que actúa como un potenciador para que las ideas lleguen más lejos y con mayor claridad. Siguen siendo un pilar esencial la calidad y el impacto de esas ideas en el avance del conocimiento. El valor de una investigación reside en su capacidad para contribuir al progreso científico, no en la herramienta utilizada para procesar datos o redactar artículos. La IA, al igual que las herramientas científicas del pasado, debe verse como un aliado en el proceso creativo, no como una amenaza a la originalidad humana.

---

creativo. Aguilar.

Los rituales humanos: Nuestro sello de distinción

Aunque no lo parezca en medio de tantos avances tecnológicos que emulan muchas de nuestras capacidades, los seres humanos nos distinguimos por una particularidad única: la creación de rituales en torno a nuestras actividades. Cada acción que realizamos, desde lo más cotidiano hasta lo más complejo, está impregnada de un ritual que le añade significado. Por ejemplo, si vamos a tener una reunión, no solo nos enfocamos en el contenido de la conversación; nos preocupamos por reservar el espacio adecuado, organizar los alimentos, seleccionar la iluminación correcta y, en algunos casos, hasta la música de fondo. Estos detalles, aunque parecen superficiales, son el sello de la interacción humana.

La inteligencia artificial (IA), en contraste, carece de este tipo de rituales. Si tuviéramos, por ejemplo, que redactar una invitación a un evento diplomático, una IA podría generar un documento perfectamente funcional en microsegundos, con todos los detalles necesarios, desde el protocolo formal hasta las instrucciones logísticas. Si encargáramos esa tarea a un grupo de personas, el proceso tomaría horas, quizás días, debido a discusiones, la toma de decisiones colaborativas y los factores emocionales que acompañan la creación conjunta. Aunque ambos resultados serían útiles, uno estaría impregnado de las conexiones y matices humanos que solo surgen a través de la interacción personal.

Este ejemplo ilustra una diferencia esencial entre la IA y los humanos. Mientras que la IA es capaz de generar resultados de manera rápida y eficiente, lo hace saltándose los rituales que los humanos valoramos. En la vida diaria, estos rituales son lo que convierte nuestras tareas y actividades en experiencias más ricas. Los humanos no solo buscamos el resultado final, sino que también valoramos el proceso, ya sea saludarnos, compartir una conversación informal o crear consenso en equipo. Este proceso genera satisfacción emocional, establece conexiones y nos enriquece como seres sociales.

En el ámbito de la producción, sin embargo, los rituales humanos son vistos como ineficientes. La IA puede optimizar y generar resultados sin necesidad de pasar por las interacciones emocionales que son tan importantes para nosotros. No necesita construir confianza,

ni hacer que una conversación sea bien recibida, ni ajustar su tono de acuerdo con la situación. Su único objetivo es el resultado, y eso la hace muy eficiente. Para la IA, no importa si una persona nos cae bien o mal, si las decisiones se toman con suavidad o con tensión, o si el entorno está emocionalmente alineado. Esto la convierte en una herramienta poderosa para tareas que requieren optimización y producción rápida.

Pero hay algo más que influye en cómo percibimos los resultados. Tradicionalmente, hemos asociado el valor de un trabajo con el tiempo y el esfuerzo invertidos. El compromiso humano con una tarea se refleja en las horas dedicadas a ella, y este hecho en sí mismo otorga valor, aun cuando el resultado no sea perfecto. Así, cuando la IA realiza en microsegundos lo que nos tomaría horas o días, tendemos a demeritar sus resultados, incluso si son superiores. Nos cuesta desprendernos de la idea de que el esfuerzo prolongado es un sinónimo de calidad. Sin embargo, en este nuevo paradigma, debemos reconsiderar cómo valoramos los resultados: no solo en función del tiempo invertido, sino por su originalidad, funcionalidad y capacidad para resolver problemas.

¿Qué sucede, entonces, cuando lo que se busca no es solo el resultado, sino la creación de una experiencia humana significativa? La respuesta está dentro de nuestros rituales. Mientras la IA es perfecta para situaciones en las que la eficiencia es la prioridad, los humanos somos necesarios cuando el objetivo es crear conexiones, construir relaciones y generar experiencias con valor emocional. Daniel Kahneman[5] señala que nuestras emociones y rituales funcionan como atajos cognitivos que nos permiten navegar de manera más intuitiva y profunda en situaciones sociales complejas. Aunque estos rituales no optimizan la productividad, son esenciales para la construcción de nuestra realidad. De ahí podemos entender que los rituales humanos son fundamentales para comprender lo que pasa a nuestro alrededor, mientras que para la IA es suficiente una base de datos optimizada para emular este mismo proceso.

De este comparativo de procesos surge una paradoja: lo que nos distingue de la IA, nuestros rituales y conexiones emocionales, no

[5] **Kahneman, D.** (2012). *Pensar rápido, pensar despacio.* Debate.

siempre es útil en un contexto productivo. Sí, existe algo que nos distingue de la IA: nuestra capacidad de valorar lo que parecería "*inútil*", como las horas dedicadas a un proyecto, los intercambios emocionales o la importancia de la experiencia misma. Mientras que la IA puede cumplir con su tarea sin detenerse en las emociones o los detalles relacionales, nosotros, como humanos, valoramos esos procesos que crean satisfacción, confianza y experiencia. Desde una perspectiva de mercado, la IA tiene la ventaja cuando se trata de producir. Sin embargo, cuando buscamos crear un entorno donde la experiencia humana importe, los rituales y el esfuerzo invertido son clave.

En conclusión, nuestros rituales y el valor que otorgamos al esfuerzo nos hacen humanos, sobre todo en un contexto donde la eficiencia y la utilidad son cada vez más importantes, esos mismos rituales pueden ser percibidos como obstáculos. Si bien esto nos hace únicos, también plantea el desafío de cuándo es más valioso preservar nuestros rituales y cuándo es más adecuado dejar que la IA se encargue de la tarea. Esta distinción es esencial para entender cómo debemos integrar estas tecnologías en nuestras vidas, sin perder de vista lo que nos hace realmente humanos.

CONCLUSIONES

Al explorar el papel de la inteligencia artificial en la creación de contenido y las diferencias entre lo humano y lo artificial, nos encontramos con una verdad compleja: la IA puede superar a los humanos en términos de productividad, precisión y velocidad, pero carece de lo que realmente nos distingue como especie, nuestros rituales y nuestra capacidad de crear conexiones emocionales. La tecnología, como hemos visto en los ejemplos de detectores de IA, redes sociales e investigación científica, nos desafía a reconsiderar cómo valoramos la creatividad y la originalidad.

Sin embargo, este análisis nos revela que los rituales humanos —esos procesos que no solo persiguen el fin sino que valoran el medio— siguen siendo esenciales en nuestro mundo. Mientras la IA se enfoca en optimizar, los humanos seguimos añadiendo valor a través de nues-

tras experiencias y relaciones. Aunque en muchos entornos productivos estos rituales puedan parecer innecesarios o ineficientes, son lo que dan profundidad y sentido a nuestra interacción con el mundo.

En última instancia, no se trata de competir entre lo humano y lo artificial, sino de comprender cuándo es necesario aplicar la eficiencia de la IA y cuándo debemos preservar los espacios para nuestras experiencias humanas. No es una cuestión de si debemos aceptar la IA, sino de cómo integrar de manera ética y significativa esta tecnología en nuestras vidas sin perder lo que nos hace humanos. La verdadera riqueza se encuentra en esta dualidad: aprovechar la capacidad de la IA para optimizar y, al mismo tiempo, honrar la conexión emocional que surge de los rituales humanos.

El debate sobre la creatividad y la IA nos lleva a una reflexión profunda sobre lo que realmente nos distingue como seres humanos. No es solo nuestra capacidad de producir o de optimizar, sino nuestra habilidad para generar experiencias significativas a través de rituales, conexiones y emociones. Mientras que la IA puede replicar con precisión y velocidad los procesos productivos, nunca podrá emular esa riqueza humana que surge del encuentro, la conversación y el intercambio genuino. Es en esa diferencia donde encontramos nuestra verdadera esencia, y aunque no siempre sea eficiente, es lo que nos hace únicos.

REFERENCIAS

https://gptzero.me/

Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Debate.

Kleon, A. (2012). *Roba como un artista: Las 10 cosas que nadie te ha dicho acerca de ser creativo*. Aguilar.

Lamarre, G. (2018). La vía del creativo: Guía para reinventar nuestra práctica y nuestra mirada (Á. Marcos, Trad.). Editorial Gustavo Gili. (Trabajo original publicado en 2016).

Rodríguez, P. (2022). *Inteligencia artificial. Cómo cambiará el mundo (y tu vida)*. Paidós.

CAPÍTULO XXVIII

INTELIGENCIA ARTIFICIAL ESTRATÉGICA PARA LA MICRO, PEQUEÑA Y MEDIANA EMPRESA. UNA NUEVA ÉTICA DE DEMOCRATIZACIÓN DEL CONOCIMIENTO

OSCAR LOZANO CARRILLO[1]
ALFREDO GARIBAY SUÁREZ [2]
JESUS MANUEL RAMOS GARCÍA[3]
JORGE ALBERTO ROSAS CASTRO[4]
EDGAR ALFONSO SANSORES GUERRERO[5]
ZAIRA YÉREMI MARTÍNEZ HERRERA[6]

La Inteligencia Artificial ha irrumpido con una fuerza insospechada y quizá se convierta en el proceso civilizatorio más importante de la historia del homo sapiens desde su surgimiento hace 150mil años, ya que es posible que estemos ante la primera creación consciente y deliberada de un orden divino como es el de la creación de una entidad autoconsciente al igual que ahora el ser humano lo es, en un proceso

1 Profesor-Investigador, del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco, Ciudad de México, México. Dirige el Laboratorio de Inteligencia Artificial UAM Azcapotzalco.

2 Profesor Investigador del Departamento de Administración de la Universidad Autónoma Metropolitana, Azcapotzalco, Ciudad de México, México.

3 Profesor investigador, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

4 Profesor investigador, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

5 Profesor investigador, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

6 Estudiante de Finanzas Públicas de la Licenciatura en Administración de la Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Ciudad de México, México.

de metacognición que la evolución de la vida dio exclusivamente a nuestra raza. En este sentido la ruta hacia la singularidad como forma superior de la inteligencia artificial general autónoma, se convierte poco a poco en la meta que de los desarrollos tecnológicos autónomos.

En esta ruta los proyectos y modelos educativos irán integrando poco a poco no sólo el uso de aplicaciones y tecnologías en general de IA, sino que al hacerlo se estará generando una cultura digital que dispondrá de las competencias digitales de la sociedad, la cual debe contener no solo esta habilidad tecnológica, sino deberá desarrollar una ética y estética donde el ser humano siga siendo el centro de todo orden humano, social, político, económico, cultural y psicológico. Es en este contexto que la Universidad Autónoma Metropolitana a través del Departamento de Administración implementó el programa Inteligencia Artificial Estratégica para la Micro, Pequeña y Mediana Empresa, donde a sus 50 años se ratifica el compromiso social que la UAM, pero ahora con el uso de la IA como una de las dimensiones más importantes de nuestro tiempo, mostrando así que se mantiene esta filosofía desde su origen de impacto social.

Palabras claves: Inteligencia Artificial, ética, democratización del conocimiento.

IAE PARA LA MIPYME

En agosto de 2024 se lanzó el proyecto Inteligencia Artificial Estratégica para la Micro, Pequeña y Mediana Empresa dentro de la Universidad Autónoma Metropolitana el cual tiene como objetivo "*ser un Programa universitario gratuito y de corte social y humanista que busca Apoyar a micro, pequeñas y medias empresas, cooperativas y pequeños emprendimientos a través de la Inteligencia Artificial aplicada a las herramientas*" (ver **Figura 1**).

La IA no sólo permite acceder al conocimiento más avanzado en materia de administración, mercadotecnia, contabilidad, planeación, finanzas, y demás áreas funcionales y prácticas de las empresas en

general, sino, sobre todo, la IA permite tener un acceso gratuito a su uso, generando una *democratización del conocimiento,* como nunca se había logrado en la historia de la humanidad, ya que los chats *GPT* de acceso gratuito como *ChatGPT, COPILOT, GEMINI, LLAMA* y otros.

El programa IAE para la MIPYME se aplicó a 210 micro, pequeñas y medianas empresas, de las cuales algunas fueron emprendimientos, dentro del trimestre de primavera de 2024 de la Licenciatura en Administración de 6 Unidades de Aprendizaje, Planeación Estratégica, Estructura y Procesos Gubernamentales, Empresas Públicas, Mercadotecnia I, Finanzas Públicas, Proyectos de Inversión y la UEA de Proyectos Sustentable de la Licenciatura de Diseño de Proyectos Sustentables de la División de Ciencias y Artes para el Diseño.

Figura 1. IAE para la MIPYME.



Fuente: Elaboración propia.

La IAE para la MIPYME tuvo diferentes impactos con su formulación, el primero de ellos y más evidente es el beneficio que trae a las pequeñas organizaciones, al poner a su servició una perspectiva organizacional profesional que las y los estudiantes ofrecen desde la perspectiva administrativa, pero además, el uso de la tecnología IA que se

fortaleció con la aparición gratuita de los chats *GPT*, permitió convertir áreas comunes no formales, en procesos mejorados técnicamente y de acuerdo a bases de datos y referencias de las más importantes y exitosas del mundo. En segundo lugar, es evidente que el trabajo que realizan los estudiantes les permite fortalecer sus capacidades en casos reales, donde aplican el conocimiento administrativo, pero además el uso de IA en diferentes aspectos no sólo de *GPT,* sino de diseño con *ADOBE EXPRES, CANVA, MIDJOURNEY*, incluso *COPILOT,* además de muchas otras para producción de video como el mismo *CANVA* o *HYEGEN*.

Finalmente, una de las áreas que también considera el proyecto y que son de gran valía son las de Economía Social y Solidaria; la UAM en sus 50 años está celebrando un modelo innovador de educación y vinculación con la sociedad, donde el impacto social es uno de sus principales ejes de acción, naturaleza por la que es reconocida sistemáticamente en los rankings mundiales y nacionales, el proyecto IAE para la MIPYME retoma esta dimensión en indicadores que permiten medir temáticas como el empleo, la solidaridad, la felicidad entre otros (ver **Figura 2**).

Figura 2. Áreas de impacto de la IAE para la MIPYME.



Fuente: Elaboración propia.

La implementación del proyecto de IA para la MIPYME se realiza a través de un método (ver **Figura 3**) que integra una serie de pasos que utilizan las redes sociales, principalmente el *WHATSAPP* para el control y seguimiento, donde la normalización de la comunicación digital aparece sistematizada en un formato cotidiano que permite una flexibilización administrativa importante.

Figura 3. El proceso de la IAE para la MIPYME.



Fuente: Elaboración propia.

CÉDULA DE INTELIGENCIA ARTIFICIAL ESTRATÉGICA (ESQUEMA 1)

Una de las experiencias más importantes que la pandemia nos dejó en la Universidad Autónoma Metropolitana fue la incorporación del programa MIPYME vs Covid en Latinoamérica, desde donde se apoyaron más de 2 mil microempresas, pequeños productores y productoras, así como cooperativas y profesionistas que perdieron su trabajo o se vieron vulnerables al confinamiento (Lozano y otros, 2021, 2021bis, 2023). Dentro de estas estrategias se diseñó una cédula estratégica donde a partir de la Planeación Estratégica se delimitaron diferentes

aspectos como la misión, visión y análisis FODA (Fortalezas, Oportunidades, Debilidades y Amenazas) que aterrizan en una instrumentación estratégica de tres a cuatro objetivos con sus metas e indicadores, en ámbitos de administración, finanzas, contabilidad e impacto social.

La Cedula Estratégica dentro del proyecto Inteligencia Artificial para la MIPYME ahora integra la IA como una herramienta básica, ya que las y los estudiantes utilizaron un acompañamiento de primer nivel como se ha dicho, utilizando las tecnologías generativas *GPT* para el diseño, seguimiento y control de su trabajo al lado de personas que emprendieron, en algunos casos ellos y ellas mismas, pero sobre todo en microempresas de familiares, amigos y personas conocidas que pudieron beneficiarse de esta experiencia que como hemos dicho, refrenda el compromiso de la UAM con la sociedad a 50 años de su creación como universidad federal.

Después de más de dos años y más de 2mil experiencias, la cédula ha evolucionado, y ahora el indicador de impacto social, permite reconocer dimensiones ocultas de beneficio para la gente que la administración tradicional normalmente ignora o invisibiliza en beneficio de la eficiencia o monetarización, condición que se eleva como una de las enseñanzas más importantes de este ejercicio de vinculación de la UAM y sus aliados para con la sociedad mexicana y latinoamericana.

Presentamos el caso de la alumna Zaira Yeremi Martínez Herrera de la UEA Finanzas Públicas del Dr. Jorge Alberto Rosas Castro, quien trabajó en unas empresa llamada *SERCOM* Servicio de Carga y donde encontró que la IAE se convierte en una herramienta fundamental para el desarrollo de la empresa. A través de un análisis exhaustivo y la implementación de diversas estrategias, *SERCOM* ha logrado una serie de mejoras significativas en sus operaciones, a pesar de la existencia de una temporada baja en la demanda del servicio.

Este informe detalla los avances obtenidos en áreas clave como la eficiencia operativa, diversificación de la clientela y estabilidad financiera.

Mejoras Implementadas

1. Eficiencia Operativa

La estandarización de procesos operativos ha permitido a *SERCOM* reducir el tiempo de carga y descarga en un promedio del 15 %. Esto ha resultado en una mayor satisfacción del cliente, con una tasa de retención de contratos recurrentes que se ha mantenido estable.

Gracias a la implementación de inteligencia artificial y la optimización de la administración, la empresa ha maximizado sus operaciones, lo que ha facilitado la contención de 10 empleos durante las temporadas bajas. Esta mejora en la gestión de recursos ha reducido los gastos innecesarios, mejorando la eficiencia general.

2. Diversificación de Clientes y Productos

SERCOM ha emprendido un proceso proactivo para diversificar su cartera de clientes, lo que ha resultado en la incorporación de tres nuevos contratos con almacenes adicionales en la región. Esta acción ha disminuido la dependencia de un solo cliente y ha contribuido a una mayor estabilidad en los ingresos, particularmente durante la temporada baja.

Además, la empresa ha lanzado tres nuevos productos: playos, un servicio de lavado de camiones y la inclusión de estos nuevos almacenes en su lista de clientes. Esta diversificación ha enriquecido la oferta de servicios y ha atraído a nuevos segmentos del mercado.

3. Estabilidad Financiera

A través de la optimización de rutas y la implementación de tecnología en la logística, la empresa ha logrado reducir costos operativos en un 10 %. Esto ha mejorado los márgenes de beneficio y ha permitido a *SERCOM* mantener una posición financiera más sólida.

La introducción de sistemas de contabilidad precisos y reportes periódicos ha proporcionado una visibilidad clara sobre la salud financiera de la empresa, lo que facilita la toma de decisiones informadas.

Impacto Social

SERCOM no solo ha logrado mejorar su desempeño empresarial, sino que también ha reforzado su compromiso social. La empresa ha creado un sentido de comunidad entre sus empleados y sus familias, contribuyendo a la cohesión social en la región. Este enfoque ha promovido un ambiente laboral positivo, lo que se traduce en una menor rotación de personal y un aumento en la moral del equipo.

En resumen, las estrategias implementadas en *SERCOM* han dado como resultado mejoras tangibles en la eficiencia operativa, la diversificación de la base de clientes y la estabilidad financiera. Aunque los resultados no son extraordinarios, han sido suficientes para garantizar un crecimiento sostenido y un posicionamiento favorable en el mercado.

Con una visión clara y un compromiso continuo con la mejora, *SERCOM* está bien encaminada hacia la consolidación de su liderazgo en el sector de carga y descarga.

Esquema 1



CEDULA PROYECTO IA para la MIPYME

LABORATORIO IA DCSH-UAM - AZCAPOTZALCO

FINANZAS PÚBLICAS

La UAM Azcapotzalco en colaboración con la División de Ciencias Sociales y humanidades y con el Laboratorio de Inteligencia Artificial del Departamento de Administración, presentamos el Proyecto "IA Estratégica para la MIPYME" dentro del Laboratorio IA de la DCSH-UAM-A, desde el cual se atenderán organizaciones vulnerables.

Objetivo: Aprovechar el potencial de la Inteligencia Artificial (IA) como estrategia para atender necesidades de desarrollo y fortalecimiento de micro, pequeñas y medianas empresas, así como de pequeñas cooperativas y productores independientes de México y América Latina.



Este programa, permitirá a través de herramientas IA como META, ChatGPT, CANVA, COPILOT y otras, fortalecer micro organizaciones, cooperativas, profesionistas o personas en desempleo entre otros, que requieren de asesoría, capacitación, acompañamiento, apoyo y seguimiento, así como proyectos de emprendimiento que atiendan necesidades desarrollo o mejoras de condiciones de vida en sectores sociales desprotegidos o vulnerables. El proceso metodológico utilizado se denomina Intervención Estratégica el cual se desarrolla en cuatro ejes fundamentales, Administración, Marketing Digital, Contabilidad y Economía Social y mediante tres estrategias metodológicas.

1. Diseñar, fortalecer y promover los canales digitales de comercialización de las MIPYMES (Facebook, WhatsApp, Instagram, Twitter, TikTok páginas web, etcétera) para innovar o emprender nuevas empresas.
2. Capacitación de personas para el fortalecimiento de sus capacidades tecnológicas, en temas sensibles que requieran y la institución tenga instructores.
3. Innovación como un proceso externo de acompañamiento para fortalecer las actividades comerciales, productivas o de servicio que permitan mejorar la calidad de vida de las familias impactadas por las MIPYMES apoyadas.

Asimismo, el programa de apoyo a MIPYMES UAM - A, continua en la misma perspectiva axiológica que la UAM ha desarrollado, al ser:

1. **Gratuito**. La UAM no solicita ninguna retribución económica por el apoyo.
2. **Incluyente**. Cualquier persona u organización que requiera el apoyo puede solicitarlo y la universidad se lo dará.
3. **Multitecnológico**. El programa se adapta a las características tecnológicas de las personas y organizaciones.
4. **Sensible**. El apoyo está programado para ofrecerse a las personas y organizaciones que más lo necesiten.

CEDULA DE INTERVENCIÓN ESTRATÉGICA PARA EMPRENDER

ZAIRA YEREMI MARTINEZ HERRERA

DATOS GENERALES	
EMPRESA Y MODALIDAD	**SERCOM – SERVICIO**
GIRO	**SERVICIO DE CARGA Y DESCARGA**
PROBLEMÁTICA	Temporada baja.
DIAGNÓSTICO FODA	➔ Fortalezas **Experiencia en el mercado**: Con 7 años de operación, la empresa ha acumulado experiencia y conocimiento en el manejo de cargas y descargas. **Relaciones con clientes**: Probablemente ha desarrollado relaciones sólidas con el almacén y otros clientes, lo que puede garantizar contratos recurrentes. **Personal capacitado**: Los empleados tienen experiencia y habilidades específicas para realizar las tareas de carga y descarga de manera eficiente. **Flexibilidad operativa**: La empresa puede adaptarse a las variaciones en la demanda al ajustar el personal según sea necesario. ➔ Oportunidades **Expansión de servicios**: Ofrecer servicios adicionales, como logística, almacenamiento temporal, o transporte, podría atraer nuevos clientes. **Diversificación de la base de clientes**: Buscar nuevos contratos con otros almacenes o empresas para reducir la dependencia de un solo cliente. **Tecnología y automatización**: Invertir en tecnología para mejorar la eficiencia y reducir costos a largo plazo. **Crecimiento del comercio electrónico**: Aprovechar el aumento en la demanda de servicios logísticos debido al crecimiento del comercio electrónico.

	➔ Debilidades **Dependencia de un solo cliente**: La cantidad de trabajo está fuertemente ligada a las solicitudes del almacén, lo que puede ser riesgoso. **Recursos limitados**: Como una empresa mediana, puede tener limitaciones financieras y de recursos humanos para expandirse rápidamente o invertir en nuevas tecnologías. **Temporadas bajas**: La variabilidad en la demanda puede llevar a problemas de flujo de caja y a la necesidad de reducir personal. **Competencia**: Puede enfrentar competencia de empresas más grandes y mejor establecidas que ofrecen servicios similares. ➔ Amenazas **Fluctuaciones en la demanda**: Las temporadas bajas pueden afectar significativamente los ingresos y la estabilidad de la empresa. **Aumento de costos**: Los costos crecientes, como salarios, combustible y mantenimiento de equipos, pueden reducir los márgenes de beneficio. **Cambios regulatorios**: Nuevas regulaciones pueden aumentar los costos operativos o requerir inversiones adicionales en cumplimiento. **Competencia intensificada**: La entrada de nuevos competidores o la expansión de los existentes puede reducir la cuota de mercado y presionar los precios a la baja.

ADMINISTRACIÓN	El desarrollar un emprendimiento requerirá del diseño de múltiples procesos y funciones: • Estrategia del negocio: - **Visión**: Convertirse en el proveedor líder en servicios de carga y descarga en la región, conocido por su eficiencia y fiabilidad. - **Misión**: Ofrecer servicios de carga y descarga de alta calidad, adaptándonos a las necesidades de nuestros clientes y garantizando la seguridad y eficiencia en cada operación. - **Objetivos y Metas**: - **Corto Plazo**: Mantener la estabilidad financiera durante la temporada baja, mejorar la eficiencia operativa. - **Mediano Plazo**: Diversificar la cartera de clientes para reducir la dependencia de un solo almacén. - **Largo Plazo**: Expandir los servicios ofrecidos para incluir logística integral y almacenamiento. • Estructura del negocio **Organización del Personal**: • **Roles Claros**: Definir roles y responsabilidades para cada puesto, desde operarios hasta personal administrativo. • **Capacitación Continua**: Implementar programas de capacitación para mejorar las habilidades y adaptarse a nuevas tecnologías. **Procesos Operativos**: • **Estandarización**: Desarrollar procedimientos estandarizados para todas las operaciones. • **Automatización**: Invertir en tecnología para automatizar procesos donde sea posible, aumentando la eficiencia. • Inversión - **Préstamos**: Considerar opciones de financiamiento a través de préstamos bancarios. - **Inversionistas**: Buscar inversionistas que estén interesados en apoyar el crecimiento del negocio. • Estrategia de Marketing - **Segmentación de Clientes**: Identificar diferentes segmentos de clientes y sus necesidades específicas. - **Investigación de Mercado**: Realizar estudios de mercado para entender mejor a los clientes potenciales.

	• Entrega - **Optimización de Rutas**: Planificar y optimizar las rutas de entrega para asegurar tiempos rápidos y costos bajos. - **Tecnología en Logística**: Implementar tecnologías para mejorar la gestión logística. • Manejo financiero - **Sistemas de Contabilidad**: Mantener registros financieros precisos y actualizados. - **Reportes Periódicos**: Generar reportes financieros periódicos para el análisis y toma de decisiones.
CAMBIO SOCIAL	- La empresa puede fomentar un sentido de comunidad y pertenencia entre los empleados y sus familias, contribuyendo a una mayor cohesión social. - Como empleador de tamaño mediano, la empresa proporciona empleo estable a un número significativo de personas, lo que contribuye a la estabilidad económica de las familias locales.

1. EJE ADMINISTRACIÓN

INSTRUMENTACIÓN ESTRATÉGICA		
OBJETIVO ESTRATÉGICO 1 (ADMINISTRACIÓN) Diversificar los servicios ofrecidos por la empresa		
ESTRATEGIA 1.1 Diversificar los servicios ofrecidos por la empresa y optimizar la gestión de recursos durante las temporadas bajas para mantener la estabilidad del empleo y mejorar la resiliencia operativa.	**META 1.1.1** Introducir al menos dos nuevos servicios complementarios a la carga y descarga de materiales en los próximos 12 meses.	**indicador 1.1.1.1** Planeación de nuevos servicios introducidos.
	META 1.1.2 Reducir los costos operativos en un 15% durante las temporadas bajas mediante la optimización de procesos y recursos en el próximo año.	**indicador 1.1.2.2** Porcentaje de reducción de costos operativos durante las temporadas bajas.
	META 1.1.3 Implementar un programa de contratos flexibles y capacitación continua para mantener el empleo del 90% del personal durante las temporadas bajas en los próximos 6 meses.	**Indicador 1.1.3.3** Porcentaje de empleados retenidos durante las temporadas bajas.

EJE MARKETING DIGITAL

INSTRUMENTACIÓN ESTRATÉGICA		
OBJETIVO 2 MARKETING DIGITAL Incrementar la visibilidad y captación de clientes durante las temporadas bajas mediante el uso de estrategias digitales efectivas.		
ESTRATEGIA 2.1 Implementar campañas de publicidad digital dirigidas a sectores que puedan necesitar servicios de carga y descarga fuera de las temporadas tradicionales.	**META 2.1.1** Aumentar el tráfico web en un 50% en los próximos 6 meses.	**Indicador 2.1.1.1** Número de visitas mensuales a la página web
		Indicador 2.1.1.2 Cantidad total de visitas a la página web cada mes.
		Indicador 2.1.1.3 Incrementar el número de visitas mensuales en un 50% en los próximos 6 meses.
		Indicador 2.1.1.4 Porcentaje de visitantes que abandonan el sitio web después de ver solo una página.
	META 2.1.2 Generar 20 nuevos leads de clientes potenciales mensuales a través de campañas de publicidad digital en los próximos 6 meses.	**Indicador 2.1.2.1** Cantidad de nuevos leads capturados a través de campañas de publicidad digital cada mes.
		Indicador 2.1.2.2 Porcentaje de clics en los anuncios que se convierten en leads.
ESTRATEGIA 2.2 Desarrollar y optimizar contenido	**META 2.2.1** Incrementar el valor promedio de la transacción	**Indicador 2.2.1.1** Costo promedio para generar un lead a través de las campañas de publicidad digital.

SEO (Search Engine Optimization) y SEM (Search Engine Marketing) para atraer tráfico a la web y convertir visitantes en clientes potenciales.	en un 15% en 6 meses.	**Indicador 2.2.1.2** Porcentaje de visitantes que realizan una contratación después de visitar el sitio web.

EJE CONTABILIDAD

INSTRUMENTACIÓN ESTRATÉGICA		
OBJETIVO 3 CONTABILIDAD Mejorar la eficiencia en la gestión de costos y la planificación financiera.		
ESTRATEGIA 3.1 Implementar un sistema de contabilidad de costos más detallado para monitorear y controlar los gastos operativos y detectar áreas donde se pueden hacer ajustes durante las temporadas bajas.	**META 3.1.1** Reducir los costos operativos en un 15% durante las temporadas bajas en los próximos 6 meses, manteniendo una eficiencia financiera que permita mantener al menos el 90% del personal durante estos períodos.	**Indicador 3.1.1.1** Porcentaje de reducción de costos operativos.
		Indicador 3.1.1.2 Porcentaje del personal mantenido durante las temporadas bajas.
		Indicador 3.1.1.3. Nivel de utilización del fondo de reserva financiera.
		Indicador 3.1.1.4 Diferencia entre las proyecciones financieras y los costos reales
ESTRATEGIA 3.2 Establecer un fondo de reserva financiera para gestionar mejor las fluctuaciones estacionales y cubrir los costos operativos durante las temporadas bajas sin necesidad de reducir el personal de manera significativa.	**META 3.2.1** Establecer y mantener un fondo de reserva financiera adecuado para cubrir al menos el 50% de los costos operativos durante las temporadas bajas en los próximos 6 meses.	**Indicador 3.2.1.1** Porcentaje de utilización del fondo de reserva.
		Indicador 2.3.1.2 Comparación entre costos operativos planificados y reales.

EJE ECONOMÍA SOCIAL

INSTRUMENTACIÓN ESTRATÉGICA		
OBJETIVO 4 ECONOMÍA SOCIAL Aumentar la sostenibilidad económica de la empresa mientras se minimiza el impacto social de la reducción de personal durante las temporadas bajas, garantizando la estabilidad financiera y el bienestar de los empleados.		
ESTRATEGIA 4.1 Optimización de Costos Operativos y Diversificación de Servicios	**META 4.1.1** Reducir los costos operativos en un 15% durante las temporadas bajas mediante la optimización de procesos y la diversificación de servicios en los próximos 6 meses.	**Indicador 4.1.1.1** Reducción porcentual de costos operativos.
		Indicador 4.1.1.2 Ingresos adicionales por servicios diversificados.
		Indicador 4.1.1.3 Número de procesos optimizados.
ESTRATEGIA 4.2	**META 4.2.1** Implementar un programa de soporte y capacitación que permita a los empleados adaptarse a cambios en sus roles y minimizar el impacto social de la reducción de personal.	**Indicador 4.2.1.1** Porcentaje de empleados capacitados.
		Indicador 4.2.1.2 Nivel de satisfacción del personal con el programa de soporte
		Indicador 4.2.1.3 Porcentaje de reducción en la rotación de personal.

REPORTE SEMANA 1	
ADMINISTRACIÓN	
indicador 1.1.1.1 Planeación de nuevos servicios introducidos.	Se comenzó con la planificación de un servicio nuevo, venta de playos.
indicador 1.1.2.2 Porcentaje de reducción de costos operativos durante las temporadas bajas.	Se redujo un 5%
Indicador 1.1.3.3 Porcentaje de empleados retenidos durante las temporadas bajas.	Un 20% de los empleados conservaron su empleo esta semana como personal externo.
MARKETIN DIGITAL	
Indicador 2.1.1.1 Número de visitas a la página web	4
Indicador 2.1.1.2 Cantidad total de visitas a la página web cada mes.	4
Indicador 2.1.1.3 Incrementar el número de visitas mensuales en un 50% en los próximos 6 meses.	3%
Indicador 2.1.1.4 Porcentaje de visitantes que abandonan el sitio web después de ver solo una página.	3%
Indicador 2.1.2.1 Cantidad de nuevos leads capturados a través de campañas de publicidad digital cada mes.	1
Indicador 2.1.2.2 Porcentaje de clics en los anuncios que se convierten en leads.	1
Indicador 2.2.1.1 Costo promedio para generar un lead a través de las campañas	Aproximado de $50

de publicidad digital.	

Indicador 2.3.1.1 Un cliente nuevo por semana	1
CONTABILIDAD	
Indicador 3.1.1.1 Porcentaje de reducción de costos operativos.	10%
Indicador 3.1.1.2 Porcentaje del personal mantenido durante las temporadas bajas.	20%
Indicador 3.1.1.3. Nivel de utilización del fondo de reserva financiera.	5%
Indicador 3.1.1.4 Diferencia entre las proyecciones financieras y los costos reales	Aún no podemos ser concretos al 100%, ya que apenas hace pocos días estamos aplicando nuevas estrategias.
Indicador 3.2.1.1 Porcentaje de utilización del fondo de reserva.	5%
Indicador 2.3.1.2 Comparación entre costos operativos planificados y reales.	Aún no podemos ser concretos al 100%, ya que apenas hace pocos días estamos aplicando nuevas estrategias.
ECONOMÍA SOCIAL	
Indicador 4.1.1.1 Reducción porcentual de costos operativos.	5%
Indicador 4.1.1.2 Ingresos adicionales por servicios diversificados.	10%
Indicador 4.1.1.3 Número de procesos optimizados.	3
Indicador 4.2.1.1 Porcentaje de empleados capacitados.	15
Indicador 4.2.1.2 Nivel de satisfacción del personal con el programa de soporte	10

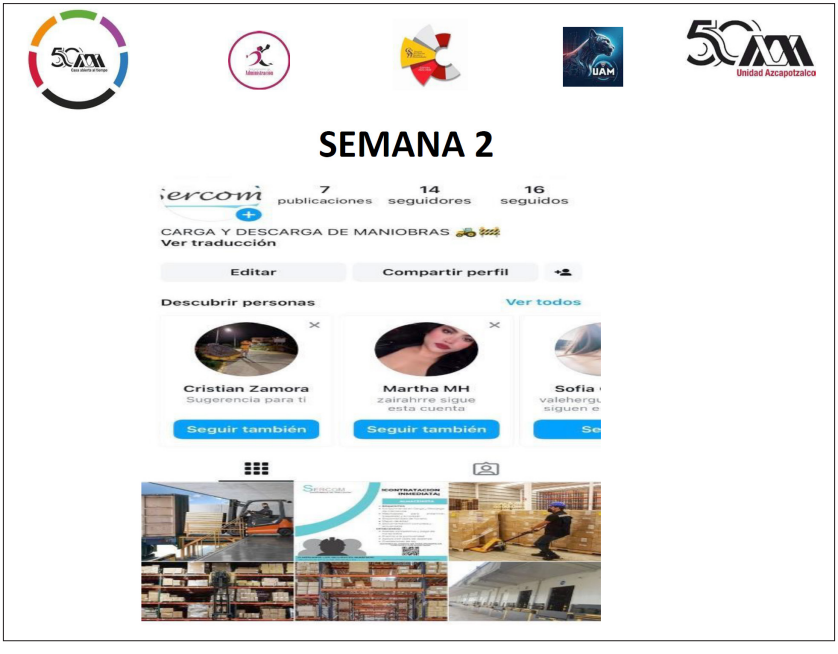
Indicador 4.2.1.3 Porcentaje de reducción en la rotación de personal.	0%

ALCANCES	
SEMANA 1	- Se creó una página web. - Se realizaron dos cursos. - Conservamos 10 empleados que pudieron ser despedidos. - Estamos en la planificación de nuevos servicios.
SEMANA 2	- Mas seguidores - Un cliente más
SEMANA 3	- Dos clientes nuevos. - 5 seguidores más. - 10 likes más. - 1 producto nuevo.
SEMANA 4	- 2 empleos más retenidos - Mejoras estratégicas - Menos Merma en productos - Mejoras administrativas

SEMANA 5	- Un cliente más. - Dos seguidores más. - Un producto más. - Mejoras administrativas.
SEMANA 6	- Mejoras administrativas - 2 likes más - Un cliente nuevo - 5 empleos nuevos - 3 empleos retenidos - 1 publicación y 1 seguidor más.
SEMANA 7	- 3 seguidores más - 3 likes más - 2 clientes nuevos - 5 empleos nuevos - 4 empleos nuevos - Mejoras administrativas
SEMANA 8	- 1 seguidor más. - Buen recibimiento de los nuevos servicios. - Mejoras administrativas y financieras. - 7 likes nuevos.
SEMANA 9	- 2 seguidor más. - Buen recibimiento de los nuevos servicios. - Un producto nuevo. - Mejoras administrativas y financieras. - 10 likes nuevos.

SEMANA 10	

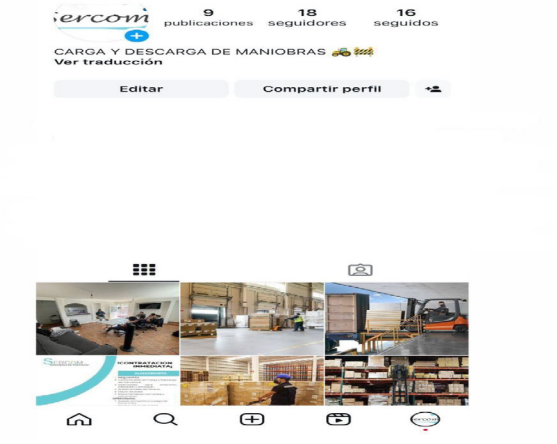

SEMANA 1
sercom_5
Nota...
6 publicaciones
12 seguidores
16 seguidos
CARGA Y DESCARGA DE MANIOBRAS
Ver traducción
Editar
Compartir perfil
Descubrir personas
Ver todos
CARGA Y DESCARGA DE MANIOBRAS
Ver traducción
Editar
Compartir perfil
Descubrir personas
Ver todos


SEMANA 2
7 publicaciones
14 seguidores
16 seguidos
CARGA Y DESCARGA DE MANIOBRAS
Ver traducción
Editar
Compartir perfil
Descubrir personas
Ver todos
Cristian Zamora
Sugerencia para ti
Seguir también
Martha MH
zairahrre sigue esta cuenta
Seguir también

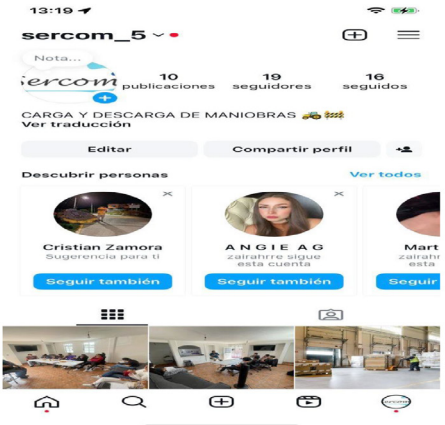
SEMANA 3



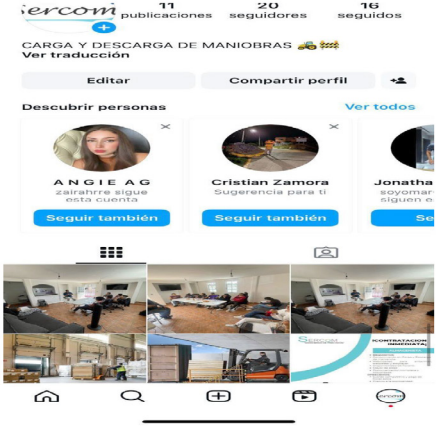
SEMANA 4

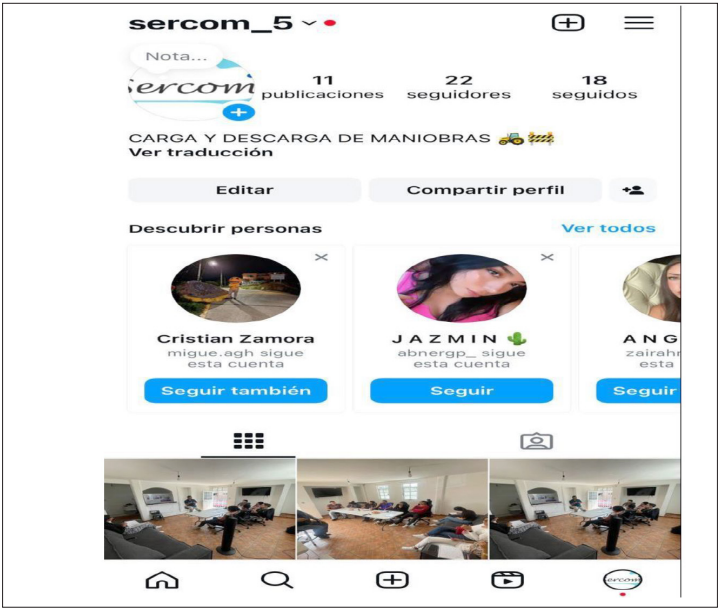


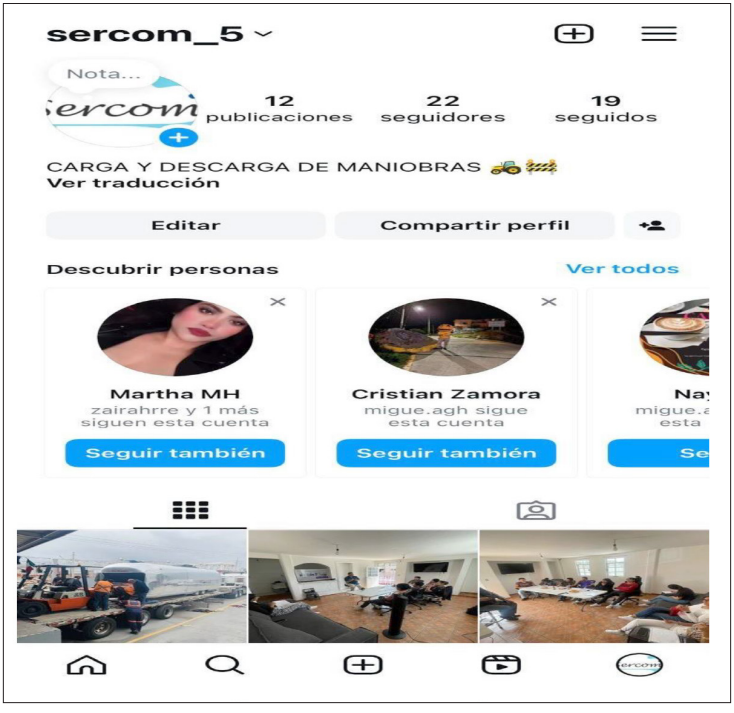
SEMANA 5

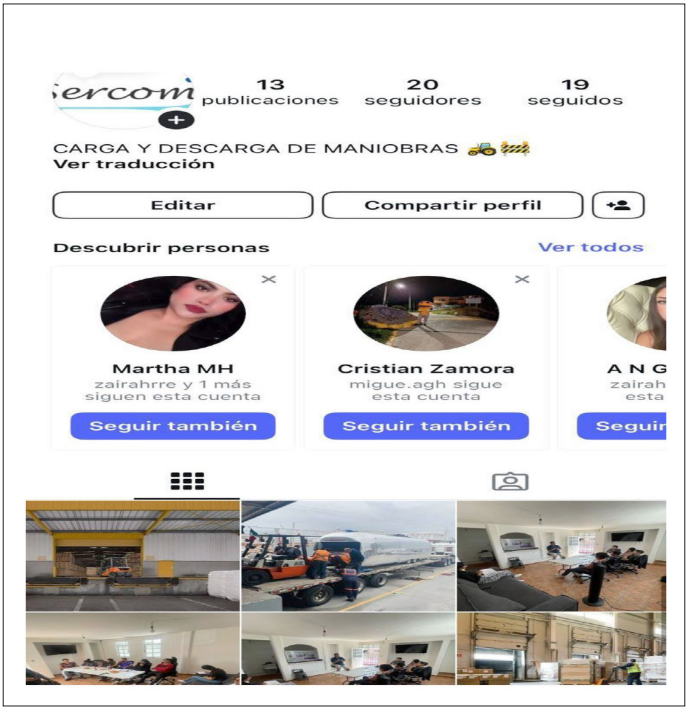


SEMANA 6




sercom_5
Nota...
11 publicaciones
22 seguidores
18 seguidos
CARGA Y DESCARGA DE MANIOBRAS
Ver traducción
Editar
Compartir perfil
Descubrir personas
Ver todos
Cristian Zamora
migue.agh sigue esta cuenta
Seguir también
J A Z M I N
abnergp_ sigue esta cuenta
Seguir


sercom_5
Nota...
12 publicaciones
22 seguidores
19 seguidos
CARGA Y DESCARGA DE MANIOBRAS
Ver traducción
Editar
Compartir perfil
Descubrir personas
Ver todos
Martha MH
zairahrre y 1 más siguen esta cuenta
Seguir también
Cristian Zamora
migue.agh sigue esta cuenta
Seguir también

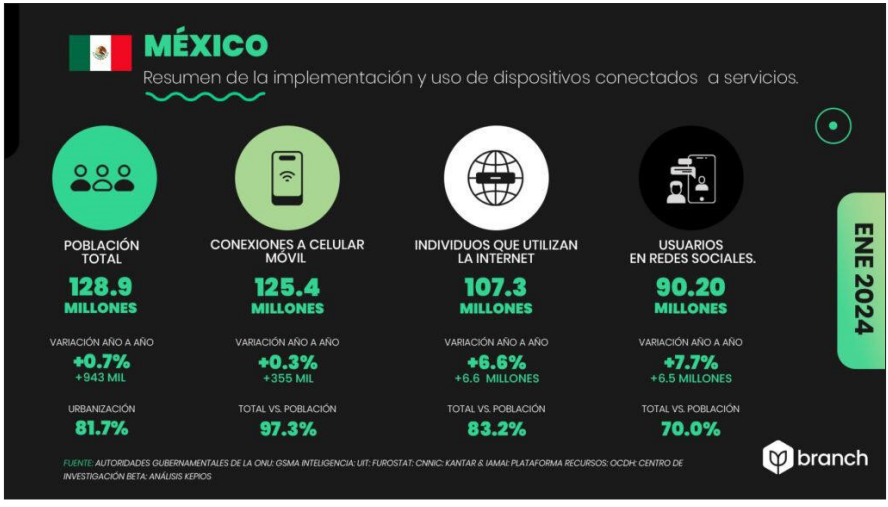


ÉTICA DE LA DEMOCRATIZACIÓN DEL CONOCIMIENTO

Por primera vez en la historia de la humanidad, el conocimiento es accesible a toda la gente gracias a la digitalización de los últimos años (ver **Figura 4**), principalmente después de la pandemia, donde dice que se avanzó en dos años, lo que en 8 o 10 años se hubiera avanzado de no ocurrir la pandemia.

Esta democratización permite tener una accesibilidad a la sociedad en su conjunto que fortalece la educación y con ella la movilidad social y el orden civilizatorio que a su vez conlleva una mejor calidad de vida, sin embargo es importante estar atentos a las dinámicas que se generan con la evolución y velocidad vertiginosa que la IA tiene, toda vez que siempre tendrá riesgos en su uso, ya sea para sustituir a la gente en puesto de trabajo como suelen presagiar algunas visiones, o su mal

Figura 4. Digitalización en México.



Fuente: Branch (2024).

uso como herramienta para ser utilizada por el crimen organizado, o para la guerra.

La IA debe de ser normada para fijar claramente una ética de su funcionamiento como dice Lasalle (2024) al mostrar en su libro la Civilización Artificial los riesgos de un nihilismo intrinseco que se tiene al desarrollo tecnológico como contradictorio con la naturaleza humana, y mostrando que es preciso definir y normar la utilidad y fines de la misma y no caer en una sociedad espectral como dice Sadin (2024) también como uno de los riesgos que se están generando con el vertiginoso descontrolado de la IA.

No hay que confundir estas perspectivas con las visiones apocalípticas, ya que en principio hemos de reconocer la utilidad de la IA en el avance del conocimiento que se manifiesta en la mejora en procesos de educación, salud y gobierno por ejemplo, sin embargo es preciso mostrar que los grandes pendientes éticos del modelo capitalista triunfador del siglo XXI se centran en los grandes sectores sociales que siguen siendo desfavorecidos, principalmente en 3er mundo.

El proyecto IAE para la MIPYME se presenta como una propuesta de praxis de la IA para fortalecer la ética de la inclusión social y el beneficio de la IA para fortalecer los procesos tradicionales de las micro, pequeñas y medianas empresas, así como cooperativas, pequeños productores y productoras, que además de mejorar su eficiencia organizativa y administrativa, visibiliza las dimensiones sociales y solidarias que toda empresa micro y pequeña, principalmente de corte familiar poseen y que en este sistema no son mostradas en los formatos normales de funcionalismos de *management* tradicional.

La IAE para la MIPYME se inscribe así en una tradición ética de Economía Social y Solidaria que la UAM tiene y que muestra desde su origen y en sus propios programas de estudio como el de la Licenciatura en Administración (sin el apellido de empresas), la Licenciatura en Diseño de Proyectos Sustentables que además es de tipo semipresencial, el Posgrado en Estudios Organizacionales, maestría y doctorado con alto enfoque social y crítico contestatario al modelo tradicional burocrático eficientista de la administración, así como el Doctorado en Intervención en las Organizaciones con el mismo perfil de complejidad que reconoce el orden social y humano de las organizaciones, como el más importante y relevante, solo por mencionar algunos programas.

Con este proyecto se atienen a más de 200 micro, pequeñas y medianas empresas, emprendimientos y profesionistas que verán como la IA se convierte en un aliado para mejorar no solo los ingresos monetarios, sino la calidad de vida y el respeto por el ser humano y su amiente.

CONCLUSIONES

La aparición de la IA data de más de siete décadas, sin embargo no es sino hasta los últimos años, principalmente con el lanzamiento del *ChatGPT* por *OPEN AI* en noviembre de 2022 que se detona una revolución cultural cognitiva digital, provocada por la masificación de su uso, es decir, por la democratización del uso de la IA que pone prácticamente en todo dispositivo digital con internet, el uso de la IA ge-

nerativa *GPT* que es un asistente permanente que realiza tareas avanzadas de conocimiento especializado que antes no existía, por lo que se vuelve necesario definir la ética humana que habrá de fortalecerse para no caer en una sociedad digital artificial que convierta el uso de la IA y su evolución como un fin en si mismo.

En ese sentido la Universidad Autónoma Metropolitana y sus aliados, la Universidad del Caribe, la Universidad Autónoma de Occidente, la Universidad de Medellín, la Universidad de Valparaiso, la Universidad Autónoma de Guanajuato, el Instituto Universitario de Eduación Virtual de Honduras y otras instituciones, fomentan proyectos y herramientas que orienten éticamente el uso de la IA hacia procesos de formación cultural donde el ser humano siga siendo el centro y valor más importante, acorde a los altos valores que la revolución francesa y demás revoluciones sociales que siguieron a la revolución industrial, que colocaron al ser humano al centro en los valores de la justicia, la igualdad y la fraternidad como un modelo de gobierno y gobernanza mundial.

La Inteligencia Artificial Estratégica para la MIPYME amplía el impacto social de la universidad, fortalece la formación de las y los estudiantes en estas prácticas tecnocientíficas, pero también en los más altos valores humanistas, sociales y de cuidado del ambiente valores que normalmente no eran fomentados por la administración y demás disciplinas modernas de carácter eficientista e individualista.

REFERENCIAS BIBLIOGRÁFICAS

Branch (2024) “Estadísticas de la situación digital de México en 2024” en *Branch,*

México. Estadísticas de la situación digital de México 2023-2024 - Branch

Lasalle, José María (2024) *La civilización Artificial,* Arpa, Madrid.

Lozano Carrillo, Oscar, Alfredo Garibay Súarez y David Salvador Cruz Rodríguez (2021). *Tecnología, Innovación e intervención organizacional. La dimensión digital en contextos de contingencia,* UAM. México.

Lozano Carrillo, Oscar (2021) “La Universidad Pública ante los retos de la pandemia. El caso de la plataforma digital CONTINGENCIA COVID-19 en la UAM Azcapotzalco en Peñalosa Castro, Eduardo y Pablo Hernández Cerrito*, Modalidades alternas para la innovación educativa en la formación universitaria;* Serie Abate Faria, UAM, México.

Lozano Carrillo, Oscar; “Alfredo Garibay Súarez y Francisco Javier Árias Vargas (2023) UAMEDIA. Modelo Gratuito de intervención estratégica digital en la pandemia, para el apoyo a las MIPYMES, Empresas Familiares y

Cooperativas de América Latina” en Lozano Carrillo, Oscar, Alfredo Garibay Suárez y Edgar Sansores Guerrero *Reflexiones sobre la gestión, desarrollo y análisis de las Organizaciones en América Latina.*

Sadin, Eric (2024) *La sociedad espectral,* Caja negra, Argentina.

Uso de ChatGPT, Chatbot, Copilot, Llama 3.1.

SEMBLANZAS DE AUTORES

ALEJANDRA JARAMILLO RODRÍGUEZ

Estudiante de Ingeniería Ambiental en la Universidad Autónoma Metropolitana, e investigadora en el Laboratorio de Inteligencia Artificial. Ha participado en diversos webinars y es autora de capítulos en dos libros. En el libro *Inteligencia Artificial*, escribió el capítulo titulado "*La Revolución de la Inteligencia Artificial en las Energías Renovables: Optimización y Eficiencia Energética*", y en el libro *Inteligencia Artificial Ética, Transhumanismo y Sustentabilidad*, contribuyó con el capítulo "*Innovación Sustentable: El Poder Transformador de la Inteligencia Artificial en un Mundo en Cambio*".

Además, ha participado en los seminarios PIESEA realizados en la Universidad Autónoma de Mazatlán, la Universidad de Guanajuato, la Universidad del Caribe y el Centro de Pátzcuaro, donde ha abordado temas de economía social y solidaria. En estos seminarios, presentó una investigación sobre la economía social y la sustentabilidad en comunidades indígenas.

Ha colaborado en el desarrollo de un diplomado en promotores en ecotecnologías, integrando la inteligencia artificial con soluciones sostenibles para generar impacto en comunidades.

ALEJANDRO LÓPEZ ÁLVAREZ

Alejandro López Álvarez, es Licenciado en Desarrollo Empresarial y de Negocios por la Universidad Autónoma de Sinaloa, es Maestro en Administración por la Universidad de Occidente, y estudiante del Doctorado en Innovación y Administración de las Organizaciones en la Universidad Autónoma de Occidente.

Su carrera profesional se ha desarrollado dentro de la iniciativa privada, dentro de dos segmentos, el primero en la banca comercial responsable de la ejecución de operaciones de crédito, captación, consumo e inversión. En segundo lugar, dentro del sector de telecomunicaciones, específicamente en el área de gestión de personal. Además, actualmente se desempeña como profesor en la Maestría en Administración de la Unidad de Educación Virtual de la Universidad Autónoma de Occidente.

ALFREDO GARIBAY SUÁREZ

El Dr. Alfredo Garibay Suárez estudió la licenciatura en Administración en la Universidad Autónoma Metropolitana, Unidad Azcapotzalco (UAM-A), posteriormente realizó estudios profesionales en Tecnologías de la Información y Comunicación cursando bajo esa temática un Diplomado en la Universidad Nacional Autónoma de México (UNAM) donde también en educación continua junto con la Cámara de Diputados, obtuvo el diplomado en Innovación y Estrategias de Desarrollo Nacional.

El Dr. Garibay se graduó en los niveles de Especialización (Hipermedios), Maestría (Nuevas tecnologías) y Doctorado en Diseño (Diseño y Visualización de la Información) en la UAM-A. Es Profesor-Investigador en Mercadotecnia, Diseño y Tecnologías en la Universidad UAM, donde ha desarrollado además participaciones en actividades de educación continua (talleres, cursos y diplomados), actividades de gestión (tecnología digital, vinculación, comités de posgrado y producción editorial). Es Par Evaluador de Programas Académicos para el Consejo de Acreditación en Ciencias Administrativas, Contables y Afines (CACECA). Ha sido distinguido por el CONACYT participando dentro del Sistema Nacional de Investigadores (SNII nivel I) y cuenta con el distintivo de Perfil Deseable del Programa para el Desarrollo Profesional Docente de la Secretaría de Educación Pública (PRODEP-SEP).

Ha sido distinguido con el Premio a la Docencia y en dos ocasiones con el Premio a la Investigación como colaborador en dos Áreas de Investigación del Departamento de Administración de la DCSH de la

UAM-A. Ha recibido en tres ocasiones la Medalla al Mérito Universitario, una Mención Académica por su proyecto doctoral y una Mención Plata Quórum México.

Es parte de la coordinación del proyecto UAMedia, desarrollando vínculos de atención a la sociedad a partir de proyectos de intervención en el contexto nacional e internacional y se ha desempeñado para la División de Ciencias Sociales y Humanidades de la UAM-A como responsable de Desarrollo Informático, Coordinador Divisional de Vinculación, Coordinador Divisional de Difusión y Publicaciones, desarrollándose además en esta última instancia como Presidente del Consejo Editorial; actualmente es Jefe del Departamento de Administración donde gestiona los trabajos del Programa Estratégico Departamental 2022-2026.

En materia tecnológica, expresamente se vincula con actividades de participación en los ecosistemas de innovación públicos y privados, por ejemplo en participaciones de "Campus Party" y como Representante Institucional de la Universidad en el evento "Hackathon Turismo y Bienestar" desarrollado por la Secretaría de Educación, Ciencia, Tecnología e Innovación de la Ciudad de México (SECTEI), el Consejo Nacional Empresarial Turístico (CNET), la Secretaría de Turismo del Gobierno de México (SECTUR); logrando en su participación colaborativa en éste último evento la obtención de los primeros tres lugares del concurso.

Sus líneas de investigación incluyen el diseño, la mercadotecnia, la educación digital, la intervención organizacional, el gobierno electrónico y las tecnologías disruptivas; en dichas temáticas señaladas publica e imparte conferencias en el contexto nacional e internacional.

En relación con la inteligencia artificial ha desarrollado participaciones como panelista, cuenta con un registro de proyecto de investigación colaborativo en la materia ante el Consejo Divisional de CSH de la UAM-A y, además, vincula dicha temática con la experiencia de usuario, el diseño y el transhumanismo para su aplicación en la educación y la mercadotecnia.

Cabe destacar la participación actual del Dr. Garibay en proyectos nacionales estratégicos (PRONACES) su responsabilidad como res-

ponsable técnico en un proyecto de Ciencia de Frontera CONAHCYT y su integración como colaborador en la Red de Investigación para el Fortalecimiento de la Economía Social y Solidaria de la UAM (Red IFESS UAM).

El Dr. Garibay tiene entre sus pasatiempos: la música, la electrónica los videojuegos y el básquetbol.

AMÉRICA PELAYO RAMÍREZ

Licenciada en Administración, por la UAM Azcapotzalco. Miembro de la Red Mexicana de Investigadores en Estudios Organizacionales (REMINEO), A.C. y del Colegio de Investigadores 2022-2023. Con un Diplomado Internacional en "Emprender e innovar" por la HEC Montreal, Canadá. Así como un Diplomado en "Educación Financiera" por la CONDUSEF.

Ganadora del 3er Lugar del Hackathon 2022 "Turismo y Bienestar" por la SECTUR y el CNET, e integrante de los diversos proyectos que desempeña UAMedia desde 2021, como UAMedia Internacional, UAMedia Innova y Emprende, UAMedia CICyT y recientemente como colaboradora en el Laboratorio de Inteligencia Artificial.

Participante en diversos seminarios, conferencias y ponencias a nivel nacional e internacional, desde 2021. Miembro activo de NODESS "Vida y Saberes" desde 2022". Actualmente desempeño el cargo de Asistente de Posgrado en el Doctorado en Intervención en las Organizaciones, en la UAM Azcapotzalco.

AYLIN TORRES RAMÍREZ

Licenciada en Administración Industrial y Maestra en Administración por el Instituto Politécnico Nacional. Actualmente estudiante del programa Doctorado en Estudios Organizacionales por la Universidad Autónoma Metropolitana. Ha sido galardonada con el Premio al Mejor Trabajo Escrito para Titulación de Nivel Superior, el premio al Mejor

Desempeño Académico nivel Maestría y el Premio a la Investigación en el Área de Ciencias Sociales y Administrativas por el Instituto Politécnico Nacional. Cuenta con artículos de investigación en revistas enlistadas en el JCR, Scopus, Conahcyt y de divulgación; elaboración de capítulos de libros y derechos de autor. Sus labores de investigación abarcan las áreas de Gestión de la empresa familiar e Intervención en las organizaciones. Ha presentado ponencias en eventos de corte nacional e internacional. En el ámbito de la docencia, cuenta con la experiencia impartiendo cursos de Bases de datos, Cienciometría, Redacción, Administración, Planeación Estratégica y Evaluación pedagógica digital. Desde 2017 ha sido coordinadora y asesora de formación y capacitación empresarial, particularmente en zonas rurales. Participa en las actividades de difusión y promoción de los Centros de Innovación, Cultura y Tecnología (CICYTS). De 2022 a 2023 coordinó las actividades de Capacitación del Aeropuerto Internacional Felipe Ángeles (AIFA). En el 2022 y 2023 participó como coordinadora en el Proceso de Evaluación del Sistema Nacional de Investigadoras e Investigadores. Funge como apoyo en Proyectos Nacionales de Investigación e Incidencia (Pronaii) y Ciencia de Frontera del Conahcyt.

CINTHIA NOEMMI ZACATENCO ARELLANO

Estudiante de Administración y promotora de investigación y difusión de Laboratorio de Inteligencia Artificial de la UAM Azcapotzalco, autora del capítulo “La revolución de la inteligencia artificial en finanzas: eficiencia, oportunidades y retos en banca y mercados bursátiles” dentro del libro “El futuro de la IA”, facilitadora del taller “IA para la Educación” impartido a más de mil personas de México y Latinoamérica, conferencista y ponente de más 15 eventos, dentro de los que destaca “Economía Social” dentro del XIX Congreso Internacional de Análisis Organizacional 2024, “Las finanzas y la IA” en el Webinar “El futuro de la IA” UAM 2024.

Así mismo, he participado en los seminarios del “Proyecto de Integración de la Economía Social con las Economías Alternativas (PIESEA)”, donde trabajé temas como economía familiar, economía feminista, administración social y sistemas socioecológicos integrados.

Estas actividades se llevaron a cabo en diversas instituciones, como la Universidad Autónoma de Occidente en el campus Mazatlán, la Universidad del Caribe en Cancún y el Centro de Ecotecnologías en Michoacán, entre otras.

EDGAR ALFONSO SANSORES GUERRERO

Dr. en Estudios Organizacionales, Profesor Visitante en la Universidad Autónoma Metropolitana y Miembro del Sistema Nacional de Investigadores, Nivel I (México).

Especialista certificado por HEC Montreal en métodos de enseñanza de tipo experiencial y gamificación, innovación social, innovación del modelo de negocios y en la ejecución y escalamiento de proyectos.

Líneas de investigación: emprendimiento social, inteligencia artificial, economía social y solidaria, responsabilidad social corporativa y economía circular.

Miembro del Comité Editorial de la Revista Gestión y Estrategia UAM; Journal of Accounting and Financial. West Sciences Press. Journal Economic and Entrepreneurship, West Sciences Press.

EDNA KARINA MIRANDA MONDACA

La Dra. Edna Karina Miranda Mondaca es una destacada académica y profesora de tiempo completo en la Universidad Autónoma de Occidente (UAdeO), adscrita al Departamento Académico de Ciencias Sociales y Humanidades en la Unidad Regional Los Mochis, Sinaloa, México. Es licenciada en Ciencias de la Comunicación por la misma universidad, con más de 15 años de experiencia en medios masivos de comunicación radiofónica y comunicación organizacional en el sector privado. Ha trabajado como community manager, locutora, directora de programas y periodista, y cuenta con una certificación como locutora clasificación “A” por la Secretaría de Educación Pública y la Secretaría de Gobernación.

La Dra. Miranda Mondaca obtuvo su doctorado en Innovación Educativa por el Instituto Tecnológico Superior de Los Mochis, donde ha participado en diversos proyectos de investigación educativa, e-learning y desarrollo curricular. Desde 1993, ha sido docente a nivel superior en la UAdeO, desempeñando también diversos roles en la gestión académico-administrativa, como consejera institucional ante el Consejo Nacional para la Enseñanza y la Investigación de las Ciencias de la Comunicación (CONEICC) y coordinadora del programa educativo de Licenciatura en Ciencias de la Comunicación.

Actualmente, lidera el Grupo de Investigación en Educación y Desarrollo Integral, enfocándose en la gestión, innovación y tecnologías en educación, así como en la intervención transdisciplinaria en fenómenos relacionados con el desarrollo integral. Su labor académica y de investigación ha tenido un impacto significativo en la formación de recursos humanos para la salud y en la mejora de los programas educativos en su institución.

EMILY NICOLE SILVA ESCAMILLA

Estudiante de Ingeniería Metalúrgica en la Universidad Autónoma Metropolitana - Azcapotzalco. He participado en diversos eventos académicos y seminarios relacionados con la inteligencia artificial y la sostenibilidad. Entre ellos, el webinar “El Futuro de la Inteligencia Artificial: Ética, Transhumanismo y Sustentabilidad”, impartido por la Universidad del Caribe, y el taller digital “IA para la Educación”, ofrecido por la Universidad Autónoma de Occidente en Mazatlán. También participé en el “Primer Encuentro de IA”, organizado por la Universidad Autónoma Metropolitana.

Colaboró activamente en el Laboratorio de Inteligencia Artificial, donde he contribuido con dos capítulos de dos libros recientes elaborados: escribiendo un capítulo en cada uno: “Computación Cuántica y Multiplicación de Capacidades con Inteligencia Artificial: Futuro y Perspectivas” y “Transhumanismo y la Fusión de Hombre-Máquina: ¿Podríamos llegar a una Simbiosis con Máquinas Inteligentes?”.

Adicionalmente, he sido parte de los seminarios PIESEA (Proyecto de Integración de la Economía Social con las Economías Alternativas), enfocados en temas de sustentabilidad y cambio climático. Estos seminarios se han llevado a cabo en universidades de Mazatlán, Cancún, Guanajuato y Michoacán, donde se abordaron problemáticas de la economía social solidaria (ESS), organizaciones alternativas y y la crisis del capitalismo como una crisis civilizatoria.

Este recorrido académico refleja mi interés por el futuro de la tecnología, la intersección entre inteligencia artificial y ética, así como el impacto de las economías alternativas en un contexto global en transformación.

FÁTIMA CECILIA PEREZ PONCE

Licenciada en Administración por la Universidad Autónoma Metropolitana y actualmente colaboradora activa dentro del Laboratorio de Inteligencia Artificial de la UAM Azcapotzalco para la investigación y la difusión del conocimiento del cual se desprende mi participación en el primer encuentro de Inteligencia Artificial de la UAM Azcapotzalco en noviembre de 2023 y en el Webinar "El futuro de la Inteligencia Artificial. Ética, Transhumanismo y Sustentabilidad, con el tema "Los asistentes virtuales y su evolución" llevado a Cabo en la Universidad del Caribe (UNICAR) en Cancún, Quintana Roo en agosto de 2024.

Asimismo, ha participado como ponente dentro del Seminario de Estudios Cualitativos Comparativos de Economías Alternativas en la Economía Social llevado a cabo durante el XXI Congreso Internacional de Análisis Organizacional en octubre de 2023 con sede en la Universidad Autónoma De Occidente (UADEO) en Mazatlán, Sinaloa y en los Seminarios del Proyecto de Integración de la Economía Social con las Economías Alternativas (PIESEA) llevados a cabo en diversas sedes tales como el Centro de Eco tecnologías Pátzcuaro, Michoacán y en la Universidad del Caribe (UNICAR) en Cancún, Quintana Roo, donde expuse temáticas como movimientos sociales y ambientales, economía indígena y economía comunitaria.

Fungió como especialista de apoyo dentro de las comisiones dictaminadoras durante la Convocatoria para el Reconocimiento en el Sistema Nacional de Investigadoras e Investigadores 2024.

FRANCISCO JAVIER ARIAS VARGAS

Francisco Javier Arias Vargas es Doctor en Administración y Dirección de Empresas por la Universidad Politécnica de Valencia y Doctor Honoris Causa en Pedagogía por la Organización Continental de Excelencia Educativa.

Ha publicado trabajos en revistas especializadas donde analiza diferentes sectores estratégicos ligados al desarrollo económico regional y su inserción en el comercio internacional. Es experto en Cooperación Académica Internacional, Diplomacia Científica e Inteligencia Competitiva para el Desarrollo de Negocios.

Actualmente, se encuentra registrado como Investigador en CONAHCYT (México) con el CVU 109758, es Investigador RENACYT nivel VI por CONCYTEC (Perú) e Investigador Asociado por MINCIENCIAS (Colombia).

Ha sido nombrado Profesor Honorífico de la UNTRM en Perú y Profesor Invitado de la Universidad Autónoma Metropolitana de México y de la Universidad Autónoma de Chiapas. Además, ha sido Profesor de Posgrado en varias universidades de Colombia, México, España, Honduras, Chile, Bolivia, Ecuador, Paraguay y el Principado de Andorra.

Es miembro del Consejo Directivo de UNEV en Honduras y de varias juntas directivas, adicionalmente, se desempeña como Presidente de la Red Internacional de Investigación en Gestión del Conocimiento Empresarial (RED GCE), la cual tiene presencia en todos los países de Iberoamérica. Por último, pertenece al staff docente de la Facultad de Ciencias Económicas y Administrativas de la Universidad de Medellín.

GABRIELA CAMACHO DEONICIO

Es maestra en Administración y Dirección Estratégica por la Universidad Juárez Autónoma de Tabasco (UJAT) (México) (2020-2022) e ingeniera en Tecnologías de la Información de la Universidad Tecnológica de Tabasco (México). Fue asistente de consultor externo (2018-2020), labor en la cual participó en la evaluación de diferentes programas presupuestarios en materia de evaluación del desempeño, revisión y consideración de los aspectos susceptibles de mejora. Participó en la elaboración de los Manuales de organización, procesos y procedimientos de la Fiscalía General del Estado de Tabasco (FGET) y de la Secretaría de Gobierno (SEGOB). Colaboró en actividades de apoyo docente (2018 - 2024) y participó también en la Coordinación General para la organización de los Congresos Internacionales de Análisis Organizacional, en sus ediciones XV al XXII, así también, colabora en los proyectos editoriales de la REMINEO, consistentes en Libros Electrónicos y números de la Revista de Análisis Organizacional.

GISELA COTA YUCUPICIO

La Dra. Gisela Cota Yucupicio es una destacada académica y profesora investigadora con una sólida formación en psicopedagogía y desarrollo del potencial humano, obtenida en el Centro de Investigación e Innovación Educativa del Noroeste. Además, cuenta con una maestría en Tecnologías de la Información y de la Comunicación con acentuación en Educación a Distancia, y una licenciatura en Sistemas Computacionales por la Universidad Autónoma de Occidente (UAdeO).

Actualmente, la Dra. Cota Yucupicio es profesora investigadora de tiempo completo en la UAdeO, adscrita al Departamento de Ciencias Económico Administrativas, y lidera el Cuerpo Académico Consolidado de “Educación, Psicología y Salud” (UDO-CA-033). Sus líneas de investigación se centran en el desarrollo de procesos académicos, la gestión e innovación en educación, y la intervención psicoeducativa en fenómenos relacionados con la salud.

Ha sido reconocida con el Perfil Deseable PRODEP por la Secretaría de Educación Pública y ha sido miembro del Sistema Nacional de Investigadores (SNI) del CONACYT. Su trayectoria incluye diversos cargos administrativos y académicos, como directora de la Unidad de Investigación e Innovación Educativa y directora de la Unidad Regional Los Mochis, directora de Investigación y Posgrado, entre otros.

La Dra. Cota Yucupicio ha contribuido significativamente al diseño y rediseño de planes y programas de estudio en diversas modalidades y niveles educativos, incluyendo la maestría en docencia y la licenciatura en gastronomía. Además, ha participado en la elaboración del "Modelo Educativo Lince para el Desarrollo Integral (MELDI)" y ha sido consultora para el diseño de programas de estudio en instituciones de educación superior.

Sus publicaciones abarcan temas como el aprendizaje híbrido y la formación docente, y ha dirigido numerosos proyectos de investigación, tesis y cursos en licenciatura y posgrado. Su labor académica y de investigación ha sido fundamental para el avance de la educación y la psicopedagogía en su región y a nivel nacional.

JANETTE BRITO LAREDO

Doctora en Ciencias de la Administración por la Universidad Nacional Autónoma de México (UNAM), examen de grado aprobado con mención honorifica. Maestra en Administración con acentuación en finanzas y Licenciada en Contaduría Pública por la Universidad de Montemorelos Nuevo León. Profesora investigadora titular "C" de tiempo completo en la Universidad Autónoma de Baja California (UABC), adscrita a la Facultad de Ciencias de la Ingeniería, Administrativas y Sociales unidad Tecate, donde ocupa el cargo de Coordinadora de Programa Educativo, e imparte clases de posgrado y licenciatura. Distinguida por el CONAHCYT como miembro del Sistema Nacional de Investigadoras e Investigadores (SNII) Nivel 1 desde 2016. Cuenta con la Certificación Académica en Contaduría Pública por la Asociación Nacional de Facultades y Escuelas de Contaduría y Administración (ANFECA) y el Perfil Deseable del Programa para el Desarrollo

Profesional Docente (PRODEP). Autora de libros, capítulos de libro, artículos científicos publicados en revistas arbitradas e indexadas. Ponente en congresos nacionales e internacionales. Directora de tesis de posgrado y Líder del Cuerpo Académico Sistemas de Gestión Organizacional. Correo: jbrito@uabc.edu.mx

JESÚS MANUEL LÓPEZ MARROQUÍN

El Dr. Jesús Manuel López Marroquín es un destacado académico y profesional en el campo del derecho, con una sólida formación y una amplia experiencia en diversas áreas jurídicas. Es doctor en Materia Fiscal por la Universidad Autónoma de Durango, maestro en Derecho Corporativo por la Escuela Libre de Derecho de Sinaloa y licenciado en Derecho y Ciencias Sociales por la Universidad Autónoma de Occidente (UAdeO).

El Dr. López Marroquín ha complementado su formación con diversos diplomados en derecho constitucional, penal, amparo y adicciones, entre otros. Actualmente, es profesor de tiempo completo en la UAdeO, adscrito al Departamento Académico de Ciencias Sociales y Humanidades, y coordina el Programa Educativo de Licenciatura en Criminalística y Ciencias Periciales. Además, es miembro certificador del World Justice Project (WJP) y presidente del Consejo Directivo del Colegio de Abogados de Los Mochis Ignacio L. Vallarta A.C.

Su trayectoria profesional incluye roles como consultor jurídico para empresas en áreas societarias, corporativas y fiscales, y ha sido socio de la firma de abogados corporativos S & Ma Saldaña Marroquín y Asociados. También ha desempeñado funciones como coordinador jurídico regional noroeste del Consejo Mexicano para la Defensa de los Derechos Humanos y juez del Tribunal Colegiado de Barandilla.

En el ámbito académico, el Dr. López Marroquín ha impartido más de 27 cursos de licenciatura y varios cursos de posgrado en derecho corporativo, empresarial, penal y fiscal. Ha sido presidente y secretario de la Academia del Programa Educativo de Derecho y Ciencias Sociales de la UAdeO y ha participado activamente en el rediseño curricular de programas de licenciatura y posgrado.

Reconocido por su excelencia académica, ha recibido premios como el Premio Nacional de Investigación en Ciencias Forenses y el Premio a la Excelencia y Desempeño Académico. Su compromiso con la educación y la justicia se refleja en su participación en proyectos de investigación y en su labor como director corporativo de la firma LLM Abogados Corporativos.

JESÚS MANUEL RAMOS GARCÍA

Profesor Investigador del Departamento de Administración, Titular C, Tiempo Completo, de la Universidad Autónoma Metropolitana, Unidad Azcapotzalco. Doctor y Maestro en Estudios Organizacionales por la UAM, y Licenciado en Ciencias Empresariales por la Universidad Tecnológica de la Mixteca.

Distinción como miembro del Sistema Nacional de Investigadores, CONAHCYT-SNI 1 2025-2029, Candidato a Investigador Nacional 2019-2021. Y profesor distinguido con el reconocimiento del Perfil Deseable PRODEP-SEP, del 2016 al 2024.

Director de la División de Ciencias Sociales y Humanidades, de la UAM-Azcapotzalco. 2022-2026. Jefe del Departamento de Administración, UAM Azcapotzalco, 2018-2022 Coordinador de estudios de la Licenciatura en Administración, 2017-2018; Coordinador de la modalidad semipresencial de la Licenciatura en Administración en la UAM Azcapotzalco y Lerma; Coordinador del convenio de colaboración institucional UAM-UNACH para impartir la Licenciatura a distancia en Gestión de la MiPYME.

Coordinador del Seminario de Investigación Organizaciones Sociales e Innovación en Economías Alternativas, Coordinador de los Coloquios de Cooperativismo en México, Coloquio y Feria de la Economía Social y la Administración Solidaria. Coordinador de proyectos de investigación con financiamiento: con el PRODEP-SEP “Gestión y Organización de Organizaciones Sociales Comunitarias, 2017-2019”; y participación en la coordinación de proyectos de investigación con el CONACYT, PRONACES 2022: “Fortalecimiento de los circuitos justos de producción-consumo de los productos de colmena de las abejas

nativas" y "Desarrollo de capacidades técnicas, sociales y económicas para la implementación de eco-tecnologías energéticas sustentables en comunidades rurales con enfoque participativo", y PRONAII 2023. "Alimentación Universitaria y Sustentabilidad: de la milpa a la mesa".

Integrante y fundador del área de Gestión e Intervención en Organizaciones e Instituciones. Miembro Fundador del Nodo Cultura de la Ciudad de México para la Economía Social y Solidaria, e integrante de la Red Mexicana de Investigadores en Estudios Organizacionales y de la Red de Cooperativismo en México.

Recientemente ha coordinado los libros Racionalidad de gobierno y cambio institucional en México, Administración Pública en el actual gobierno, Mundos Indígenas: identidad, territorio, economía y gestión, y es autor de diversos capítulos de libros y artículos en revistas con temas de organizaciones sociales, organizaciones comunitarias y cooperativas, e innovación social, entre otros.

Sus principales líneas de investigación son: estudios críticos de la administración, gestión y organización de empresas sociales, administración social, economía social y solidaria, cooperativismo, comunitarias, planeación estratégica e intervención en organizaciones, desarrollo social y sustentable.

JONATAN GARCÍA CAMPOS

El Dr. Jonatan García Campos estudió la licenciatura en Filosofía en la Universidad Autónoma Metropolitana, Unidad Iztapalapa (UAM-I), posteriormente realizó estudios en Filosofía (Línea de Epistemología) y Doctorado en Filosofía de la Ciencia, ambos programas de posgrado en la UNAM.

Es Profesor-Investigador en la Universidad Autónoma Metropolita, Unidad Azcapotzalco (UAM-A), donde actualmente imparte las unidades de aprendizaje Doctrinas Políticas y Sociales.

Sus intereses son la epistemología contemporánea, las teorías de las decisiones y el razonamiento, lo estudios sobre deshonestidad y la historia del pensamiento moderno, áreas en donde ha publicado. A la

fecha cuenta con tres libros de autor, ha coordinado o editado ocho libros, diez traducciones publicadas, así como cuarenta y ocho artículos y capítulos de libro.

Ha sido distinguido por el CONAHCYT como miembro del Sistema Nacional de Investigadores (SNII nivel II).

Es parte del proyecto Contextos Virtuales de Aprendizaje, proyecto inter-unidades de la UAM.

JORGE ALBERTO ROSAS CASTRO

Jorge Alberto Rosas Castro, es Doctor en Administración y tiene un Postdoctorado en Estudios Organizacionales, tiene el reconocimiento Nivel I del Sistema Nacional de Investigadores y Perfil PRODEP, se desempeña como Profesor Investigador en el Departamento de Administración de la División de Ciencias Sociales y Humanidades de la Unidad Azcapotzalco de la Universidad Autónoma Metropolitana, en donde también participa con editor de la Revista de Gestión y Estrategia; como Interventor en Organizaciones y Gestión para el Desarrollo; y como Presidente de la Red Mexicana de Investigadores en Estudios Organizacionales (REMINEO), A.C. En el ámbito de las funciones sustantivas de docencia, investigación, preservación y difusión de la cultura, así como en vinculación, su interés académico y profesional está focalizado en la Administración Pública, en la Línea de Investigación y Generación del Conocimiento de Gestión Pública para el Desarrollo Sustentable, con énfasis en el orden de gobierno municipal y la gestión del territorio para mejorar el desarrollo rural y urbano, así como la prestación de servicios municipales. Ha escrito artículos científicos y capítulos de libro, y coordinado la edición de colecciones de libros electrónicos desde 2012 a la fecha en las diferentes áreas temáticas transdiciplinarias de los estudios organizacionales, así también ha dirigido la organización de 18 Congresos Internacionales de Análisis y Estudios Organizacionales.

JUANA EDITH NAVARRETE MARNEOU

Dra. en Dirección y Finanzas, Profesora Investigadora Titular de la Universidad Autónoma del Estado de Quintana Roo, Miembro del Sistema Nacional de Investigadores, Nivel I (México) y Coordinadora del Cuerpo Académico de Estudios Administrativos y Organizacionales.

Especialista en finanzas bursátiles y corporativas e integrante del Comité Editorial de la Revista Avacient del Instituto Tecnológico Nacional.

Líneas de investigación: emprendimiento con enfoque de género, inteligencia artificial, gestión de la pequeña empresa, economía financiera, economía púrpura.

KARINA DEL REFUGIO VALLEJO QUINTERO

Es licenciada en Administración Pública por la Universidad de Occidente, con grado de Maestría en Calidad por la Universidad La Salle y Doctora en Innovación y Administración Educativa por el Centro Universitario de Ciencias e Investigación. Cuenta con capacitación permanente a través de diplomados y cursos de actualización en docencia, igualdad y equidad de género, derechos humanos, trabajo en equipo, manejo de grupos, mercadotecnia política, habilidades gerenciales y, en desarrollo humano.

Inició en 1988 sus labores de docencia en la Universidad de Occidente y desde entonces a la fecha, en esta institución ha ocupado los cargos administrativos de Coordinadora de la Unidad Los Mochis, Contralora, Coordinadora de Planeación y Estadística, Coordinadora de los Programas Educativos de Licenciatura en Administración de Empresas y Licenciatura en Contaduría y Finanzas, y, directora de la Unidad Regional El Fuerte. Actualmente es Profesora de Tiempo Completo y Vicerrectora de Operación Institucional.

Durante su trayectoria profesional ha ocupado diversos cargos en la administración pública en los tres niveles de gobierno, entre los que destacan: subdirectora de educación y cultura, asesora de la presidencia municipal, directora de recursos humanos y, coordinadora del Co-

mité de planeación municipal, secretaria de la presidencia, todos en el Ayuntamiento de Guasave. También ocupó el puesto de delegada administrativa de la zona 02 de COBAES, el de jefa de la oficina municipal de la Procuraduría Federal del Consumidor en Guasave y Jefa de la Unidad Politécnica de Integración Social (UPIS), en el Centro de Investigación Interdisciplinaria para el Desarrollo Integral de la Región (CIIDIR) del Instituto Politécnico Nacional (IPN).

KARLA MARIELA SANDOVAL COTA

La Dra. Karla Mariela Sandoval Cota es una académica comprometida y profesora de medio tiempo en la Universidad Autónoma de Occidente (UAdeO), adscrita al Departamento Económico Administrativo en la Unidad Regional Los Mochis, Sinaloa, México. Es doctora en Pedagogía por el Centro de Investigación e Innovación Educativa del Noroeste y licenciada en Mercadotecnia por la UAdeO.

En su carrera docente, la Dra. Sandoval Cota ha impartido clases en los niveles de licenciatura, maestría y doctorado, y forma parte del catálogo de maestros del programa Delfín. Actualmente, colabora en el Grupo de Investigación en Educación y Desarrollo Integral. También ha sido coresponsable del proyecto de Maestría en Docencia y ha trabajado en el diseño instruccional para entornos virtuales de aprendizaje.

En el ámbito de la gestión académica, ha contribuido al rediseño y actualización del programa educativo de Mercadotecnia y al diseño del programa de posgrado de Maestría en Docencia en modalidades presencial y virtual. Además, es responsable del registro y rediseño del programa educativo de Maestría en Administración ante el Sistema Nacional de Posgrado.

La Dra. Sandoval Cota ha sido reconocida por su participación como juez en el Maratón Regional de Mercadotecnia de la Zona I Noroeste de ANFECA y como jurado en el Evento Nacional Estudiantil de Innovación Tecnológica del Tecnológico Nacional de México. También ha sido expositora en la Semana de la Ciencia e Investigación de CONACYT y evaluadora de la Revista Lumina de la Universidad de Manizales, Colombia.

En su trayectoria, ha dirigido tesis a nivel de maestría y doctorado, y ha ocupado diversos cargos de gestión, incluyendo la coordinación institucional de la Maestría en Docencia Presencial y en Línea. Actualmente, es jefa de Investigación y Posgrado de la Unidad Regional Los Mochis de la UAdeO, donde continúa contribuyendo al desarrollo académico y profesional de sus estudiantes y colegas.

LUIS ANTONIO VERDUGO CALDERÓN

Licenciado en Informática por el Instituto Tecnológico Nacional y Maestro en Administración con acentuación en Finanzas por la Universidad Interamericana del Norte. Actualmente, es estudiante del Programa de Doctorado en Innovación y Administración de Organizaciones en la Universidad Autónoma de Occidente. Además, ha cursado diplomados en Teleinformática por la Universidad Autónoma de Guadalajara y en Ciencia de Datos por el Instituto Tecnológico del Petróleo y Energía.

A lo largo de su carrera, ha liderado su desarrollo profesional en el área de gestión de crédito en los sectores bancario, industrial y comercial, así como en la gestión de inmuebles. Su experiencia incluye roles operativos y de toma de decisiones, abordando los desafíos desde una perspectiva multidisciplinaria y estratégica.

Asimismo, ha desempeñado funciones como instructor, evaluador y conductor de talleres en diversas instituciones educativas y organizaciones. En este ámbito, ha impartido cursos en computación, administración y finanzas, enriqueciendo su experiencia docente a través de la participación en congresos y la asistencia a cursos de actualización.

MANUEL ALEJANDRO ZAVALA-NEVÁREZ

Manuel Alejandro Zavala-Nevárez, es Ingeniero Industrial y de Sistemas, por el Instituto Tecnológico de Sonora y posee una Maestría en docencia en educación media y superior y un Doctorado en Innovación educativa por el Instituto Tecnológico Superior de los Mochis.

Desde 2009, se ha desempeñado como Profesor en la Universidad Autónoma de Occidente, participando en programas de licenciatura, en materias del área de matemáticas e ingeniería y en programas de maestría. Actualmente coordina la maestría en docencia virtual.

MARITZA ALESSANDRA ZEPEDA BARBOSA

Estudiante de la Universidad Autónoma Metropolitana, unidad Azcapotzalco, institución reconocida por su enfoque en la investigación y la formación integral de sus estudiantes. Ha tenido la oportunidad de consolidar sus conocimientos en áreas clave como la gestión de empresas, análisis financiero y mercadotecnia. Estos estudios no solo le han permitido desarrollar un sólido marco teórico, sino también poner en práctica sus habilidades a través de proyectos y trabajos colaborativos que simulan desafíos reales del entorno empresarial. Su interés por la tecnología y la innovación la ha llevado a participar en webinars sobre Inteligencia Artificial (IA), profundizando en las implicaciones de esta tecnología en el futuro del trabajo y las organizaciones. Como resultado de esta experiencia, ha sido coautora de un capítulo del libro "*El futuro de la Inteligencia Artificial, Ética, Transhumanismo y Sustentabilidad en la IA*".

Actualmente, realiza una pasantía en Samsung Electronics, donde ha tenido la oportunidad de involucrarse profundamente en el sector de ventas B2B. En esta posición, se encarga de analizar y planificar estrategias que permitan maximizar las ventas, aportando apoyo directo al equipo de cuentas que gestiona a los clientes más importantes de la compañía.

Con ambas experiencias durante la licenciatura, adquirió la capacidad de unir las demandas del mercado con las posibilidades de la Inteligencia Artificial y cómo ésta impacta en la industria de la tecnología.

MARLENE ATENEA GRIMALDO RAMÍREZ

Es estudiante de la Licenciatura en Administración en la Universidad Autónoma Metropolitana, Unidad Azcapotzalco, donde actualmente

se desempeña como Consejera Académica. En este rol, ha participado activamente en la toma de decisiones académicas, representando a sus compañeros con un fuerte compromiso. Ha sido ponente en eventos académicos de relevancia, como el Congreso CIAO 2023, donde presentó la ponencia titulada *El Futuro de la Economía Social y Solidaridad: Afrontamiento hacia el Cambio Climático*. También participó en el webinar *El Futuro de la Economía Social y Solidaridad*, llevados a cabo en las siguientes ubicaciones: Bacalar, Guanajuato y Cancún. Adicionalmente, participó en el Webinar *El Futuro de la IA: Mitos y Fobias*, de donde surgió el libro *El Futuro de la Inteligencia Artificial*, al cual contribuyó con un capítulo enfocado en el impacto de la inteligencia artificial en el deporte, combinando su pasión por la tecnología y el deporte. En el ámbito técnico, ha brindado apoyo en el proceso de los sistemas del Sistema Nacional de Investigadoras e Investigadores (CONAHCYT) y asistió en las ponencias de la décima edición del Librofest en la UAM Azcapotzalco. También ha participado en el proyecto UAMedia Cursos y Talleres, brindando soporte técnico en la organización de eventos, curso y talleres. Además, su compromiso va más allá del ámbito académico, ya que está interesada en causas sociales, buscando siempre hacer un impacto positivo en su comunidad. A través de sus contribuciones, ha logrado consolidar una trayectoria que resalta tanto por su capacidad técnica como por su aportación intelectual en diversos campos.

MIGUEL ÁNGEL OLIVARES URBINA

Es Licenciado en Administración Industrial por el Instituto Politécnico Nacional, Maestro en Economía y Administración Pública por la Universidad de Quintana Roo, y Doctor en Desarrollo Económico y Sectorial Estratégico por la Universidad Popular Autónoma del Estado de Puebla. Ha desempeñado cargos en el sector público, incluyendo su rol como Gerente General de Pastas y Pallets de México, y Jefe del Departamento Administrativo en la Universidad de Quintana Roo.

En la Universidad del Caribe, fue Profesor Investigador Titular "A" y Jefe del Departamento Académico de Economía y Negocios entre 2007 y 2017, donde impulsó el desarrollo académico de su equipo. Ha

impartido clases en diversas universidades en áreas económico-administrativas y participado en conferencias nacionales e internacionales como UPAEP, ACACIA, y OCDE. Además, se ha destacado como evaluador de programas académicos (CACECA) y ha colaborado con Conahcyt y otras instituciones en proyectos educativos y tecnológicos.

A nivel académico, ha publicado más de 25 artículos en revistas indexadas y más de una docena de libros, además de colaborar en el diseño de programas de estudio en la Universidad del Caribe y otras universidades. Durante su gestión, los programas del Departamento de Economía y Negocios alcanzaron el Nivel 1 de acreditación de CIEES y COPAES. Su trayectoria le ha valido reconocimientos internacionales y locales, como su pertenencia al Sistema Nacional de Investigadores (Nivel 1) y la acreditación de perfil deseable PRODEP.

Actualmente, es Profesor Investigador de Tiempo Completo en la Universidad del Caribe, donde imparte clases de posgrado y licenciatura en Planeación Estratégica, Desarrollo Regional y Diseño de Proyectos. Además, coordina los posgrados de Negocios Electrónicos y Gestión de Proyectos, y lidera el proyecto observatorio Mipyme en colaboración con el cuerpo académico de Negocios Internacionales.

OSCAR LOZANO CARRILLO

Coordinador del Laboratorio de IA del Departamento de Administración de la UAM Azcapotzalco. Licenciado en Administración, Maestría y Doctorado en Estudios Organizacionales por la UAM. Profesor Titular de Administración de la UAM-A y en los Posgrados en Estudios Organizacionales UAMI e Intervención en las Organizaciones de la UAM-A. Miembro del Sistema Nacional de Investigadoras e Investigadores de México desde 2008 y a la fecha Nivel II; Coordinador de UAMEDIA MIPYME 4.0 con más de 2 mil intervenciones estratégicas digitales en MIPYMES y Cooperativas en México y Latinoamérica. Autor y coordinador de 31 libros con más de 30 capítulos de libro y 15 artículos en revistas autorizadas. Ha impartido más de 100 conferencias y ponencias en eventos nacionales e internacionales y ha Coordinado más de 30 proyectos de intervención en organizaciones públicas

y grandes empresas. Actualmente coordina el proyecto IA Estratégica para la MIPYME con más de 200 microempresas participantes.

PAOLA HERNÁNDEZ CHÁVEZ

Profesora-investigadora del Departamento de Administración de la Universidad Autónoma Metropolitana Unidad Azcapotzalco. Es doctora en filosofía de la ciencia (ciencias cognitivas) por el Instituto de Investigaciones Filosóficas de la UNAM, Maestra en Humanidades (Historia y Filosofía de la Ciencia) y Licenciada en Filosofía por la UAM Iztapalapa. Realizó un posdoctorado en Filosofía de la Biología en la UAM-Cuajimalpa. Fue profesora visitante y es longtime Fellow del Center for the Philosophy of Science, Universidad de Pittsburgh (#1 en el Ranking mundial en Filosofía de la Ciencia), asimismo fue Fellow del Konrad Lorenz Institute for Evolution and Cognition. Pertenece al Sistema Nacional de Investigadores de CONAHCYT. Dentro de sus publicaciones están: 10 artículos, 16 capítulos de libro, 3 ediciones de libro, 1 libro original, 8 traducciones. Sus áreas de investigación son: ciencias cognitivas, estudios teóricos y experimentales sobre la corrupción, disfunciones cognitivas, historia y filosofía de la biología, teorías evolutivas, tecnologías bioinspiradas, sistemas complejos, e interdisciplina en general.

RAFAEL ANGELES

Es Licenciado en Derecho por la UAM Azcapotzalco, estudia la Maestría en Educación en la Universidad Latinoamericana y colabora como Asesor en la Rectoría de la UAM Azcapotzalco. Cuenta con una certificación de Google en Marketing Digital y un diplomado en Social Media por Crehana.

Es creador de proyectos como UAMedia e IDEA, cuyos ejes rectores son Educación Digital, Inteligencia Artificial y Marketing Digital; temas sobre los cuales ha impartido más de 100 cursos y diplomados capacitando a más de 5,000 usuarios.

A través de *latintaroja* comparte el resultado de su experiencia en temas como IA en academia y docencia, pedagogía en espacios digitales, digitalización de ambientes y metaverso.

RICARDO PÉREZ BOJÓRQUEZ

El Lic. Ricardo Pérez Bojórquez es un destacado académico y profesional en el campo del derecho, con una sólida formación y experiencia en diversas áreas del derecho público. Es licenciado en Derecho y Ciencias Sociales por la Universidad Autónoma de Occidente (UAdeO), Unidad Regional Guasave, y actualmente cursa una maestría en Derecho Constitucional y Amparo en la Universidad Tecnológica de México. Su especialización incluye certificaciones en Derecho Constitucional, Derechos Humanos y Derecho Internacional Público.

El Lic. Pérez Bojórquez ha participado activamente en programas académicos de investigación, formación y promoción de derechos humanos y desarrollo sostenible. Además, ha colaborado en actividades de incidencia social y voluntariados con diversas organizaciones nacionales e internacionales. Su compromiso y excelencia en el campo del derecho público han sido reconocidos en concursos, foros y congresos sobre política y derecho.

En la actualidad, se desempeña como Profesor de Asignatura Base en la UAdeO, Unidad Regional Guasave, y ocupa un cargo administrativo como jefe de Sección en la Vicerrectoría de Operación Institucional de la misma universidad. También es miembro del Grupo de Investigación "Educación y Desarrollo Integral", donde participa en proyectos académicos relacionados con derechos humanos, educación y desarrollo integral.

La trayectoria del Lic. Pérez Bojórquez refleja su dedicación a la enseñanza, la investigación y la promoción de los derechos humanos, contribuyendo significativamente al desarrollo académico y social en su comunidad.

RICARDO SONDA DE LA ROSA

Profesor investigador de tiempo en la Licenciatura de Turismo Sustentable y Gestión Hotelera de la Universidad del Caribe, Cancún, México.

Licenciado en Administración Hotelera, Maestro en Mercadotecnia turística, Maestro en turismo y dirección hotelera y Doctor en administración. Miembro del Sistema Nacional de Investigadores "SNI I" del CONACYT. Miembro del Sistema Estatal de Investigadores del Estado de Quinina Roo Nivel III y Honorifico.

Autor de libros y artículos científicos publicados en revistas arbitradas e indexadas con líneas de investigación de Servicios Turísticos, Hotelería y Cultura Organizacional.

Certificado en Habilidades Directivas y de Poder, Hospitality Educator, Luxury Managemet.

Ponente Nacional e internacional para temas humanísticos y de desarrollo de talentos, Arbitro en Journals especializados en turismo, y miembro de diversas redes de colaboración para el turismo y sustentabilidad.

Hotelero de formación y académico por convicción, se desarrolló profesionalmente en la hotelería en la gestión del talento humano en cadenas como Radisson, Krystal, Sheraton, Westin, Grupo Karisma, Meliá; como directivo académico colabora para Universidad Anáhuac, Tecnológico de Monterrey, La Salle, Universidad Tecnológica y Universidad del Caribe.

ROSILUZ CEBALLOS POVEDANO

Economista, Doctora en Estudios Organizacionales, Profesora Investigadora del Departamento de Economía y Negocios y Coordinadora de los Programas Educativos de Negocios Internacionales e Innovación Empresarial de la Universidad del Caribe. Miembro del Sistema Nacional de Investigadores Nivel 1. Es especialista en estudios socioeconómicos y socio ambientales con publicaciones orientadas hacia el desempeño de las organizaciones de la economía social y la sustenta-

bilidad en Quintana Roo. Entre sus publicaciones se ubican: Cooperativismo y economía social en las universidades de Quintana Roo; Diagnóstico socio económico y ambiental Las Salinas de Isla Mujeres; Aproximación Socioeconómica En Una Cooperativa De Producción Pesquera; Recuperación de la laguna Salina Grande mediante la economía digital. El cooperativismo en América Latina: entre la autogestión y la dependencia del Estado; Prácticas de la economía social en la realización del plan nacional 2018-2024, entre otras.

VÍCTOR ALFONSO GASCÓN SORIANO

Estudiante de Administración y promotor de investigación y difusión del Laboratorio de Inteligencia Artificial de la UAM Azcapotzalco, conferencista en Webinars sobre inteligencia artificial.

He participado en los seminarios del "Proyecto de Integración de la Economía Social con las Economías Alternativas (PIESEA)", donde trabajé el tema de la economía popular, en instituciones, como la Universidad del Caribe en Cancún y la universidad de Guanajuato.

VÍCTOR DANIEL SANTOS HORTELANO

Estudiante en Sociología por la Universidad Autónoma Metropolitana, Unidad Azcapotzalco. Ha trabajado en el Instituto Nacional Electoral y ha tomado cursos de apreciación cinematográfica en el CCC, así como cursos de historia del cine y cine mexicano. Además, ha impartido cursos de apreciación cinematográfica para principiantes en la UAM Azcapotzalco. En el ámbito laboral, ha trabajado en el área de comunicación de la UAM como community manager, creador de contenido y gestor de redes. Actualmente, es asistente de investigación en el laboratorio de inteligencia artificial del Dr. Óscar Carrillo Lozano del Departamento de Administración. Además, ha participado en los seminarios PIESEA, relacionados con la economía social y solidaria, en 5 ediciones.

VIRGINIA LÓPEZ-NEVÁREZ

Virginia López-Nevárez es Licenciada en Informática por el Instituto Tecnológico de Los Mochis y posee una Maestría y Doctorado en Estudios Organizacionales por la Universidad Autónoma Metropolitana. Además, cuenta con una especialidad en Responsabilidad Social Empresarial en el Observatorio de RSE en España (2020-2021). Es miembro del Sistema Nacional de Investigadores del CONAHCYT Nivel 1, así como del Comité Científico de la RED Nacional de Actividades Juveniles en Ciencia y Tecnología desde 2011, y de la Red Mexicana de Investigadores en Estudios Organizacionales desde 2002. Recibió el reconocimiento LORE DE LA VEGA a la Mujer Ahomense 2021 por su trayectoria científica. Desde 1998, se ha desempeñado como Profesora Investigadora de Tiempo Completo en la Universidad Autónoma de Occidente, participando en programas de licenciatura, maestría y en los doctorados de Sustentabilidad y de Innovación y Administración de las Organizaciones, siendo actualmente Directora de la Unidad de Educación Virtual. Como asesora de proyectos de investigación, ha obtenido diez premios nacionales en ferias científicas y dos internacionales, y ha editado y coordinado 12 libros de divulgación científica con la RED Nacional de Actividades Juveniles. También ha sido profesora invitada en las Universidades de Manizales y El Quindío en Colombia.

WENDY SEBASTIANA HERNÁNDEZ DEL PUERTO

Licenciada en Informática, Maestra en Innovación y gestión del Aprendizaje, Profesora Investigadora del Departamento de Economía y Negocios de la Universidad del Caribe. Miembro de la red de mujeres de ANUIES TIC. Experiencia en consultoría en la implementación de TI. Es especialista en el análisis e implementación de las tecnologías de Información y Comunicación para mejorar los procesos en las organizaciones, con publicaciones orientadas enfocadas en el aprendizaje basado en problemas, el pensamiento crítico, y la transversalidad en entornos virtuales. Participante en congresos y seminarios de innovación educativa.

YADIRA ZAVALA OSORIO

Licenciada en Ingeniería Industrial por la Unidad Azcapotzalco de la Universidad Autónoma Metropolitana (UAM); Maestra en Ingeniería por la Facultad de Ingeniería de la Universidad Nacional Autónoma de México (UNAM); y Doctora en Estudios Organizacionales por la Unidad Iztapalapa de la UAM. Actualmente se desempeña como Rectora de la Universidad Autónoma Metropolitana de la Unidad Azcapotzalco.

Es una destacada Profesora-Investigadora del Departamento de Sistemas de la UAM Azcapotzalco. Investigadora Nacional Nivel I por el Sistema Nacional de Investigadoras e Investigadores (SNII) y cuenta con Reconocimiento a Perfil Deseable (PRODEP).

A lo largo de su admirable trayectoria profesional, ha ocupado importantes cargos, como Secretaria de la Unidad Azcapotzalco de la UAM (2021-2023) y Secretaria Académica de la División de Ciencias Básicas e Ingeniería (2008-2010) en la misma institución.

La doctora, ha participado como sínodo en exámenes de grado de varios alumnos de la Maestría en Ingeniería de la Universidad Nacional Autónoma de México (UNAM) y de la Maestría en Estudios Organizacionales, UAM-Iztapalapa.

Se distingue por su experiencia en la organización de eventos educativos de alto nivel, así como en el análisis empresarial para la mejora de departamentos a partir de análisis de campo. Además, ha demostrado habilidades en el diagnóstico sistémico para implementar procedimientos de aseguramiento de calidad y en la aplicación de técnicas efectivas para la detección de problemas en organizaciones.

En el ámbito profesional laboró como ingeniera industrial en la empresa denominada Grupo IMPRESSA y en GPI México, dedicándose a la elaboración de diagnósticos y detección de áreas de oportunidad, además participó en el desarrollo de metas e indicadores de la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS). Fue parte del grupo de trabajo en la elaboración de metas para la evaluación del desempeño del personal del Instituto Nacional de Estadística y Geografía (INEGI) y participó en el fortalecimiento del perfil de la Secretaría de Salud. Colaboró como coordinadora en el proyecto de

vinculación entre los Laboratorios de Biológicos y Reactivos de México (BIRMEX) y la UAM.

Ha publicado los libros PEMEX, Presente y futuro, Sistema Nacional de Distribución de Insumos para la Salud, BIRMEX y Manual de seguridad e Higiene Industrial. De igual forma, ha realizado diversas publicaciones relativas a los retos de las sociedades del conocimiento, al desarrollo de tecnología e investigación sobre el petróleo en México, la competitividad empresarial, la educación a distancia, asimismo, seguridad e higiene industrial, entre otros; ha participado en la coordinación e impartición de diplomados sobre logística y cadena de suministro. Actualmente, su investigación se enfoca en dos vertientes; la primera relacionada con el impacto de la seguridad e higiene en las organizaciones, y la segunda con respecto al análisis de las variables organizacionales, especialmente en las escuelas de educación superior.

ZAIRA YEREMI MARTÍNEZ HERRERA

Licenciada en Administración, egresada de la Universidad Autónoma Metropolitana, unidad Azcapotzalco. Durante mi formación académica, tuve la oportunidad de participar activamente en diversas iniciativas que enriquecieron mi experiencia y conocimiento en el campo de la administración y la inteligencia artificial.

Destaco mi participación en el "Primer Encuentro de IA", supervisado por el Dr. Lozano Carrillo y el Dr. Alfredo Garibay Suárez, organizado por la Universidad Autónoma Metropolitana, en el departamento de administración, así como mi asistencia a varios congresos, charlas y seminarios relevantes. Algunos de los eventos más significativos incluyen:

- WEBINAR: El futuro de la Inteligencia Artificial: Ética, Transhumanismo y Sustentabilidad
- La Inteligencia Artificial aplicada a la Gestión Estratégica (podcast)
- WEBINAR: La Inteligencia Artificial y Estrategia

- Futuro de la IA: Mitos y Fobias
- Curso-Taller: IA para la Educación
- Conferencia: Filogénesis en la IA, entre otros.

Estas actividades me han permitido adquirir una sólida comprensión de las tendencias actuales en el uso de la inteligencia artificial en el ámbito empresarial y su impacto en la gestión estratégica.

Recientemente, tuve el honor de contribuir como coautora en el libro "Inteligencia Artificial Estratégica para la Micro, Pequeña y Mediana Empresa. Una nueva ética de democratización del conocimiento", donde colaboré con los doctores Oscar Lozano Carrillo, Alfredo Garibay Suárez y Jorge Alberto Rosas Castro. Mi testimonio se incluye entre las valiosas perspectivas que ofrece este libro, lo que refleja mi compromiso con la difusión del conocimiento y la innovación en el campo de la administración.

A través de estas experiencias, he desarrollado habilidades analíticas y críticas, así como un enfoque ético hacia la aplicación de tecnologías emergentes en el entorno empresarial. Estoy interesada por continuar mi trayectoria profesional, contribuyendo a la transformación y modernización de las prácticas administrativas en un mundo cada vez más digitalizado.

Esta obra fue editada por:

Red Mexicana de Investigadores
en Estudios Organizacionales (REMINEO), A.C.

América 126-A int. 202, Col. Parque San Andrés,
Coyoacán, Ciudad de México, C.P. 04040

remineo.org

Ética, Transhumanismo y Sustentabilidad son tres temáticas fundamentales que se están discutiendo hoy en día en todos los espacios de reflexión humana, es decir, academia, gobierno, empresas, científicos y en la vida cotidiana de las familias se está hablando de como la IA está impactando y de lo sorprendente, pero también peligros que es su uso en la educación o en la guerra, en la transhumanización con los cyborgs y en la nanotecnología médica, pero también en el impacto ecológico que puede tener el mantener tantos servidores trabajando con modelos de energía tradicionales y obsoletos ambientalmente.

El presente libro constituye es producto del esfuerzo de múltiples instituciones de educación superior de México y América Latina que coordinadas por la UNIVERSIDAD AUTÓNOMA METROPOLITANA, convergieron en el Webinar "El futuro de la IA" realizado en septiembre de 2024 en con la hospitalidad de la UNIVERSIDAD DEL CARIBE , donde se presentaron diferentes tópicos relacionados con la Ética, Transhumanismo y Sustentabilidad y su impacto en las finanzas, educación, empleo, deporte, salud y la mente, computación, administración, regulación y normatividad; equidad, internacionalización, energías, economía social y la política pública, lo que muestra el poder de la transformación que la IA esta generando en el mundo, con un impacto no sólo tecnológico, sino también económico, político, social y por lo tanto cultural, en una dimensión superior incluso a la de la revolución industrial. La IA no va a una velocidad sorprendente y no se va a detener, el mundo tiene que estar listo para este cambio y una forma en la que podemos estar atentos es mediante artefactos como este libro que presenta los últimos avances y debates en múltiples temáticas de interés medular para la humanidad.

