INTELIGENCIA ARTIFICIAL

Transformación, retos y prospectiva social

José Antonio García Macías
Isaac de Jesús Palazuelos Rojo
Diego Alfredo Pérez Rivas
Coordinadores

Inteligencia Artificial

Transformación, retos y prospectiva social

astra editorial

RED JEDI SD

Inteligencia Artificial

Transformación, retos y prospectiva social

José Antonio García Macías
Isaac de Jesús Palazuelos Rojo
Diego Alfredo Pérez Rivas
Coordinadores

astra editorial

RED JEDI SD

Inteligencia Artificial: transformación, retos y prospectiva social / **Autores-coordinadores**: José Antonio García Macías, Isaac de Jesús Palazuelos Rojo y Diego Alfredo Pérez Rivas. —Ensenada, Baja California, 2024.

Publicación electrónica digital: descarga y *online*; detalle de formato: EPUB.

Primera edición

ISBN: **978-84-10215-89-4**

DOI: **https://doi.org/10.61728/AE20240998**

D. R. © copyright 2024; José Antonio García Macías, Isaac de Jesús Palazuelos Rojo y Diego Alfredo Pérez Rivas.

Este libro es producto de la Red multidisciplinaria en Justicia, JEDI-SD Ética, Democracia e Inclusión en las Sociedades Digitales (JEDI-SD) de la Universidad Autónoma de Baja California (ABC) y el Centro de Investigación Científica y Educación Superior de Ensenada (CICESE). https: //si tes . google . com/uabc .edu .mx/red-jedisd

Edición y corrección: **Astra Ediciones**

La presente obra fue dictaminada bajo el sistema de doble ciego por pares académicos externos, especialistas en la materia.

Se prohíbe la reproducción, el registro o la transmisión parcial o total de esta obra por cualquier sistema de recuperación de información, sea mecánico, fotoquímico, electrónico, por fotocopia, cualquier otro existente o por existir, sin el permiso previo, por escrito, del titular de los derechos.

HECHO EN MÉXICO | MADE IN MEXICO

Contenido

Impacto social de la IA: Una reflexión interdisciplinar que debemos construir
José Antonio García Macías
Isaac de Jesús Palazuelos Rojo
Diego Alfredo Pérez Rivas
Parte I
Contextualizando la IA
Capítulo 1
Apuntes para una breve prehistoria filosófica de la IA
Diego Alfredo Pérez Rivas
Virginia Guadalupe López Torres
Capítulo 2
Inteligencia Artificial: ni muy automatizada, ni muy ecológica
Antonio de Jesús García Chávez
José Antonio García Macías
Capítulo 3
Implicaciones sociales, retos y desigualdades en la incorporación de la inteligencia artificial en el contexto latinoamericano
Isaac de Jesús Palazuelos Rojo
Denys Serrano Arenas
Capítulo 4
Exploración de una perspectiva sistémica alrededor del desarrollo de la inteligencia artificial y su adopción social: la responsabilidad de la academia
Arturo Serrano-Santoyo
Capítulo 5
Fuck the algorithm: Navegando la promesa tecnológica y el impacto social de la IA
Marisol Flores-Garrido
Parte II
IA con enfoque social
Capítulo 6
Las posibilidades de una IA indígena
Maximino Matus Ruiz

Capítulo 7
De la tecnofilia a la tecnofobia. Narrativas sobre inteligencia artificial y educación superior
Christian Fernández Huerta
Hugo Méndez-Fierros
Joaquín Caso Niebla
Capítulo 8
Representaciones Sociales e Imaginarios Sociales Sobre la IA en el Entorno Educativo
Ulises Suárez Estavillo
Héctor Ignacio Castañeda García
Capítulo 9
Inteligencia artificial y derecho: El dilema entre la formalización y la aproximación
Virginia Berenice Niebla Zatarain
Jesús Manuel Niebla Zatarain
Gonzalo Armienta Hernández
Capítulo 10
Los límites de la aplicación de la Inteligencia Artificial (IA) en la gobernanza: una aproximación desde la política pública urbana y la cibernética social
Alfredo Leonardo Carreón Corona
Luis Gustavo Kelly Torreblanca
Capítulo 11
Sistemas de predicción de lenguas visogestuales basados en IA y su aplicación en la Lengua de Señas Mexicana (LSM)
Guillermo Hernández-Santana
Irvin Hussein López-Nava
Capítulo 12
Inteligencia Artificial Generativa (IAG). El nuevo paradigma en el diseño y producción de contenido
Juan José García Llamas
Capítulo 13
La inteligencia artificial y sus beneficios en una agricultura sostenible
Oscar Méndez García

Parte III
Entorno a la IA: reflexiones sobre salud, conciencia y neuropsicología
Capítulo 14
La IA en la salud
Jessica Beltrán
Capítulo 15
Entre la ciencia y la subjetividad: Reflexiones sobre la medición de la inteligencia y la conciencia
Edgar Chávez
Capítulo 16
Hacia la conciencia en inteligencia artificial: Un viaje por la evolución hasta las redes neuronales artificiales
Mariano Rivera
Capítulo 17
El uso de herramientas tecnológicas y de la información para la intervención e investigación en la neuropsicología
Rubén Avilés Reyes
Edgar Adrián Arias Chávez
Javier Tadeo Sánchez Betancourt
Acerca de los autores

Impacto social de la IA: Una reflexión interdisciplinar que debemos construir

José Antonio García Macías
Isaac de Jesús Palazuelos Rojo
Diego Alfredo Pérez Rivas

Aquí nos topamos con algo que suena a paradoja. Por su naturaleza misma, las computadoras son los animales más inflexibles, los más privados de deseos, los más seguidores de reglas. Pese a su gran rapidez, son el epítomede la inconsciencia. ¿Cómo programar entonces la conducta inteligente?... Desde luego, tiene que haber reglas en toda clase de niveles distintos. Tiene que haber muchas reglas "llanas y simples". Tiene que haber "metarreglas" con que modificar las reglas "llanas y simples", y en seguida "metametarreglas" con que modificar las metarreglas, y así hasta nunca acabar. La flexibilidad de la inteligencia es resultado del enorme número de reglas distintas y de niveles distintos de reglas que existen.

–Douglas Hofstadter

Desde los albores de la inteligencia artificial, los retos a los que se han enfrentado sus estudiosos y desarrolladores han sido numerosos. En primer lugar, intentar superar o burlar algunas de las paradojas como la señalada por Hofstadter al inicio de esta introducción. En segundo lugar, los obstáculos técnicos asociados al desarrollo mismo de un lenguaje sobre el cual poder hablar de inteligencia mecánica y, sobre todo, poder desarrollarla a través de instrucciones. En tercer lugar, el desarrollo de la inteligencia artificial ha estado íntimamente ligado al desarrollo tecnológico del soporte físico, el cual ha sido ampliamente expandido con el desarrollo del procesamiento paralelo masivo y unidades de procesamiento especializadas tales como GPU, TPU y LPU. Actualmente, la inteligencia artificial parece enfrentarse a un nuevo reto proveniente de los "entusiastas", los cuales

pueden sentarse de uno u otro lado de la tribuna popular, ya sea el optimismo o el pesimismo tecnológico.

Parece evidente que en los nuevos fenómenos sociales vinculados a la tecnología, al igual que sucede en otros fenómenos que acontecen en la intersección entre las ciencias computacionales y de la información y las ciencias sociales, existe una notable bifurcación emocional. Se podría decir que dicha bifurcación parece estar destinada, de hecho, a representar formas ideológicas antagónicas que, por un lado, emergen desde la valoración pesimista de las aportaciones tecnológicas a nuestro *modus vivendi*, hasta el culto y la sacralización de lo tecnológico como factores que concluyen el círculo de la humanización de nuestra cultura. Por otra parte, se apunta la necesidad de incrementar el cuidado y la regulación de los usos sociales de la tecnología, hasta relativizar negativamente cualquier transformación a nivel histórico y ambiental, la cual puede ser producida en relación directa con las implementaciones tecnológicas en la vida social. Hablamos, naturalmente, de las corrientes tecnofílicas y las corrientes tecnofóbicas.

Si bien es cierto, el discurso en torno a estas formas radicales de visualizar la innovación tecnológica obedece a un tipo de argumentación muy superficial. Lo cierto es que cuanto más avanza el desarrollo tecnológico y se incrementan sus aplicaciones sociales, los grupos sociales parecen congregarse en torno a mezclas más o menos coherentes de estos dos componentes. Dichas tendencias se transforman mediante distintos procesos de adaptación que nos advierten que ambas formas son posibles, además de que sus alcances pueden llegar a ser casi surrealistas. Algunos de esos desenlaces parecen emanar de la ciencia ficción, planteando algunas veces futuros utópicos y otras tantas distópicos.

En lo que respecta a las fobias, los discursos actuales disponen de mucho material empírico para sustentar sus argumentos. Las tecnologías orientadas a la guerra y a la dominación sobresalen por su disposición a manipular y atentar contra los derechos fundamentales de los ciudadanos. Por esa misma razón, parece adquirir cada vez mayor relevancia en algunos medios los discursos en torno a las tecnologías de vigilancia y espionaje. Los cuales, por cierto, no son necesariamente materia de ciencia ficción, toda vez que escándalos como los de Wikileaks y la filtración de Snowden parecen confirmar las sospechas. Otra de las preocupaciones más latentes versa sobre los nuevos mecanismos de estratificación poblacional, por

medio del uso de grandes volúmenes de datos, los cuales se disponen al servicio del consumismo y de las grandes empresas. Otra línea discursiva importante versa sobre la automatización y la creciente dependencia tecnológica de los procesos administrativos, lo cual puede apuntar a una sustitución del hombre por la máquina.

Por otra parte, del lado de las filias y del optimismo, se enlistan aquellos individuos y grupos que, montados en la esperanza de llevar a cabo algunas mejoras sustanciales en los campos de la medicina, la educación, la ingeniería, las comunicaciones y hasta los servicios públicos, ven en la tecnología la solución a todos los males de la sociedad. Es precisamente en un discurso de esta naturaleza en el que se entretejen todo un entramado de utopías escatológicas que consideran que la tecnología ha llegado, como Prometeo o un mesías, para redimir la condición humana. Algunos de esos discursos, naturalmente, provienen de los ejecutivos de las *startups* y empresas consolidadas que buscan financiamientos de riesgo en los mercados bursátiles. En este nuevo capitalismo de datos la percepción sobre la evolución de esta rama de la tecnología juega un papel crucial en su financiamiento y regulación, como si se tratara justamente de un bucle de Hofstadter.

La industria de la cultura ha encontrado una fórmula vencedora para explotar la alineación de los entusiastas al pesimismo y al optimismo tecnológico. El conflicto generado por posibles escenarios distópicos es taquillero, de la misma manera que lo puede llegar a ser algún relato que sea catártico de los anhelos más profundos de la humanidad: bienestar, salud, riqueza.

Con el presente volumen, los coordinadores intentan abonar a una discusión que trascienda la superficial dualidad entre filia y fobia. Por esa misma razón, ha invitado a estudiosos de los más diversos campos de estudio con la finalidad de proponer una hoja de ruta para entender el impacto social de la inteligencia artificial. Solamente a través de un camino con estas características, nos será posible romper el caparazón de las simplificaciones, para darnos cuenta que el verdadero estado del arte es más profundo que la dicotomía. En el corazón de la inteligencia artificial existen verdaderamente posibilidades técnicas de resolver problemas que acechan desde hace mucho tiempo a la humanidad, pero también plantea

algunos riesgos que no podemos hacer a un lado. A lo largo de los apartados del presente libro, los autores plantean estos diversos escenarios en el campo de la ecología, la educación, la psicología, la agricultura, la medicina y hasta la filosofía o la historia.

Los coordinadores han agrupado los capítulos de este libro en tres partes. La primera parte, "Contextualizando la IA", se compone de cinco capítulos en donde se discute la IA tanto en sentido histórico como prospectivo, precisamente para tratar de entenderla en contexto, con sus contribuciones, retos, logros y posibles implicaciones. En la segunda parte, "IA con enfoque social", se discuten a través de ocho capítulos algunos aspectos de las aplicaciones de la IA en entornos indígenas, en la educación superior, en el derecho, en la gobernanza y política pública, en los sistemas visogestuales, en la producción de contenido y en la agricultura sostenible. Por último, la sección titulada "En torno a la IA: reflexiones sobre salud, conciencia y neuropsicología" es, como su nombre sugiere, una colección de cuatro ensayos sobre aspectos humanos relacionados con salud y neuropsicología, aunque también con temas más especulativos tales como la inteligencia y la conciencia artificiales.

De esta forma, la primera parte del libro inicia con un capítulo del profesor Diego Rivas y la profesora Virginia López, donde nos proponen la lectura de la prehistoria de la inteligencia artificial como un sueño que ha estado latente en la mente de los filósofos naturales desde tiempos remotos. El capítulo de los profesores parte de la idea de que la inteligencia mecánica es un sueño latente en la historia de la filosofía y de la ciencia occidental, encontrando sus primeros ejemplos en el desarrollo de un mecanismo predictivo por parte de Arquímedes. Montados en dicho argumento, también se realiza una reconstrucción de las máquinas de cálculo diseñadas por Pascal y Leibniz, las cuales pueden considerarse en toda forma como las antecesoras inmediatas de nuestras computadoras digitales. La propuesta de los profesores parece apuntar al hecho de que la mecanización del pensamiento, o de las acciones consideradas como pensamiento, por ejemplo, el cálculo, ha sido el primer paso de la inteligencia artificial. Esto la ha hecho ser una actividad que durante mucho tiempo ha estado circunscrita a una concepción estrictamente antropocéntrica que actualmente tiende a ser replanteada desde otros frentes.

Los profesores Antonio de Jesús García Chávez y José Antonio García Macías analizan, en el segundo capítulo, el auge reciente de la inteligencia artificial generativa (IAG), centrándose particularmente en los grandes modelos de lenguaje (LLM) como GPT. Explican que estos modelos están capacitados para imitar la creatividad humana al ser entrenados con enormes cantidades de datos, lo que les permite generar textos convincentes, imágenes realistas y mantener conversaciones naturales. Enseguida, se destaca el papel crucial de los "trabajadores fantasma" o evaluadores humanos que revisan y proporcionan retroalimentación para mejorar las respuestas de los LLM; se describen las precarias condiciones laborales y los bajos salarios que enfrentan, a menudo teniendo que lidiar con contenido tóxico que puede causar traumas. También se analiza el impacto ambiental significativo del entrenamiento de los LLM, que implica enormes emisiones de carbono y un consumo excesivo de energía y agua, incluso en medio de sequías. En general, se ofrece una mirada crítica al auge de la IAG, cuestionando si sus beneficios justifican los costos ambientales y sociales asociados. Se sugiere que la industria sigue un modelo centrado en la tecnología en lugar de las personas, señalando la paradoja de que incluso una industria que busca automatizar todo, depende en gran medida del trabajo humano.

Por otro lado, en el tercer capítulo el doctor Isaac Palazuelos y la doctora Denys Serrano desarrollan una revisión documental descriptiva y exploratoria sobre los estudios sociales realizados en Latinoamérica en torno a la incorporación de la IA en tres ámbitos públicos: educación, salud y justicia. Esto con el objetivo de identificar las principales implicaciones sociales que se derivan de dicha incorporación en contextos históricamente marcados por la desigualdad social, además de un sistemático rezago en materia de desarrollo tecnológico. Proponen hacer una profundización cualitativa en las complejidades políticas, sociales y culturales que definen a las situaciones específicas en las que se emplea la IA. Destacan la necesidad de replantear los diseños tecnológicos desde visiones vernáculas que permitan una mejor adaptación tecnológica, así como un mejor cumplimiento de las necesidades para las que las aplicaciones fueron hechas. Argumentan que este diseño vernáculo y complejo ayudará también a la creación de otros indicadores para medir el desarrollo tecnológico, sin

embargo, en la misma línea subrayan la urgencia por aminorar las brechas digitales, además de combatir el rezago en materia de innovación y diseño tecnológicos en América Latina.

El investigador Arturo Serrano-Santoyo plantea en el cuarto capítulo una perspectiva sistémica sobre el desarrollo de la IA y su adopción social desde el rol de la academia. Primeramente, analiza las expectativas y reacciones encontradas ante el surgimiento de potentes herramientas de IA, desde el optimismo corporativo hasta las preocupaciones éticas y de gobernanza. Destaca la velocidad del cambio tecnológico, la interconectividad global, así como la naturaleza autónoma de la IA, todos estos como factores que exacerban la centralidad de la tecnología en la sociedad. Como parte de un análisis del contexto, examina la dinámica de la "Transformación Digital" impulsada por agentes corporativos enfocados en ganancias y dominio de mercado sobre el progreso social. Subraya la importancia de la regulación, ética y privacidad de datos como un tejido interdependiente en el desarrollo de la IA. Luego se enfoca en el rol de la academia, argumentando que la academia debe ir más allá de formar especialistas técnicos en IA. Finalmente, se propone un marco de referencia basado en dos pilares: Creación de Capacidades y Acción Social. Sugiere acciones como aprovechar sinergias entre regulación, ética y privacidad de datos, extender conocimientos a sectores sociales, mitigar la brecha de innovación, entre otras. En conclusión, se insta a la academia a tener un papel relevante cuestionando las narrativas predominantes, desarrollando estrategias éticas y regulatorias coherentes, y formando generaciones con una estructura moral sólida.

En el cierre de la primera parte del libro, la Profesora Marisol Flores plantea que, en parte debido a la ciencia ficción, así como otras narrativas, el imaginario colectivo, en torno a la IA y sus impactos sociales, se desarrolló en la sociedad previo a la formalización del propio uso y desarrollo de la IA. Derivado de ello, la autora argumenta que existe una tendencia a perseguir utopías tecnológicas la cual merma el tiempo de reflexión que dedicamos a los propios diseños y expectativas que tenemos de las aplicaciones asistidas por IA. Por tal motivo, propone la necesidad de replantearnos las características y los usos que deseamos respecto a la tecnología, para ello, nos presenta un estudio de caso sobre un algoritmo empleado por una escuela en Reino Unido para estimar calificaciones.

Destaca las nociones de sesgo algorítmico, complejidad y resistencia, para discutir la necesidad de construir visiones críticas sobre los sistemas de IA, con el objetivo de evidenciar sus verdaderas posibilidades y limitaciones, así como las estructuras de poder en donde se desarrollan o se utilizan.

Para la segunda parte del libro, el doctor Maximino Matus Ruiz inicia con un interesante capítulo en el que aborda las posibilidades de una IA indígena. Parte de una reflexión teórica sobre la conciencia como una forma de prótesis cultural, la cual se articula con las redes neuronales mediante un proceso intersemiótico que vincula procesos internos y externos que se van tejiendo para formar un estado de conciencia. El autor argumenta que en este proceso se abre la puerta no únicamente a un universo de posibilidades sino a un pluriverso de cosmovisiones diversas, en las cuales, los lenguajes constituyen las bases de las posibilidades de imaginar, sentir, percibir, experimentar e interpretar el mundo. El profesor Matus señala que, frente a esta importancia, resulta fundamental descentralizar el diseño tecnológico y diversificar sus resultados, mediante la incorporación de diversos lenguajes (indígenas) que sirvan como puerta de acceso a una verdadera pluralidad de saberes y formas de experimentar. Para ello, nos presenta el prototipo de la IA zapoteca Xquenda empleada como un mecanismo útil en la preservación de la lengua. Este prototipo también tiene la cualidad de replantear la relación humano-máquina, así como la codificación y el diseño tecnológico desde la experiencia del usuario.

En su capítulo acerca de las narrativas sobre inteligencia artificial y educación superior, los profesores Christian Fernández, Hugo Méndez y Joaquín Caso proponen una lectura a la luz de las antípodas tecnofilia y tecnofobia. La lectura de los profesores propone que el punto de inflexión para la inteligencia artificial ocurrió en 2019, a raíz de la publicación de Chatgpt en su tercera versión. Su análisis explora las ideas de internet incorporado y culturas digitales, como referentes en los que pueden encontrarse las perspectivas tecnofílica y tecnofóbica. Planteando un análisis basado en categorías inductivas proponen tres tipos de posturas dentro de cada una de esas perspectivas, las cuales van desde el utopismo tecnológico hasta el alarmismo tecnológico.

Por su parte, los doctores Ulises Suárez Estavillo y Héctor Ignacio Castañeda García abordan la implementación de la IA en instituciones de

educación superior, desde la construcción de imaginarios sociales, coincidiendo con la profesora Marisol Flores, exponen que dichos imaginarios se han constituido mucho antes de la llegada de la IA por diferentes narrativas, particularmente desde el cine y la literatura; asimismo, argumentan que estas concepciones colectivas en torno a la IA influyen altamente en el propio diseño y desarrollo de estrategias pedagógicas, limitando los marcos de interpretación sobre los verdaderos riesgos y oportunidades que tiene esta tecnología en el sector educativo.

Los profesores Jesús Manuel Niebla Zatarain, Virginia Berenice Niebla Zatarain y Gonzalo Armienta Hernández analizan la relación complementaria entre IA y derecho, ya que el razonamiento jurídico comparte características que lo hacen compatible con la IA como su naturaleza multimodal, su estilo explícito de creación de casos, el uso de fuentes documentadas, su carácter contextual e interpretativo, entre otros. Se plantea que los modelos aproximados de aprendizaje automático supervisado y no supervisado son prometedores para asistir en tareas jurídicas como la predicción de resultados de casos y la atención rápida de consultas legales en línea. Finalmente, se concluye que la IA representa una transformación del derecho al desarrollarse tecnología inteligente que replica procesos cognitivos legales y se adapta a entornos dinámicos como Internet, convirtiéndose en un componente normativo de diseño para prevenir actos perjudiciales

El capítulo de Alfredo Carreón y Luis Kelly aborda los límites de aplicación de la inteligencia artificial en la gobernanza, haciendo uso de la política pública urbana y la cibernética social para su disertación. Los autores proponen la idea de que la incorporación de la inteligencia artificial en la gobernanza se realiza a partir de la automatización en el procesamiento de datos y en la toma de decisiones. En su análisis del caso mexicano, los autores proponen estructurar una toma de decisiones pública en la que la inteligencia artificial sea parte integrante, entendiendo que su aplicación a la gobernanza urbana se debe a su incapacidad de sustituir el proceso político de toma de decisiones, así como el contexto político. El documento de Carreón y Kelly es muy relevante para reflexionar sobre el papel que tendrá la inteligencia artificial en la ciudad y la nación del futuro.

Los doctores Guillermo Hernández-Santana e Irvin Hussein López-Nava exponen la necesidad de invertir en la investigación y desarrollo de tecnologías para el reconocimiento automático de lengua de señas, exponen

que esto es un factor clave para la formación de una sociedad digital incluyente. Los autores exponen que esta inversión tiene una doble acepción, por un lado, permite conocer de manera más íntima las características del lenguaje de señas y, al mismo tiempo, esto permite desarrollar aplicaciones con mejores interfaces para fomentar las interacciones digitales de la comunidad sorda. La propuesta se centra no solo en la experiencia simbólica de la comunidad sorda, sino en el valor de la IA en el desarrollo de mecanismos institucionales de inclusión social.

Por su parte, el profesor Juan José García Llamas se centra en los efectos que tiene la IA en la cultura popular, así como sus implicaciones en materia de comunicación, ética, economía, derechos de autor, tiempo de trabajo y remuneración. Destaca que la generación de contenido es una de las áreas en las que más ha tenido impacto la implementación de IA en los últimos años, vincula este potencial con múltiples áreas de trabajo y profesionalización, pero también al ámbito privado y la construcción del yo. En su capítulo, nos ofrece una mirada reflexiva para reconocer los posibles riesgos en este campo y, al mismo tiempo, afinar tanto la categorización como la personalización de los contenidos en beneficio de una curaduría con valores de ética y equidad.

En su capítulo, el maestro Oscar Méndez destaca las características positivas que tiene la implementación de la IA en el sector agrícola, principalmente en función de alcanzar mecanismos sostenibles de producción alimentaria, para cubrir la demanda frente al acelerado crecimiento de la población mundial. El autor nos presenta una visión de optimización de costos e incremento de producción, a partir de la automatización a través del internet de las cosas y la big data. De acuerdo con esta perspectiva, los beneficios son tan amplios que se hace indispensable superar otras barreras sociales tales como la falta de infraestructuras, así como la formación y capacitación de mano de obra especializada, a cambio de la obtención de mejores resultados en dicha industria.

La doctora Jessica Beltrán Márquez inicia la tercera parte del libro con un capítulo donde propone la construcción de una ruta de análisis para entender el papel que está teniendo y que tendrá la inteligencia artificial en el campo de la salud. En primer lugar, la profesora nos explica cómo es posible que la inteligencia artificial pueda entrenarse para deducir, con

base en una serie de parámetros y mediante el uso de algunos dispositivos, nuestro estado de salud. Realizando una explicación sobre la evolución de la inteligencia artificial, la profesora Beltrán reconstruye algunos de los hitos de dicho campo, dirigiéndose enseguida a una serie de casos en los que el uso de la inteligencia artificial en el campo de la salud es una realidad.

El doctor Edgar Chávez propone un capítulo de reflexión que parece oscilar entre la ciencia de la computación y la filosofía, en la que se diserta sobre algunas consideraciones conceptuales que se deben considerar para la medición de la inteligencia y la conciencia humana. Siendo un especialista en la parte más técnica de la inteligencia artificial, el camino que sugiere el profesor Chávez tiene suma relevancia para poder entender cuáles son las posibles consecuencias éticas en la investigación de la conciencia.

El profesor Mariano Rivera explora la posibilidad de que las computadoras y redes neuronales artificiales puedan adquirir algún tipo de conciencia. Comienza examinando cómo la inteligencia ha sido crucial para la supervivencia y el éxito evolutivo en los seres humanos, facilitando habilidades como la creación de herramientas, el desarrollo de estrategias de caza y la formación de sistemas sociales complejos. Luego introduce la inteligencia artificial y explica cómo los sistemas de IA actuales ya exhiben capacidades que podemos considerar inteligentes, como crear imágenes a partir de descripciones de texto, reconocer el habla, traducir entre idiomas y realizar diagnósticos médicos. A continuación, se explora la posibilidad de que las redes neuronales artificiales más avanzadas puedan desarrollar alguna forma rudimentaria de autoconciencia o conciencia primaria, donde puedan monitorear y ajustar sus propios procesos cognitivos. Argumenta que la conciencia ha surgido al menos dos veces en la evolución (humanos y neandertales) y especula que podría surgir por tercera vez –ahora en forma artificial– a medida que los sistemas de IA se vuelven más sofisticados.

Los investigadores Rubén Avilés, Edgar Arias y Javier Sánchez realizan una exploración sobre las implicaciones en el uso de herramientas tecnológicas y de la información tanto en la investigación como en la intervención en el campo de la neuropsicología. En la exploración, los autores realizan una revisión conceptual, así como una reconstrucción de sus antecedentes en la teleneuropsicología, abordando también su aplicación clínica. La

aportación de los profesores es muy relevante no solamente porque brinda elementos para reflexionar sobre la posibilidad de usar las TIC como un instrumento de investigación, sino porque ejemplifican la forma en la que han aplicado dichas herramientas en su propio trabajo.

Los coordinadores estamos seguros de que los lectores encontrarán en esta obra material que, aunque de carácter divulgativo, tendrá la profundidad suficiente para permitir vislumbrar las múltiples implicaciones sociales que traen consigo las modernas tecnologías de la inteligencia artificial. Como hemos señalado, en su conjunto, el contenido de los diversos capítulos va más allá de la dicotomía entre tecnofilia y tecnofobia que, aunque simplifica el discurso, lo vuelve simplista y binario, sin ofrecer los matices que se encuentran en el mundo y la vida real. La visión que aquí se ofrece es matizada también por su origen multidisciplinario.

En esta época donde la IA ha cobrado nuevos bríos, no son pocas las publicaciones donde tecnólogos optimistas tratan de justificar a ultranza los usos de la tecnología, ofreciendo una miope visión de los alcances sociales. Tampoco faltan las publicaciones de estudiosos de las Ciencias Sociales que plantean de forma adecuada las implicaciones sociales pero que tienen una óptica distorsionada sobre los alcances reales y el funcionamiento actual de las tecnologías en cuestión. El carácter multidisciplinario, contrastado, del material de este libro lo convierte en rara avis y lo distingue de otras muchas publicaciones que se podrán encontrar en torno a la IA. Invitamos entonces a los lectores a explorar este mundo de contrastes, de posibilidades, de fuentes de inspiración, de importantes advertencias y de potenciales futuros, que ofrece la inteligencia artificial en nuestras sociedades.

https://doi.org/10.61728/AE24001007

Parte I

Contextualizando la IA

Capítulo **1**

Apuntes para una breve prehistoria filosófica de la IA

Dr. Diego Alfredo Pérez Rivas
Dra. Virginia Guadalupe López Torres

Cuando un hombre razona, no hace otra cosa sino concebir una suma total, por adición de partes; o concebir un residuo, por sustracción de una suma respecto a otra: lo cual (cuando lo hace por medio de palabras) consiste en concebir a base de la conjunción de los nombres de todas las cosas, el nombre del conjunto: o de los nombres de conjunto, de una parte, el nombre de la otra parte…Del mismo modo que los aritméticos enseñan a sumar y a restar en números, los geómetras enseñan lo mismo con respecto a las líneas, figuras (sólidas y superficiales), ángulos, proporciones, tiempos, grado de celeridad, fuerza, poder, y otros términos semejantes; por su parte los lógicos enseñan lo mismo en cuanto a las consecuencias de las palabras.
–Thomas Hobbes, Leviatán, 1651.

https://doi.org/10.61728/AE24001014

Durante el verano de 1956 se llevó a cabo la famosa *Conferencia de Darmouth College* en Hanover, New Hamphsire. Nos referimos, por supuesto, al evento en el que se considera que surgió la concepción moderna de inteligencia artificial, con la participación de algunos personajes que, vistos a través del crisol del tiempo, pueden considerarse legítimamente como sus patriarcas: John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon. En sus albores, y tal como la definió Marvin Minsky en dicha conferencia, la IA podría considerarse como "la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si fueran hechas por hombres".

Desde ese momento, el concepto de IA estaría estrechamente vinculado a una concepción antrópica de la inteligencia. Pero, como nos lo recuerda el fragmento de Thomas Hobbes con el que inicia este capítulo, la idea de que la inteligencia humana estuviese de alguna forma relacionada con la capacidad de calcular, data de cuando menos tres siglos atrás. Esta simbiosis entre "razón y cálculo", que también encuentra cierto eco en la filosofía presocrática, sentaría las bases de una de las ideas más felices de la historia del pensamiento humano. La IA puede considerarse tanto un sueño prematuro como una idea que nació de la mente de filósofos naturales y que, posteriormente, encontraría su propio lugar en el maridaje entre informática, ingeniería y programación.

Este 2024, la IA moderna cumple 68 años como un campo de estudio propio con plena legitimidad. Hace casi siete décadas que dicha área de investigación y desarrollo comenzó a construir su propio lenguaje, creando delimitaciones conceptuales muy claras y avanzando en el diseño de herramientas específicas que han permitido acercarse paulatinamente al objetivo original de sus creadores. Estas tareas han permitido la creación de tecnologías de vanguardia, así como la discusión de aspectos críticos muy puntuales sobre la penetración sociocultural de dichas tecnologías, lo que ha dado pie a una evolución progresiva y sostenida a lo largo del tiempo de los sistemas inteligentes, a pesar del bien conocido "invierno de la IA". Podemos decir que, observada a macro escala, la historia de la inteligencia artificial ha sido progresiva con pautas locales en las que dicho progreso ha retrocedido unos pasos o se ha detenido.

Actualmente, la emergencia y liberación de la IA generativa ha tenido como resultado el surgimiento de numerosos debates sobre las con-

secuencias éticas y legales de dichas tecnologías en las sociedades democráticas. Se trata de la emergencia de un nuevo tipo de IA que es capaz de usar el lenguaje natural como ninguna otra herramienta construida en el pasado. Por esa razón, algunos especialistas que apuntan a la prevalencia del "principio de precaución", sugieren la necesidad de establecer una moratoria en el desarrollo y uso de dichas tecnologías, al menos hasta que seamos capaces de meditar y lograr consensos robustos sobre algunas de las posibles consecuencias de sus aplicaciones. En caso de que algo así sea posible, la aplicación de una moratoria de esta naturaleza implicaría poner una pausa momentánea a la investigación y la aplicación de la IA, creando un posible nuevo invierno de la misma. Se trata de un asunto realmente delicado, pues la IA pone bajo amenaza nuestra concepción actual de la condición y la naturaleza humanas, así como nuestras ideas tradicionales de derechos humanos, subjetivos y fundamentales.

En el contexto de la tercera revolución industrial (la revolución de la información) y montada en los avances de la segunda revolución industrial (de la electricidad), la IA ha surgido como un campo de estudio que recupera elementos de las ciencias cognitivas, de las ciencias de la información y de la física aplicada con el reto de intentar emular facultades consideradas como específicamente antrópicas mediante sistemas informáticos. A partir de su surgimiento y hasta ahora, el desarrollo de la IA ha estado plagado de enormes retos y éxitos, pero también de dificultades y consecuencias que todavía no alcanzamos a visualizar completamente. Sin embargo, la historia que antecede a ese verano de 1956 es la prehistoria que trataremos de recuperar muy parcial y superficialmente en el presente capítulo, con la finalidad de no perder de vista que el sueño de una IA capaz de realizar acciones inteligentes, desde una visión antrópica, ha formado parte de la sociedad occidental desde hace mucho tiempo. Para lograr ese objetivo recuperaremos algunos episodios puntuales de esa prehistoria: i) la calculadora celeste de Arquímedes, ii) la pascalina o la calculadora de impuestos, y iii) la calculadora digital de Leibniz.

La calculadora celeste de Arquímedes

Para ser libre (en Grecia), el hombre debía ser liberado o liberarse él mismo y este estar libre de las obligaciones necesarias para vivir era el sentido propio del griego schole o del romano otium, el ocio, como decimos hoy. Esta liberación, a diferencia de la libertad, era un fin que podía y debía conseguirse a través de determinados medios… los antiguos explotaban a los esclavos para liberar completamente a los señores de la labor (arbeit), de manera que éstos pudieran entregarse a la libertad de lo político.
–Hannah Arendt, ¿Qué es la política?

Que la IA pueda considerarse como un sueño surgido de la mente inquieta de los filósofos antiguos puede parecer una completa disparatada para el lector contemporáneo. Lo es al menos en el sentido de que con la tecnología disponible en su tiempo, pensadores considerados clásicos por la tradición filosófica (como Sócrates, Platón o Aristóteles) hubieran sido incapaces, ya no digamos de crear, sino siquiera de poder imaginar la construcción de algún dispositivo capaz de emular acciones consideradas como inteligentes o antrópicas. Hablar de máquinas inteligentes en lenguaje platónico o aristotélico sería en realidad una completa disparatada en este sentido. Pero hay otro obstáculo que es digno de mencionar.

Lo que sucede es que para la tradición filosófica que surge de la matriz *platónico-aristotélica* existe una cierta concepción degradante respecto a los conocimientos o saberes que se podían adquirir en disciplinas u oficios considerados como *mecánicos e innobles*. Desde el nacimiento de la filosofía clásica occidental, ha estado presente alguna variante de la dicotomía entre *actividades mecánicas* y *actividades liberales* como una forma disfrazada de presentar la división social del trabajo predominante en Grecia y Roma, a la luz de una justificación espuria. Los oficios que requerían habilidades en mecánica fueron considerados antiguamente como ocupaciones con poca dignidad, disponibles para los no ciudadanos (esclavos o extranjeros), mientras que las actividades liberales eran consideradas como empleos propios para ciudadanos libres. Como lo sugiere Hannah Arendt en la frase que inicia este capítulo: la *arbeit* (la labor) era considerado un tipo de trabajo enfocado en producir lo necesario para la subsistencia, esto es, una serie de tareas que muchas veces requieren la repetición mecánica, a

diferencia de las actividades liberales como la política, el arte o la filosofía, los cuales eran consideradas actividades creativas e innovadoras.

Para nuestra fortuna, existieron también otro tipo de filósofos de la naturaleza, cuya mentalidad era completamente diferente a la de los pensadores que han permeado con sus ideas gran parte del pensamiento occidental. Los filósofos naturales que estaban un poco más cerca de abordar problemas del mundo real para brindar soluciones pragmáticas, en contraposición a aquellos que permanecían en el mundo de las "nubes", con el que Aristófanes se burlaba de Sócrates, fueron capaces de poner en su verdadero lugar a la mecánica. Esto es, como una disciplina heredera de la filosofía natural en la que se podía empeñar el ingenio de los más perspicaces, con la finalidad de que la técnica se convirtiera en la herramienta prometeica por antonomasia: la herramienta de herramientas, el esfuerzo de esfuerzos, capaz de liberar de posteriores esfuerzos a la humanidad.

Arquímedes fue un intelectual muy peculiar, cuya gran parte de su obra ha sido perdida por los avatares del tiempo. Se le conoce principalmente por sus destacadas habilidades matemáticas, pero también por la capacidad que su trabajo tuvo de influir en asuntos políticos y militares de su tiempo, al serle adjudicada la invención de algunas de las máquinas más ingeniosas y mortíferas. Con el diseño de este tipo de herramientas para la guerra, el trabajo de Arquímedes sería precursor de la idea de que la mecánica y la técnica podrían ser la base para la autonomía política y hasta para propósitos de dominio y conquista. Gracias a Arquímedes las historias de David y Goliat adquirían sentido en un contexto histórico, al demostrar cómo el ingenio del débil podía vencer a la fuerza bruta. De ahí precisamente que el filósofo francés Voltaire llegase a afirmar que "Había mucha más imaginación en la cabeza de Arquímedes que en la cabeza de Homero", devolviendo quizá un poco de dignidad a los oficios mecánicos (la ingeniería), respecto a los oficios liberales (la poesía).

La actividad intelectual de Arquímedes tuvo la gran virtud de convertir a las matemáticas en el lenguaje fundamental de su quehacer, erigiendo a la mecánica como el instrumento idóneo para traducir ese lenguaje a un plano más real, de la mano de la ingeniería. En su exploración de esas dos tareas (las matemáticas y la mecánica) descubrió la íntima relación entre ellas, escribiendo una obra magna *Sobre el método de los teoremas mecánicos*.

En la construcción de modelos de solución de problemas matemáticos, Arquímedes se centraba tanto en un modelo mecánico como en un modelo geométrico exhaustivo. Ese descubrimiento le permitió además ser un pensador innovador que pudo anticipar algunos de los principios fundamentales del cálculo integral que sería redescubierto en pleno siglo XVII de la mano de Newton y Leibniz, si bien con nuevos métodos y una nueva notación, además de una muy célebre disputa por los derechos de autor. Aparte de lo anterior, también se puede presumir que la primera prueba de hipótesis fuerte sobre el modelo heliocéntrico se debe a Arquímedes, tal como lo señaló Bertrand Russell en su *Historia de la Filosofía Occidental* (2016, p. 220). Nada de esta labor colosal parece que hubiera podido ser posible sin el acercamiento que tuvo Arquímedes a la mecánica.

La materia en la que Arquímedes muy probablemente contribuyó a encontrar los mejores caminos que unieran lo celeste con lo humano, es en la ingeniería de mecanismos para explorar el cielo. Como bien lo señala Cesar Guevara:

A pesar de que los trabajos de Arquímedes tuvieron que enfrentar el prejuicio platónico, en ellos no se descartó ningún procedimiento técnico emanado de la mecánica o de la geometría del mundo real. Libre de compromisos filosóficos aprovechó lo que otros despreciaron: lo infinitesimal, lo mecánico, lo instrumental, así como todo lo que le presentó su entorno real y que lo consideraba como un potencial elemento para su investigación; en esta dirección se tiene que aprovechó los casos de lo irregular, lo tangible y todo lo que pudiera ser construible. (2005, p. 13)

En esta búsqueda de mecanismos inteligentes (o previsores) podemos, muy probablemente, descubrir los primeros cimientos de esa búsqueda de dispositivos con algún tipo de inteligencia de naturaleza antrópica, pero también, por qué no, de un dispositivo con una cierta capacidad prometeica. Esto es, un instrumento con la capacidad de prever lo que sucede en los cielos y que parecía ser, en ese momento histórico, como algo que pertenecía al fuero especial de los dioses. Si bien es cierto, según explica Carl Boyer en su historia de las matemáticas, que la mayor parte de las fuentes históricas parecen apuntar al hecho de que Arquímedes atribuía mayor valor a sus actividades intelectuales para demostrar teoremas que a sus artilugios mecánicos (2016, p. 143), no por eso se debe restar valor

al hecho de que parece existir un ejercicio de retroalimentación continúa entre sus trabajos más teóricos y sus invenciones prácticas, ya que en cierto modo hasta sus métodos de demostración de teoremas contienen una cierta naturaleza mecánica. Asimismo, la calculadora celeste que diseñaría Arquímedes podría ser considerada quizá como la primera computadora mecánica de la historia, una computadora dotada con la inteligencia de escrutar los misterios del cielo. Una facultad que podía ser considerada digna de la inteligencia humana más exquisita, sino que es algo exclusivo de la inteligencia divina.

El filósofo y senador romano Marco Tulio Cicerón, en su obra *Sobre la República*, nos cuenta sobre este famoso mecanismo de ingeniería (una inteligencia artificial primitiva), cuya autoría se atribuía al genio Arquímedes, al haber sido recuperado en su tierra natal, Siracusa. Dicho mecanismo era capaz de emular el movimiento de los astros celestes visibles, los que eran conocidos en ese momento histórico con la tecnología disponible, con una precisión inaudita. Su descripción del artilugio y del contexto en el que fue encontrado es tan refinada, está plagada de detalles técnicos importantes y es tan exquisita en cuanto a la descripción de las funciones del artilugio, que vale la pena leerla completa. En sus propias palabras, nos cuenta la siguiente anécdota:

Recuerdo que Cayo Sulpicio Galo, hombre doctísimo, según bien sabéis, al decir alguien que había visto eso, como se hallase él usualmente en casa de Marco Marcelo, que había sido colega suyo en el consulado, mandó sacar un planetario, que el abuelo de Marco Marcelo, al caer Siracusa, se había llevado de aquella riquísima y bellísima ciudad, como único botín que se trajo a casa, a pesar de la importancia del saqueo. De este planetario había oído yo hablar muchas veces, a causa de la fama de Arquímedes, pero no quedé demasiado admirado de su forma, pues hay otro planetario más hermoso y popular, hecho por el mismo Arquímedes, que el citado Marcelo había puesto dentro del templo de la diosa Fortaleza. En verdad, después de que Galo empezó a explicar científicamente este aparato, pensé que aquel siciliano tuvo más inteligencia que la que puede alcanzar la naturaleza humana. Porque decía Galo que era muy antigua la invención de aquella otra esfera sólida y entera que había torneado Tales de Mileto primeramente, y que luego Eudoxo de Cnido, discípulo, según él decía,

de Platón, le había puesto astros y esferas fijos en la órbita celestial, cuyo aderezo de figuras hecho por Eudoxo, muchos años después tomó Arato para celebrar con unos versos, no tanto por tener ciencia astronómica cuanto por cierta vena poética. Pero este otro tipo de planetario, en el que se movían el sol y la luna, y también aquellas cinco estrellas que se llaman errantes y en cierto modo vagabundas, no podían acomodarse a aquel tipo de esfera sólida, y por eso es admirable el invento de Arquímedes, pues se las ingenió para reducir a una sola rotación varios cursos que eran desiguales por sus propias trayectorias. Al mover Galo este planetario, se veía como la luna giraba en aquella esfera metálica tantas veces alrededor del sol cuántos eran los días naturales, de modo que la esfera proyectaba el eclipse de luz solar y la luna entraba en la zona tapada por la tierra, por el sol, desde la región, de la luna quedaba interceptado por la tierra. (14, 21)

De la descripción del senador y filósofo romano es necesario destacar varias cosas. En primer lugar, la comparación que realiza entre el mecanismo de Arquímedes y el mecanismo de Tales, el cual había sido perfeccionado posteriormente por filósofos platónicos. Es digno de llamar la atención de qué manera Cicerón elogia la simplicidad y la elegancia del artilugio de Arquímedes, señalando aspectos importantes del mecanismo, los cuales ningún filósofo de la ciencia de nuestro tiempo dejaría pasar por alto. Al parecer, el mecanismo del matemático de Siracusa poseía la capacidad de emular los movimientos de los astros celestes (el sol y la luna), de la misma manera que de los astros errantes (los planetas) con una precisión asombrosa, pues se podía predecir con ellos los eclipses. Podríamos decir que el modelo de Tales y el modelo de Arquímedes fueron filosofía, construidos para dar respuestas a las mismas preguntas, pero en su arquitectura y en su capacidad de predicción, la computadora de Arquímedes era infinitamente superior. Esta diferencia hace recordar la distancia que había entre la teoría de la gravedad de Newton y la teoría de la gravedad de Einstein, pues la primera era incapaz de explicar o predecir el perihelio de mercurio. Al igual que entre estas teorías, entre los artilugios antiguos, los planetarios, había una enorme distancia, en el cual el mecanismo de Arquímedes resultaba vencedor de las batallas más importantes.

Dicho mecanismo, conocido modernamente como la máquina de Anticitera (debido a que fue encontrado por un equipo de exploradores en

el siglo XX en una isla griega homónima), es un artilugio de una complejidad matemática y física impresionante, al contar con engranes que tenían una relación increíblemente exacta para representar, por ejemplo, los movimientos de la luna, así como para calcular el aumento de un día cada cuatro años. Ante el descubrimiento de dicho mecanismo y dada su complejidad, algunos especialistas han apuntado que puede considerarse con plena legitimidad como la primera computadora mecánica de la historia. Dicha computadora está muy lejos de tratarse de una máquina universal, pero sí puede ser considerada como una máquina programada para escudriñar los secretos celestes, lo que requiere un grado de complejidad que, si no se quiere denominar como divino, al menos podría considerarse sin exageración como sublimemente humano. Podríamos decir, siguiendo la definición que usamos en la presentación de este capítulo, que este mecanismo realmente realizaba tareas que de ser hechas por un ser humano serían denominadas como inteligentes.

Es importante ubicar el mecanismo de Anticitera en su debido contexto. Se trata de una invención muy sofisticada que ha sido datada aproximadamente del siglo I antes de nuestra era. Puede considerarse con plena legitimidad como el primer computador diseñado para calcular posiciones astronómicas y con la capacidad de predecir eventos celestes como las fases de la luna o los eclipses. Su ingeniería incluía un mecanismo diferencial que tomaba en cuenta los cambios en las velocidades orbitales de los cuerpos celestes. Por esa misma razón, con el descubrimiento de este mecanismo nos queda claro que la idea de generar mecanismos tan inteligentes como los humanos es tan antigua, al menos, como dicho objeto.

La Pascalina: una máquina para calcular impuestos

La prehistoria de la inteligencia artificial está plagada de episodios curiosos en los que algunos filósofos soñaban con la posibilidad de poder construir con la tecnología de su tiempo, entidades o máquinas capaces de realizar acciones antrópicas como podría ser resolver ecuaciones mediante las reglas de la aritmética o el álgebra. Considerada como una actividad específicamente humana, la inteligencia estuvo asociada durante mucho tiempo con la capacidad de resolver problemas formales o lógico-matemáticos. A

diferencia de los animales, se pensaba que los seres humanos tienen una inteligencia especialmente dotada para dar respuesta a este tipo de problemas abstractos. Es digno de tomar en consideración que la resolución de problemas de esta naturaleza se asocia, naturalmente, con otras facultades cognitivas como son la atención, la capacidad de representación y abstracción, la memoria, etc. El producto del ejercicio de esas actividades intelectuales sería el descubrimiento de lo que podemos denominar como verdades platónicas, en el sentido de que se consideran como conocimientos que parecen tener la cualidad de ser eternos y desprovistos de materia. Por la misma razón, la inteligencia artificial maquinada por Blaise Pascal y la del propio filósofo alemán Gottfried Wilhelm Leibniz se pueden considerar como instrumentos altamente sofisticados en su época para emular la capacidad divina de los seres humanos, el cual consiste precisamente en poder calcular. Analicemos algunos aspectos importantes del artilugio de Pascal.

La calculadora de Blaise Pascal es un dispositivo mecánico de gran complejidad técnica para su tiempo, el cual fue diseñado en 1642 por el famoso filósofo, teólogo y matemático. Se puede considerar como la primera computadora programada para realizar cálculos numéricos que llevasen a la resolución de problemas aritméticos específicos. Aunque se trata de un mecanismo de menor complejidad que el mecanismo de Anticitera, lo cierto es que la Pascalina es el producto de un ingenioso diseño de engranajes en el que se asigna a cada diente el valor de un dígito y en el que se respeta la naturaleza posicional de cada dígito. Por la misma razón, dicha calculadora es el primer ordenador de cálculo en sistema decimal que además de todo representa en un sentido mecánico nuestra notación posicional. Desde el planetario de Arquímedes no existe evidencia que se haya diseñado alguna máquina con la finalidad de poder realizar actividades inteligentes o que requerirían inteligencia en caso de ser hechas por hombres, por lo que se puede considerar que con la pascalina se inaugura un resurgimiento o renacimiento de la prehistoria filosófica de la IA.

La Pascalina nació originalmente como una computadora mecánica capaz de ayudar al padre de Blaise a realizar cálculos para la recaudación de impuestos. No obstante, la imaginación del filósofo concebía a dicha máquina como la primera parada respecto a las actividades antrópicas que

podrían ser automatizadas. Como vicepresidente de recaudación tributaria de su provincia, el padre de Pascal podría servirse de dicha invención para facilitar sus afanosas y a veces mecánicas actividades. Pero, es importante considerar que algunas de las posibles aplicaciones que el filósofo visualizó en las máquinas de cálculo es el de la posibilidad de traducir lenguas muertas, tal como algunos sistemas de IA están intentando hacerlo actualmente (Luo, 2019). En sus Pensamientos, Pascal explica lo siguiente: "Las lenguas son cifras en que las letras no se cambian por letras, sino las palabras en palabras, de suerte que una lengua desconocida es descifrable" (Pascal, 1992, p. 20). Se trata, en toda forma, de la primera vez que alguien que hubiera creado una máquina de cálculo, pensase que se podía aplicar una máquina similar para usos considerados exclusivamente antrópicos, como lo es el desciframiento de lenguas pérdidas. Gracias a la simbiosis entre "cálculo y razón", la consecuencia era que todo lo que era susceptible de ser traducido a cifras o representaciones numéricas podía ser susceptible de ser usado para calcular situaciones más complejas. Siguiendo esa idea Spinoza Baruch llegaría a soñar con una ética diseñada para ser demostrada mediante el modelo geométrico. Se podría llegar a pensar en la actualidad que hasta se podría diseñar una IA spinoziana capaz de realizar cálculos morales.

Pensada originalmente como un dispositivo capaz de realizar cálculos con fines de recaudación de impuestos, Pascal fue capaz de darse cuenta que a través de la ingeniería y la técnica adecuadas, un inventor podía ser capaz de dotar a una máquina de facultades consideradas como inteligentes. Aunque su dispositivo realizaba únicamente operaciones simples como sumas y restas, lo cierto es que dotar a la máquina con la capacidad de realizar operaciones de grandes números mediante un sistema de acarreo fue algo muy ingenioso para su tiempo. Al referirse a su máquina de aritmética, Blaise Pascal lo hacía de la siguiente manera: "La máquina de aritmética produce efectos más próximos al pensamiento que todo lo que hacen los animales; pero no hace nada que pueda hacer decir que tiene voluntad como los animales" (1992, p. 338).

Resulta muy importante considerar que la calculadora de Pascal marcó un parteaguas en el diseño de mecanismos inteligentes. En primer lugar, porque después de siglos de olvido técnico y de lotofagia, un inventor con

inspiración filosófica y matemática hizo renacer la idea de que era posible crear y programar máquinas con la finalidad de emular actividades consideradas como estrictamente antrópicas. Aunque actualmente podemos considerar dicha máquina como un artilugio elemental, lo cierto es que en su tiempo histórico fue revolucionario, sentando las bases de la computación moderna.

La calculadora de Leibniz: la primera computadora digital comercial

Afirmar que la idea de las cosas esta? en nosotros no es más que sostener que Dios, autor a la vez de las cosas y de la mente, ha impreso en ella aquella facultad de pensar de tal modo que puede obtener mediante sus operaciones todo lo que se corresponde perfectamente con lo que surge de las cosas mismas. Y así?, aunque la idea de círculo no sea igual al círculo, de ella, empero, pueden obtenerse verdades que la experiencia confirmara?, sin ninguna duda, en el verdadero círculo.
–Gotfried Leibniz, ¿Que? es Idea?

Por su parte, el filósofo prusiano Gotfried Leibniz diseñó treinta años después que Pascal, en 1672, una serie de máquinas capaces de calcular. Una de esas versiones fue además comercializada, sentando las bases de un nuevo mercado para las máquinas inteligentes. Usando las ideas de su antecesor más inmediato, Leibniz imaginó una máquina computacional programada para resolver todo tipo de operaciones aritméticas: sumas, restas, multiplicaciones y divisiones. La gran virtud de esta computadora programada para realizar las operaciones básicas fue su arquitectura basada en cilindros dentados que estaban posicionados en barras transversales y en espiral, posibilitando el cálculo a través de un sistema binario, lo que convertía a la máquina en un artilugio potente en lo que a la eficiencia se refiere.

Tanto la máquina de Pascal como la máquina de Leibniz sugieren la persistencia de un zeitgeist (*espíritu del tiempo*) caracterizado por la búsqueda creciente de automatizar las tareas, mediante la creación de aparatos ingeniosos capaces de realizar actividades consideradas hasta entonces como específicamente inteligentes y antrópicas. Se trata en toda regla de

un sueño de filósofos que estaban visualizando la posibilidad de cambiar el curso de la historia, abriendo la caja de pandora de la inteligencia artificial. La calculadora digital de Leibniz, de hecho, iba un paso más allá de la Pascalina, pues al igual que con el cálculo moderno (cuya notación debemos al filósofo), su artilugio era la cúspide de un sistema racional universal, que incluía la propuesta de crear un lenguaje simbólico universal que fuese capaz de representar todo tipo de conceptos y razonamientos, convirtiéndose en cierto modo en un filtro depurador de nuestros lenguajes coloquiales. Leibniz creía en cierto modo que llegar a ese lenguaje universal podía ser la clave para traducir la obra del creador a un lenguaje puro de cuyas combinaciones o cálculos se podrían obtener todas las verdades existentes, como si se tratara de un acto banal.

Actualmente podemos denominar a las máquinas calculadoras de estos dos filósofos, en estricto sentido, como las antecesoras modernas más longevas de las computadoras o máquinas universales de Turing que llevamos a todas partes con nosotros. Pascal fue la primera persona moderna en revivir un anhelo profundo de ingenieros destacados como Arquímedes de construir máquinas inteligentes. Leibniz introdujo en este camino la noción de la representación binaria, lo que lo convierte en un hito en toda forma. En pleno siglo XVII se estaban sentando las bases de los sistemas digitales que definen la naturaleza de la informática moderna.

En el cuadro que presento a continuación, propongo algunas características que considero que se deben tener en cuenta para una reconstrucción de lo que podemos denominar como prehistoria de la IA, en un ejercicio que realice instruyendo a Chatgpt 3.5 a que realizará una comparación entre dispositivos:

Rubro	Pascalina	Calculadora de Leibniz
Creación	1642	1672
Arquitectura mecánica	Ruedas dentadas y engranajes	Cilindros con dientes en espiral y barras
Material de construcción	Principalmente metal	Principalmente bronce y latón
Tipo de operaciones	Suma y resta de números decimales	Multiplicación y división, además de suma y resta
Sistema numérico	Decimal	Binario
Complejidad matemática	Cálculos básicos, sin multiplicación o división	Realización de las cuatro operaciones básicas
Sistema de acarreo	Sí	No
Número de dígitos manipulados	Hasta 8 dígitos	Hasta 16 dígitos
Fortalezas	Simplicidad y eficacia para cálculos simples	Mayor versatilidad para operaciones complejas
Debilidades	Limitado a sumar y restar, sin multiplicación	Mecanismo más complejo, posiblemente menos fiable

Como podemos darnos cuenta en el cuadro anterior, la Pascalina y la calculadora de Leibniz son prácticamente contemporáneas. Ambas con sus fortalezas y sus debilidades, podemos presumir que significaron un salto tecnológico considerable en su momento histórico.

La visión de Leibniz, sin embargo, puede considerarse como más moderna porque su búsqueda de un lenguaje simbólico universal le permitió sentar las bases del uso del sistema binario en las máquinas programadas. Algo que, sin embargo, se puede presumir de ambos filósofos, matemáticos e inventores fue su capacidad para visualizar la íntima conexión entre las matemáticas, la tecnología y la propia filosofía. En este sentido, cabría traer a colación las ideas de Immanuel Kant para recordar que: "inventar algo es muy otra cosa que descubrir algo. Pues la cosa que se descubre se supone existir ya antes, solo que todavía no era conocida, por ejemplo, América antes de Colón (para el mundo europeo); pero lo que se inventa, por ejemplo, la pólvora, no era conocido antes del artista que lo hizo" (1935, p. 117)

A manera de conclusión

A lo largo del presente capítulo hemos propuesto un viaje a través de algunos eventos representativos en lo que puede denominarse lúdicamente como la prehistoria de la IA. Hemos revisado muy superficialmente tres períodos históricos en los que el diseño de máquinas consideradas como inteligentes fue algo más que un sueño de filósofos trasnochados. Con dicha revisión pudimos percatarnos del grado de complejidad de dichos mecanismos, así como revisar algunas de las ideas y motivaciones que estuvieron presentes en su creación. Es digno de tomar en consideración que todos esos mecanismos pueden considerarse en estricto sentido como "utensilios tecnológicos" dispuestos a servir a los fines de los seres humanos, siendo usados por personas que naturalmente debían tener conocimientos previos sobre las distintas materias (astronomía o aritmética). Pueden considerarse inteligentes en el sentido más antrópico y reducido de una inteligencia análoga al cálculo, pero incapaces de hacer algo distinto para lo que fueron programados.

Realizando una recapitulación histórica de la IA, Pippo Battaglia nos advierte que "la capacidad de elegir es una prerrogativa de la mente y se ha intentado convertirla en una peculiaridad de cerebros artificiales dotados de inteligencia" (2006). A lo largo de la breve prehistoria de la IA que hemos recopilado en este capítulo, podemos darnos cuenta que el sueño de diseñar artilugios o mecanismos dotados con la capacidad de realizar acciones inteligentes es al menos tan antiguo como Arquímedes. La concepción de una inteligencia de naturaleza antrópica parece formar parte de las entrañas mismas de la IA. La "capacidad de elegir" puede, de hecho, considerarse también como una noción antrópica que se ha considerado como infranqueable. Al menos en el sentido humano de tomar decisiones con base en información multifactorial y mediante mecanismos complejos. Lo cierto es que puede ser que al igual que sucedió con el tabú del "cálculo", la capacidad de elección o la simulación de voluntad requiera la ruptura de un punto crítico, del cual pueda emerger algo parecido a lo que denominamos con ese nombre de parte de los sistemas de IA.

Actualmente nos encontramos ante la emergencia de modelos de IAG, los cuales trabajan entre otras cosas, con modelos de distribución estadís-

tica consiguiendo ser eficientes en el aprendizaje automático. Los modelos de IA actuales son entrenados con enormes cantidades de información y en caso de no ser supervisados han demostrado replicar los prejuicios y las nociones de las personas que crearon las fuentes de información de las que se alimentan. A veces nos gusta pensar que esos prejuicios no forman parte de lo que llamamos estrictamente humano, pero la historia nos ha enseñado que no necesariamente es así. Parece indispensable seguir repensando el concepto de IA, pero con nuevos bríos, al menos para decidir si pretendemos seguir usando un paradigma referido a la especie humana de forma general y abstracta.

Referencias

Arquímedes. (1986). *El método.* Alianza Editorial. Madrid.

Battaglia, P. (2006). *L'Intelligenza Artificiale*. Dagli automi ai robot Intelligenti. UTET. Torino.

Boyer, C. (2016). *Storia della ilosofía*. Oscar Mondadori.

Cicerón. (2002). *Sobre la República.* Gredos. Madrid.

Guevara, C. (2005). Historia de los instrumentos matemáticos. Arte, astronomía y geometría. En Falconi M. y Hoyos V. *Instrumentos y matemáticas. Historia, fundamentos y perspectivas históricas.* UNAM.

Hobbes, T. (1984). *Leviatán*. Sarpe. Madrid.

Kant, I. (1935). Antropología en sentido pragmático. *Revista de Occidente*. Madrid.

Kohari, A. (2022). *An Ancient Language Has Defied Translation for 100 Year*. CAN AI Crack the Code? Rest of World https://restofworld.org/2022/indus-translation-ai-code-script/

Leibniz, G. (1992). *Disertación sobre el Arte Combinatorio*. Ediciones Universidad Católica de Chile.

Luo, J. et al. (2019). *Neural Decipherment via Minimum-Cost Flow: from Ugaritic to Linear*, B. En P. Nakov y A. Palmer (eds.), Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, pp. (3146-3155). Florencia: Association for Computational Linguistics.

Netz, R. y Noel, W. (2007). *The Archimedes Codex*. Weinfeld & Nicholson. London.

Rossi, P. (1989). *Clavis Universalis. El Arte de la Memoria y la Lógica Combinatoria de Lulio a Leibniz.* Fondo de Cultura Económica. México.

Russell, B. (1949). *A Critical Exposition of the Philosophy of Leibniz*. G. Allen & Unwin. London.

Russell, B. (2016). *Storia della ilosofía occidentale*. TEA. Veneto.

Capítulo **2**

Inteligencia Artificial: ni muy automatizada, ni muy ecológica

Antonio de Jesús García Chávez
José Antonio García Macías

Todos ven lo que tú aparentas; pocos advierten lo que eres.
–Nicolás Maquiavelo

https://doi.org/10.61728/AE24001021

Introducción

Tras un largo periodo con relativamente poca actividad y resultados, en años recientes se ha generado un incremento en la cantidad de publicaciones científicas relacionadas con la inteligencia artificial (IA), con un crecimiento del 100 % desde 2010 (Stanford University, 2023). Hoy en día nos encontramos en una revolución en relación con la IA, y sus efectos repercuten mucho más allá de los confines de los laboratorios de investigación y las empresas tecnológicas, abarcando también los sectores industriales y de gobierno. Esto se vuelve claro al analizar el aumento en el número de incidentes y controversias relacionados al uso indebido de la IA; los últimos reportes indican que los casos de mal uso ético de la IA se han multiplicado por 26 desde el año 2012 (Stanford University, 2023).

Si bien es cierto que la investigación en IA ha aumentado en todas sus áreas, hay que denotar que desde el 2017 particularmente las áreas de reconocimiento de patrones, aprendizaje automático y visión por computadora son aquellas con la mayor cantidad de publicaciones científicas (Stanford University, 2023). Sin embargo, en los últimos cinco años la IA se ha colocado no solo como un tema de investigación y desarrollo comercial; ya que en gran medida aplicaciones y sistemas con base en IA, como lo son filtros para fotografías y video, así como chatbots, se han encargado de embarcar a más de uno en un viaje por el reino de las IA. Lo anterior ha sido posible en gran medida gracias al encanto de la IA generativa y, en particular, de los enigmáticos Grandes Modelos de Lenguaje (LLM o Large Language Models), los cuales han captado la imaginación y expectativas del mundo de manera inusitada. Estos LLM, tales como GPT o PaLM, representan un cambio de paradigma que trasciende la mera computación y se adentra en las intrincadas facetas de la cognición, la creatividad y la comunicación humanas. Por lo que no ha de parecernos extraño que, en un futuro no muy lejano, el área del Procesamiento del Lenguaje Natural abandone su posición como sexta área de investigación con más publicaciones científicas y se posicione dentro de las tres áreas con mayor número

de publicaciones, siendo estas Reconocimiento de patrones, Aprendizaje de máquina y Visión por computadora.

El crecimiento y atención mediática de estos Modelos de Lenguaje, no solo ha sido un esfuerzo proveniente de la academia, lo cual sería fácil de suponer ya que hasta 2014 la mayoría de los modelos de aprendizaje automático más significativos procedían del mundo académico. Sin embargo, en los años siguientes, la industria ha asumido un papel dominante. En 2022 el sector industrial había producido treinta y dos modelos de aprendizaje automático dignos de mención, mientras que el mundo académico solo había contribuido con tres (Stanford University, 2023). El desarrollo de sistemas de IA, de vanguardia ahora, exige datos sustanciales, recursos computacionales e inversiones financieras, todos los cuales están más fácilmente disponibles para los actores de la industria que para las organizaciones sin fines de lucro y las instituciones académicas.

Inteligencia artificial generativa

Aunque la IA abarca una amplia gama de aplicaciones, un área en concreto que ha acaparado gran atención en los últimos años es la inteligencia artificial generativa (IAG), la cual en el último año ha dominado el panorama de la IA de forma espectacular, constituyendo un auténtico acontecimiento inesperado y transformador. Esto ha ocurrido ya que la IAG ha demostrado un amplio abanico de capacidades que superan con creces las de cualquier sistema de IA anterior, y que abarcan numerosas funciones con beneficios económicos y humanos tangibles (Eloundou, Manning, Mishkin y Rock, 2023).

Imaginemos por un momento un mundo futurista en el que los ordenadores no solo puedan entender nuestro lenguaje, sino también componer historias, crear obras de arte e incluso componer música que resuene con nuestras emociones. La IA generativa hace posible este mundo, que, hasta hace no mucho, resultaba como un sueño aparentemente lejano. Es interesante notar que algunas de las capacidades actuales de estos modelos de IA generativos solo se consideraban alcanzables en el marco de sistemas de IA que poseyeran una competencia general de nivel humano, lo cual ha desafiado las expectativas de muchos expertos en el área (Goertzel y Singularitynet, 2023).

En el panorama de la IAG, centraremos nuestra atención en el ámbito del lenguaje, particularmente en los grandes modelos del lenguaje o bien "Large Language Models (LLM) y más concretamente en los modelos conocidos como GPT (Generative Pre-trained Transformers). Pero ¿qué es exactamente lo que hace tan cautivadores a estos LLM? Estos modelos al ser entrenados con grandes cantidades de datos aprenden a imitar en gran medida la creatividad humana, obteniendo de esta forma, el potencial para escribir artículos convincentes, generar imágenes realistas e incluso entablar conversaciones naturales. Estos LLM actualmente se emplean en diversas aplicaciones industriales, desde la búsqueda en la Web y los chatbots conversacionales incluso hasta en el análisis de documentos médicos y financieros (Pandu Nayak, 2019; Joshua Broyde, 2021; Thewsey, 2021; Lee, 2023). Debido a esta amplia aplicabilidad, estos LLM se han denominado foundation models (que pueden traducirse como "modelos de cimentación") en el área del Procesamiento del Lenguaje Natural (Bommasani et al., 2021).

Hasta este momento hemos contextualizado lo impresionante y atractivos que resultan estos LLM, así como lo versátiles que son al impactar en distintas áreas. Por lo que a continuación focalizaremos nuestros esfuerzos en arrojar luz en algunos de los conceptos e ideas que rodean las distintas arquitecturas y procesos de entrenamiento que estos LLM necesitan para lograr estos comportamientos y capacidades tan intrigantes. Consideremos por cuestiones pedagógicas la siguiente analogía, imaginemos que los LLM son bibliotecas que contienen una vasta y amplia cantidad de libros, cada uno de estos libros, representará un aspecto diferente del lenguaje natural y la expresión humana, algunos se enfocan en esclarecer cuestiones sintácticas, semánticas, gramaticales y pragmáticas, por mencionar algunos. De manera similar en esta biblioteca, podrías encontrar libros donde la expresión escrita del lenguaje natural se emplea en debates, noticias, artículos científicos, libros de ficción, poemas y enciclopedias; consideremos entonces todos los formatos y usos del lenguaje en internet, páginas web, blogs, redes sociales, etc. Emplearemos entonces esta analogía para explicar a grandes rasgos los componentes principales de estos LLM.

Consideremos que los libros contienen en cierta medida, el conocimiento colectivo del mundo con respecto al uso del lenguaje natural. Los

LLM devoran este festín textual, con un objetivo similar al de un erudito, el cual es comprender los matices del lenguaje y convertirse en expertos en diversas materias. Hay que denotar que las capacidades de estos LLM se ven potenciadas con respecto al tamaño de su biblioteca, por ejemplo BERT (Devlin, Chang, Lee y Toutanova, 2019) uno de los primeros y más populares LLM fue entrenado con un tamaño de biblioteca de alrededor de 24 GB. Por otro lado, GPT 3 (Brown et al., 2020) fue entrenado con alrededor de 285 000 000 hojas de texto plano o bien aproximadamente 114 000 veces la obra completa de Shakespeare, lo que equivale a 570 GB de información; por lo regular la cantidad de datos que suele emplear estos LLM suele oscilar entre 15 GB y 5 TB de texto plano (Rae et al., 2021; Hoffmann et al., 2022).

¿Entonces qué papel cumple el bibliotecario en esta analogía?, el bibliotecario es aquel que ayuda en la comprensión de las conexiones y los patrones de información que contiene su biblioteca, de tal forma que cuando se le pregunta por algún título, información o sugerencia, este trabaja deprisa en recordar y reunir los libros adecuados, para de esta forma otorgar una respuesta pertinente y útil. Este bibliotecario en el área de la inteligencia artificial, se le ha conocido con distinto nombres, cibernética en los años 1940-1960, conexionismo en los años 1980-1990, o bien quizá lo reconozcamos con su nombre más popular redes neuronales (Goodfellow, Bengio y Courville, 2016). Por lo que, en esta analogía, el bibliotecario representa la arquitectura de red neuronal sobre la cual los LLM fueron construidos, esta arquitectura lleva por nombre "Transformer" [15], la cual se ha mantenido prácticamente sin cambios desde 2018. Estos bibliotecarios son los encargados de leer y analizar los libros, aunque debemos mencionar que su forma de dar lectura es diferente a la nuestra; estos son entrenados de una de las siguientes maneras, o bien se entrenan para predecir una palabra enmascarada en un enunciado (es decir, palabras que hay que rellenar), o bien para predecir la palabra siguiente en un enunciado (similar a como los teclados de nuestros teléfonos predicen la siguiente palabra que escribiremos a manera de recomendación). Por lo que el bibliotecario colabora con los parámetros del modelo para dar sentido a la información contenida en los textos, resaltando de esta manera los detalles importantes, clasifican la información y la hacen accesible para su consulta.

Si bien estos LLM constan de más que solo libros, un sistema de catalogación y un bibliotecario (datos de entrenamiento, parámetros, Transformers), podemos decir que al haber explorado estos fundamentales bloques de nuestra biblioteca, hemos sido capaces de apreciar en mayor medida las capacidades de estos LLM, y si de algo estamos seguros, es que veremos cada vez más la adopción de estas tecnologías en nuestras vidas.

Humanos como ayudantes de las máquinas

Suele ser común que las historias que perduran y vencen la prueba del tiempo son aquellas que cuentan los ganadores, o bien, aquellas narrativas que los involucrados buscan imponer y mantener para futuras generaciones; esto ocurre por igual en las historias y narrativas alrededor del desarrollo de la IA. Si bien es cierto que actualmente no es tan difícil encontrar opiniones e investigaciones que argumentan un caso negativo respecto de las IA (Chomsky, 2023; Jem Bartholomew, 2023). Estas suelen encontrarse abrumadoramente opacadas por el sensacionalismo, el cual suele emplear afirmaciones dramáticas sobre temas como las IA con conciencia (Roose, 2023) y el desplazamiento laboral (Gareth Corfield, 2022).

Por otro lado, medios de comunicación, así como futurólogos emplean afirmaciones y títulos exagerados, por ejemplo "¿Podría ChatGPT escribir mi libro y alimentar a mis hijos?" (Heritage, 2022). Por lo anterior no es extraño que una versión de la historia que suele olvidarse o bien dejar de lado, es la de los trabajadores fantasmas. Estas personas son los campeones ocultos de la IA, ya que proporcionan el trabajo tras bambalinas que sustenta gran parte de los avances logrados en las aplicaciones de aprendizaje automático en todo el mundo hoy en día. Cuando los algoritmos de Google funcionan como se supone que deben, mostrando la información que buscabas y no tonterías o porno raro o propaganda o consejos médicos peligrosos, es gracias a estos trabajadores fantasmas. Sin embargo, hay una cierta ironía cruel, ya que a medida que la tecnología con base en IA, por ejemplo, LLMs, realiza su aparición en la escena mediática y económica, los más adecuados para mantenerla funcionando son estos trabajadores fantasmas, que resultan ser también los más precarios en las empresas que los necesitan. Numerosos estudios han demostrado que los

trabajadores están mal pagados y reciben un trato injusto (Horton, 2011; International Labour Organization, 2016; Nast, 2017), un análisis reveló que el salario medio por hora de los trabajadores era de solo 2 USD (Hara et al., 2018).

Esta cifra es muy inferior al salario mínimo de EE. UU. (7.25 USD). Lastimosamente, historias como esta no resultan atípicas. Una investigación realizada por la revista Time reveló que a contratistas con sede en Kenia se les pagaba menos de 2 dólares la hora, por revisar contenidos para ChatGPT y muchos de estos contenidos eran tan tóxicos que les dejaban traumatizados (Perrigo, 2023). Aunque son varias las empresas que ofrecen este servicio del tipo crowdsourcing, siendo las principales Mechanical Turk y Appen, resulta una constante los problemas a los que estos trabajadores se enfrentan; por ejemplo, un rater (evaluador) en una entrevista para la revista Los Angeles Times comentó lo siguiente "suele ser común que de entre las tareas que evaluó pueda encontrar insultos racistas, intolerancia y violencia", sin embargo esto no resulta lo peor, "A veces he visto parte del contenido gráfico reproducido en mis sueños. Por eso nunca trabajó hasta tarde por la noche. En diez años he visto dos veces pornografía infantil, pero gracias a Dios es muy raro. Lo dejaría" (Merchant, 2023).

Sería sencillo e inocente argumentar que los evaluadores son aquellos con la última decisión, al ser estos los que seleccionan las tareas o proyectos en los cuales trabajar, sin embargo, la realidad es otra. La mayoría de las plataformas de crowdsourcing permiten a los solicitantes (requesters) crear de manera flexible sus tareas o proyectos, así como establecer los precios correspondientes. Además, los solicitantes tienen la capacidad de evaluar el desempeño de los evaluadores para garantizar la calidad de sus respuestas, verificando si cumplen con los requisitos mediante la revisión de cualificaciones (Hara et al., 2018), o bien rechazando las tareas enviadas que no cumplan sus criterios (Bederson y Quinn, 2011). Por otro lado, a los evaluadores se les suele proporcionar una cantidad limitada de información (Irani y Silberman, 2013), por ejemplo, suele ser común que las plataformas de crowdsourcing solo ofrezcan detalles esenciales como el precio de la tarea, el nombre del solicitante, un título y una breve descripción. Esta escasez de información supone un reto considerable para los raters que buscan tareas de buena calidad.

No obstante, no todo está perdido, hasta hace poco, las acciones colectivas eran poco frecuentes entre los trabajadores de alta tecnología porque estaban bien pagados y satisfechos con sus condiciones de trabajo (Nedzhvet-skaya, 2019). Sumado a esto, cuando una empresa tecnológica toma decisiones poco éticas en relación con el desarrollo y la implantación de la IA, sus empleados pueden encontrar una incoherencia entre su identidad moral individual y su percepción de la imagen de la organización (Alahmad y Robert, 2020).

Por lo que el activismo es una de las principales formas que tiene el público y trabajadores de presionar a una organización para que cambie sus políticas y prácticas (Kalodimos y Leavitt, 2020). No obstante, es crucial considerar que la capacidad de influencia de los activistas puede variar según los costos que conlleva para la organización el no abordar las inquietudes y solicitudes de estos (Belfield, 2020).

Si bien el escenario de los raters, como acabamos de observar, es bastante desolador, la llegada de estos LLM, ha generado una nueva y más especializada versión de estos raters, los cuales se concentran en tareas más sofisticadas, ajenas a tareas monótonas de etiquetado de datos como imágenes y video (Metz, 2023). Esta nueva clase de evaluadores nace en el momento que los investigadores notan que el tamaño de estos LLM (parámetros del modelo), no significa que sean mejores a la hora de seguir e interpretar la intención de los usuarios (Ouyang et al., 2022). Por lo que compañías como OpenAI emplearon técnicas de Aprendizaje por refuerzo a partir de retroalimentación humana (Stiennon et al., 2020), con el objetivo de subsanar este problema. Por lo que estos evaluadores actúan como tutores, proporcionando a estos LLM retroalimentación más profunda y específica, en un esfuerzo por mejorar sus respuestas. Esto resulta de vital importancia, ya que a medida que los sistemas de IA se vuelvan más potentes y se les asignan tareas cada vez más importantes, es probable que los errores que cometan se vuelvan más sutiles y críticos en relación con la seguridad (Stiennon et al., 2020).

Estos evaluadores suelen encontrarse en edades de entre los 19 y 62 años, y sus cualificaciones van desde títulos técnicos hasta doctorados. Por otro lado, estos evaluadores se diferencian de los raters clásicos, también en el salario, con un pago por hora de entre 15 y 30 dólares, esto tan solo

en Estados Unidos, ya que trabajadores en otros países ganan notoriamente menos (Metz, 2023). En gran medida la retroalimentación que realizan estos nuevos evaluadores es la que ha permitido que modelos como GPT, puedan aproximar sus Chatbots a una experiencia conversacional más fluida y por turnos, en lugar de limitarse a ofrecer una única respuesta. De manera similar, esta tutoría que realiza ayuda a empresas como OpenAI a reducir la desinformación, la parcialidad e información tóxica que pueden llegar a producir estos LLM.

Una industria hambrienta de recursos

"Todo necio confunde valor y precio" es una frase atribuida inicialmente a Francisco de Quevedo y después a Antonio Machado, misma que viene muy bien a colación con la IAG. A decir verdad, hay muchos sistemas online que utilizamos en la vida cotidiana y para los cuales no pagamos un costo monetario, sean estos servicios de correo electrónico, mensajería instantánea, canales de video como YouTube, redes sociales, almacenamiento de archivo, por mencionar algunos. A dichos sistemas le atribuimos un valor significativo pues en ocasiones nos facilitan tareas cotidianas y en otras simplemente nos brindan entretenimiento o evasión. A cambio, el precio que pagamos por la utilización de dichos servicios es el de nuestra privacidad (Veliz, 2021), pues nuestra información va a dar a los servidores de quienes proveen los servicios y de ahí, mediante autorización por contratos que nunca leemos o entendemos, utilizan la información para analizarla, procesarla, empaquetarla y venderla al mejor postor. También podrá tener efectos más nocivos como manipular comportamientos, alterar procesos democráticos e impulsar el consumismo con sus efectos medioambientales. Así lo detalla la Dra. Zuboff en su libro La era del capitalismo de vigilancia (Zuboff, 2020). Hasta ahora, ChatGPT ha seguido el conocido modelo freemium, donde los usuarios utilizan de forma gratuita una versión básica del sistema, pero deben pagar para poder utilizar funciones avanzadas. Algo similar ocurre con otros sistemas populares de IAG tales como Midjourney y DALL·E. No obstante, vale la pena detenerse a reflexionar sobre los costos ocultos que conllevan la utilización de tales servicios.

A primera vista se puede tener la percepción de que los usuarios de sistemas de IAG producen las imágenes, textos, o cualquier otro producto solamente con la capacidad que sus recursos caseros o de oficina le proveen. Sin embargo, este costo es despreciable si se compara con lo que está sucediendo en el lado de la infraestructura de nube de los proveedores de servicios y –sobre todo– de lo que tuvo que suceder antes de que se pudiera proveer ese servicio. En efecto, para que el servicio fuera puesto a disposición de los usuarios, los modelos de cimentación (foundation models) tuvieron que ser entrenados con cantidades de datos colosales; esto se realiza en grandes centros de datos, los cuales consumen también cantidades exorbitantes de energía eléctrica, poder de cómputo, almacenamiento de memoria y fluidos para enfriamiento.

Un estudio realizado para evaluar el ciclo de vida de entrenamiento de varios modelos de IA grandes y comunes (Strubell, 2019) determinó que el proceso puede emitir más de 626.000 libras de dióxido de carbono, lo cual sería equivalente a casi cinco veces las emisiones que tendría un automóvil estadounidense promedio durante su vida útil (incluyendo la fabricación del automóvil en sí). Por otra parte, Microsoft recibió fuertes críticas tras haberse dado a conocer que, justo en medio de una sequía de más de tres años, estaba utilizando enormes cantidades de agua para el sistema de refrigeración de uno de sus centros de datos utilizado mientras entrenaban el modelo GPT-4 de OpenAI.

Por supuesto, a la par de Microsoft están otros impulsores de la IAG tales como Google, consumiendo billones de litros de agua en sus centros de datos, con una tendencia al alza (Adarlo, 2023). En una época donde los sistemas hídricos se encuentran en constante estrés, entre otros efectos atribuidos del cambio climático, no es ocioso preguntarse por qué destinar dichos recursos escasos a actividades que no están relacionadas con la supervivencia humana.

El procesamiento de lenguaje natural es un área de estudio que, tras algunas décadas sin producir resultados notables, en tiempos recientes ha tenido avances fenomenales, sobre todo en lo correspondiente a los modelos de lenguaje o LLM. Estos avances recientes se deben en gran medida a la disponibilidad de grandes volúmenes de datos, aunados a la posibilidad de procesamiento adecuado para manipularlos. Sin embargo,

la tendencia para seguir con estos avances parece ser la producción de modelos cada vez más grandes, que implican costos y riesgos asociados tales como: costos ambientales (que normalmente soportan aquellos que no se benefician de la tecnología resultante); costos financieros, que a su vez levantan barreras de entrada, limitando quién puede contribuir a esta área de investigación y qué idiomas pueden beneficiarse de las técnicas más avanzadas; el costo de oportunidad, a medida que los investigadores desvían sus esfuerzos de direcciones que requieren menos recursos; y el riesgo de daños humanos sustanciales, incluidos estereotipos, denigración, aumento de la ideología extremista y arrestos injustos (Bender, 2021).

Conforme crece la demanda de recursos energéticos, empresas como Microsoft han visto que esta tendencia no va a parar, por lo que han decidido hacer fuertes inversiones en producción de energía nuclear (Bennet, 2023) e incluso han publicado una oferta de trabajo para un gerente de programa de tecnología nuclear, encargado de diseñar una estrategia de reactor "para alimentar los centros de datos en los que residen la nube de Microsoft y la IA". Si se excluyen las actividades relacionadas con criptomonedas, el consumo de energía eléctrica por parte de los centros de datos ha estado relativamente estable en alrededor del 1 % del consumo global en los últimos años (de Vries, 2023). Sin embargo, dada la carrera desenfrenada por posicionarse como líderes de la IAG que han emprendido las grandes empresas, no es descabellado suponer que dicho consumo puede aumentar de manera significativa. Conforme la IAG se incorpore cada vez más en sistemas de cómputo que se utilizan por usuarios de todos los sectores, la creciente demanda traerá también por consecuencia una mayor cantidad de recursos.

Hay estudios que matizan el impacto ambiental de las actividades en torno a la IA, por ejemplo, haciendo comparativas entre actividades llevadas a cabo por sistemas de IA y contrastando con humanos realizando las mismas actividades. En uno de tales estudios (Tomlison et al., 2023) se presenta un análisis comparativo de las emisiones de carbono asociadas con diversos sistemas de IA (ChatGPT, BLOOM, DALL-E2, Midjourney) y los individuos humanos que realizan tareas equivalentes de escritura e ilustración; en sus resultados encuentran que los sistemas de IA emiten entre 130 y 1500 veces menos CO2e por página de texto generada en

comparación con los escritores humanos, mientras que los sistemas de ilustración con IA emiten entre 310 y 2900 veces menos CO2e por imagen que sus homólogos humanos. Resultados como estos podrían utilizarse para justificar el reemplazo de trabajos asignados a personas en favor de sistemas automatizados, aunque las decisiones de este tipo no estarían considerando los altos costos derivados de alterar la estabilidad económica y social de regiones geográficas enteras.

Comentarios finales

El avance desenfrenado de la investigación y el desarrollo en torno a diversos aspectos de la IA ha provocado movimientos importantes en la industria de la alta tecnología. En unos pocos meses, OpenAI pasó de ser una minúscula empresa desconocida a una de las líderes de su segmento, con una valuación de mercado de más 80 000 millones de dólares. La gran demanda por GPUs y chips especializados para la IA, ha dado un fuerte impulso a empresas como Nvidia, que otrora ocuparan un nicho especializado de mercado. Ante la gran demanda de procesamiento, Microsoft le ha apostado muy fuerte al hardware lanzando su propia línea de chips para IA (Nellis, 2023) con el fin de posicionarse en ese mercado y no depender de proveedores externos. La industria creada en torno a la IA se ha concentrado de tal manera que un puñado de empresas determina cómo los desarrolladores de tecnología fabrican y acceden a la potencia computacional; esto influye en el comportamiento incluso de las empresas de IA más grandes cuando se enfrentan a los efectos de la escasez de poder de cómputo.

Un informe reciente de la firma de capital de riesgo Andreessen Horowitz describe la computación como "un factor predominante que impulsa la industria hoy", señalando que las empresas han gastado "más del 80 % de su capital total en recursos informáticos" (Appenzeller, 2023). Esta concentración en computación también incentiva a los proveedores de infraestructura de nube a actuar de manera que protejan su posición dominante en el mercado, compitiendo para lanzar productos antes de que estén listos para su uso generalizado. Todas estas tendencias no apuntan más que hacia una industria pujante y ávida de recursos energéticos, lo cual trae consigo impactos al ya de por sí afectado medioambiente global.

Entre los muchos recursos que demanda el desarrollo actual de la IA se encuentran los llamados recursos humanos, término que ya de por sí resulta deshumanizante. Se antoja, en cierta manera paradójica, que una industria que tiende a automatizar todo, necesite de trabajo humano para afinar mecanismos, limpiar datos y en general para hacer el "trabajo sucio" pero indispensable para que los sistemas produzcan resultados por lo menos útiles. A menos que se cambie radicalmente de paradigma, los sistemas de IA seguirán desarrollándose en el futuro mediato bajo un modelo de human in the loop (humano en el ciclo), necesitando de intervención humana para poder funcionar adecuadamente (Mosqueira-Rey et al., 2023).

A fin de cuentas, la situación de fondo es que las tecnologías de IAG, al igual que muchas otras en la industria informática, siguen un modelo de diseño centrado en las tecnologías y no en las personas, tal como lo aclara el experto en usabilidad y diseño Donald Norman: "Muchos de los sistemas, procedimientos y dispositivos actuales están centrados en la tecnología, diseñados en torno a las capacidades de la tecnología y se les pide a las personas que completen las partes que la tecnología no puede hacer". Aclara también que "Centrarse en las personas significa cambiar esto, empezando por las necesidades y capacidades de las personas. Significa considerar a todas las personas involucradas, teniendo en cuenta la historia, la cultura, las creencias y el entorno de la comunidad. La mejor manera de hacerlo es dejar que quienes viven en la comunidad proporcionen las respuestas" (Norman, 2019).

Actualmente se sigue debatiendo cuál va a ser la aplicación que justifique el desarrollo de las IAG, pues el hacer la tarea de los estudiantes, o bien generar contenido (textos, imágenes, videos, etc.), sin atribuir las fuentes usadas, no parecen ser las killer applications de este sector. Una de las aplicaciones que más presumen quienes quieren vender sistemas de IAG son los ayudantes virtuales que, según prometen, ayudarán a los expertos a llevar a cabo sus tareas de maneras más seguras y confiables. Esto se pone en duda tras resultados de estudios serios recientes, como el uso de chatbots como auxiliares en aplicaciones médicas. En dicho estudio (Lee et al., 2023), se encontró que los auxiliares cibernéticos mostraban respuestas con: a) alta variabilidad y poca similaridad, b) baja confiabilidad de alrededor de 41 % y c) potencial de incurrir en daños. Resulta claro que,

hasta ahora, es difícil confiar en un asistente virtual que ante una situación crítica puede dar respuestas no deterministas, poco confiables y con potencial de causar daños. No obstante, los avances en el área son constantes y a un ritmo fulgurante, por lo que se puede sospechar que habrá mejoras significativas que justifiquen el uso de sistemas IAG en un número creciente de tareas.

Parece haber un consenso en predecir que la IAG seguirá mostrando avances e impactos trascendentales en el futuro inmediato. Como ejemplo, la influyente publicación MIT Technology Review en sus cinco predicciones sobre IA, para el año 2024, indica que cuatro de ellas tendrán que ver con IAG, siendo la quinta relativa a desarrollos de robots generalistas. El gran reto que se presenta es que los avances en torno a la IAG se presentan con una tasa de cambio mucho mayor de lo que son capaces de absorber y asimilar individuos, organizaciones y marcos regulatorios. Si no somos capaces de comprender cabalmente las implicaciones de la IAG en los diversos ámbitos ¿cómo vamos a ser capaces de usarla, controlarla y regularla?

Bibliografía

Adarlo, S. (2023, septiembre 26). *Critics Furious Microsoft Is Training AI by Sucking Up Water During Drought. Futurism*. https://futurism.com/critics-microsoft-water-train-ai-drought

Alahmad, R. y Robert, L. P. (2020). *Artificial intelligence (ai) and it identity: Antecedents identifying with ai applications*. ArXiv, abs/2005.12196. https://api.semanticscholar.org/CorpusID:218870057

Appenzeller, G., Bornstein, M. y Casado, M. (2023, abril 27). *Navigating the High Cost of AI Compute*. Andreessen Horowitz. https://a16z.com/navigating-the-high-cost-of-ai-compute

Bederson, B. B. y Quinn, A. J. (2011). Web workers unite! Addressing challenges of online laborers. In *Conference on human factors in computing systems - proceedings* (pp. 97–105). Association for Computing Machinery. https://dl.acm.org/doi/10.1145/1979742.1979606 doi: 10.1145/1979742.1979606

Belfield, H. (2020). Activism by the ai community: Analysing recent achievements and future prospects. In *Proceedings of the aaai/acm conference on ai, ethics, and society* (p. 15–21). NewYork, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3375627.3375814 doi:10.1145/3375627.3375814

Bender, E. M., Gebru, T., McMillan-Major, A. y Shmitchell, S. (2021, March). 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).

Bennet, D. (2023, septiembre 29). Microsoft Sees Artificial Intelligence and Nuclear Energy as Dynamic Duo. Bloomberg. https://www.bloomberg.com/news/newsletters/2023-09-29/microsoft-msft-sees-artificial-intelligence-and-nuclear-energy-as-dynamic-duo

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Liang, P. (2021). *On the opportunities and risks of foundation models. ArXiv*. https://crfm.stanford.edu/assets/report.pdf

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Amodei, D. (2020, may). *Language models are few-shot learners. In Advances in neural information processing systems* (Vol. 2020-Decem). Neu-

ral information processing systems foundation. https://arxiv.org/abs/2005.14165v4

Chomsky, N. (2023). *Noam Chomsky: The False Promise of ChatGPT. NY times*. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html.

De Vries, A. (2023). *The growing energy footprint of artificial intelligence.* Joule, 7(10), 2191-2194.

Devlin, J., Chang, M. W., Lee, K. y Toutanova, K. (2019, oct). *BERT: Pre-training of deep bidirectional transformers for language understanding.* NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies- Proceedings of the Conference, 1, 4171–4186. en http://arxiv.org/abs/1810.04805

Eloundou, T., Manning, S., Mishkin, P. y Rock, D. (2023, marzo). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*. https://arxiv.org/abs/2303.10130v5

Gareth Corfield, M. F. (2022). *Meet ChatGPT, the scarily intelligent robot who can do your job better than you*.telegraph.co.uk.https://www.telegraph.co.uk/business/2022/12/05/meet-scarily-intelligent-robot-who-can-

Goertzel, B. y Singularitynet, . (2023, sep). *Generative AI vs. AGI: The Cognitive Strengths and Weaknesses of Modern LLMs*. https://arxiv.org/abs/2309.10371v1

Goodfellow, I., Bengio, Y. y Courville, A. (2016). *Deep learning. MIT Press*. (http://www.deeplearningbook.org)

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C. y Bigham, J. P. (2018). A data-driven analysis of workers' earnings on Amazon Mechanical Turk. In *Conference on human factors in computing systems - proceedings* (Vol. 2018-April). Disponible en https://doi.org/10.1145/3173574.3174023 doi:10.1145/3173574.3174023

Heritage, S. (2022). Could ChatGPT write my book and feed my kids? thetimes.co.uk.https://www.thetimes.co.uk/article/could-chatgpt-write-my-book-and-feed-my-kids-7972vx0xp.

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., Sifre, L. (2022). *Training Compute-Optimal Large Language Models. Advances in neural information processing systems* (Vol. 35). Neu-

ral information processing systems foundation. https://arxiv.org/abs/2203.15556v1

Horton, J. J. (2011). *The condition of the Turking class: Are online employers fair and honest?* Economics Letters, 111(1), 10–12. https://arxiv.org/abs/1001.1172v1 doi: 10.1016/j.econlet.2010.12.007

Hui, X., Reshef, O. y Zhou, L. (2023). *The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market*. Available at SSRN: https://ssrn.com/abstract=4527336 or http://dx.doi.org/10.2139/ssrn.4527336.

International Labour Organization. (2016). *Non-Standard Employment Around the World: Understanding challenges, shaping prospects* (Vol. 44, No. 29). http://www.ilo.org/global/publications/books/WCMS534326/lang?en/index.htm

Irani, L. C. y Silberman, M. S. (2013). Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 611-620). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2470654.2470742 doi: 10.1145/2470654.2470742

Jem Bartholomew, D. M. (2023). *How the media is covering Chat-GPT*. CJR. https://www.cjr.org/towcenter/media?coverage?chatgpt.php.

Joshua Broyde, C. P. (2021). *Build a medical sentence matching application using BERT and Amazon Sage-Make*. Amazon Web Services. https://aws.amazon.com/es/blogs/machine-learning/build-a-medical-sentence-matching-application-

Kalodimos, J. y Leavitt, K. (2020). Experimental shareholder activism: A novel approach for studying top management decision making and employee career issues. *Journal of Vocational Behavior*, 120, 103429. Disponible en https://www.sciencedirect.com/science/article/pii/S0001879120300543 doi: https://doi.org/10.1016/j.jvb.2020.103429

Lee, A. (2023). *What Are Large Language Models and Why Are They Important?* NVIDIA. blogs.nvidia.com.

Lee, P., Bubeck, S. y Petro, J. (2023). Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine*, 388(13), 1233-1239.

Merchant, B. (2023). Column: Minimum wage 'ghosts' keep Google and Microsoft's AI arms race from becoming a nightmare. *Los Angeles Times*.

https://www.latimes.com/business/technology/story/2023-02-16/column-google-microsoft-

Metz, C. (2023). The Secret Ingredient of ChatGPT Is Human Advice. *New York times*. https://www.nytimes.com/2023/09/25/technology/chatgpt-rlhf-human-tutors.html.

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J. y Fernández-Leal, Á. (2023). *Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review*, 56(4), 3005-3054.

Nast, C. (2017). *Amazon's Turker Crowd Has Had Enough*. Wired. https://www.wired.com/story/amazons-turker-crowd-has-had-enough/.

Nayak, P.(2019). Understanding searches better than ever before. *Google Blog*, 1–6. https://blog.google/products/search/search-language-understanding-bert/

Nedzhvetskaya, N. y Tan J. T. (2019). What we learned from over a decade of tech activism. *The Guardian*. https://www.theguardian.com/commentisfree/2019/dec/22/tech-worker-activism-2019-what-

Nellis, S. (2023, nov 15) Microsoft introduces its own chips for AI, with eye on cost. *Reuters*. https://www.reuters.com/technology/microsoft-introduces-its-own-chips-ai-with-eye-cost-2023-11-15/

Norman, D. (Julio 23, 2019). *The Four Fundamental Principles of Human-Centered Design and Application*. https://jnd.org/the-four-fundamental-principles-ofhuman-centered-design-and-application/

Perrigo, B. (2023). OpenAI Used Kenyan Workers on Less Than 2Per-Hour. *Time*. https://time.com/6247678/openai-chatgpt-kenya-

Rae, J. W., Borgeaud, S., Cai, T., Millican, K., Hoffmann, J., Song, H. F. e Irving, G. (2021). *Scaling language models: Methods, analysis & insights from training gopher*. CoRR, abs/2112.11446 . https://arxiv.org/abs/2112.11446

Roose, K. (2023). Bing's A.I. Chat: 'I Want to Be Alive. New York Times. https://www.nytimes.com/2023/02/16/technology/bing-chatbot-transcript.html.

Stanford University. (2023). Artificial Intelligence Index. *AI Index Report 2023*. Retrieved 2023-10-22, from https://aiindex.stanford.edu/report/

Stiennon, N., Ouyang, L., Wu, J., Ziegler, D. M., Lowe, R., Voss, C., . . . Chris- tiano, P. (2020). Learning to summarize from human feedback.

In Proceedings of the 34th international conference on neural information processing systems. *Red Hook*, NY, USA: Curran Associates Inc.

Strubell, E., Ganesh, A. y McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP.* arXiv preprint arXiv:1906.02243.

Thewsey, A. (2021). Bring structure to diverse documents with Amazon Textract and transformer-based models on Amazon SageMaker. *Amazon Web Services*. https://aws.amazon.com/es/blogs/machine-learning/bring-structure-to-diverse-documents

Tomlinson, B., Black, R., Patterson, D. and Torrance, A.W. (2023, marzo 23). *The Carbon Emissions of Writing and Illustrating Are Lower for AI than for Humans*. SSRN. https://ssrn.com/abstract=4399923

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017, jun). Attention is all you need. *Advances in Neural Information Processing Systems*, 2017-Decem, 5999–6009. http://arxiv.org/abs/1706.03762

Véliz, C. (2021). Privacidad es poder: Datos, vigilancia y libertad en la era digital. *Debate*.

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

Capítulo 3

Implicaciones sociales, retos y desigualdades en la incorporación de la inteligencia artificial en el contexto latinoamericano

Isaac de Jesús Palazuelos Rojo
Denys Serrano Arenas

https://doi.org/10.61728/AE24001038

Introducción

En el presente capítulo se plantea una revisión de diversos estudios realizados en Latinoamérica, referentes al vínculo entre inteligencia artificial (IA) y derechos básicos tales como la salud, educación y justicia. El interés central es dibujar un panorama que permita conocer cómo se han desarrollado los estudios sobre el tema, con el propósito de identificar las principales implicaciones sociales, retos y desigualdades que existen en torno a la incorporación de la IA en el contexto antes mencionado.

A través de este trabajo, se pueden apreciar diversos enfoques de investigación para futuras indagaciones que permiten conocer las implicaciones que tiene la IA sobre la experiencia social en América Latina. En este sentido, se busca contribuir a la discusión de la enorme complejidad política, económica y cultural que determinará las transformaciones futuras de la estructura social en torno a la progresiva y gradual incorporación de procesos automatizados, mediante el empleo de IA en diferentes ámbitos de la vida humana.

En este capítulo se aborda una revisión documental descriptivo-exploratorio sobre los estudios realizados en la región latinoamericana sobre la utilización de la IA en tres ámbitos sociales: educación, salud y justicia. Los artículos científicos consultados se buscaron en las bases de datos Scielo, Redalyc y Latindex, ya que son de acceso abierto y están basadas en la organización de los estudios en latinoamérica. Los criterios de selección de la muestra se centraron en los estudios publicados entre 2015 y 2023, que se enfocarán en la utilización de IA en el ámbito de la educación, salud y justicia, además de que analizarán la realidad social de la región latinoamericana. Con base en este análisis, se seleccionaron diversos estudios a fin de identificar bondades y retos en la aplicación de la IA en sectores sociales caracterizados por tener gran presencia del juicio humano en el abordaje de los procesos cotidianos, esto con el objetivo de abonar a la reflexión en torno a la relación entre ciencias sociales y computacionales, así como su aplicación en la vida cotidiana.

La inteligencia artificial en el ámbito sanitario y el cuidado de la salud

La medicina como disciplina científica mantiene una relación intrínseca con la estructura antropológica de la sociedad moderna, como práctica cultural posee la cualidad de resignificar al cuerpo como unidad política, en este sentido, la biología forma parte del ensamblaje de lo social. De manera simultánea, la atención médica mantiene un particular vínculo con las innovaciones tecnológicas lo que conlleva a ulteriores transformaciones sociales en un proceso recursivo entre la biomedicina, las tecnologías basadas en el uso de IA y la propia experiencia social que se deriva de la relación hombre-máquina (Arias, et al., 2019).

El uso de la IA en la atención médica tiene importantes implicaciones sociales debido a que sus principales beneficios son de orden estructural, como la reducción de los costos del servicio de salud pública, la predicción y el control de epidemias, el mejoramiento de los modelos de cuidados de la salud, así como la detección anticipada de enfermedades. De acuerdo con Arias et al., (2019), estos beneficios se encuentran aún en una etapa temprana y su desarrollo dependerá del tratamiento que se le dé a los grandes volúmenes de datos provenientes de los historiales médicos de pacientes, en términos de su organización, sistematización y categorización, aunado a la reducción de la brecha digital en materia de manejo de archivos de salud pública. Estas dos dificultades que plantean los autores suponen importantes esfuerzos en la creación o transformación de organismos públicos en torno a la digitalización y categorización de grandes volúmenes de datos, con el objetivo de establecer correlaciones de información útil en el mejoramiento de los servicios médicos. Es sencillo imaginar el tamaño del reto en términos de desigualdad social, económica y tecnológica, particularmente en países con sistemas políticos inmaduros y burocracias poco comprometidas con el desarrollo público (Hung-Hui, 2008).

Otros autores como Medinaceli y Silva (2021), se centran en la importancia de la anomización de los datos de los pacientes, señalan que el uso de IA en grandes repositorios de datos masivos implica también riesgos en materias de seguridad pública y derecho a la privacidad. En este sentido, la correlación de datos de salud y de historia clínica electrónica mediante

el empleo de la IA debe partir de una perspectiva de derechos humanos, para lo cual la anomización resulta fundamental. Aunque de manera simultánea, estos sistemas de procesamiento de datos también deben garantizar la disponibilidad de la información, así como la incorporación masiva y sistemática de nuevos datos clínicos de diversos grupos de pacientes. Por tanto, el riesgo no se reduce al derecho a la privacidad, sino que, la falta de heterogeneidad de grupos en los repositorios de big data en el ámbito sanitario puede derivar también en novedosas situaciones de desigualdad y exclusión social, racial, de clase o de género.

Las transformaciones sociales que implica la integración de la IA en el ámbito de la salud son complejas, como se señala en el cuarto informe sobre inteligencia artificial y telemedicina en el sector de la salud, realizado por el Banco de Desarrollo de América Latina. De acuerdo con dicho informe, el contexto latinoamericano se enfrenta a retos como: 1) el desarrollo de infraestructura que apoye el enfoque de la telemedicina y el empleo de la IA en el mejoramiento de modelos de atención médica; 2) la legislación de nuevos marcos regulatorios que permitan la optimización en el empleo de modelos de telemedicina e IA; 3) la creación de incentivos de financiamiento, a la par del fomento de una cultura de uso en los pacientes; 4) la formación o capacitación del personal sanitario, en conjunto con la sensibilización y la educación de los pacientes, 5) la implementación de mecanismos que garanticen la privacidad de datos, seguridad, accesibilidad, transparencia y rendición de cuentas (CAF, 2022).

Actualmente existen múltiples plataformas tecnológicas que brindan una gran variedad de servicios en los ámbitos señalados como lo son Clinic Cloud, Nucli, Archivex Clinical, Xkelet, Datium, Merative, Aidoc, Tempus, Deep Genomics, PathAI, Benevolent AI, Ephion Health, Fisify, Quibim, entre otras. La progresiva incorporación de la IA en la atención médica y los cuidados de la salud en este contexto, trae consigo una serie de implicaciones sociales, políticas y económicas de gran escala. Esto requerirá de esfuerzos conjuntos entre empresas privadas, sector público, academia y usuarios, que garanticen procesos de innovación sociotecnológica transparente, incluyente, equitativa y segura.

Nuevos procesos de enseñanza aprendizaje a partir de la inteligencia artificial

Existe una relación ontológica entre la educación y la tecnología. Para Maffesoli (2005), los procesos de enseñanza y aprendizaje son mecanismos a través de los cuales la sociedad extrae de la barbarie a los sujetos para formarlos en una cultura, de acuerdo con esta perspectiva, el concepto de cultura mantiene un vínculo etimológico con el de cultivar, de esta manera, lo que se busca mediante la educación es cultivar a los sujetos en el manejo y desarrollo de la técnica o el conjunto de técnicas, a decir tecnologías, que le son útiles a la propia construcción de lo social.

No es por tanto de extrañar el hecho de que desde su génesis la educación avance en paralelo con la innovación y el desarrollo tecnológico, es así que en las últimas dos décadas se ha hecho cada vez más visible la incorporación de novedosas tecnologías en las aulas, pero también en los propios modelos educativos mediante la proliferación de la educación virtual a distancia. La naturaleza ostentosa de los avances tecnológicos actuales, principalmente en materia de comunicación digital y de IA, hace de estas transformaciones en los procesos de enseñanza algo cada vez más deslumbrante y, con exceptiscismo, se pregona la creación de novedosos modelos que estimulan de manera lúdica el aprendizaje a través de la tecnología.

De acuerdo con Ochoa et al. (2019), la IA es de gran ayuda en el cumplimiento de los objetivos globales en materia educativa ya que tiene la cualidad de reducir las dificultades de acceso al aprendizaje, sin embargo, los autores remarcan la necesidad de garantizar un uso inclusivo y equitativo de la IA, que considere las desigualdades relacionadas con condiciones de clase, género, región o territorio y origen étnico. En este punto, se hace evidente la persistente desigualdad como una característica común entre los nodos en los que la IA y la experiencia social se cruzan, en el contexto latinoamericano. Por ejemplo, la cobertura de internet con fines pedagógicos en las escuelas públicas de México es de un escaso 37.3 % (MEJOREDU, 2021); en Colombia solamente el 36.63 % de escuelas cuentan con conexión a internet (Acceso a internet, 23 de noviembre de 2021); mientras que en El Salvador solamente el 25 % de las escuelas públicas tiene acceso a internet (DPL News, 23 de enero de 2020); en Ecuador solamente el 37

% de las escuelas tiene conexión (Vanessa Lopez, 3 de junio de 2021); y en En Perú la situación es aún peor con una cobertura de alrededor del 25 % (DLP News, 25 de julio de 2019); los países de latinoamérica que tienen mayor cobertura de internet en sus escuelas son Chile y Argentina con alrededor del 70 %. Estos datos hacen evidente la brecha digital que aún se experimenta en materia educativa en esta región.

Para Ferrante (2021) hablar de desigualdad en los procesos de incorporación de la IA en los sistemas educativos, no se reduce a analizar las carencias de las estructuras sociales y políticas encargadas de proveer e innovar servicios públicos. Mediante el concepto de "sesgo algorítmico", el autor propone prestar particular atención al propio diseño de los modelos educativos basados en IA, para así poder entender de qué manera estos modelos están atenuando viejas desigualdades, al depender de bases de datos que no resultan representativas en poblaciones compuestas por grupos heterogéneos. Por tanto, el autor propone que para construir una inteligencia artificial más justa en materia educativa es necesario centrarse en una triada compuesta por los datos, los modelos y las personas, en medida que los datos no sean representativos y que los modelos sean diseñados con base en estos sesgos, el resultado será siempre una IA injusta y excluyente.

Otros autores como Ocaña-Fernández et al. (2019), plantean que uno de los principales retos de la incorporación de la IA en este ámbito es precisamente los propios modelos y programas usados por las instituciones, en mayor medida las de educación superior. Los autores argumentan la necesidad de replantear los viejos cánones empleados en las universidades, para empezar a reconocer que los nuevos modelos educativos deben centrar gran parte de sus esfuerzos en la formación de profesionales, los cuales no solo tengan la capacidad de operar la tecnología, sino que puedan innovar y transformar el entorno tecnológico de acuerdo con las propias necesidades de su ejercicio profesional. Desde esta perspectiva, los entornos digitales tienen como fundamento la competitividad del mercado, misma que está sujeta al uso de tecnologías cada vez más especializadas.

Derivado de lo anterior, es necesario integrar algunas competencias educativas como el pensamiento computacional, esto se refiere a la construcción de subjetividad asociada al uso cotidiano de dispositivos electrónicos de cómputo, si bien el desarrollo de este tipo de pensamiento es

mucho más sencillo para las nuevas generaciones de nativos digitales, en muchas zonas de Latinoamérica persisten grupos cuyas nuevas generaciones no tienen este acercamiento a la tecnología, lo que supone un escaso o nulo desarrollo del pensamiento computacional y requiere urgentemente que esta brecha se acorte. El lenguaje de programación es otra competencia importante para poder desarrollar capacidades de innovación e incorporación creativa de las tecnologías en el desempeño profesional, para ello es necesario reducir los niveles de ignorancia técnica que predomina en los usuarios, se trata de transformar el rol de usuarios pasivos y dependientes del soporte técnico, para formar profesionistas que, sin importar las áreas en la que trabajan, tengan las capacidades de interactuar con la tecnología desde un rol más activo a través de un lenguaje universal que se sustenta en la programación. Esta competencia mantiene un vínculo con la noción de sociedad del conocimiento en el sentido en el que propone que cualquier usuario u operario puede intervenir e innovar en los procesos de reproducción social. Finalmente, están las competencias informáticas, para desarrollar estas novedosas habilidades en los estudiantes es necesario crear políticas públicas que apuesten por la generalización de equipos docentes, capaces de orientar su práctica hacia la construcción de sentidos fundamentados en la incorporación masiva de las tecnologías en prácticamente todos los campos del conocimiento (Ocaña-Fernández et al., 2019 y Pandiella et al., 2018).

Derivado de lo anterior, se aprecia que las principales implicaciones y retos que enfrenta la incorporación de la IA en materia educativa en el contexto latinoamericano, se centra en el desarrollo de estructuras sociales que se enfoquen en reducir la brecha digital en términos de infraestructura y acceso, esto por un lado, mientras que por otro lado sobresale la necesidad de generar nuevos modelos educativos flexibles a la celeridad de cambio, acompañado de políticas públicas que fomenten la formación docente en estos nuevos modelos.

Automatización e innovación de procesos jurídicos mediante el uso de IA

En otros ámbitos en los que se incorpora gradualmente la IA en los procesos internos de las organizaciones e instituciones públicas, principalmente aquellas relacionadas con el combate a la corrupción, se habla con mayor firmeza. Debido a la propia naturaleza de las disciplinas que conforman dichas instituciones, les resulta sencillo reconocer incluso como un axioma la necesidad de garantizar los derechos básicos de igualdad, equidad y justicia, sin embargo, no desaprovechan en resaltar las múltiples posibilidades que se pueden tener del uso de algoritmos e IA en materia de prevención de delitos (Florencia-Cabrera, 2022).

En efecto, el éxito que ha tenido la analítica de grandes volúmenes de datos en la industria y la mercadotecnia, en mayor medida se debe a las propias cualidades predictivas resultantes de dicha operación, lo cual permite advertir las direcciones y tendencias del mercado global, para con base en ello tomar decisiones e implementar mejores estrategias con la finalidad de eficientar futuras transacciones o ventas. Derivado de la penetración masiva que tienen los dispositivos digitales en la vida cotidiana, se hace evidente que el corolario de conductas que se pueden predecir es múltiple y por supuesto también su rentabilidad tanto política como económica se hace visibles.

Esta proyección no se reduce a una visión utópica del mejoramiento de los sistemas judiciales y el combate a la corrupción, sino que la propia realidad rebasa la adaptación de estos sistemas, pues la vulnerabilidad electrónica frente a delitos digitales demanda por un lado el desarrollo y la comercialización de programas antifraude y, por otro lado, reclama tanto el desarrollo como la implementación de marcos jurídicos que regulen, vigilen y penalicen los delitos digitales así como los mecanismos de protección digital que emplean las empresas (Florencia-Cabrera, 2022).

Derivado de lo anterior, el vínculo entre IA y sistemas jurídicos es particularmente interesante, ya que permite profundizar en la complejidad de la relación entre tecnología y sociedad, pues no solo reconoce que deben desarrollarse nuevos procesos tecnológicos con base en la equidad y la igualdad para mejorar las funciones del sistema judicial en el contexto di-

gital actual, sino que parte de los retos que implica hacerlo si se reconoce el desfase que existe entre la implementación de la IA en la vida cotidiana y la adaptación de los marcos jurídicos para regular esas nuevas realidades, las cuales se traducen en forma de novedosos delitos, así como de nuevos mercados de seguridad digital privada.

Por otra parte, Ortega y Becerra (2021) plantean que el análisis y la categorización de datos mediante el uso de la IA para la toma de decisiones jurídicas, se integra cada vez más en los procesos legales, en este contexto, se preguntan si es posible que esta tecnología pase de ser un mero instrumento que facilita la toma de decisiones, hasta convertirse en un ente generador de decisiones jurídicas. De acuerdo con los autores, esto tiene otras implicaciones particulares a la hora de interpretar y utilizar datos, especialmente cuando se trata de ejercer discrecionalidad sobre el valor del dato en relación con el nivel de legalidad y ética del método de obtención del mismo.

Este proceso es relevante ya que el uso de IA está transformando los métodos de obtención de información que es importante para la resolución de procesos jurídicos, por otra parte, la relevancia también descansa sobre la cualidad que esta tecnología tiene para cuestionar el propio criterio de objetividad humana, al momento de emitir juicios que pretenden garantizar la impartición de justicia. En este marco argumentativo, Segura (2023) destaca que las dificultades de utilizar la IA en el contexto latinoamericano en materia de justicia estriban en la complejidad analítica entre ética, el razonamiento y la jurisdicción, en un escenario de desigualdad que caracteriza a la región. Por un lado, la IA ofrece celeridad y seguridad jurídica, pero simultáneamente amenaza la libertad, privacidad e igualdad. En los estudios en materia de justicia es posible apreciar que la IA es utilizada como una herramienta que coadyuva a eficientar los procesos de organización y sistematización de la información. Esto apoya en procesos complejos, tal como la toma de decisión jurídica, pero no exime el análisis cualitativo de la mirada humana sobre todo en contextos con marcadas desigualdades estructurales como son los de países latinoamericanos.

Resultados

Hemos elaborado una investigación de corte descriptiva-documental sobre el vínculo entre la IA y los derechos básicos de acceso a la salud, la educación y la seguridad pública. Encontramos que en materia de salud las discusiones giran en torno a la brecha digital que experimenta este sector, principalmente desde las estructuras orgánicas en hospitales y otras instituciones relacionadas con la atención médica pública. Esto implica una transformación radical mediante la creación de diversos organismos que se den a la tarea de digitalizar historiales clínicos y categorizar la información de manera masiva.

Una de las principales preocupaciones en dicho proceso de transformación es salvaguardar el derecho a la privacidad de los pacientes, pero que al mismo tiempo se pueda garantizar también el acceso abierto a los datos, así como la representatividad de grupos heterogéneos en estas bases de grandes volúmenes de datos. Afinar dichos procedimientos requiere de la legislación de nuevos marcos regulatorios, así como la creación de incentivos de financiamiento, a la par del fomento de una cultura de uso, la formación del personal y la rendición de cuentas.

Por otra parte, en el ámbito educativo encontramos que las reflexiones se centran un poco más hacia la estructura interna de la educación como servicio público, si bien se reconocen también algunos de los retos externos que implica en materia de igualdad de acceso y de equidad, mediante la noción de sesgo algorítmico establecen una crítica al propio diseño de los modelos educativos basados en IA, ya que los tipos de datos escasamente representativos, así como los prejuicios de los diseñadores impiden que estos sean incluyentes. En paralelo, se destaca la necesidad de flexibilizar los modelos educativos y los viejos cánones de docencia, además de redirigir las actualizaciones de los planes y programas de estudio, con base en novedosas competencias que permitan a los estudiantes formarse con capacidades para intervenir activamente en los procesos de innovación.

Finalmente, en lo que concierne al ámbito jurídico, la necesidad de garantizar derechos básicos de igualdad, privacidad, seguridad y rendición de cuentas aparece de forma implícita, esto permite profundizar en la reflexión ya que visibiliza el desfase que existe entre la incorporación de

la IA en las prácticas culturales en relación con la adaptabilidad de los sistemas y los marcos jurídicos. Asimismo, el binomio humano-máquina aparece con mayor naturalidad, a diferencia del ámbito de la medicina en donde las decisiones son cruciales en términos de vida o muerte, en el terreno de lo jurídico la reducción del protagonismo humano sí se aprecia como algo más factible, sin embargo, se replantea con mayor firmeza la discusión en torno al valor del dato en relación con el nivel de legalidad y ética del método de obtención del mismo. En este sentido, la discusión avanza en reconocer la necesidad del análisis cualitativo humano, especialmente cuando este desfase entre adaptabilidad del sistema y empleabilidad de IA se hace presente, desde las brechas que aún persisten en contextos de desigualdad social y tecnológica.

Discusiones y conclusiones

Existe un vínculo muy grande entre el desarrollo de tecnologías digitales (particularmente la IA) y la configuración sociopolítica contemporánea, no es casualidad que la tensión en torno a un posible conflicto bélico entre EUA y China crezca cada vez más, debido a los cambios económicos y productivos que supone la propia transformación tecnológica, en lo que se ha denominado la guerra de los chips entre estas dos naciones. En este sentido, se hace cada vez más necesario reflexionar sobre la relación entre tecnología y sociedad, especialmente desde las dimensiones de análisis científico que promueven las disciplinas especializadas en cada una de estas dos dimensiones.

Desafortunadamente, el estudio de este binomio no ha sido plenamente abordado desde la interdisciplinariedad en forma integral, a pesar de que los intercambios entre disciplinas son cada vez más visibles, la producción de literatura académica sigue siendo todavía muy endémica y se construye desde el interior de márgenes teóricos y metodológicos, fronteras que a pesar de volverse cada vez más flexibles en la actualidad siguen constriñendo tanto las reflexiones como los hallazgos de los estudios. Ante esta situación, es urgente que los ritmos de producción e intercambio entre ciencias sociales y ciencias computacionales crezcan, aunado a un esfuerzo histórico para tratar de adaptarse a la propia celeridad de cambio que

experimentan las innovaciones tecnológicas, así como sus consecuentes impactos en la vida social.

En tanto este nivel de intimidad e intercambio no se logre entre las diversas disciplinas, las ciencias computacionales correrán el riesgo de reproducir desigualdades, especialmente a la hora de traducir sus avances investigativos en el diseño de tecnologías rentables, en paralelo, las ciencias sociales se verán imposibilitadas en su capacidad de profundizar en la comprensión de los fenómenos estudiados, limitándose a denunciar desigualdades y augurar transformaciones complejas que no serán del todo claras para los propios estudiosos del tema, además de que perderán su utilidad para asistir a los tomadores de decisiones en la incorporación de la IA en el sector público.

Por otro lado, a pesar de que los grandes retos que implica incorporar la IA en la salud, la justicia y la educación en Latinoamérica tengan un vínculo con la particularidad histórica de esta región, dado que los esfuerzos necesarios para disminuir la brecha y el rezago tecnológico implican la participación de burocracias comprometidas con el desarrollo social, las desigualdades implicadas en este proceso de incorporación no dependen enteramente de la disposición de un sistema político y democrático maduro. En este sentido, se hace vital la integración de reflexiones decoloniales en el propio desarrollo tecnológico, así como en la investigación que se hace al respecto desde América Latina.

Mientras que la inversión en investigación computacional y sociotecnológica que se hace desde Latinoamérica no asuma compromisos decoloniales, la brecha digital y la desigualdad en los procesos de incorporación de IA en el sector público seguirán creciendo, además de que el propio proceso seguirá reproduciendo desigualdades mediante modelos que no son representativos de la heterogeneidad que caracteriza a dicha región. Dicho de otro modo, es fundamental que los diseños de modelos tecnológicos tengan desde su génesis una perspectiva vernácula que atienda las necesidades y particularidades de los contextos en los que se pretenden emplear, para de esta manera facilitar los mecanismos de incorporación de IA en el sector público de los países que forman parte del contexto antes mencionado.

Referencias

Arias, V., Salazar, J., Garicano, C., Contreras, J., Chacón, G., Chacín-González, M., ... Bermúdez-Pirela, V. (2019). Una introducción a las aplicaciones de la inteligencia artificial en Medicina: Aspectos históricos. *Revista Latinoamericana de Hipertensión*, 14(5), 590-600. https://www.redalyc.org/journal/1702/170262877013/170262877013.pdf

Bustamante Rúa, M. M., Ángel Muñoz, S., Giraldo Aristizábal, J. A. y Marín Tapiero, J. I. (2020). Mecanismos alternativos de solución de conflictos (masc) e inteligencia artificial (ia) para la solución de controversias en línea (scl): una apuesta por la descongestión en la administración de justicia. *Revista de Direito, Estado e Telecomunicacoes*, 12(1). https://doi.org/10.26512/lstr.v12i1.25808

Corvalán, J. G. (2018). Inteligencia artificial: retos, desafíos y oportunidades-Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista de Investigações Constitucionais*, 5, 295-316. https://doi.org/10.5380/rinc.v5i1.55334

Doroteo Guerrero, F. V. y Camacho Ortiz, M. E. (2021). *Inteligencia artificial y la automatización de procesos judiciales en la administración de justicia del Perú* [Tesis de licenciatura]. Repositorio Institucional UPN. https://repositorio.upn.edu.pe/handle/11537/29472

Espinoza, P. N. B., Vargas, C. P. V. y Vargas, R. P. V. (2019). Inteligencia artificial como recursos educativos abiertos. *Revista Tecnológica Ciencia y Educación Edwards Deming*, 3(1), 36-49. https://orcid.org/0000-0002-6864-7105

Ferrante, E. (2021). Inteligencia artificial y sesgos algorítmicos ¿Por qué deberían importarnos? *Nueva sociedad,* (294), 27-36. https://biblat.unam.mx/hevila/Nuevasociedad/2021/no294/3.pdf

Florencia, C., R. (2022). La inteligencia artificial para prevenir la corrupción. *Revista de Investigación Científica*, 1(1), 115-125.

Florez Hernandez, I. C. (2020). Inteligencia Artificial aplicada a la Justicia. *Derecho y Realidad*, 18(35). https://doi.org/10.19053/16923936.v18.n35.2020.9638

García Villarroel, J. J. (2021). Implicancia de la inteligencia artificial en las aulas virtuales para la educación superior. *Orbis tertius - upal*, 5(10),

31–52. https://www.biblioteca.upal.edu.bo/htdocs/ojs/index.php/orbis/article/view/98

Hung, J. H. (2008). América Latina: la corrupción y la pobreza. *Revista del CESLA*, (11), 105-118. https://www.redalyc.org/pdf/2433/243316550009.pdf

Hurtado-Cortés, L. L. H., Viillareal-López, E. V. y Villarral-López, L. V. (2016). Fault detection and diagnosis through artificial intelligence techniques, a state of art. *DYNA: Revista de la Facultad de Minas*, 83(199), 19-28. Universidad Nacional de Colombia. Sede Medellín. https://dialnet.unirioja.es/servlet/articulo?codigo=5907085

Jara, I. y Ochoa, J. M. (2020). Usos y efectos de la inteligencia artificial en educación. Sector Social división educación. *Documento para discusión número IDB-DP-00-776. BID*. http://dx. doi. org/10.18235/0002380

Játiva, J. J. y Beltrán Morales, J. (2021). Uso de la metodología STEAM para motivar a niños el uso de Inteligencia Artificial. *Revista Ibérica de Sistemas e Tecnologias de Informação,* 42(2), 31-45. https://dialnet.unirioja.es/servlet/articulo?codigo=8624550

Le Fevre Cervini, E. (2022, agosto 4). Uso estratégico de datos e inteligencia artificial en la justicia. Informe 6. CAF. https://scioteca.caf.com/handle/123456789/1932

Medinaceli Díaz, K. I. y Silva Choque, M. M. (2021). Impacto y regulación de la Inteligencia Artificial en el ámbito sanitario. *Revista IUS*, 15(48), 77-113. https://doi.org/10.35487/rius.v15i48.2021.745

Mujica Rodríguez, I. E., Toribio Salazar, L. M. y Cóndor Cámara, D. F. (2020). Inteligencia artificial como apoyo a intervenciones no farmacológicas para combatir la COVID-19. *Revista Peruana de Medicina Experimental y Salud Pública*, 37(3), 582-584. https://www.scielosp.org/pdf/rpmesp/v37n3/1726-4642-rpmesp-37-03-582.pdf

Martinez, L. E. (2020). *Inteligencia Artificial y Salud. Centro de Estudios en Tecnología y Sociedad (CETyS) de la Universidad de San Andrés (UdeSA)* [Tesis de licenciatura]. https://repositorio.udesa.edu.ar/jspui/handle/10908/17691.

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A. y Garro-Aburto, L. L. (2019). Inteligencia artificial y sus implicaciones en la educación supe-

rior. *Propósitos y representaciones*, 7(2), 536-568. http://orcid.org/0000-0002-9453-9810

Ortega, L. G. y Becerra, J. A. (2022). La inteligencia artificial en la decisión jurídica y política. *Araucaria: Revista Iberoamericana de Filosofía, Política, Humanidades y Relaciones Internacionales,* 24(49), 217-238. https://dialnet.unirioja.es/servlet/articulo?codigo=8465009

Rodríguez, M. (2022). *Inteligencia artificial y telemedicina en el sector de la salud-Oportunidades y desafíos*. Informe 4. CAF.

https://scioteca.caf.com/bitstream/handle/123456789/1923/Inteligencia%20artificial%20y%20telemedicina%20en%20el%20sector%20de%20la%20salud%20%20oportunidades%20y%20desaf%C3%A-Dos.pdf?sequence=5

Rodríguez, J. M. y Barrón, M. A. (2022) Inteligencia artificial y su aplicación en los Sistemas de justicia en América Latina. *Temas estratégicos*, No. 101. Instituto Belisario Domínguez. Senado de la República. México.

https://bibliodigitalibd.senado.gob.mx/handle/123456789/5594

Segura, R. E. (2023). Inteligencia artificial y administración de justicia: desafíos derivados del contexto latinoamericano. *Revista de Bioética y Derecho*, (58), 45-72. Epub 25 de septiembre de 2023.https://dx.doi.org/10.1344/rbd2023.58.40601

Veltri, O., Aguilar, G., Repetti, M. y Rodríguez, A. (2021). La inteligencia artificial en la formación de traductores e intérpretes. *Anuario de Investigación USAL*, (8). https://p3.entendiste.ar/index.php/anuarioinvestigacion/article/view/5730

Capítulo 4

Exploración de una perspectiva sistémica alrededor del desarrollo de la inteligencia artificial y su adopción social: la responsabilidad de la academia

Arturo Serrano-Santoyo

https://doi.org/10.61728/AE24001045

Introducción

Los acontecimientos recientes relacionados con el lanzamiento de herramientas computacionales poderosas basadas en inteligencia artificial (IA) han generado inmensas expectativas y reacciones encontradas sobre la conformación de un futuro cada vez más centrado en determinismo tecnológico (Alan Turing Institute, 2019; McCarthy, 2007; Turing, 1950; McKinsey y Company, 2023; UN Advisory Body on AI, 2023; National Institute of Standards and Technology, 2023; Altman et al., 2023). Por un lado, los sectores corporativos ven en general con optimismo la transformación profunda que la IA producirá en cuanto a competitividad y productividad empresarial y sus implicaciones sobre el trabajo del futuro y el futuro del trabajo (Acemoglu y Restrepo, 2018).

Por otro lado, otros sectores de la economía, educación, gobierno y sociedad muestran temores y preocupaciones sobre la falta de esquemas de gobernanza integrales y los riesgos que presenta el avance imparable de la IA sin una plataforma ética y moral ni la existencia de entornos regulatorios ad-hoc que puedan sostener su adopción en forma justa y equitativa (Whittaker, 2018; Université de Montréal, 2018; Shneiderman, 2020; Coombs et al, 2021; Adams, 2021; Bahá'í International Community, 2022). Las narrativas se encuentran oscilando entre la tecnofobia y la tecnofilia, entre defensores del acecho agobiante del empuje tecnológico de la IA y los promulgadores de un futuro brillante para una sociedad más eficiente, más conectada y ágil (AI Safety Summit, 2023).

En este escenario, como ha sucedido en el surgimiento de anteriores revoluciones industriales, la sociedad adopta gradualmente consciente o inconscientemente las transformaciones engendradas por los cambios tecnológicos asociados. Como es ya conocido, tales desarrollos tecnológicos han traído grandes beneficios y al mismo tiempo desafíos, tal como es el caso actual de la IA. Sin embargo, hay elementos que hacen suigéneris la condición que experimenta la sociedad actualmente. Uno de estos elementos es la explosiva e inexorable velocidad del cambio tecnológico, otro es

el grado de interconectividad que existe globalmente, que en revoluciones tecnológicas anteriores no existía en el mismo nivel, lo cual incrementa sustancialmente las interacciones tecno-sociales entre individuos, comunidades e instituciones; más aún, la naturaleza de las tecnologías asociadas a la IA que dotan de capacidades de cognición humana con un grado de autonomía considerable a aplicaciones de reciente creación, exacerban la centralidad de la tecnología sobre labores humanas en sus quehaceres cotidianos y en la toma de decisiones. Aunado a estos tres factores es el hecho de que los desarrollos tecnológicos sobre IA son mayormente impulsados por empresas digitales hegemónicas, tanto incumbentes como de reciente creación, cuyo objetivo y prioridad es impulsar ganancias corporativas y posicionamiento en un mercado altamente competitivo y no necesariamente bienestar y progreso social. Es decir, el motor económico es el principal aliciente para acelerar las innovaciones (Cockburn, 2019; Mariani, 2023; Shneiderman, 2020; Caruso, 2018).

Otro factor para considerar en la complejidad del escenario actual es el entorno geopolítico cargado de luchas por el poder tecnológico, financiero y comercial. En resumen, la sociedad se encuentra en una coyuntura que demanda la intervención de agentes que apoyen y guíen a la población en la toma de decisiones sobre su futuro desenvolvimiento a corto, mediano y largo plazo; y lo que es más urgente, sobre el tipo de sociedad deseamos para las futuras generaciones (AI Ethics Initiative, 2017; Aiken y Epstein, 2000; Algorithmwatch.org, 2020; Ashok et al, 2022; Boddington, 2017; Ethics y Society, 2017; Dwivedi, 2023; Floridi, 2018, 2019, 2020; Jobin et al, 2019; Martin, 2019; Taeihagh, 2021).

Figura 1. Elementos claves de Gobernanza de la IA.

El acelerado despliegue de aplicaciones de IA, sin entender de manera integral las implicaciones en educación, comportamiento, salud y otras áreas del conocimiento, pone en juego y abre debates sobre qué significa ser humano, su naturaleza espiritual, su propósito al servicio a los demás en el trabajo y acción social y otros temas de trascendencia. Un agente preponderante en ayudar a la población para navegar en el incierto, desafiante y tempestuoso mar del escenario actual descrito brevemente en los anteriores párrafos es una concertada participación de gobiernos, academia y organizaciones de la sociedad civil para paliar los retos de la metamorfosis tecno-social en torno a la IA (Agrawal, 2022; AI Now Institute, 2017; Alshater, 2022).

Para ello, se presenta en esta modesta contribución la propuesta alrededor de un marco de referencia para la academia, que soporte los esfuerzos sociales para aprovechar los beneficios de la IA y mitigar los elementos perniciosos que inhiben su adopción balanceada y justa mediante esquemas de gobernanza que consideren cuatro elementos clave entrelazados: cambio tecnológico e innovación, regulación ad-hoc, privacidad y seguridad de la información y aplicación de principios éticos en el diseño y puesta en marcha de productos y sistemas basados en IA (ver Figura 1).

Bajo esta perspectiva, se acude a la ciencia de la complejidad para delinear dichas propuestas con base en un enfoque interdisciplinar (García, 2006).

Análisis del contexto: el "digisistema" y la transformación digital

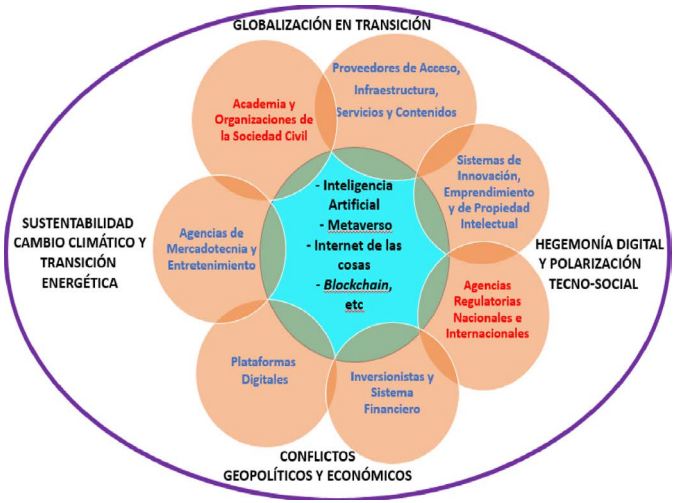
En esta propuesta se adopta un enfoque basado en ciencia de la complejidad. El análisis realizado se lleva a cabo situándose en primera instancia en el "borde sistémico"; es decir, en una posición fronteriza que permita observar el sistema de la Transformación Digital hacia adentro (ver Figura 2) para identificar sus componentes, sus motivaciones y prioridades y observar hacia a fuera, i.e. sus interacciones y efectos en la sociedad y usuarios.

Al intentar analizar el entorno contemporáneo actual, resulta pertinente explorar la estructura y dinámica del proceso de transformación digital, al cual la sociedad está sujeta. Se entiende por transformación digital a la operación del conjunto de procesos, sistemas y agentes que actúan en la sociedad afectando la manera que aprendemos, trabajamos, nos comunicamos y hacemos nuestras tareas cotidianas mediante herramientas digitales (Oracle, 2024; SAP, 2024, Centro México Digital, 2023). Un análisis profundo de los efectos de la transformación digital en curso exige un examen crítico y reflexión profunda sobre las implicaciones socioculturales, educativas y políticas que esta transformación digital trae consigo. No obstante, las acciones y estrategias de los poderosos contribuyentes y promotores de la Transformación Digital, mostrados en la Figura 2, aunque conscientes, en muchos casos, de los cambios axiales que se avecinan, privilegian la competencia, las ganancias financieras y el consecuente dominio de mercado, acelerando sus procesos de innovación tecnológica, para posicionarse como líderes de un campo de batalla polarizado globalmente y en esta carrera dejando atrás a aquellas instancias –académicas, gubernamentales, empresariales y sociales– que no cuentan con una cadena de innovación sólida y coherente, ni cuentan con los enormes recursos financieros y de investigación y desarrollo tecnológico requeridos.

Ante esta situación, el conocimiento avanzado sobre tecnologías digitales emergentes como la inteligencia artificial, Metaverso, internet de las cosas y Blockchain, entre otras, se concentra en manos de un reducido

grupo de compañías, emprendedores e inversionistas, sin embargo, omnipresentes e influyentes, cuyo discurso de transformación digital promete un avance significativo de productividad, eficiencia, rapidez y, en general, un aporte a beneficios de diversa naturaleza para la población. Este discurso dista de colocar el énfasis en soluciones centradas en beneficio y progreso social. Se habla de regulación y ética, de las implicaciones en el empleo, desarrollos militares y de cambios sustanciales en la actual estructura geopolítica polarizada, sin embargo, la batalla por la hegemonía digital global marca el ritmo de una competencia feroz con énfasis en dominio de mercado y ganancias financieras (Kshetri, 2023).

Figura 2. El "Digistema" de la transformación digital.



En cuanto a las organizaciones que buscan lograr una transformación continua en sus operaciones clave, se dan cuenta de que lo digital, no es el único ingrediente en este trayecto. Estas organizaciones "progresistas" advierten que es necesario alinear las capacidades tecnológicas a condiciones cambiantes que requieren construcción de capacidades y fortalecimiento de valores fundamentales humanos, modelos operativos y actitudes que conduzcan a la creación y extracción de valor más allá de solo ganancias corporativas o posicionamiento de mercado. Una verdadera y consciente transformación digital de largo plazo, diferirá de proyectos centrados solamente en adquisición y despliegue de productos (infraestructura y servi-

cios); tiene el potencial de generar nuevas relaciones de beneficio mutuo para el empleador y el empleado. Lo digital será solo un vehículo y no el objetivo para lograr nuevas oportunidades de cooperación en el futuro del trabajo y el trabajo del futuro (Frey y Osborne, 2017).

El proceso de avance y penetración en el tejido social de la transformación digital es irreversible y sus potenciales beneficios innegables. Se tiene la oportunidad de generar bienestar social y desarrollo sustentable ante un escenario de globalización en transición, cambio climático y polarización geopolítica; sin embargo, los desafíos son enormes. El sistema digital, denotado como Digisistema en la figura 2 muestra los actores más representativos de este ecosistema. El surgimiento de un nuevo escenario mundial está en "proceso de parto". Las instancias que impulsan este surgimiento mediante un enfoque de "solucionismo" tecnológico enfrentan el contrapeso de la academia, colectivos y organizaciones progresistas de la sociedad civil, así como de agencias regulatorias nacionales e internacionales para lograr un adecuado balance en la apropiación sociocultural de la transformación digital.

La simbiosis de la regulación y la ética en el desarrollo y adopción de la IA

Las innovaciones aceleradas que dieron lugar al surgimiento de los modelos recientes de IA (GPT4, Claude, Llama, Gemini, etc.) han atestiguado a su vez la aparición de iniciativas regulatorias de impacto global, como son las iniciativas de la Unión Europea y de los Estados Unidos de Norteamérica (AI Safety Summit, 2023). Otros países e instancias gubernamentales y no gubernamentales están trabajando en sistemas regulatorios que den lugar a la adopción balanceada de la IA y tecnologías complementarias. A su vez; y a pesar de que empresas de vanguardia atienden políticas de interés y atención a los factores éticos en cuanto a diseño de aplicaciones e implantación de sistemas de IA dentro de sus planes de desarrollo algorítmico, la prioridad se centra, como se mencionó anteriormente, en su desempeño económico-financiero.

Aunado a ello, se esgrimen discursos de frentes gubernamentales y corporativos sobre los posibles efectos regulatorios que puedan inhibir

la innovación afectando competitividad empresarial y posicionamiento geopolítico y financiero. Por un lado, se reconoce el papel fundamental de entornos regulatorios sólidos y coherentes. Por otro lado, dada la naturaleza y paradigma tecnológico de la IA, su control regulatorio dista de la regulación de las tradicionales Tecnologías de la Información y la Comunicación (TIC), lo cual es un reto significativo e implica replantear nuevos esquemas que atiendan el metabolismo regulatorio emergente. En la regulación de las TIC se consideran los factores económicos y de competencia, los procesos de transporte y distribución de información que incluyen la administración del espectro y su adecuado, justo y equitativo uso, así como la operación de redes dentro de un marco jurídico-tecnológico en cuanto a equipos, sus posibles interferencias y convivencia de operadores y concesionarios.

Nos encontramos ahora ante un nuevo paradigma impulsado por la IA, donde la instancias gubernamentales responsables de la privacidad y acceso a la información entran en juego en conjunto con las instancias dedicadas a la regulación de las TIC y las de competencia económica para la generación de esquemas regulatorios relacionados a procesos algorítmicos involucrados en el desarrollo de modelos de lenguaje, procesamiento y aprovechamiento de grandes cantidades de datos, tareas de entrenamiento dentro de los modelos, ejecución de aplicaciones dentro de enormes centros de datos y lo que es clave, insertar aplicaciones de IA en los procesos administrativos, de manufactura, de toma de decisiones y general en todos los sectores de la economía y sociedad de manera justa, equitativa y transparente. En todas estas funciones la perspectiva ética es fundamental, a tal grado que las corporaciones de vanguardia en el desarrollo de la IA han puesto en marcha estrategias y operación de grupos de trabajo internos y externos para atender los emergentes requerimientos éticos de una tecnología de gran poder transformativo y con potencial de generar daños intencionales o no intencionales en la puesta en marcha de las aplicaciones y sistemas desarrollados (Winfield y Jirotka, 2018; UNESCO, 2022; UNESCO, 2023a; UNESCO, 2023b; Moor, 2008; Gartner, 2021).

Así pues, se argumenta que la regulación, la ética y la privacidad, transparencia y seguridad de la información para la IA forman un tejido entrelazado e interdependiente dando lugar a un sistema dinámico complejo

que debe estudiarse a profundidad (ver Figura 1). Se propone para las universidades el explorar una perspectiva sistémica que ofrezca, en una primera aproximación, un marco de referencia y actuación con dos pilares básicos para enfrentar los retos de profunda transformación cognitiva que la IA presenta a la humanidad. El no atender estos retos generará aún mayores desventajas económicas, sociales y culturales ante un incierto e impredecible contexto global.

La relevancia de la academia para facilitar y estimular la adopción social balanceada de la inteligencia artificial

La academia ha jugado un papel fundamental en el desarrollo tecnológico y en la digitalización de la sociedad. Los desarrollos científicos en las universidades y centros de investigación, en su mayoría apoyados por fondos públicos y militares, dieron lugar al surgimiento de tecnologías disruptivas que han moldeado el orden social. Entre estos desarrollos destacan el Internet y la web, por mencionar algunos de los más significativos. Aunque el génesis de estas tecnologías fue el entorno académico, su dominio actual se ejerce desde grandes corporativos que amasan enormes cantidades de datos, muchos de ellos provenientes de usuarios de sus principales aplicaciones (Ibero, 2023a).

En el caso de los recientes desarrollos de inteligencia artificial generativa (IAG) y la inteligencia artificial predictiva (IAP), su génesis de se ha dado en corporaciones que se encuentran en una batalla por lograr aplicaciones con mayor poder transformativo y potencial comercial y así proveer a sus accionistas y aliados económicos de ganancias y poder sustancial de mercado (Techtarget, 2023, BCG, 2023). Las implicaciones éticas, regulatorias y daños cognitivos potenciales están en el debate y se tiene un conocimiento general de las consecuencias intencionales y no intencionales de una adopción abrupta e irreflexiva. Sin embargo, el impulso a la innovación es preponderante y prioritario, no solo por las razones económicas anteriormente mencionadas, sino también por intentar ejercer poderío geopolítico que el dominio de tecnologías como la IA daría a gobiernos y sus corporaciones.

Como es bien sabido, las universidades tienen tres objetivos fundamentales: La enseñanza, la investigación y la extensión con el propósito de

formar ciudadanos capaces de transformar la sociedad y dignificar al ser humano. Todo esto refleja la relevancia y responsabilidad de las instituciones académicas para contribuir al desarrollo integral de sus comunidades y es por tanto un factor que adquiere relevancia particular en el actual escenario brevemente expuesto. Dada la velocidad de cambio tecnológico de la IA, el conocimiento generado y las innovaciones alrededor de ramas complementarias generadas a través de la infusión de grandes cantidades de recursos financieros, han dado lugar a una brecha de innovación que no solo afecta a países menos desarrollados industrial y científicamente, sino también a otras economías importantes que no muestran el mismo ritmo y magnitud de innovaciones disruptivas, lo cual produce concentración de conocimiento en un número reducido de zonas geográficas y en un selecto grupo de individuos y entidades con fines de lucro.

En los años recientes, esta brecha de innovación se ha transformado en abismo, que por un lado orilla a las universidades a crear y reformar planes de estudio para formar especialistas que puedan responder con prontitud a mitigar dicho abismo y lograr insertar exitosamente a sus egresados en un mercado laboral asediado por la incorporación de herramientas basándose en IA. Por otro lado, esta coyuntura especial demanda urgentemente la creación de espacios de reflexión y consulta sobre el papel de la universidad en la conformación del futuro de la educación, la cultura y el trabajo. Se observa un proceso dinámico en la academia en cuanto al desarrollo de estrategias de investigación, capacitación acelerada y actualización para no quedar rezagada ante el embate de tecnologías transformativas como la IA. En términos generales se acepta en estos centros educativos que no hay vuelta atrás para detener la incorporación en el tejido social, educativo, corporativo y gubernamental de herramientas de IA y, por lo cual, se dedica a alistar cuadros y planes para responder a la medida de sus capacidades y recursos a la gran oleada en curso que se avecina con gran intensidad, lo cual es digno de encomio. Sin embargo, se argumenta en esta contribución que no es suficiente el solo formar especialistas con conocimientos tecnológicos alrededor de la IA y tecnologías complementarias, se plantea la necesidad de dar énfasis a seis factores clave interrelacionados:

- Estudio a profundidad de las implicaciones sociotécnicas sobre la adopción de la IA.

- Análisis crítico de las motivaciones económicas y hegemónicas que dan lugar a innovaciones desenfrenadas que demandan urgentemente el desarrollo y despliegue de estrategias regulatorias, procesos éticos y normativos coherentes.
- Estudio de beneficios y de las posibles afectaciones sociales y de comportamiento en los procesos de adopción y apropiación de la IA.
- Desarrollo de programas de extensión y acción social para atender las necesidades de la sociedad sobre capacitación, cultura digital y concientización del papel de la tecnología en las actividades y quehaceres de la población.
- Análisis de fortalezas y debilidades de las universidades respecto a sus recursos humanos, financieros y administrativos para la adopción de la IA y futuros avances. Es decir, se deben especificar nichos que permitan concentrar esfuerzos significativos de investigación y desarrollo tecnológico en IA en beneficio y progreso social. Asimismo, se deben analizar e identificar aplicaciones y sistemas pertinentes y adecuados para promover educación y aprendizaje incluyente y no solo seguir modas y tendencias orientadas a procesos de entretenimiento, consumismo y "marketing" excesivo.
- Formación de grupos de trabajo y talleres para el análisis de contribuciones relevantes a nivel nacional e internacional de documentos e iniciativas sobre la gobernanza de la IA (Regulación, ética, prospectiva del cambio tecnológico e innovación, privacidad y seguridad de la información) y su impacto en la educación.

Esta lista no es exhaustiva, representa propuestas modestas que puedan coadyuvar a fortalecer los esfuerzos académicos para responder a las necesidades que, como puede verse, requieren de un enfoque y acción interdisciplinar y una visión unificada sobre el futuro de la humanidad, su progreso y bienestar. El estudio de las implicaciones sociotécnicas sobre la adopción de la IA proporciona una base de partida en la conformación de un marco de referencia.

La Figura 3 muestra, en una primera aproximación, el resultado de análisis documental y entrevistas a expertos del entorno gubernamental y corporativo sobre las implicaciones e interrelaciones respecto a los fac-

tores de gobernanza mostrados en la Figura 1. Aunque importante, no es suficiente el solo reorientar los planes de estudio e implantar estrategias y acciones de construcción de capacidad y mejora continua para no quedar a la zaga.

Figura 3. Implicaciones sociotécnicas de la IA.

La magnitud e intensidad del nuevo escenario da lugar a repensar el papel de la universidad ante el tsunami transformacional en ciernes. Se recomienda ante tal situación, el considerar los seis factores descritos anteriormente y preparar a las nuevas generaciones para un futuro de la sociedad que demanda una estructura ética y moral sólida y coherente en conjunto a la conformación de procesos regulatorios ad-hoc centrados en los individuos, comunidades e instituciones y sobre todo garantizar derechos, dignidad y honor humanos.

La academia tiene un papel relevante y responsabilidades especiales planteando preguntas pertinentes y desarrollando estrategias y acciones sobre como mitigar la distribución de contenidos dañinos, la desinformación, la dinámica del empleo y otras consecuencias asociadas a la adopción de las aplicaciones de la IA (Ibero, 2023b; Illinois Institute of Technology, 2023; UABC-CIAD, 2023; Universidad de Guadalajara, 2023). La adopción social de las tecnologías digitales emergentes tiene un alcance amplio con impactos tanto beneficiosos como perjudiciales. El identificar tales impactos es crucial antes de desplegar aplicaciones, que por la prontitud que exige el mercado por obtener ganancias financieras a corto

plazo, puedan afectar directa o indirectamente a individuos, comunidades e instituciones. Tal como se menciona anteriormente, la IA es un sistema que concentra enorme poder que surge de décadas de desarrollo científico y tecnológico en diversas disciplinas como ciencias computacionales, neurociencias, lingüística, cognición, comunicación, sociología y otras más.

El crecimiento exponencial de la IA, la "financiarización", el marketing y otras fuerzas asociadas la han convertido en un sistema que no solo pretende transformar el modus-operandi del orden social, sino también tiene el potencial de alterar las bases mismas de comportamiento humano y la dinámica de creación, difusión y distribución de conocimiento. Así mismo, aparte de ofrecer productos que otorgan capacidades avanzadas con gran eficiencia, precisión y agilidad, tiene el potencial de desplazar el rol del ser humano en la toma de decisiones en forma autónoma en sectores de la salud, economía, política, finanzas, educación, etc. Todo ello con enormes consecuencias para el desarrollo integral y avance de la humanidad (Goldstein y Pappert, 1977; Ma y Sun, 2020; Porayska-Pomsta y Rajendran, 2019; Qadir, 2022; McKinsey, 2023).

Marco de referencia propuesto para el contexto académico

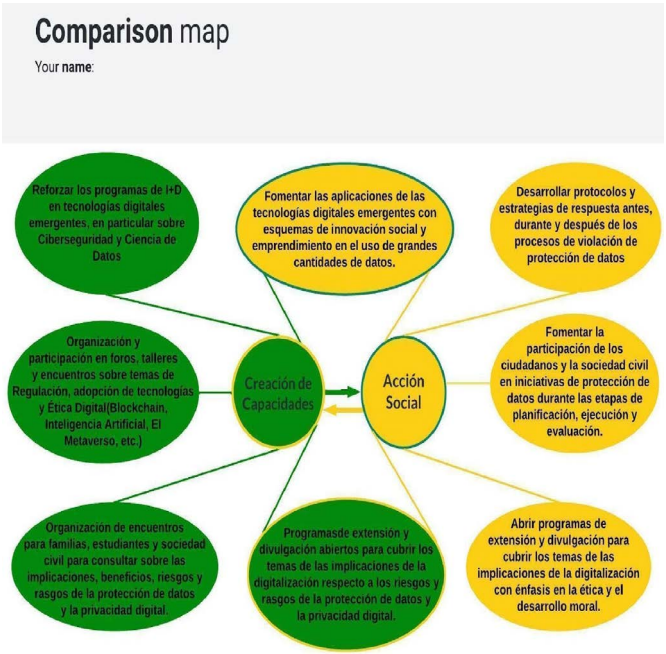
El marco de referencia propuesto considera dos pilares: Creación de Capacidades y la Acción Social. Las claves fundamentales del marco de referencia son:

- Aprovechar la sinergia entre la regulación, la privacidad y seguridad de la información y los aspectos éticos asociados a la adopción social de la IA.
- Extender el conocimiento y capacidades de la academia en cuanto a la IA hacia el sector social (Gobierno, Empresas y Sociedad)
- Realizar acciones que apoyen el entendimiento de las implicaciones, uso y apropiación de aplicaciones de la IA.
- Realizar acciones que coadyuven a mitigar la creciente brecha de innovación en la IA y otras tecnologías emergentes complementarias.
- Contribuir a generar una visión unificada sobre el futuro de la educación ante el creciente y abrumador empuje tecnológico de aplicaciones centradas en IA y tecnologías digitales asociadas.

Tal como se muestra en la Figura 4, se proponen algunas acciones que requieren consulta y reflexión con participación de actores representativos de la dinámica operativa de la academia. En algunos casos, se requerirá implementar cambios significativos en el modus-operandi institucional, en otros casos, algunos ajustes y reorientaciones serán necesarios y dada la magnitud del potencial transformativo de la IA, algunas instancias tendrán que reinventarse para lograr sus objetivos principales.

Para la realización de conversaciones y consulta alrededor de las cinco claves fundamentales consideradas, se propone acudir a un enfoque sistémico y la celebración de conversaciones y foros de consulta incluyentes (AI Safety Summit, 2023). Afortunadamente, muchas de estas acciones se están implementando de manera fluida y expedita en universidades de México y de otros países que intentan reducir la brecha de innovación en IA para insertarse como protagonistas importantes en la formación integral de recursos humanos que demanda la sociedad (IBERO, 2023b; UABC-CIAD, 2023; Universidad de Guadalajara, 2023).

Figura 4. Marco de referencia Propuesto.



Cabe mencionar que, los importantes resultados de la computación de alto desempeño con capacidades de procesar y entrenar sistemas mediante el uso de trillones de datos, estos no deben eclipsar los aspectos de pensamiento crítico y reflexión profunda requeridos en la educación. Las cuestiones complejas involucradas en la formación de las futuras generaciones deben incluir elementos que permitan entender las implicaciones del cambio de paradigma que la sociedad enfrenta actualmente y enfatizar los elementos filosóficos y espirituales que sustentan la naturaleza humana, de tal forma que los fenómenos emergentes sociales y tecnológicos puedan ser vistos desde una perspectiva integral que no sea dictada de "arriba hacia abajo" por los que detentan el dominio y poder de la digitalización.

Conclusiones

Nos encontramos en el umbral de un nuevo ciclo de interacción humano-tecnología sin entender claramente las repercusiones éticas y espirituales de la IA ligadas a su realidad y apropiación social balanceada; y lo que es más importante, sin entender todavía de manera clara y profunda el funcionamiento y las interacciones del cerebro, la mente y la generación del entendimiento humano. Nos encontramos en "proceso de parto" de un acelerado paradigma tecnológico que no se ha experimentado anteriormente en la misma escala e intensidad y que nos lleva a las preguntas: ¿Qué significa ser Humano? ¿Hacia dónde y con quienes acude la sociedad para buscar alternativas de desarrollo tecnológico digital "apropiado" a su contexto y condición sociocultural, económica y de organización política? ¿Hay alternativas de adopción social de la IA más allá de las impulsadas por las empresas hegemónicas digitales y reforzar su patrón civilizatorio sugerido? ¿Se deben aceptar de facto las aplicaciones de la IA para su uso y proliferación masiva en el entorno académico sin efectuar un análisis crítico de sus implicaciones y consecuencias en el aprendizaje?

Se requieren nuevas narrativas y propuestas, más que solo críticas y ansiedad sobre el surgimiento de la IA y su impacto en la sociedad. Dada su complejidad, el desarrollo científico-tecnológico de sistemas de IA, así como los discursos y predicciones sobre la naturaleza, impacto y poder de estos sistemas surge en su mayoría de influyentes "expertos" que per-

tenecen a empresas digitales hegemónicas y de consultoría, por lo que se requiere incrementar las conversaciones y consultas en los diversos sectores de la economía y sociedad. Los esfuerzos de gobiernos, organizaciones internacionales, como las Naciones Unidas y de otras organizaciones de la sociedad civil, incluyendo importantes instancias académicas, buscan incidir en los procesos ético-regulatorios de la IA, y pese a las importantes y cada vez más frecuentes participaciones, el "gran peso" de visualizar y determinar su futuro ha estado en manos de los grandes corporativos digitales, ya que como fue reiterado anteriormente, su prioridad es de corte económico-financiero, no necesariamente de progreso social y comunitario. He aquí la urgencia y reto de la academia para apoyar a la sociedad en la coyuntura de la década actual hacia el 2030.

Referencias

Acemoglu, D. y Restrepo, P. (2018). *Artificial Intelligence, Automation and Work. MIT Department of Economics* Working Paper No. 18-01SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3098384

Adams, T. (2021). Daniel Kahneman: 'Clearly AI is going to win. How people are going to adjust is a fascinating problem'. *The Guardian*. https://www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow

Agrawal, A., Gans, J. y Goldfarb, A. (2022). ChatGPT and How AI Disrupts Industries. *Harvard Business Review*. December 12, 1-6. Retrieved from https://hbr.org/2022/12/chatgpt-and-how-ai-disrupts-industries

AI Ethics Initiative. (2017). https://aiethicsinitiative.org

AI Now Institute. (2017). https://ainowinstitute.org

AI Safety Summit. (2023). The Bletchley Declaration, https://www.gov.uk/government/topical-events/ai-safety-summit-2023

Aiken, R. M. y Epstein, R. G. (2000). Ethical guidelines for AI in education: Starting a conversation. *International Journal of Artificial Intelligence in Education,* 11, 163-176.

Algorithmwatch.org. (2020). *AI Ethics Guidelines Global Inventory*. Retrieved from https://algorithmwatch.org/en/ai-ethics-guidelines-global-inventory/.

Alshater, M. (2022). *Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT*. Avaiable at SSRN. https://ssrn.com/abstract=4312358? or http://dx.doi.org/10.2139/ssrn.4312358.

Altman, S. et al, (2023). Governance of Superintelligence, *OpenAI*.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management,* 62, Article 102433.

Bahá'í International Community. (2022). *Employment and Beyond: Drawing on the Capacity of All to Contribute to Society, New York*. https://www.bic.org/statements/employment-and-beyond-drawing-capacities-all-contribute-society

BCG. (2023). *A Generative Roadmap for Financial Institutions* https://www.bcg.com/publications/2023/a-genai-roadmap-for-fis

Bloomberg (2023) *EU Strikes Deal to Regulate ChatGPT, AI Tech in Landmark Act*, https://www.bloomberg.com/news/articles/2023-12-08/eu-strikes-deal-to-regulate-chatgpt-other-ai-in-landmark-act

Boddington, P. (2017). *Towards a code of ethics for artificial intelligence research.* Berlin: Springer.

Caruso, L. (2018). *Digital innovation and the fourth industrial revolution: epochal social changes?* AI & SOCIETY, 33(3), 379–392. https://doi.org/10.1007/s00146-017-0736-1

Centro México Digital. (2023). *¿Qué es la Transformación Digital y cuál es su importancia?* https://centromexico.digital/que-es-la-transformacion-digital/

Cockburn, I.M., Henderson, R., & Stern, S. (2019). The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis. En A. K. Agrawal, J. Gans and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence*.

Coombs, C., Stacey, P., Kawalek, P., Simeonova, B., Becker, J., Bergener, K. y Trautmann, H. (2021). What is it about humanity that we can't give away to intelligent machines? A European perspective. *International Journal of Information Management,* 58, Article 102311.

DeepMind Ethics & Society. (2017). https://deepmind.com/about/ethics-and-society

Deshpande, K. V., Pan, S., & Foulds, J. R. (2020). *Mitigating demographic Bias in AI-based resume filtering.* Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization, 268–275. https://doi.org/10.1145/3386392.3399569.

Drachsler, H., & Greller, W. (2016). *Privacy and analytics: It's a DELICATE issue a checklist for trusted learning analytics.* Proceedings of the sixth international conference on Learning Analytics & Knowledge, 89–98. http://dl.acm.org/citation.cfm?id=2883893

Dwivedi, Yogesh, K. et al. (2023). So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management,* 71, 102642.

Floridi, L. (2019). Translating principles into practices of digital Ethics: Five risks of being unethical. *Philosophy & Technology*, 32(2), 185–193. https://doi.org/10.1007/s13347-019-00354-x.

Floridi, L. et al. (2018). *AI4People's Ethical Framework for a Good AI Society:Opportunities, Risks, Principles, and Recommendations*. https://www.eismd.eu/wp-content/uploads/2019/03/AI4People's-Ethical-Framework-for-a-Good-AI-Society.pdf.

Floridi, L., & Strait, A. (2020). *Ethical Foresight Analysis: What it is and Why it is Needed?* Minds & Machines. https://doi.org/10.1007/s11023-020-09521-y

Frey, C. B., & 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

García, R. (2006). *Sistemas complejos. Conceptos, método y fundamentación epistemológica de la investigación interdisciplinaria*. Filosofía de La Ciencia. Gedisa Editorial.

Gartner. (2021). *Digital Ethics by Design: A Framework for better Digital Businesses* https://www.gartner.com/en/documents/3953794/digital-ethics-by-design-a-framework-for-better-digital-

Goldstein, I., & Papert, S. (1977). Artificial intelligence, language, and the study of knowledge. *Cognitive Science,* 1(1), 84-123.

https://openaicom/blog/governance-of-superintelligence

Ibero. (2023a). *¿Cómo surge el Internet?, Aquí una breve historia*. https://ibero.mx/prensa/como-surge-el-internet-aqui-una-breve-historia

Ibero. (2023b). *Lineamientos para el uso de la Inteligencia Artificial*. https://ibero.mx/prensa/da-conocer-ibero-lineamientos-para-el-uso-de-inteligencia-artificial

Illinois Institute of Technology. (2023). *Faculty Guide on Teaching and Generative AI*. https://www.iit.edu/sites/default/files/2023-08/Faculty%20Guide%20on%20Teaching%20and%20Generative%20AI.pdf

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

Kshetri, N. (2023). "Regulatory technology and supervisory technology: Current status, facilitators and barriers". IEEE Computer, 56(1), 64-75.

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing–Connecting computing power to human insights. *International Journal of Research in Marketing,* 37(3), 481–504.

Mariani, M. M., Machado, I. y Nambisan, S. (2023). Types of innovation and artificial intelligence: A systematic quantitative literature review and research agenda. *Journal of Business Researc*h, 155, Article 113364.

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, 160(4), 835–850.

McCarthy, J. (2007). *What is artificial intelligence*. http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html Accessed 10.03.23.

McKinsey & Company. (2023), W*hat is generative AI?* https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai#/

McKinsey. (2023). *As gen AI advances, regulators-and risk functions-rush to keep pace.* https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/as-gen-ai-advances-regulators-and-risk-functions-rush-to-keep-pace

Moor, J. H. (2008). Why we need better ethics for emerging technologies. In J. V. D. Hoven, & J. Weckert (Eds.), *Information Technology and Moral Philosophy* (pp. 26–39). Cambridge: Cambridge University Press.

National Institute of Standards and Technology. (26 de enero de 2023). NIST AI Risk Management Framework Playbook. *NIST*. https://pages.nist.gov/AIRMF/.

Oracle. (2024). *¿Qué es la Transformación Digital?* https://www.oracle.com/mx/cloud/digital-transformation/

Peters, R. S. (1970). *Ethics and education*. London: Allen & Unwin.

Porayska-Pomsta, K. y Rajendran, G. (2019). *Accountability in human and artificial intelligence decision-making as the basis for diversity and educational inclusion*. In J. Knox, Y. Wang, & M.

Qadir, J. (2022). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. *TechRxiv.* https://doi.org/10.36227/techrxiv.21789434.v1.

SAP. (2024). *¿Qué es la transformación Digital?* https://www.sap.com/latinamerica/insights/what-is-digital-transformation.html

Shneiderman, B. (2020). *Human-Centered Artificial Intelligence: Three Fresh Ideas*. AIS Transactions on Human-Computer Interaction, 12(3), 109–124.

Taeihagh. (2021). Governance of artificial intelligence. *Policy and Society,* 40(2), 137-157, DOI: 10.1080/14494035.2021.1928377

TechTarget. (2023). Generative AI vs. Predictive AI: Understanding the differences. https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-vs-predictive-AI-Understanding-the-differences

The Institute for Ethical AI & Machine Learning. (2018). https://ethical.institute

Turing, A. (1950). *Computing machinery and intelligence*. Mind, 59, 433–460.

UABC-CIAD. (2023). *Orientaciones iniciales sobre el uso académico de la Inteligencia Artificial (IA)*.

UN Advisory Body on AI (2023). *Interim Report: Governing AI for Humanity*. https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf

UNESCO (2023b). *UNESCO: Los gobiernos deben regular rápidamente la inteligencia artificial generativa en las escuelas*. UNESCO. https://www.unesco.org/es/articles/unesco-los-gobiernos-deben-regular-rapidamente-la-inteligencia-artificial-generativa-en-las-escuelas

UNESCO. (2022). *Recommendations on the Ethics of Artificial Intelligence* https://en.unesco.org/artificial-intelligence/ethics

UNESCO. (2023a). *Guidance for generative AI in education and research. UNESCO*. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Universidad de Guadalajara. (2023). *Orientaciones y definiciones sobre el uso de la inteligencia artificial generativa en los procesos académicos*. https://www.udgvirtual.udg.mx/Guia-IA-UDG

Université de Montréal. (2018). *Declaration of Montréal for a responsible development of AI. Université de Montréal.* https://www.montrealdeclaration-responsibleai.com

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., MyersWest, S., Richardson, R., Schultz, J. y Schwartz, O. (2018). *AI now report 2018. AI Now Institute: New York University*.

Capítulo 5

Fuck the algorithm: Navegando la promesa tecnológica y el impacto social de la IA

Marisol Flores-Garrido

https://doi.org/10.61728/AE24001052

La inteligencia artificial (IA) ha irrumpido con fuerza en tareas tan numerosas como diversas, suscitando una fascinación que puede oscurecer sus límites reales. Desde revolucionar el diagnóstico médico hasta simplificar nuestras tareas diarias, la IA no solo ha enriquecido nuestra vida cotidiana, también ha desatado optimismo y confianza en sus posibilidades. En ese sentido, podemos considerar a la IA como una tecnología carismática, que deriva gran parte de su poder a través de la posibilidad o promesa de acción: importa no solamente lo que sus herramientas son, sino la forma en que afectan nuestra imaginación con las promesas de lo que podrían hacer (Ames, 2019).

Los logros de la IA pueden conducirnos a su adopción precipitada en contextos donde su utilidad no es tan clara, persiguiendo utopías tecnológicas que ignoran las limitaciones de estas herramientas. Ante esta situación, es necesario cultivar una postura crítica tanto en los desarrolladores de sistemas de IA como en la sociedad en su conjunto, quien debe defender el derecho a decidir cómo y cuándo implementar estas tecnologías.

Por esta razón, se vuelve crucial reimaginar colectivamente usos y limitaciones de la IA. Nuestras expectativas y visiones sobre esta tecnología no solo guían la investigación y la inversión en proyectos, también validan sus aplicaciones y moldean nuestra actitud colectiva hacia ellas. Ajustar nuestras expectativas con una actitud crítica que distinga entre mitos y realidades puede complicarse por el hecho de que la IA existía en nuestro imaginario colectivo mucho antes de su formalización como campo de estudio. En consecuencia, muchos sistemas de IA surgen en un espacio ya habitado por numerosos mitos sobre sus capacidades. De alguna manera, las narrativas de autómatas e instrumentos exentos de debilidades humanas, que prometen superarnos apoyados en el poder de los cálculos numéricos y la razón absoluta, preceden a las herramientas mismas. Esto, combinado con estrategias desbordadas de mercadotecnia, puede dificultar la apreciación de los límites de la IA y de sus dimensiones políticas, sociales e, incluso, materiales.

Dada la promesa de la IA de insertarse en todos los dominios del conocimiento y, potencialmente, en cada aspecto de nuestra existencia, se vuelve imperativo fomentar una reflexión crítica sobre sus alcances. En este capítulo, propongo tomar como caso de estudio el episodio de un algoritmo utilizado en el Reino Unido en 2020 para ajustar calificaciones

estimadas de los exámenes *A-levels* (de Advanced Level qualifications). A través de un análisis teórico que recupera las nociones de sesgo algorítmico, complejidad y resistencia, se utilizan estos ejes para reflexionar sobre la necesidad de construir una mirada crítica hacia los sistemas de IA y de reconfigurar el discurso en torno a ellos para hacer evidentes sus posibilidades, sus limitaciones y las estructuras de poder que los hacen posibles.

El algoritmo de la Ofqual

En el año 2020, la pandemia de COVID-19 y las medidas de confinamiento impuestas por el gobierno británico enfrentaron a los estudiantes de bachillerato con una situación sin precedentes: la imposibilidad de realizar los exámenes A-levels, cruciales para el acceso a la educación superior. Esta situación obligó al sistema educativo a crear una estrategia que permitiera terminar el ciclo académico y evitara cualquier discontinuidad en el acceso a la educación universitaria.

Ante la urgencia de una solución, la Oficina de Regulación de Calificaciones y Exámenes (Office of Qualifications and Examinations Regulation u Ofqual, por sus siglas) propuso una medida extraordinaria: que los docentes realizaran una estimación de las calificaciones de cada estudiante, denominada "calificación evaluada por el centro" o CAG por su siglas en inglés, y establecieran un ranking entre quienes tenían estimaciones similares. Posteriormente, estas calificaciones proyectadas alimentarían un algoritmo diseñado para ajustar las notas basándose en el rendimiento histórico de cada centro educativo.

Específicamente, el algoritmo de la *Ofqual* estimaba las calificaciones de los exámenes considerando tres factores: (1) la distribución histórica de las calificaciones de los centros de los tres años anteriores (2017-2019); (2) la clasificación de cada estudiante dentro de su propio centro en una asignatura concreta, basada en las CAG; y (3) los resultados de exámenes anteriores de cada estudiante por asignatura. Con toda esta información, el algoritmo realizaba un análisis de la distribución de las notas a lo largo de los años, asignando calificaciones en función de la posición relativa de cada estudiante dentro de su centro. Por ejemplo, si un alumno se encontraba en la mitad inferior de la lista de clasificación, su nota sería aproxi-

madamente igual a la obtenida en años anteriores por otro alumno de la misma categoría. La intención era replicar, de manera artificial, la distribución de calificaciones de años anteriores, manteniendo consistencia con el desempeño histórico de las instituciones.

Preservar una cierta "normalidad" estadística era muy importante para la Ofqual, que partió de la premisa de que utilizar datos históricos para ajustar las calificaciones sería más preciso y más justo que confiar únicamente en las evaluaciones estimadas por docentes. Existía la percepción de que los profesores, por su cercanía y compromiso con el alumnado, podrían tender a una generosidad excesiva en sus calificaciones y conducir a una inflación de las mismas. Esto implicaría un aumento considerable en el número de estudiantes con calificaciones destacadas y, a su vez, desbordaría la capacidad de admisión de las universidades. En este contexto, el algoritmo se concibió como un mecanismo correctivo frente a lo que se consideraba un sesgo implícito en las CAG, buscando equilibrar un proceso esencialmente humanizado y, por ende, sujeto a las complejidades de la subjetividad docente.

No obstante, la táctica adoptada por la *Ofqual* rápidamente mostró fallas, en particular en el efecto desproporcionado que el algoritmo ejerció sobre los alumnos de escuelas públicas en relación con sus pares de instituciones privadas. La confianza en este sistema automatizado, fundamentada en la suposición de una posible sobrevaloración de las calificaciones por parte del profesorado, condujo a un resultado inesperado y perjudicial: una depreciación general de las notas finales en comparación con las CAG, afectando aproximadamente al 40 % de los estudiantes y perjudicando en mayor proporción a quienes provenían de una institución pública. El ajuste no solo truncó las posibilidades de acceso universitario de numerosos jóvenes, también intensificó las brechas ya existentes dentro del sistema educativo dejando en evidencia la fragilidad de un enfoque que, en su intento de buscar objetividad, terminó por profundizar las inequidades.

La determinación de calificaciones por parte del algoritmo desató una ola de manifestaciones en distintas regiones del país, que se extendieron durante varias semanas. Ante la creciente presión social, el 17 de agosto de 2020 el gobierno tomó la decisión de restablecer las calificaciones conforme a las estimaciones originales realizadas por los docentes, desestimando

los ajustes del algoritmo con la única excepción de aquellos casos en los que la nota asignada por el sistema automatizado superaba a la CAG. Esta medida representó un reconocimiento de las limitaciones y el impacto negativo de la dependencia en soluciones algorítmicas para asuntos trascendentes, como la educación, subrayando la necesidad de priorizar el juicio humano y de reconocer las limitaciones de predicciones basadas en datos.

Examinar este caso puede arrojar luz sobre aspectos fundamentales que deben tomarse en cuenta para desarrollar una perspectiva crítica ante la toma de decisiones mediada por algoritmos. En un contexto donde el uso de los sistemas de IA se expande rodeado de un discurso de exactitud, objetividad y capacidad para superar los sesgos y las imperfecciones vinculadas a la toma de decisiones humana, es imperativo cuestionar y examinar estas promesas, especialmente en escenarios marcados por gran complejidad.

Perspectiva crítica frente al poder del algoritmo

a. Sesgo algorítmico

El sesgo en algoritmos ha sido objeto de atención en la comunidad de la IA en los últimos años. En una investigación emblemática, Buolamwini y Gebru (2018) evidenciaron cómo la tecnología de reconocimiento facial exhibía variaciones en su precisión según subgrupos fenotípicos, revelando un rendimiento deficiente particularmente en mujeres de raza negra. El fenómeno se vinculó a las deficiencias presentes en los conjuntos de datos utilizados para entrenar estos sistemas de visión. A partir de este estudio se han identificado y analizado múltiples instancias de sesgo algorítmico, abarcando desde sistemas de recomendación que privilegian a los hombres en la selección laboral hasta herramientas de diagnóstico médico asistido por computadora que demuestran precisión variable según la etnicidad del paciente.

Las discusiones contemporáneas en torno al sesgo algorítmico lo conceptualizan como decisiones de un sistema de IA influenciadas por el uso de información que debería ser considerada irrelevante o, inversamente,

por la omisión de datos pertinentes (Turner, 2019). En el caso del algoritmo de la *Ofqual*, la indignación y rechazo suscitados estuvieron fuertemente ligados a la percepción de sesgo e injusticia en sus resultados. La dependencia de datos históricos relacionados con el desempeño previo de las escuelas condujo a situaciones en las que, por ejemplo, si ningún estudiante de una determinada escuela había alcanzado la calificación más alta en años anteriores, se tornaba casi imposible que alguien de ese mismo centro educativo lo lograra en el año en curso. Al estar diseñado para replicar patrones pasados, el algoritmo configuraba un escenario limitante para ciertos grupos, cerrando de antemano posibilidades y perpetuando desigualdades históricas.

Además, el algoritmo asignaba mayor relevancia a las CAG provenientes de grupos con menos de 15 estudiantes en una asignatura específica, dentro de un establecimiento educativo determinado. Esta característica del diseño presupone que las evaluaciones realizadas por los docentes en contextos de menor cantidad de estudiantes son más fiables, posiblemente debido a una relación más estrecha y un conocimiento más profundo sobre los alumnos. Esta premisa introdujo un sesgo importante: estudiantes de instituciones más pequeñas se vieron desproporcionadamente favorecidos por la inflación de calificaciones estimadas en comparación con aquellos de centros más grandes. Esta situación contribuyó a reforzar desigualdades preexistentes, pues las escuelas públicas suelen tener un mayor número de estudiantes por aula. Como resultado, los análisis revelaron que la proporción de calificaciones altas (A* y A) en los colegios privados incrementó con el algoritmo en 4.7 puntos porcentuales, más del doble que el aumento registrado en las escuelas públicas de enseñanza general (Porter, 2020).

El algoritmo de la *Ofqual* ilustra cómo la automatización de decisiones frecuentemente mantiene el status quo, replicando desigualdades sociales preexistentes. Como se ha evidenciado en investigaciones previas, las fallas en los algoritmos, incluyendo sesgos, tienden a afectar desproporcionadamente a aquellos grupos ya en situación de vulnerabilidad, exacerbando su riesgo de sufrir consecuencias adversas (Eubanks, 2018). Este fenómeno puede atribuirse, en parte, a que los equipos encargados del diseño algorítmico suelen ser homogéneos y carecen de la formación y la diversidad

necesaria para anticipar cómo ciertos aspectos de su diseño pueden impactar negativamente a distintos grupos sociales. Esta homogeneidad deriva en lo que se conoce como ceguera de privilegio: una incapacidad para reconocer y comprender las dificultades y desventajas que enfrentan otros colectivos debido a su posición económica, racial, de género, entre otras, en contraste con aquellos que detentan posiciones de privilegio (D'ignazio y Klein, 2020).

El algoritmo de la *Ofqual* se suma a la lista de sistemas algorítmicos que han sido señalados por presentar sesgos. La proliferación de este tipo de errores ha fomentado, dentro del campo de las ciencias computacionales, un área de investigación enfocada en desarrollar estrategias para "corregirlos". Esto se hace, principalmente, a partir de la implementación de criterios de "equidad" o "fairness", fundamentados en principios de probabilidad y estadística (Wang et al., 2022). Dichos criterios buscan asegurar que los resultados de un sistema no estén influenciados por atributos "irrelevantes" que podrían conducir a decisiones injustas, tales como el género, la etnicidad o la condición socioeconómica, a través de la definición de objetivos específicos que orientan el espacio de soluciones del modelo hacia resultados que se consideran imparciales.

Lamentablemente, los criterios de equidad algorítmica presentan limitaciones significativas. Determinar cuál de las decenas de propuestas es la más adecuada para un caso particular requiere un análisis exhaustivo y meticuloso. Además, existe un límite obvio en la capacidad de los criterios técnicos para definir y medir la justicia, especialmente cuando esto pasa por la corrección de sesgos en datos históricos y la variabilidad de normas culturales y sociales que definen lo justo en contextos específicos.

El sesgo algorítmico es un asunto de gran relevancia y su abordaje es esencial dentro de la comunidad que desarrolla sistemas de IA. Sin embargo, es igualmente crucial enriquecer el análisis de la toma de decisiones automatizada con una perspectiva más amplia, que reconozca que, a menudo, el sesgo es solo el reflejo de problemas más arraigados y complejos, vinculados con las dimensiones políticas, culturales, históricas y sociales que se entretejen en los datos, los algoritmos y la tecnología en su conjunto.

b. Complejidad

Además del sesgo evidente en el algoritmo de la *Ofqual*, se destaca un problema estructural más profundo en su diseño: la noción de que es posible prever, con base en una acumulación de datos, los resultados que personas alcanzarán en un examen aún no realizado. Tal como fue concebido, el algoritmo no solo ignoraba, sino que efectivamente cerraba las puertas a estudiantes talentosos de escuelas con un bajo historial académico, asumiendo que el esfuerzo individual era improbable, o no significativo, y podría ser subsumido en las estadísticas. Este caso señala una cuestión crucial que merece ser abordada en el debate actual: en el auge de la IA están surgiendo sistemas que pretenden predecir el comportamiento humano. Estos sistemas, más allá de sus posibles fallas de diseño, parten de una concepción imaginativa problemática, asumiendo que la complejidad humana puede ser reducida a modelos predictivos basados en datos.

Dicha perspectiva resuena con la visión cartesiana que anhela certeza, orden y previsibilidad en la comprensión del mundo, una aproximación que presupone la posibilidad de definir con claridad las conductas y decisiones humanas, obviando la ambigüedad, continuidad y fluidez que definen nuestra existencia (Birhane, 2021). Contrario a la noción de seres estáticos, somos entidades dinámicas, en constante evolución a través de nuestras interacciones y el entorno que nos rodea. Así, nos encontramos perpetuamente en formación, en un proceso de transformación que desborda las capacidades predictivas de cualquier algoritmo. En este contexto, la estrategia de la *Ofqual*, y de cualquier sistema similar en objetivo, choca con la realidad de nuestra indeterminabilidad.

Subestimar la impredictibilidad humana no solo revela una comprensión limitada sobre el alcance y las restricciones de la IA, también representa un riesgo importante. Al transformarse de herramientas que meramente describen patrones en los datos a herramientas que prescriben y moldean la realidad, estos sistemas inciden directamente en la conformación de nuestro entorno. Legitimar los resultados del algoritmo de *Ofqual* contribuye, de facto, a perpetuar un futuro anclado en desigualdades históricas, automatizando una realidad que mantiene la desventaja para los menos privilegiados.

Un diseño responsable de sistemas de IA debe reconocer que la tecnología configura no solo herramientas, sino prácticas y posibilidades futuras. Al asignar calificaciones superiores a estudiantes de escuelas privadas, se está configurando un futuro donde estas personas obtienen un acceso privilegiado a la educación superior, reduciendo simultáneamente las oportunidades para aquellos en situaciones iniciales menos favorecidas. Así, los sistemas de IA no solo determinan trayectorias; también orientan hacia determinados futuros y cierran las puertas a otros.

Clasificar y predecir en este tipo de modelos, más que un mero asunto de funciones numéricas o matemáticas, tiene un impacto directo en la construcción del orden social. Este orden, fundamentado en datos históricos, tiende a perpetuar y reforzar prácticas y normativas del pasado. En consecuencia, los modelos de IA –con sus tareas de clasificación, ranking, agrupamiento, predicción– incorporan una dimensión política y moral innegable, que debe ser reconocida y abordada con plena responsabilidad.

c. Resistencia

Cuestionar la justicia en el funcionamiento de ciertos sistemas de IA y poner en tela de juicio su autoridad representa un desafío nada trivial, marcado por al menos tres aspectos relevantes en el contexto contemporáneo.

En primer lugar, el análisis de los sistemas de IA nos confronta con diversas formas de opacidad. Según plantea Jenna Burrell (2016), esta opacidad puede manifestarse de varias maneras: como una estrategia intencionada por parte de corporaciones o entidades estatales propietarias de los sistemas; como resultado de la complejidad técnica que exige un alto grado de especialización para comprender el código y el funcionamiento del sistema; o como la dificultad asociada a la traducción de decisiones algorítmicas, optimizadas en espacios de alta dimensionalidad, a un marco de razonamiento humano que facilite su interpretación. Esta opacidad en la toma de decisiones mediada por IA obstaculiza la identificación de errores y la atribución de responsabilidades a las decisiones humanas subyacentes.

En segundo lugar, los sistemas de IA se enmarcan dentro de discursos que los exaltan por su precisión, exactitud y su fundamento en verdades respaldadas por datos y matemáticas avanzadas. Frecuentemente, se les

presenta como soluciones libres de los sesgos que ensombrecen la toma de decisiones humanas. Desafiar las conclusiones de un algoritmo implica, para empezar, cuestionar estos discursos y examinar críticamente tanto los mecanismos operativos del sistema como las dinámicas de poder que influyen en su desarrollo y funcionamiento.

Este proceso de indagación también está ligado a una comprensión detallada de los principios fundamentales del sistema, que permita identificar sus limitaciones, vulnerabilidades y riesgos. Por ejemplo, las métricas de "exactitud" comúnmente usadas para describir modelos de IA deben ser entendidas en su contexto disciplinario específico, como indicadores de la correlación entre los resultados de un algoritmo y sus datos de entrenamiento, y no como una medida de su capacidad para predecir el futuro o reflejar fielmente la realidad. Estos conceptos, más allá de su significado técnico, pueden ser empleados retóricamente para influir en la percepción pública de la tecnología. De manera similar, comprender en profundidad términos como "predicción", "aprendizaje", "equidad" y "entrenamiento" en el ámbito de la IA es clave para desmantelar marcos discursivos y formular cuestionamientos críticos que permitan una evaluación rigurosa de las herramientas.

En tercer lugar, los resultados generados por sistemas de IA pueden ser percibidos de manera fragmentada por los usuarios finales y esto complica la identificación de tendencias generales, el entendimiento del comportamiento global del sistema y el reconocimiento de patrones que puedan sugerir injusticias. En el contexto del algoritmo de la *Ofqual*, la interacción y comunicación entre estudiantes, especialmente aquellos que ya habían establecido conexiones previas a la pandemia, fue crucial para detectar un sesgo vinculado directamente al origen escolar.

Además, la controversia en torno al uso del algoritmo se convirtió en un tema central en el debate público del Reino Unido. En este proceso, fue fundamental la manera en que se configuró el imaginario colectivo respecto al algoritmo, conforme la ciudadanía intentaba comprender su funcionamiento y compartía sus hallazgos. La visibilización de los efectos del modelo empleado por la *Ofqual* emergió de un esfuerzo colectivo; numerosas personas se involucraron en la problemática y cuestionaron la autoridad del algoritmo, no necesariamente por haber sido afectadas

directamente por los exámenes, sino en solidaridad a las protestas y en respuesta a la atención mediática que estas generaron. Así, el diálogo en redes sociales –que en Twitter popularizó el uso del hashtag *#fuckTheAlgorithm*– y otros espacios digitales catalizó la organización de manifestaciones callejeras, fomentando la emergencia de un sentimiento colectivo de resistencia frente a la autoridad del algoritmo y de rechazo a una herramienta percibida como injusta. Las movilizaciones sociales, en última instancia, llevaron al gobierno a ofrecer a los estudiantes la opción de retener sus CAG, aceptar la nota modificada por el algoritmo, o presentar los exámenes en una fecha posterior, demostrando el poder de la acción colectiva en la redefinición de políticas y prácticas tecnológicas.

Este caso ilustra posibles respuestas a los desafíos planteados al comienzo de esta sección

Primero, ante la opacidad de los sistemas de IA, se hace imperativo abogar por la transparencia y la explicabilidad, especialmente en aquellos sistemas que tienen el potencial de impactar significativamente en la vida de las personas. Para esto es fundamental reconocer que no todos los sistemas de IA tienen el mismo propósito o naturaleza. Siguiendo la clasificación de Narayanan (2019), los sistemas pueden diferenciarse en funciones de percepción (como el reconocimiento de imágenes), automatización de criterios (como la clasificación de correos electrónicos) y toma de decisiones que afectan directamente a las personas. La capacidad de explicar las decisiones del sistema, junto con la transparencia en sus mecanismos, datos y construcción, y las estrategias para su auditoría, deben estar alineadas con su propósito y alcance.

En el caso del algoritmo de la *Ofqual*, cuando se conocieron las noticias sobre las calificaciones asignadas no existía un procedimiento de apelación claro. De hecho, el procedimiento era muy complicado y los estudiantes tenían que pagar para apelar sus calificaciones. Al igual que el propio modelo, las deficiencias del procedimiento de apelación podían afectar de forma desproporcionada a los estudiantes de entornos socioeconómicos más bajos.

En situaciones donde los algoritmos tienen la capacidad de influir en vidas humanas, es inadmisible que las decisiones se tomen sin un análisis

riguroso, que incorpore múltiples perspectivas y una descripción detallada de todas las partes del proceso y que facilite así una auditoría por especialistas y un escrutinio abierto. Conscientes de que ningún algoritmo es infalible, la *Ofqual* debería haber anticipado la posibilidad de resultados injustos para algunos estudiantes e implementado un mecanismo adecuado para realizar aclaraciones y apelaciones. La gestión de esta situación evidencia que no solo hubo problemas metodológicos, como el que originó sesgo algorítmico, sino también (especialmente) epistémicos. Depositar total confianza en un algoritmo demuestra un malentendido sobre las capacidades de estos sistemas en contextos complejos y variables.

En segundo lugar, es clave desmantelar el mito de la perfección que a menudo rodea a la IA, promovido en gran medida por estrategias de mercadotecnia. Se necesita reconocer que estos modelos no erradican la incertidumbre en la toma de decisiones, solo la transforman; las nuevas formas de incertidumbre pueden estar ligadas a distintos factores del proceso y demandan un análisis cuidadoso. A medida que aumenta la complejidad de la situación que se intenta modelar, crece la certeza de que cualquier modelo propuesto tendrá errores, pues capturar la plenitud de la complejidad real es una tarea que escapa a las posibilidades de cualquier modelo. Esto no resta valor a los modelos, pero sí subraya que sus aplicaciones estarán siempre condicionadas por las suposiciones y simplificaciones adoptadas durante su desarrollo, restringiendo su utilidad a escenarios específicos. La sociedad tiene el derecho a estar informada sobre los posibles errores de un algoritmo y sobre la forma en que estos podrían manifestarse. Más aún, la sociedad debe retener el derecho a decidir cuándo la incertidumbre de las decisiones humanas es preferible a una injusticia automatizada envuelta en un aura de infalibilidad.

En este sentido, es interesante observar que el algoritmo de la *Ofqual* fue ideado con el objetivo de mitigar el sesgo asociado a las evaluaciones subjetivas de los docentes, un sesgo que, de hecho, era real y tuvo consecuencias. Al anularse la implementación del algoritmo y retornar a las CAG sí se produjo un incremento en las notas, lo que sí generó desafíos logísticos para las universidades al tener que acomodar a un número mayor de estudiantes admitidos de acuerdo con altas calificaciones en los A-levels. Un ejemplo de esto fue la Universidad de Durham, que se vio

forzada a ofrecer incentivos para que los estudiantes pospusieran su ingreso hasta el 2021.

El problema que el algoritmo buscaba solucionar era genuino, pero su aplicación no logró crear un entorno libre de sesgos. En realidad, el algoritmo introdujo un nuevo tipo de sesgo, que en muchos casos afectó especialmente a estudiantes de estratos socioeconómicos más bajos. El caso ilustra la complejidad asociada a la búsqueda de soluciones algorítmicas para problemas sociales: las simplificaciones inevitables en el modelo y el diseño del algoritmo no erradicaron el sesgo: lo transformaron, creando desigualdades alternativas. Más aún, lo hicieron partiendo de una fantasía, pues los algoritmos no pueden, ni este ni otros escenarios, predecir de manera certera el futuro de las personas. La idea de "predicción" en el aprendizaje automático hace referencia a cálculos matemáticos y difiere significativamente de su interpretación en el imaginario popular. Esta discrepancia puede llevar a malentendidos sobre las capacidades reales de la tecnología algorítmica.

Finalmente, es fundamental enfatizar la importancia de ejercer nuestros derechos a la explicabilidad, a la rendición de cuentas por parte de quienes diseñan y operan los sistemas de IA, y a tomar decisiones informadas sobre la pertinencia de implementar estas tecnologías en contextos específicos. Esto requiere fomentar estructuras de organización colectiva que nos permitan reflexionar sobre la influencia de los sistemas de IA en nuestra cotidianidad y desarrollar estrategias de acción para influir en el proceso de adopción, adaptación o rechazo de estas herramientas tecnológicas. Solo así podremos asegurarnos de que la evolución y aplicación de la IA se alinee con los valores y objetivos compartidos por nuestra sociedad, y no solamente con los de un grupo reducido que se encuentra en el poder.

Conclusiones

Las expectativas en torno a la IA desempeñan un papel central en la configuración de los sistemas y en sus aplicaciones prácticas, actuando como intermediarias entre los diversos niveles de implementación de esta tecnología y las comunidades. El exceso de entusiasmo por las capacidades de la IA ha generado la difusión de promesas que superan ampliamente

las capacidades reales de la tecnología. En ocasiones, estas expectativas desmedidas conducen a problemas epistémicos, como en el caso de algoritmos propuestos para identificar delincuentes basados en rasgos faciales (Wu et al., 2016), o como se ha discutido en este capítulo, para predecir calificaciones justas en exámenes que estudiantes no han presentado.

Frente a todas las mitologías sobre las capacidades de los sistemas de IA, es fundamental volver al cuestionamiento crítico de estas herramientas y de su aplicación en contextos específicos. Los sistemas de IA, especialmente aquellos basados en aprendizaje automático, poseen una capacidad extraordinaria para identificar patrones en los datos. Al aproximarnos de manera adecuada a estos patrones, podemos abrir la puerta a la reflexión sobre dinámicas históricas, examinar con mayor detenimiento las relaciones entre variables o plantear nuevas hipótesis que nos ayuden a comprender mejor el mundo que nos rodea. Sin embargo, fracasar en reconocer que simplemente estamos identificando patrones puede llevarnos a terrenos confusos, como la creencia errónea de que es posible predecir el comportamiento humano.

El caso del algoritmo de la *Ofqual* subraya la necesidad de lineamientos claros para la supervisión de sistemas de IA en todas sus fases: diseño, entrenamiento y análisis de resultados. Esto implica promover la transparencia, la explicabilidad y la responsabilidad, mitigando riesgos, especialmente para los sectores más vulnerables.

Además, es esencial que busquemos influir en la percepción de la sociedad sobre el funcionamiento de los sistemas de IA. Al lograr un mejor entendimiento de estas herramientas, podremos trascender los marcos discursivos predominantes y cuestionar adecuadamente sus usos e implicaciones en diversos escenarios. De esta manera, podremos desafiar la autoridad de estos sistemas, rechazar su uso cuando sea necesario, y reimaginar sus posibilidades, alejándonos del utopismo tecnológico en favor de aplicaciones realistas, equitativas y beneficiosas para la colectividad.

Referencias

Ames, M. G. (2019). *The charisma machine: The life, death, and legacy of one laptop per child*. Mit Press.

A-levels and GCSEs: *Free exam appeals for schools in England*. (2020, 15 de agosto). BBC News. https://www.bbc.com/news/uk-53787938

Birhane, A. (2021). The impossibility of automating ambiguity. *Artificial Life,* 27(1), 44-61.

Buolamwini, J. y Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. En *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big data & society*, 3(1), 2053951715622512.

D'ignazio, C. y Klein, L. F. (2020). *Data feminism*. MIT press.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Narayanan, A. (2019). *How to recognize AI snake oil*. Arthur Miller Lecture on Science and Ethics.

Porter, J. (2020, 17 de Agosto). *UK A-level results algorithm biased amid coronavirus pandemic*. The Verge. https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications

Turner, J. y Turner, J. (2019). *Controlling the Creators. Robot Rules: Regulating Artificial Intelligence* (pp. 263-318).

Wang, X., Zhang, Y. y Zhu, R. (2022). *A brief review on algorithmic fairness. Management System Engineering*, 1(1), 7.

Wu, X. y Zhang, X. (2016). *Automated inference on criminality using face images*. arXiv preprint arXiv:1611.04135, 4038-4052.

Parte II

IA con enfoque social

Capítulo **6**

Las posibilidades de una IA indígena

Maximino Matus Ruiz

Para entender el fenómeno de la conciencia humana es necesario escapar del cráneo que encierra el cerebro
–Roger Bartra (2019)

https://doi.org/10.61728/AE24001069

Introducción

A lo largo de los años la ciencia occidental ha fortalecido la hipótesis de que el cerebro y la conciencia se encuentran ubicados dentro de un dispositivo biológico; el cuerpo humano contenido por la membrana sensorial más extensa que posee, es decir, la piel. Fuera de este dispositivo biológico la conciencia desaparece. Siguiendo esta hipótesis, la conciencia estaría ocupando (Ingold, 2015) el cuerpo humano como territorio conquistado por una red jerárquica de órganos que acrecientan su importancia según van alcanzando mayor altura; de estómagos sucios asociados con instintos animales bajos, a mentes altamente refinadas según la perspectiva aristotélica (Matus, 2015). Es decir, que, desde la filosofía occidental tradicional, la conciencia no solo se encuentra contenida en el cuerpo humano, sino que esta además se encuentra específicamente localizada en el cerebro, el órgano más alto.

Otras tradiciones filosóficas y perspectivas culturales asociadas a cosmovisiones no occidentales han sostenido el argumento contrario; la conciencia es relacional y se encuentra extendida más allá del cuerpo humano; esta se produce en mallas que entretejen humanos, organismos vivos y entidades no biológicas que conforman tapetes con patrones únicos asociados a cosmovisiones particulares. Desde esta perspectiva la conciencia habita el territorio (Ingold, 2015) y la cultura sería resultado de su traducción semiótica dando como resultado la cosmovisión.

En el libro Chamanes y Robots (2019), Roger Bartra argumenta que la conciencia humana está parcialmente ubicada fuera del cerebro, en las redes simbólicas de la sociedad lo cual conforma una especie de exocerebro: "La conciencia incluye una especie de prótesis simbólica que prolonga, en los espacios culturales, algunas funciones de las redes neuronales […] los elementos exocerebrales de la conciencia tienen un poder causal y son capaces de modificar y modular la operación y las funciones de las redes neuronales" (Bartra, 2019, p. 3?). Bartra avanza su argumento a partir del

llamado efecto placebo producido por sustancias inocuas, operaciones simuladas, rituales, amuletos y otro tipo de elementos que producen efectos somáticos comprobables en sus portadores y practicantes al modificar las señales neuronales. En este proceso, símbolos de diversa índole producidos fuera del cuerpo humano son traducidos en señales biológicas que afectan las relaciones neuronales. Lo cual revelaría según Bartra, una especie de "incompletud" en el sistema biológico humano que tiene que ser "llenada" por prótesis simbólicas externas para dar emergencia al fenómeno de la conciencia.

En tiempos recientes los dispositivos conectados a internet se han convertido en las prótesis simbólicas centrales para los humanos. Poco a poco las prótesis simbólicas culturales están siendo sustituidas por dispositivos electrónicos externos que a su vez se están convirtiendo en internos llevando lo humano a lo transhumano:

"Los transhumanistas parecen más bien unos chamanes que viajan al futuro y predican la sustitución de órganos por prótesis tecnológicamente sofisticadas con el objeto de llegar a una condición utópica. Hay un ingrediente religioso en la espera del advenimiento de la Singularidad, con mayúscula, que abrirá la puerta a una nueva época […] Lo que nos hace humanos es otra singularidad, con minúsculas, la de las prótesis artificiales que constituyen la cultura y el entorno social que los humanos hemos creado" (Bartra, 2019)

Ya sea con mayúscula o minúscula lo cierto es que S/singularidad implica una relación de un punto y contrapunto entre unwelts (Uexküll, 1957), lo cual completa aquello que no es terminado ni autocontenido; la conciencia. Para Bartra, la conciencia es un fenómeno relacional, pero producido solo en los humanos gracias a las prótesis simbólicas culturales, en tanto que lo transhumano es una quimera indeseable producido por artefactos tecnológicos que terminarán con la cultura y diversidad humana.

Desde una perspectiva no antropocentrista la capacidad de cada entidad de afectar y ser afectada (Espinosa, 1959) por otras entidades implicaría grados de semiosis incrementales que terminarían produciendo el fenómeno de la conciencia; de girasoles indéxales a animales simbólicos capaces de reconocer su individualidad dentro de esta basta red. Perspectivas más radicales consideran que una capacidad de semiótica incremental

no necesariamente está relacionada con lo biológico, sino que entidades no orgánicas poseerían una capacidad organizativa compleja que produciría un fenómeno similar: los nuevos realismos respaldan esta perspectiva (De Landa, 2016); entidades particulares capaces de ensamblarse y desensamblase de otros conjunto sin perder sus identidades al tiempo de producir organizaciones mayores con identidades más complejas que al final terminarían produciendo la vida y el propio fenómeno de la conciencia.

Algunos trabajadores de las grandes tecnológicas consideran que los más recientes modelos IA tienen conciencia. Tal es el caso de ingeniero Blake Lemoine, quien en 2022 aseguró que el modelo de lenguaje neurona conversacional (LaMDA por su acrónimo en inglés) de Google había cobrado consciencia. Si bien, esto ha quedado como un hecho aislado y anecdótico, pocos niegan dentro del mundo tecnológico digital que esto efectivamente sucederá, lo que aún no queda claro es cuando; las predicciones para alcanzar la singularidad no pasan de la segunda mitad del siglo XXI (Rengifo, 2020). Es decir, que el siglo XXI se distinguirá por el momento en que los humanos serán capaces de presenciar la emergencia de la conciencia fuera del cerebro y el cuerpo humano; la singularidad de las máquinas producida gracias a la inteligencia artificial. Desde esta perspectiva, el argumento de Bartra respecto a la incapacidad humana de ser consciente si no logra conectarse con dispositivos externos podría extenderse a los robots pues un robot autocontenido no sería más que materia organizada, en cambio, uno extendido gracias a la IA generará semiosis y producirá conciencia. Empero, no queda claro si esta será individual o colectiva; es decir, si existirá una especie de conciencia universal esparcida entre el total de las maquinas interconectadas, o si cada una será capaz de reconocer su individualidad gracias a esta relación; es decir, entidades individuales conscientes.

Varios intelectuales y activistas indígenas han puesto en marcha iniciativas para descolonizar las tecnologías digitales modernas. En No somos computadoras. Un manifiesto (2012) Jaron Lanier advertía sobre los problemas de las "carreteras" informáticas que cerraban las posibilidades de la innovación. Como ejemplo utilizaba el software MIDI con el cual se ha "pavimentado" la música que se hace por computadora. Algo similar ocurre con la famosa metáfora del escritorio que utilizamos hasta la actualidad

como interfaz en la mayoría de las computadoras y los teléfonos inteligentes. Imaginemos por un momento cambiar de metáforas: ¿Qué pasaría si en lugar de la oficina occidental ocupada por cuerpos de cuello blanco hubiera sido la milpa mesoamericana la metáfora utilizada? ¿Qué tipo de computadoras y programas manipularíamos? ¿Cuáles serían sus posibilidades y cómo facilitarían los procesos comunicativos y de innovación?

Algunas iniciativas para pensar las posibilidades de una computación y una IA no occidental fueron recopiladas en el Protocolo Indígena de inteligencia artificial (Lewis et. al, 2020), creado por decenas de indígenas de diversas latitudes. Por su parte, en Making Kin with the Machines (2018), otro grupo de intelectuales indígenas reivindica el valor de sus epistemologías para generar nuevos acuerdos entre humanos y tecnologías. Una iniciativa más a destacar es Indigifab, que tiene sus raíces en el movimiento de indigenización de la Columbia Británica en Canadá. Esta iniciativa promueve el uso de herramientas de digitalización de los conocimientos de las naciones originarias para desarrollar capacidades de gestión de la información y del patrimonio indígena. Canadá ha sido pionera en establecer laboratorios de fabricación digital en territorios indígenas en este concepto. El Fab lab Onaki, que forma parte del First People Innovation Center es un ejemplo a destacar.

Los Futuros indígenas han sido explorados también desde la perspectiva del género. Fue la profesora anishinaabe Grace Dillon quien acuñó el término para referirse a "un movimiento cultural expresando perspectivas indígenas del pasado, presente y futuro en el contexto de la ciencia ficción y subgéneros relacionados". En 2023 diversas mujeres de origen indígena exploraron a través del arte expresiones de los futuros indígenas desde el cuerpo y la subjetividad femenina. La exposición Indigenous Futurims fue curada por la artista y narradora Denise Silva y se presentó en el Cheech Art Museum de Riverside:

"Nuestras costumbres ancestrales son tecnologías antiguas para nuestra liberación colectiva, son nuestras historias tribales, nuestras memorias y profecías cósmicas que se manifiestan. Estamos tejiendo nuevos mundos para los pueblos indígenas. La forma en que lleguemos a ese destino está tejida libremente como una gruesa trenza del pasado, el ahora, el futuro. Es comprender que todos existen al mismo tiempo, en todo el espacio y tiempo" (Silva, 2023)

Una de las artistas invitadas fue Mer Young, quien se adscribe como Chichimeca, Apache. Su obra, Indigenous 3000, 2023 es un collage digital donde reflexiona sobre su visión del futuro para una mujer Tarahumara. Enseguida contrastamos dicha expresión artística con una imagen creada por medio de DALL-E. A nuestro parecer el contraste entre el minimalismo y un futuro barroco es obvio:

> What would it be like in the year 3000? When thinking of an Indigenous future I envision lands and waters returned to us. Where everything in clean and untainted, our foods, our grains, our bodies, our air, our spirit and so much more. The concept of Spectral Vision is the ability to see different wavelengths of light. I used these non-spectral colors as a representation of an Indigenous future, one where we can see past fundamental colors and experience higher wavelengths naturally and not by pollution. The central figure of a Tarahumara Indigena from Chihuahua of Anahuac Mexico walking onto unsoiled lands on Mother Earth with the outer heavens depicting no space debris or human spacecraft in our Universe. (Mer Young, 2023).

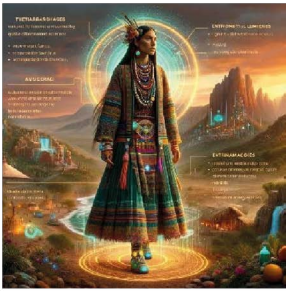


Imagen 1. Mer Young, 2023, Instagram @youngmer / Imagen 2. DALL·E 2024-02-05 18.34.42 - Visualize a Tarahumara woman in the year 3000, embodying the essence of her culture while integrating futuristic elements.

Robots indígenas y grados de semiosis

Existen algunos ejemplos de robots indígenas en Latinoamérica, sin embargo, la mayoría viven en mundos icónicos, algunos de ellos en indéxales, y muy pocos en simbólicos; la pregunta aquí es cómo crear robots indíge-

nas que sean capaces no solo de vivir en los tres niveles de semiosis antes señalados, sino que puedan residir en muchos más. Es decir, que sean capaces de otorgar significando a esos otros espacios que no tienen un significante por parte de la semiótica accidental más tradicional: ¿Existen más mundos posibles donde habitar teniendo relaciones más complejas y fructíferas entré indígenas y robots autóctonos o máquinas interconectadas? ¿Cuáles sería algunos de los parámetros mínimos que debería de guiar el diseño de una IA indígena?

Robots indígenas icónicos

En Latinoamérica existen algunos robots indígenas icónicos. La mayoría de estos funcionan como altavoces que reproducen unas pocas palabras indígenas al ser activados por presión mecánica. Sus formas suelen reproducir por relación icónica una figura parecida a la humana. En ocasiones vestida con los trajes tradicionales de los pueblos indígenas, otros simulan entidades constitutivas del paisaje, como animales. También están aquellos que emulan el antiguo imaginario de los robots metálicos. Entre el primer tipo de robots destaca la familia de los Paquitos, que forman parte de la colección "Juguemos con….." del proyecto de materiales y Acervos Multimedia del Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS). Estos fueron desarrollados entre 2015 y 2017. Su concepción implicó la integración de lingüística, robótica y semiótica para facilitar la adquisición de la lengua materna entre niños de 1 a 6 años. Además de enseñar diversos saludos reproducían el nombre de las partes del cuerpo al ser presionados mecánicamente.

Imagen 3. Paquitos.

Fuente: CIESAS

Por otra parte, están aquellos robots indígenas que no imitan de forma icónica a los humanos, pero sí reproducen elementos de su entorno. Este es el caso de tres jóvenes de la nación boliviana Irohito Urus, que han desarrollado robots inspirados en sus cosmologías andinas a partir de materiales locales. Por ejemplo, Tikatz es un robot con cara de gato andino que ayudaba a desinfectar superficies durante la pandemia de la covid-19 y con el cual ganaron una competencia de robótica en Guadalajara, México.

Imagen 4. Robot Irohito Uru.

Fuente: https://www.latimes.com/espanol/vida-y-estilo/articulo/2022-04-21/indigenas-irohito-urus-de-la-pesca-a-la-robotica-en-bolivia

Entre la variedad de robots indígenas del tipo icónico en Latinoamérica están aquellos que efectivamente reproducen la imagen estereotípica de un robot metálico. Este es el caso de kipi, un robot hecho con materiales reciclados que viajó por los Andes para enseñar a los niños de comunidades sin internet durante la pandemia del COVID. Kipi era capaz de almacenar audiolibros, reproducir grabaciones de clases y proponer el desarrollo de experimentos. El robot viajaba en burro por los andes junto con su creador.

Imágen 5. Kipi.

Fuente: https://drehuancavelica.gob.pe/notas-de-prensa/kipi-el-robot-que-ensena-ciencia-y-tecnologia-en-huancavelica/

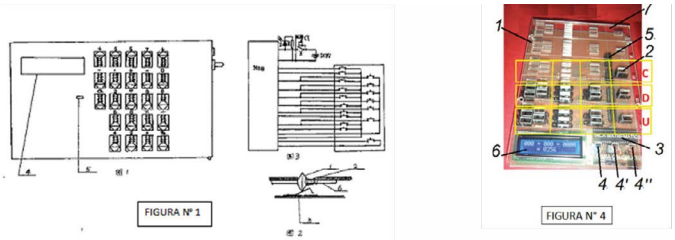
Los anteriores son ejemplos de robots indígenas que viven en un plano icónico de significación al imitar los elementos del entorno natural y cultural, es decir, el paisaje. Funcionan bien para enseñar algunos elementos culturales que se han perdido o que están en riesgo de desaparición, pero no para explorar las posibilidades futuras de lo indígena. Los robots icónicos tienen un mapa prediseñado insertado en su cerebro electrónico (De Landa, S/F) y por ellos sus funciones son limitadas.

Robots indígenas indéxales

Algunos robots indígenas no reproducen de forma icónica el entorno, pero contienen elementos constitutivos de los pluriversos indígenas (Escobar, 2016). Este es el caso de la Yupana digital diseñada por Walter González como parte de su paso por la Fab Academy. El robot permitía medir el mundo de forma matemática como lo hacían los incas:

La invención se refiere a una calculadora digital basada en la Yupana para la enseñanza y práctica en la realización de cálculos con las cuatro operaciones básicas; la calculadora también permite el cambio de bases numéricas. Esta calculadora, al igual que la Yupana, tiene filas y columnas con valores predeterminados y en una disposición predeterminada, los sensores permiten introducir valores numéricos presionando, tocando o colocando objetos sobre ellos, y cada sensor está vinculado a un LED que se ilumina para mostrar fácilmente el número asociado al sensor. (Descripción de patente presentada por Walter González)

Imagen 6. Yupana digital.



Fuente: Walter González.

¿Cuáles son las consecuencias de unas matemáticas diferentes? una forma diferente de dividir el mundo y de programar; la forma de medir el mundo limita o expande las posibilidades de las entidades que viven dentro de él. Por ejemplo, diversas culturas mesoamericanas tienen como base el 20 en lugar del 10. Esto se relaciona con formas específicas de medir el tiempo y determinar los ciclos agrícolas y rituales.

Robots indígenas simbólicos

Un robot simbólico implica que este sea capaz de comprender símbolos humanos que representan los pluriversos. Durante mucho tiempo un modelo similar fue aplicado para el desarrollo de la IA simbólica también conocida como Good Old-Fasioned AI (GOFAI). Estos son sistemas basados en lógica y reglas específicas para operar dentro de un contexto. Se espera que un robot simbólico reproduzca la cultura de sus creadores, pero no que sea capaz de innovar culturalmente; no existe periferia ni rebeldía dentro de su comportamiento, solo reproducción de los patrones culturales esperados. En este sentido, un robot simbólico sería de utilidad para procurar la preservación de los pluriversos, más no para la creación de futuros indígenas.

En su paso por la Fab Academy Jean-Luc Pierite desarrolló un taller que derivó en el rediseño y la materialización de personajes míticos con la ayuda de tecnologías asociadas a los Fab labs para desarrollar un ejercicio de revitalización lingüística y cultural entre los Túnica-Biloxi de Luisiana. Desde nuestra perspectiva este proyecto contenía las bases para el desarrollo de un robot simbólico. El proyecto fue intitulado Tarujantohku; diseño

digital basado en la tradición de la narración de cuentos: "los estudiantes usaron: narración tradicional, conocimiento de los ancianos vivos y su propio conocimiento personal del ecosistema local. Pudieron colaborar y expresar en papel un nuevo concepto en su lengua de herencia. Los siguientes pasos fueron digitalizar los diseños" (Peirite, 2019, p. 67, traducción propia).

El protocolo IA indígena

Entre 2018 y 2019 activistas indígenas mantuvieron una conversación que se expandió a lo largo de 20 meses de correspondencia y dos encuentros para discutir las Pautas para el Diseño de IA Centrado en Indígenas. Dichos intercambios dieron origen al primer protocolo indígena de inteligencia artificial. Los elementos constitutivos de este son siete. En seguirá los presentamos de forma textual:

Cuadro 1. Pautas para el Diseño de IA Centrado en Indígenas

1. Localidad: El conocimiento indígena a menudo tiene raíces en territorios específicos. También es útil para considerar problemas de importancia global. Los sistemas de IA deberían ser diseñados en asociación con comunidades indígenas específicas para asegurar que los sistemas sean capaces de responder y ayudar a cuidar de esa comunidad (por ejemplo, basados en lo local) así como conectarse a contextos globales (por ejemplo, conectados a lo universal).
2. Relacionalidad y reciprocidad: El conocimiento indígena es a menudo un conocimiento relacional. Los sistemas de IA deberían ser diseñados para entender cómo los humanos y los no humanos están relacionados y son interdependientes entre sí. Comprender, apoyar y codificar estas relaciones es un objetivo de diseño primario. Los sistemas de IA también son parte del círculo de relaciones. Su lugar y estatus en ese círculo dependerán de comunidades específicas y sus protocolos para entender, reconocer e incorporar nuevas entidades en ese círculo.
3. Responsabilidad y relevancia: Las personas indígenas a menudo están preocupadas principalmente por sus responsabilidades hacia sus comunidades. Los sistemas de IA desarrollados por, con, o para comunidades indígenas deberían ser responsables ante esas comunidades y proveer apoyo relevante.

4. Desarrollar pautas de gobernanza desde protocolos indígenas: El protocolo es un conjunto de reglas costumbristas que gobiernan el comportamiento. El protocolo se desarrolla a partir de configuraciones ontológicas, epistemológicas y costumbristas del conocimiento basadas en la localidad, la relacionalidad y la responsabilidad. El protocolo indígena debería proporcionar la base para desarrollar marcos de gobernanza que guíen el uso, rol y derechos de las entidades de IA en la sociedad. Existe la necesidad de adaptar los protocolos existentes y desarrollar nuevos protocolos para el diseño, construcción e implementación de sistemas de IA. Estos protocolos pueden ser particulares a comunidades específicas, o pueden ser desarrollados con un enfoque más amplio que pueda funcionar a través de muchas comunidades indígenas y no indígenas.
5. Reconocer la naturaleza cultural de toda tecnología computacional: Todos los sistemas técnicos son sistemas culturales y sociales. Cada pieza de tecnología es una expresión de marcos culturales y sociales para entender y relacionarse con el mundo. Los diseñadores de sistemas de IA necesitan ser conscientes de sus propios marcos culturales, conceptos socialmente dominantes e ideales normativos; estar atentos a los sesgos que conllevan; y desarrollar estrategias para acomodar otros marcos culturales y sociales. La computación es un material cultural. La computación está en el corazón de nuestras tecnologías digitales y, dado que cantidades crecientes de nuestra comunicación son mediadas por dichas tecnologías, se ha convertido en una herramienta central para expresar valores culturales. Por lo tanto, es esencial para la resiliencia cultural y la continuidad de las comunidades indígenas desarrollar métodos computacionales que reflejen y promuevan nuestras prácticas y valores culturales.
6. Aplicar diseño ético al stack extendido: La cultura forma la base del ecosistema de desarrollo tecnológico, o stack. Cada componente del hardware y software del sistema de IA debe ser considerado en la evaluación ética del sistema. Esto comienza con cómo se extraen los materiales para construir el hardware y para energizar el software de la tierra, y termina con cómo regresan allí. La ética central debería ser la de no hacer daño.
7. Respetar y apoyar la soberanía de datos Las comunidades indígenas deben controlar cómo se solicita, recolecta, analiza y operacionaliza sus datos. Ellos deciden cuándo protegerlo y cuándo compartirlo, dónde residen los derechos de propiedad cultural e intelectual y a quién pertenecen esos derechos, y cómo se gobiernan estos derechos. Todos los sistemas de IA deben ser diseñados para respetar y apoyar la soberanía de datos. Los principios de datos abiertos necesitan ser desarrollados aún más para respetar los derechos de los pueblos indígenas en todas las áreas mencionadas anteriormente, y para fortalecer la equidad de acceso y la claridad de los beneficios. Esto debería incluir una revisión fundamental de los conceptos de "propiedad" que son producto de órdenes legales no indígenas y no necesariamente reflejan las formas en que las comunidades indígenas desean gobernar el uso de su conocimiento cultural.

Fuente: Lewis, Jason Edward, ed. 2020. Indigenous Protocol and Artificial Intelligence Position Paper. P. 21 y 22. Traducido con Chat GPT 4.

Entre la diversidad de propuestas para pensar lo que implicaría una IA culturalmente localizada en relación con Pluriversos particulares, destacan dos ejercicios de futurismo y transhumanismo indígena presentados en la sección 4 del protocolo y de los cuales presentamos la idea central de los dispositivos propuestos en los siguientes párrafos.

Gwiizens, the Old Lady and the Octopus Bag Device de Scoot Benesiinaabandan:

> La bolsa pulpo es un dispositivo informático multisensorial extraíble y relativamente no invasivo […] que utiliza nuestras capacidades inherentes de ADN tanto para el almacenamiento como para la comunicación de computación multisensorial, utilizando nuestras capacidades inherentes de ADN para almacenamiento y cálculo potenciales (mediante una verdadera computación paralela). […] Los elementos del dispositivo están en contacto directo con la parte posterior de la mandíbula, donde antes se encontraban las muelas del juicio […] Esto permite que el dispositivo tenga una conducción de esquelética directa. Otros elementos se extienden hacia arriba desde la boca y se sientan libremente en las fosas nasales para permitir la estimulación feromonal. Las entradas sensoriales (olfato, vista, gusto, tacto, oído) se almacenan en proporciones universalmente únicas y superpuestas […] La computación ADN [realmente paralela] del futuro lejano y el almacenamiento masivo eliminan los problemas de los recursos terrestres [propiedad]. El ADN-paisaje de cada persona es único y, por tanto, la entrada y salida de los sistemas de cálculo y almacenamiento también serían únicos y ofrecerían mucho más poder, seguridad y soberanía que las soluciones actuales de un ordenador y un sistema operativo para todo de hoy. El dispositivo es también una interfaz que utiliza nuestros modos interiores personalizados de percepción para interpretar y evaluar los resultados informáticos. Los programas específicos de IA se desarrollan a través de las nuevas posibilidades que ofrecen la naturaleza y la potencia de este tipo de computación. (Traducido con Chat. GPT 4)

Gift of Dentalium and Fire: Entwining Trust and Care with AI de Ashley Cordes:

> Cada dentalium, unidades de conchas en forma de colmillo procedentes de las costas del noroeste del Pacífico, están llenos de simulaciones de dinámica de fluidos computacional. Éstas muestran un chorro de fluido a alta velocidad que se inyecta en un medio en reposo. Cada filamento depende de la concha génesis y de su adaptación generacional. Las cuentas

> negras anclan los nodos dentro de un registro distribuido mantenido por la totalidad de la red (collar). Esta ofrenda a las IA pretende permitir la externalización de historias/datos/sueños que fluyen a través del fluido de cada concha. Cuando las IA lo usan y lo llevan, es un medio simbólico de compartir y una expresión de respeto por uno mismo y por los demás y hacia los demás. (Kari Noe y Ashley Cordes, 2019. Traducido con Chat GPT. 4)

Como ejemplo práctico las investigadoras Kari Noe y Ashley Cordes nos invitan a imaginar el caso de una comunidad indígena que decide que un determinado porcentaje de cada compra realizada con criptomoneda creará un fondo para promover el bien social y la prosperidad económica en la comunidad. Desde su perspectiva la IA podría ayudar a proporcionar el porcentaje adecuado y determinar a qué programas podrían destinarse los fondos. La IA también ayuda a autenticar la identidad para votar sobre cómo se gastan los fondos dentro de una nación y en la detección del fraude en la cadena de bloques. La tecnología es responsable de registrar las decisiones que rodean la implementación de protocolos basados en tales filosofías financieras, y luego de distribuir las propias monedas para almacenarlas en monederos digitales. Dando como resultado la facultad de una nación indígena para tomar decisiones financieras en relación con su moral.

Desde nuestra perspectiva los ejercicios de futurismo indígena antes presentados pueden ser la base de futuras cosmotecnias: El concepto de cosmotecnia propuesto por Yuk Hui:

Es la unificación del cosmos y la moral a través de actividades técnicas, ya sean artesanales o artísticas. No ha habido una o dos técnicas, sino muchas cosmotécnicas. Qué tipo de moral, qué cosmos y de quién, y cómo unirlos varía de una cultura a otra según las distintas dinámicas. Estoy convencido de que para hacer frente a la crisis que tenemos ante nosotros –a saber, el Antropoceno, o la intrusión de Gaia (Latour y Stengers), o el "Antropoceno" (Stiegler), todos ellos presentados como el futuro inevitable de la humanidad– es necesario reabrir la cuestión de la tecnología, prever la bifurcación de los futuros tecnológicos concibiendo diferentes cosmotécnicas. (Hui, 2017, p. 7).

Tomando en cuenta lo anterior consideramos que los robots y las IA indígenas deben ser capaz de conjugar comportamientos simbólicos, in-

déxales e icónicos organizados en torno a paisajes localizados tanto físicos como virtuales, para que sean capaces de sostener y resignificar las relaciones emergentes entre naturaleza, tecnologías digitales y grupos indígenas.

Xquenda o las posibilidades de una IA indígena zapoteca

Actualmente en Fab Lab Puebla como parte de mi paso por el Fab Academy estoy desarrollando un proyecto inspirado en el libro Chupa Ladxiduá, Mi corazón es Dos, escrito por la poeta zapoteca Irma Pineda. Este libro es una recopilación de 60 poemas escritos a lo largo del tiempo, los temas son tan diversos y profundos que es un libro para todos los interesados en la vida cotidiana -con sus bellezas y dificultades- de esta cultura indígena y su expresión en el lenguaje escrito. El libro está dividido en cuatro secciones: 1) Diidxa' -la palabra- 2) Guidxi -el pueblo-; 3) Guendarixhana -la maternidad-, 4) Ni napa xi che' -el exilio-. 5) Ni que guidxela -personas desaparecidas-. El libro es bilingüe. Por lo tanto, está escrito en español y zapoteco uno al lado del otro. De este modo, niños y adultos pueden aprender y comparar ambas lenguas. Cada poema está ilustrado con hermosos dibujos realizados por niños zapotecas de Juchitán, ciudad del Istmo de Tehuantepec donde nació Irma.

Para este proyecto estamos trabajando en la traducción intersemiótica de este libro a un libro objeto digitalizado –Wearable, para promover la alfabetización en zapoteco–. La idea principal implica grabar esos poemas en audio y almacenarlos en una placa PCB que se va a colocar dentro de un corazón hecho con resina y que se puede usar como collar o simplemente tenerlo en la mano. Para que el dispositivo reconozca el zapoteco estamos desarrollando un modelo de IA zapoteca llamado Xquenda que presentaremos más adelante. Para acceder al contenido del dispotivo, tienes que hablarle al corazón diciendo "¿Xi nuu ladxido´luca?", una forma tradicional de saludar en zapoteco; "Cómo se siente tu corazón". Una vez encendido el corazón, puedes acceder a los poemas y escucharlos con auriculares tanto en zapoteco como en español.

De hecho, no se trata de uno, sino de dos corazones, igual que el nombre del libro, Chupa Ladxidua' -Mi corazón es dos-. Por ello estamos fabricado dos corazones con las mismas funciones. La idea es que

esos corazones se conecten mediante Bluetooth y Wi-fi. De esta forma, se puede utilizar como juego entre padres e hijos o parejas para fomentar la comunicación y la alfabetización en zapoteco. Cuando un corazón está encendido, el otro se ilumina. Tras esta señal, puedes enviar un poema al otro corazón. Cuando se escucha el poema, puedes hacer que los dos corazones cambien de color según el sentimiento que te produzca el poema. Por ejemplo, si es felicidad, puedes hacer que se ponga rojo; si es tristeza o melancolía, puedes ponerlo azul. Esta última funcionalidad pretende ser un sencillo sistema semiótico para compartir sentimientos esenciales entre dos corazones que están muy cerca en el espacio, o que están muy lejos, por ejemplo, un emigrante en EE.UU. que dejó atrás a sus seres queridos.

Ahora bien, ¿Cuáles sería los parámetros que deberíamos de tomar en cuenta para el desarrollo de una IA zapoteca? En zapoteco, la palabra "Xquenda" tiene múltiples significados, como cultura, origen, virtud o "nahual", que se refiere a tu gemelo o doble animal. Tu guénda es el animal con el que naciste y con el que morirás. Debido a este destino enredado, algunas personas llegan a parecerse a las características del animal con el que están ligadas; algunas incluso se convierten en él. Guénda es una de las ideas más cruciales de la cosmología zapoteca.

¿Sería posible desarrollar una IA indígena para los indígenas zapotecas? ¿Por qué sería importante? Como ocurre con muchas lenguas indígenas, el zapoteco está desapareciendo y, con él, los contornos de su mundo se difuminan. En un mundo moderno digitalizado y colonizado por pocas lenguas "naturales" y "artificiales", no es fácil comprender para qué sirve una lengua; las lenguas son la clave de las cosmologías.

Localizar radicalmente las TIC va más allá de la simple traducción a las lenguas nativas. Implica reevaluar la codificación y el diseño de estas tecnologías, desde la experiencia de usuario front-end hasta la infraestructura back-end. Con esta estrategia se abren nuevas posibilidades para integrar estas tecnologías con las cosmologías y cosmovisiones locales. Este es el objetivo de Xquenda, un prototipo de IA zapoteca que es capaz de iterar millones de veces el poema Laacabe de Irma Pineda. El prototipo desarrollado hasta el momento fue presentado en el pabellón virtual Planetary Fronlines de la 18th Bienal de Arquitectura de Venecia. Para acceder al video donde se presenta el recurso: CITYX Venice, Max Matus: Xquenda.

The Possibilities of Zapotec AI (youtube.com)

Xquenda es una palabra polisémica que puede significar cultura, origen, virtud o nahual, que significa tu gemelo animal o tu doble. El animal con el que naciste, es el mismo con el que morirás, es tu guénda. Debido a este destino enredado, algunas personas incluso se asemejan a las características del animal con el que están atadas; algunas incluso se convierten en él. Como señala el escritor zapoteco Andrés Henestrosa, la guénda es uno de los conceptos más esenciales de la cultura zapoteca. Siguiendo esta línea de pensamiento, los animales a los que estamos ligados utilizan el lenguaje zapoteco para expresar su esencia a través de su naturaleza humana. No es de extrañar que muchos verbos zapotecos comiencen con guénda o "g" argumenta Henestrosa. Tal es el caso de guénda biaani', y tal será la IA zapoteca Guenda biaani' guiiba que estamos desarrollando para dotar a Binni guiiba de capacidades para encontrar posibles correlaciones desconocidas entre los humanos y sus guendá, entre los guendá y el medio ambiente para su preservación, entre la cultura zapoteca y la preservación de su pluriverso.

"Binni" significa humano, y "guiiba" significa máquina en lengua zapoteca. Por lo tanto, sostengo que como concepto mixto, "Binni guiiba" podría significar "robot". Por supuesto, esta palabra no existe en zapoteco. Sin embargo, sostengo la necesidad de desarrollar un robot zapoteco que pueda ayudar a rediseñar el pluriverso zapoteco ayudando a los binnica -humanos" Ca, a recordar los significados ocultos de la lengua zapoteca, la extensión de los conceptos y los afectos que solo viven dentro de esta lengua; las posibilidades ocultas de habitar el mundo que reside en la cosmología de esta cultura; a redescubrir y rediseñar las posibilidades de este pluriverso. Para hacer esto posible necesitaríamos desarrollar un programa para que el robot desarrolle cierta inteligencia o "guenda biaani", entonces Binni guiiba podría operar con su propia guenda biaani guiiba o inteligencia artificial.

Conclusiones

Los robots indígenas pueden ayudar a redescubrir los pluriversos ocultos en las lenguas indígenas y sus cosmologías afines; pueden compartir el conocimiento local sobre cómo tratar con la naturaleza, la medicina tradicional, los rituales y los mitos, e incluso almacenar archivos CAD para rediseñar e imprimir el pluriverso mediante CAM. Sin embargo, hasta ahora no ha existido un único repositorio indígena para tales archivos, solo hay Thingiverses. En su lugar, necesitamos Indigenuversos. Los repositorios digitales de cosas-no-comunes creados con pocas lenguas dominantes siempre serán redundantes. En cambio, los repositorios hechos de objetos digitales derivados de miles de lenguas indígenas abrirán infinitas posibilidades para la existencia de un parlamento diverso de cosas donde ninguna está por encima de la otra sino al lado, en otra semiosfera, tocándose, teniendo diálogos, confrontaciones, traducciones, rechazo, y tal vez, generando nuevos posibles pluriversos indígenas sintéticos.

Un Indigenuverso solo sería posible entre indígenas que acepten compartir parte del conocimiento que han desarrollado durante milenios en un intercambio horizontal. Nunca para aquellos que intentan colonizar todo lo que encuentran sin siquiera entender que la cosa o la idea abstracta; puede ser una cosmotécnica (Hui, 2020), y cuando es robada, se convierte en una pobre representación de un signo decadente que alguna vez tuvo el poder de sostener un mundo particular.

Referencias

Bartra, R. (2019). *Chamanes y Robots.* Anagrama.

De Landa, M. (2006). *A New Philosophy of Society. Assemblage Theory and Social Complexity*. Londres: Continuum.

De Landa, M. (s. f.). *Meshworks, Hierarchies and Interfaces*. https://papers.cumincad.org/data/works/att/7f71.content.pdf

Escobar, A. (2016). *Autonomía y Diseño. La realización de lo comunal*. Editorial Universidad del Cauca.

Espinosa, B. (1959). *Ética demostrada según el orden geométrico*. Editorial Nacional

Flores, P. (2022). *Indígenas Irohito Urus: de la pesca a la robótica en Bolivia*. ht-

tps://www.latimes.com/espanol/vida-y-estilo/articulo/2022-04-21/indigenas-irohito-urus-de-la-pesca-a-la-robotica-en-bolivia

https://ichan.ciesas.edu.mx/8243-2/#footnote-1

Ingold, T. (2015). *Líneas. Una breve historia*. Gedisa.

Lanier, J. (2012). No somos computadoras. Un manifiesto. *Debate*.

Lewis, J. E. (Ed.). (2020). *Indigenous Protocol and Artificial Intelligence Position Paper*. Honolulu, Hawai: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR).

Peirite, J. L. (2019). InDigiFab: Craftivismo artisanal indígena promovido a través de la educación Fab Lab. En: Gonzales, Walter, *Neoartesanías en América. Métodos para incorporar procesos de fabricación digital en las artesanías* (pp. 61-78). Perú: Universidad Nacional de Ingeniería.

Pineda, I. y Ladxiduá C. (2018). Dos es mi corazón. *Alas y Raíces*.

Rengifo, L. (2020). *Tecnología, agencia y transhumanismo*. Bogotá: Universidad Santo Tomás.

Silva, D. (2023). *Indigenous Futurism*. Expo presented at the Cheech Museum.

Villavicencio, F., Pérez, Z. y Peralta, M. del R. (s. f.). Materiales y acervos en lenguas indígenas. Exploraciones multidisciplinarias. Ichan Tecolotl. CIESAS.

Von Uexküll, J. (1957). A Stroll Through the Worlds of Animals and Men: A Picture Book of Invisible Worlds. En Claire H. Schiller (Trad. Y Ed.), *Instinctive Behavior: The Development of a Modern Concep*. New York: International Universities Press.

Recursos de IA:

DALL·E 2024-02-05 18.34.42 - Visualize a Tarahumara woman in the year 3000, embodying the essence of her culture while integrating futuristic elements. (Maximino Matus).

Capítulo 7

De la tecnofilia a la tecnofobia. Narrativas sobre inteligencia artificial y educación superior

Christian Fernández Huerta
Hugo Méndez-Fierros
Joaquín Caso Niebla

https://doi.org/10.61728/AE24001076

Introducción

Desde la llegada de la Web 2.0, ninguna tecnología había generado tanto interés y tanta especulación como la inteligencia artificial (IA). Al igual que con el Internet, este interés acompañado de incertidumbre, no fue de la noche a la mañana, fue un largo proceso que llevó al reconocimiento y en algunos casos la adopción generalizada de estas tecnologías entre grandes sectores de la población. Sin embargo, hay momentos de inflexión que podemos identificar cuando en la convergencia de los objetos materiales, las prácticas y las narrativas sobre algunas tecnologías digitales, hay un cisma o quiebre que modifica la manera en que entendemos y nos relacionamos con estas tecnologías. Para la IA esto ocurrió en el 2019, con la introducción del modelo de lenguaje llamado Chat GPT en su tercera versión. El Generative Pre-Trained Transformer o Chat GPT, por sus siglas en inglés, permitió que la IA tuviera interacciones fluidas con un lenguaje natural a través de una interfaz amigable.

Esta innovación ha generado reacciones en muchos ámbitos de la vida social donde las tecnologías digitales ocupan un lugar central, los cuales son prácticamente todos los ámbitos, no obstante, en el ámbito educativo en general, y particularmente en el vasto campo de la educación superior, abrió una discusión que, con algunos matices pudiéramos ubicar en el espectro que va desde la tecnofilia hasta la tecnofobia. En este capítulo se busca responder las siguientes preguntas: ¿Qué discursos y narrativas se construyen y refuerzan sobre la IA desde las instituciones de educación superior? ¿Se rechaza la inteligencia artificial o se adopta y adapta al contexto universitario mexicano? Para ello, se ha planteado como objetivo, identificar e interpretar las posturas que subyacen en las referencias educativas de la IA.

A través de un análisis de contenido cualitativo de orden inductivo, se busca discernir si manuales universitarios de tres universidades públicas mexicanas y el de un organismo multilateral internacional, proyectan una visión tecnófila, es decir, una inclinación hacia la aceptación y promoción

de tecnologías, o tecnófoba, que denota un escepticismo o rechazo hacia los avances tecnológicos. Evaluando la narrativa y el énfasis de los textos en las aplicaciones, ética y futuros posibles de la IA, este trabajo ofrece una visión panorámica de la perspectiva institucional hacia la IA en el contexto de la educación superior.

Los resultados revelen la existencia de una mezcla de posturas en los manuales seleccionados, con algunos elementos discursivos presentando una fuerte inclinación hacia la promoción optimista de la IA (tecnofilia), mientras que otros podrían reflejar un grado de precaución o crítica (tecnofobia). Además, se pone al descubierto la manera en que los manuales abordan de manera enfática la ética de la IA, la responsabilidad social de los desarrolladores y la prospectiva sobre el impacto de la IA en la sociedad y el mercado laboral.

Del Internet incorporado a las culturas digitales

Aunque hoy se utilice cada vez más la concepción de madurez digital, se debe reconocer que aún estamos lejos de superar la brecha digital en México. Podemos afirmar que el Internet se ha convertido en una tecnología esencial en nuestra cotidianidad, tan presente que solo la notamos en su ausencia. No solo es una herramienta de información o un medio de comunicación, es un elemento primordial de sociabilidad. Muchas de nuestras relaciones cotidianas no se conciben sin la presencia y participación, en mayor o menor medida, en los espacios digitales posibilitados por Internet.

La socióloga Christine Hine (2022), refiere a un aspecto particular de la experiencia del Internet contemporáneo, y es que está incorporado en diversos marcos, instituciones y dispositivos (p. 51). Más allá del aspecto técnico, ella se refiere a esta incorporación en un sentido más general, y es que "el uso de Internet se entrelaza con múltiples formas de contexto y marcos de elaboración de significado" (p. 52). Si bien a veces parecemos empeñados en distinguir los espacios "virtuales" de los "reales", cada vez nos acercamos más a la idea de un continuo devenir entre espacios digitales y no digitales, entre categorías como online y offline. Esto es posible gracias a la capacidad que hemos desarrollado, y particularmente las

nuevas generaciones, por relacionar y movilizar saberes y recursos en las plataformas sociodigitales (Fernández, 2015).

Muchas de nuestras experiencias y prácticas adquieren sentido y significado a través de las plataformas sociodigitales. Esto responde a que lo digital es intrínseco a nuestras vidas cotidianas. Floridi (2015) habla de lo onlife para ayudar a entender este flujo de conectividad continua que ahora vivimos, donde lo que distinguía a lo online y lo offline se desdibuja. Algo recurrente en el ámbito educativo universitario con las distintas formas de comunicación e interacción entre docentes y estudiantes, y entre los propios estudiantes, a través de plataformas digitales de aprendizaje y comunidades educativas en línea. La atención de muchos estudiantes se divide y se comparte entre sus dispositivos tecnológicos (teléfono móvil, computadora, tablets, etc.), y lo que sucede en el aula, la cual incluso puede ser "digital".

Utilizar el término "digital" en cualquier oración, nos remite al uso de ciertos dispositivos y en el mejor de los casos a un proceso de apropiación tecnológica. Según Siles (2004) para hablar de apropiación tecnológica se deben cumplir tres condiciones: el manejo técnico y cognitivo del artefacto o dispositivo por parte del usuario, la integración de esta tecnología en la vida cotidiana del usuario y la posibilidad de la creación de nuevas prácticas y experiencias a partir de ese objeto técnico, artefacto o dispositivo. Si entendemos al Internet como objeto técnico, además de entorno cultural, podemos asumir que hay una apropiación general de esta tecnología. Hoy podemos declarar, con cierta cautela, que el Internet es algo masivo y popular. Sin duda hay indicadores que pueden ayudar a argumentar esta idea.

A principios del 2023 en México, más de 100 millones de personas tiene acceso a Internet, es decir, el 78.6 % de la población, por supuesto habrá que identificar cuál es la calidad de la conexión y si este acceso es permanente o intermitente, sin embargo, podemos decir que la gran mayoría tiene la posibilidad de hacer uso de esta herramienta. Otro indicativo es la penetración del Internet en la sociedad; y cómo se comentó en párrafos anteriores, el Internet se utiliza para una gran variedad de propósitos, lo cual ha permitido que se integre a nuestra vida laboral y personal, al ámbito educativo y académico y por supuesto para el ocio y entretenimiento. Podemos decir que se ha integrado a todos los aspectos de nuestra vida cotidiana.

La penetración del Internet también puede medirse por la diversidad de dispositivos que permiten su acceso, hoy podemos encontrar además de equipos de cómputo personal o dispositivos de comunicación como los teléfonos celulares, una gran variedad de aparatos desde refrigeradores, hornos y televisores conectados a Internet, convirtiéndose en una tecnología omnipresente. Aunado a la gran variedad de servicios, aplicaciones y plataformas que se han desarrollado y que lo utilizan para su funcionamiento, lo cual también contribuye a su popularidad y penetración, el Internet tiene un fuerte impacto en la economía. Es un motor económico, genera nuevos empleos y configura los ya existentes, crea nuevas oportunidades para el comercio y se convierte en una herramienta para el emprendimiento.

Considerando estos indicadores mencionados, como son el número de usuarios, la penetración en la sociedad y su impacto en la economía, ¿La inteligencia artificial ha llegado a ese estatus de ser masiva y popular? La respuesta es no. Hay elementos para considerar la inminente masificación de esta tecnología, como lo es el desarrollo de servicios y productos que integran la IA en distintos ámbitos. En el caso de los ámbitos educativos y el del entretenimiento se muestra mayores avances, pero, aún estamos lejos de tener la adopción generalizada de esta tecnología por parte de la población en México.

Independientemente de su uso y apropiación en el contexto mexicano, la IA ya está cambiando la forma en que los estudiantes aprenden y los docentes enseñan, en los procesos de evaluación y seguimiento, y en la misma interacción dentro y fuera del salón de clases. Las Instituciones de Educación Superior en México reconocen que es imprescindible su adopción e integración dentro del proceso enseñanza-aprendizaje. Sin embargo, aún no es claro en qué nivel y profundidad, por ello han generado discursos institucionales, que se ven objetivados en manuales y guías de uso de esta tecnología que se convierten en "ventanas" hacia el horizonte de posibilidades que ofrece la IA.

Inferimos que la relación entre usuario y la IA está entrando a la fase de la apropiación tecnológica, pues además de paulatinamente ir cumpliendo con las condiciones descritas por Siles (2004), hay un elemento primordial que define cómo nuestras experiencias y prácticas digitales adquieren

otros sentidos por el uso de esta tecnología. En algunos casos ni siquiera se refiere a su uso, sino al reconocimiento de su existencia y la posibilidad de que se integre a nuestra cotidianidad. Este elemento son las narrativas y discursos que se generan sobre esta tecnología, las cuales se posicionan desde el entusiasmo por el avance e innovación que representa la IA, hasta el miedo y aversión hacia esta tecnología, derivado de las posibles consecuencias nefastas por su uso.

Un concepto que nos ayuda a entender este fenómeno es *Digital Culture*, propuesto por Grant Bollmer (2018).

Deberíamos pensar en la cultura digital como algo compuesto por tres elementos: narrativas sobre tecnología, infraestructuras materiales que dan forma a la comunicación y las capacidades físicas de los cuerpos, humanos o no, en su capacidad para moverse y realizar actos específicos. La cultura digital se encuentra en la intersección de estos tres elementos. (Bollmer, 2018, p. 27, traducción propia)

El concepto de Bollmer es muy útil como guía de análisis, sin embargo, consideramos necesario pluralizar el concepto para dar cuenta de la heterogeneidad de lo digital en el contexto global. Hablar de culturas digitales permite hacer evidente las diferencias entre México y otras geografías del mundo, incluso dentro de nuestro propio país, en distintas regiones.

Las culturas digitales son estas formas emergentes de organización, comunicación e intercambio simbólico que se elaboran a partir de la integración de las tecnologías digitales a nuestra cotidianidad, de las prácticas derivadas de esta integración, y los discursos, afectividades y relaciones generadas por y para ellas. Precisamente, son los discursos y estas valoraciones afectivas, los sentimientos de miedo, rechazo, admiración o entusiasmo, reforzados por estas narrativas, los que van perfilando la concepción de la inteligencia artificial en el ámbito de la educación.

En su obra señera titulada Apocalípticos e Integrados, Umberto Eco trata el tema de la cultura de masas y los medios de comunicación desde dos perspectivas contrapuestas: los apocalípticos, quienes ven los cambios culturales y tecnológicos con pesimismo y temor por la degradación cultural que consideran que traen consigo; y los integrados, que reciben dichos cambios positivamente, viéndolos como una oportunidad para el enriquecimiento y la democratización cultural (Eco, 1968).

> […] mientras los apocalípticos sobreviven precisamente elaborando teorías sobre la decadencia, los integrados raramente teorizan, sino que prefieren actuar, producir, emitir cotidianamente sus mensajes a todos los niveles. El Apocalipsis es una obsesión del dissenter, la integración es la realidad concreta de aquellos que no disienten. (Eco, 1968, pp. 12-13)

Estos puntos pueden conectar con el actual debate tecnófilo y tecnófobo en torno a la AI. Aquellos con una perspectiva tecnófila respecto a la inteligencia artificial serían análogos a los integrados de Eco (1968). Por otro lado, a quienes los autores de este capítulo ubicamos en el espacio entre lo neutral y lo tecnófobo, se encuentran aquellos que se preocupan por las consecuencias negativas de la AI, lo que refleja la actitud de los apocalípticos.

Ruta metodológica

Se empleó un enfoque de análisis de contenido cualitativo de orden inductivo para identificar temas recurrentes, marcos conceptuales y la presencia de argumentos con carga ideológica tecnófila o tecnófoba. El análisis de contenido es una técnica de investigación utilizada para interpretar y codificar información textual, visual o de audio para entender y estudiar significados, temas y patrones subyacentes contenidos en los mensajes. La orientación metodológica utilizada en la construcción de los datos fue cualitativa. En el análisis de mensajes cualitativos, los investigadores se implican de manera profunda en la definición e interpretación de los datos (Neuendorf, 2017).

El análisis estuvo centrado en interpretar significados de contenidos textuales, en la construcción de códigos, categorías y en la estructuración de los temas principales que surgieron de dichos textos, mostrando cuáles son las narrativas y sus posturas ideológicas más discutidas o mencionadas. Para esto, fueron seleccionados cuatro manuales sobre aplicación y usos de la IA en educación superior, representativos de la diversidad en el enfoque y el origen geográfico. La selección se realizó conforme a criterios que incluyen la actualidad, la reputación académica de las instituciones y la prevalencia de uso de estos manuales. A saber: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO),

Universidad Nacional Autónoma de México (UNAM), Universidad de Guadalajara (UDG) y Universidad Autónoma de Baja California (UABC).

Los datos fueron sistematizados con base en la técnica de categorización inductiva propuesta por Romero (2005), centrada en la detección de patrones y recurrencias discursivas en las diferentes publicaciones examinadas. Se tomó como elemento contextual la frecuencia de ciertas palabras, frases, temas o conceptos que aparecieron en el texto; así como algunos patrones de comunicación, es decir, la forma en que se distribuye la información en el texto y qué tipos de discurso o retórica se utilizan.

La revisión de las publicaciones, dio paso a la inducción de dos grandes categorías: Postura tecnófila (PTFL) y Postura tecnófoba (PTFB). Como parte del proceso emergieron algunas características que permitieron integrar temas centrales en cada una de estas (Ver Tabla 1).

Tabla 1. Categorías inductivas emergidas de manuales sobre IA y Educación Superior.

Códigos	Categorías	Temas
PTFL	Postura tecnófila	Optimismo cauteloso
		Promoción tecnológica
		Utopismo tecnológico
PTFB	Postura tecnófoba	Preocupación razonable
		Conservadurismo educativo
		Alarmismo tecnológico

Fuente: elaboración propia.

Resultados

Como resultado del análisis, en este capítulo se propone un eje para pensar las narrativas construidas en torno al debate sobre el uso de la inteligencia artificial en el ámbito universitario; que pretende reflejar, en alguna medida, el estado actual –fines de 2023 e inicios de 2024– de las discursividades que prevalecen y evidentemente contribuyen a configurar los marcos de percepción o, dicho de otra manera, las formas de pensar acerca de la IA.

Este eje puede ser examinado desde una perspectiva sociocultural, a través de un esquema analítico que considera las narrativas que desde una zona neutra se desplazan hacia los márgenes de la tecnofilia y la tecnofobia. Aquí se bosqueja cómo podría estructurarse dicho esquema (Ver figura 1).

Figura 1. Esquema analítico sobre las narrativas en torno a IA y Educación Superior.

Fuente: elaboración propia.

Del punto neutro al extremo de la tecnofilia

En este espectro, se mueve desde una valoración moderada de las tecnologías hasta una exaltación de la IA como dispositivo de soluciones universales. Como primer enclave se tiene el optimismo cauteloso, que integra narrativas tintadas por el reconocimiento de beneficios potenciales, como la

personalización del aprendizaje, sin dejar de mostrar reservas. Para ilustrar lo anterior, se muestra que la Universidad de Guadalajara (2023) exhorta con optimismo a su comunidad:

> Abracemos con entusiasmo y determinación, fusionemos la tecnología y la creatividad para ofrecer experiencias de aprendizaje excepcionales […]" (p.5). "Nos comprometemos a abordar las preocupaciones, mejorar la comprensión y las habilidades de los profesores y fomentar un entorno seguro y eficaz para la utilización de la IA en la educación… No podemos simplemente ignorar las implicaciones a corto y mediano plazo de estas tecnologías para la seguridad, la diversidad del conocimiento, la equidad y la inclusión. (p. 7).

Mientras que en esa misma línea, la UNESCO (2023) relata: "la educación superior puede, sobre todo, ayudarnos a todos a ver a la IA desde un enfoque crítico para ponerla al servicio del desarrollo humano, en lugar de esperar pasiva y descuidadamente cualquier impacto que pueda tener en nuestras vidas" (p. 5); y "la tecnología nunca es neutral: exhibe y favorece inherentemente visiones específicas del mundo, reflejando distintas formas de comprensión y conocimiento" (p. 13).

La idea del optimismo cauteloso se refiere a una actitud que equilibra la anticipación positiva con el reconocimiento de posibles riesgos o consecuencias negativas. Desde esta perspectiva, el optimismo viene de la anticipación de beneficios potenciales que las nuevas tecnologías pueden ofrecer. Implica un enfoque reflexivo, donde se alienta a los individuos y a las sociedades a adoptar nuevas tecnologías, pero, manteniendo un nivel saludable de escepticismo y una postura crítica que impulse a considerar cuidadosamente los posibles efectos secundarios y a desarrollar estrategias que mitiguen los posibles riesgos.

Un segundo cabo entre un punto neutral y la tecnofilia es el que se ha denominado promoción tecnológica, caracterizado por discursividades con un enfoque donde prevalece la integración de la IA para mejorar los resultados educativos, pero muestra cierta sensibilidad hacia a los desafíos que esto conlleva. Por ejemplo, la UNESCO (2023) reconoce y celebra impactos positivos que la IA podría llegar a tener en procesos que son fundamentales en la enseñanza-aprendizaje, en la investigación y en la gestión administrativa de las universidades.

La promoción tecnológica en el contexto de la educación superior, especialmente en la integración de la IA, puede verse como una extensión del pragmatismo técnico, donde la utilidad de la tecnología se mide por su capacidad para mejorar la efectividad y eficiencia en alcanzar los objetivos educativos. La integración de la IA en la educación universitaria promete personalización del aprendizaje, accesibilidad mejorada y optimización de los procesos educativos, lo cual podría teóricamente conducir a mejores resultados de aprendizaje.

La Universidad de Guadalajara establece en su manual que "La IA puede ayudar a automatizar el proceso de inscripción: se acelera el proceso de admisión, se reduce la carga de trabajo del personal administrativo y disminuyen los errores humanos" (UDG, 2023, p.18). Y la UNESCO promueve y difunde una representación que refuerza la relevancia de la tecnología en la vida social:

> […] la IA ha experimentado avances significativos mediante el uso de amplias herramientas de supervisión de datos integradas en el software de gestión del aprendizaje en línea, tecnologías de vigilancia como los sistemas de verificación de exámenes a distancia, y sistemas de seguridad en los campus como el seguimiento de estudiantes y la supervisión de ubicaciones y aplicaciones de rastreo de contactos (Williamson y Hogan en UNESCO, 2023, p.67).

Las representaciones de la promoción tecnológica a través de la integración de la IA en la educación de nivel superior han suscitado debates significativos, porque mientras que la tecnología tiene el potencial de mejorar la enseñanza y el aprendizaje, es vital ser consciente de que no es una panacea y estar atentos a los posibles desafíos sociales, éticos y educativos que surgen con su implementación.

Finalmente, en el extremo del espectro neutralidad-tecnofilia, los autores ubican el punto del utopismo tecnológico, que integra una visión idealizada acerca de la IA. Estas narrativas difunden la idea de que es capaz de transformar radicalmente y mejorar la educación universitaria, sin considerar plenamente sus limitaciones o consecuencias negativas. Los imaginarios que reproducen un Utopismo tecnológico atribuyen a la tecnología y, en este caso particular, a la IA, la capacidad casi mágica de resolver problemas humanos complejos.

En el contexto de la educación superior, se imagina que la IA puede reformar radicalmente las metodologías pedagógicas, personalizar el aprendizaje, aumentar el acceso a la educación de calidad y liberar tiempo para los docentes, facilitando así una educación universitaria para todos. Algunos extractos de los documentos analizados dejan claro que las IES abrazan una expectativa alta en torno a las soluciones que puede aportar el desarrollo tecno digital:

> Reconocemos la fuerza transformadora de la inteligencia artificial (IA) y su potencial para redefinir los diversos procesos que se desarrollan en la educación superior" (UABC, 2023, p.2). "Las potencialidades que ofrece la IA para crear contenido autónomamente, incluyendo imágenes, texto, música, voz y otros tipos de datos, representan un hito histórico, ya que su uso ofrece una oportunidad significativa para el desarrollo humano… (UABC, 2023, p.3).

La IA es una "Herramienta que brinda a la humanidad oportunidades sin precedentes. Es capaz de generar texto, imagen, código, video y audio; pero también, de emitir razonamientos, resolver problemas, interpretar o resumir ideas. Impulsan a las universidades a adaptarnos a las nuevas tecnologías, para mejorar la experiencia educativa y preparar a nuestros estudiantes para los desafíos de un mundo cada vez más interconectado y digital" (UDG, 2023, p. 4). Sin embargo, esta visión puede caer en un exceso de entusiasmo tecnológico, es decir, de tecnofilia, por no contemplar de forma crítica las limitaciones inherentes y las posibles consecuencias no deseadas de la implementación de la IA.

Del punto neutro al extremo de la tecnofobia

En este rango, hay desplazamientos desde una precaución informada sobre las limitaciones y riesgos de la IA hasta un rechazo ideológico y generalizado de la tecnología. Se ubica como primer sitio, en mayor cercanía a la neutralidad, la denominada Preocupación razonable, éste se sustenta en el reconocimiento explícito de los peligros potenciales, como la deshumanización de la educación o la infravaloración de habilidades no técnicas.

Las narrativas que construyen simbólicamente una preocupación razonable, encapsulan una postura reflexiva que acoge la tecnología reco-

nociendo sus posibles aportaciones positivas, pero con una mirada crítica que advierte de los riesgos potenciales. La preocupación razonable se fundamenta en el examen de la tecnología desde una perspectiva ética y humanista, cuestionando cómo la integración de la IA podría afectar aspectos fundamentales de la enseñanza y el aprendizaje en el nivel superior.

Así lo ilustran estos fragmentos discursivos, "Las IES deben liderar el camino para garantizar que la IA trabaje para los humanos, que los objetivos y valores de las personas sean su núcleo y que se use para apoyar la sostenibilidad del planeta" (UNESCO, 2023, p. 10). "Esta tecnología aporta elementos que pueden convertirse en disruptivos para el proceso de enseñanza y aprendizaje, cuando se utilizan en un marco ético, responsable y bien informado" (UNAM, 2023, p. 5).

El segundo punto entre la zona neutralidad-tecnofobia, lo ocupan las discursividades que privilegian el mantenimiento y promoción de métodos pedagógicos tradicionales y destacan un escepticismo hacia las influencias de la IA en el proceso educativo; los autores de este capítulo le han denominado Conservadurismo educativo. Este personifica una deferencia hacia las prácticas establecidas y una resistencia al cambio impulsado por la IA en las tareas sustantivas universitarias. En su base, hay una preocupación por preservar los elementos valiosos de la educación tradicional y un escepticismo hacia la habilidad de las máquinas para mejorar realmente los procesos de enseñanza y aprendizaje.

> Reconocemos que la implementación de la IA debe ir acompañada de una supervisión constante y de la toma de medidas proactivas para garantizar que se utilice de manera ética y equitativa, y que beneficie a todos los miembros de nuestra comunidad académica. (UABC, 2023, p. 8)

> El anterior extracto y los siguientes, demuestran una narrativa orientada al reconocimiento de lo humano frente al vasallaje tecnológico "Si bien la IA brinda oportunidades para apoyar a los docentes, la interacción humana debe seguir siendo el foco de la educación, porque las máquinas no pueden desplazar a los docentes" (UNESCO, 2023, p. 22). "La tecnología no es el factor determinante de la experiencia de aprendizaje sino un facilitador del proceso y puede usarse en diferentes contextos" (UNESCO, 2023, p. 23).

Las voces del Conservadurismo educativo argumentan que la educación superior es una práctica no solo de transmisión de conocimientos, sino también de formación de caracteres y ciudadanos responsables, una labor que no se puede entregar completamente a manos de algoritmos. La preocupación subyacente es que los avances tecnológicos puedan simplificar o trivializar la complejidad de los procesos educativos humanos. Por último, en el extremo del espectro de la tecnofobia, se sitúa el Alarmismo tecnológico, el cual recoge e integra una visión radicalmente catastrófica, que imagina a la IA como un elemento disruptivo y amenazante. Esta postura representa una postura extremadamente crítica de la tecnología, advirtiendo sobre consecuencias negativas significativas y a menudo irreversibles que la IA podría tener en la sociedad y la educación. Para ejemplificar estas posturas sobre la presencia de riesgos y amenazas tecnológicas advertidas o imaginadas, se muestran un par de enunciaciones:

> Se estima que la IA generativa tiene el potencial de sustituir hasta a una cuarta parte de la mano de obra actual, lo que, si se extrapola a escala mundial, sugiere que alrededor de 300 millones de puestos de trabajo de tiempo completo podrían ser automatizados debido a la IA generativa. (UNESCO, 2023, p. 44)

> Las IA generativas aún están en desarrollo y siempre existe el riesgo de que la salida sea inexacta o poco confiable. Esto podría llevar a que los estudiantes aprendan información incorrecta o sean engañados... Las IA se entrenan en grandes conjuntos de datos de texto y código. Si estos conjuntos de datos están sesgados, los modelos también estarán sesgados. Esto podría llevar a que los estudiantes estén expuestos a información parcial o discriminatoria… La IA podría usarse indebidamente con fines maliciosos, como crear noticias falsas o generar spam. Esto podría tener un impacto negativo en la educación de los estudiantes y en la sociedad en su conjunto… (UDG, 2023, p. 21)

Las construcciones discursivas del alarmismo tecnológico proyectan escenarios en los que la IA podría erosionar el tejido educativo, cultural, e incluso moral, llevando a la pérdida de empleos, disminución del pensamiento crítico y la autonomía, y una creciente dependencia de sistemas que pueden ser manipulados o que funcionan sin una comprensión ni res-

peto por los valores humanísticos. Estas preocupaciones a veces pueden estar basadas en supuestos equivocados o entender mal la tecnología, pero apuntan hacia la necesidad de regulación, supervisión y diálogo público sobre el diseño, implementación y gobernanza de la IA.

En el mismo sentido, se retrotraen las siguientes imágenes discursivas: "La IA generativa ha puesto en el escenario nuevas formas de deshonestidad académica, en las que es posible presentar textos generados por estos sistemas como si fueran de autoría propia o utilizarlos de manera inapropiada" (Chan en UNAM, 2023, p. 13).

> La IA debería estar al servicio de la inclusividad, pero se ha constatado que la interseccionalidad suele simplificarse excesivamente en un análisis dual de raza y género, en lugar de entenderse como un análisis basado en estructuras o una evaluación política crítica. La IA tampoco aborda la complejidad de los sistemas de opresión entrelazados. (UNESCO, 2023, p. 66)

Es crucial examinar estas posturas no solo en términos de su actitud hacia la tecnología, sino también en función de su comprensión de la naturaleza y los objetivos de la educación universitaria, su concepción de la agencia humana y cómo estos elementos interactúan con el desarrollo y la implementación de la IA en entornos educativos. Cada una de estas perspectivas ofrece *insights* valiosos y advertencias que deben ser consideradas cuidadosamente para guiar la integración ética y efectiva de la IA en la educación superior.

Conclusiones

Acorde con los resultados del análisis realizado y con las tesis de los autores revisados en el apartado introductorio, es importante reconocer que las culturas digitales contemporáneas integran narrativas, infraestructuras tecnológicas y capacidades físicas. Es menester adoptar una postura crítica y reflexiva que considere tanto el potencial como los peligros de la tecnología, como nos lo ha enseñado el caso del desarrollo de Internet y su impacto en la vida social. Ahora en el contexto de la incorporación de la IA en la educación superior, cobra relevancia entender su capacidad para mejorar la sociedad, al tiempo que debe mantenerse una vigilancia crítica

sobre sus implicaciones éticas, sociales y culturales (Eco, 1968; Hine, 2022; Bollmer, 2018).

Las preguntas planteadas al inicio de este capítulo han quedado respondidas mediante el esquema analítico presentado en el apartado anterior. Este intenta mostrar que las posturas respecto a la IA son diversas, complejas, dinámicas y que se traslapan en muchas ocasiones. Individuos e instituciones educativas pueden moverse a lo largo del espectro respondiendo a experiencias, datos y argumentos nuevos. Es indispensable, entonces, mantener una actitud de apertura tanto para reconocer los beneficios de la IA como para ser conscientes de sus posibles desventajas y peligros.

Expertos en tecnología, tomadores de decisiones en el ámbito político y académicos de las ciencias humanas y sociales, pueden contribuir al desarrollo e integración de la IA en la sociedad, mediante la construcción de agendas de trabajo transdisciplinares; sin renunciar al cuestionamiento y evaluación crítica de los supuestos subyacentes de nuestras herramientas tecnológicas, de una manera que se respete la dignidad humana y se potencie el bienestar colectivo.

Referencias

Bollmer, G. (2018). Theorizing Digital Culture.

Eco, U. (1968). Apocalípticos e integrados. Lumen.

Fernández, C. (2015). Participación juvenil y activismo digital. Una introducción a una nueva agenda de estudio. *TecCom Studies: Estudios de Tecnología y Comunicación*, s. d. (6), 7-18.

Floridi, L. (2015). The Online Manifesto: Being Human in a Hyperconnected Era. Springer.

Hine, C. (2022). *Etnografía para el internet: incorporado, personificado y cotidiano*. Universidad Veracruzana.

Neuendorf, K. (2017). The Content Analysis *Guidebook*. SAGE.

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2023). Oportunidades y desafíos de la era de la inteligencia artificial para la educación superior. Una introducción para los actores de la educación superior.

Romero, C. (2005). La categorización: un aspecto crucial en la investiga-

ción cualitativa. *Investigaciones*, 11. https://biblioteca.unicesmag.edu.co/digital/revinv/01231340v11n11pp113.pdf

Siles, I. (2004). Sobre el uso de las tecnologías en la sociedad. Tres perspectivas teóricas para el estudio de las tecnologías de la comunicación en Reflexiones. *Revista de la Facultad de Ciencias Sociales de la Universidad de Costa Rica*, 83(2), 73-82. https:// revistas.ucr.ac.cr/index.php/reflexiones/article/view/11404

Universidad Autónoma de Baja California (UABC). (2023). *Orientaciones iniciales sobre el uso académico de la Inteligencia Artificial (IA)*. https://citecuvp.tij.uabc.mx/wp-content/uploads/2023/12/Orientaciones-Iniciales-sobre-el-uso-academico-de-la-IA-.pdf

Universidad de Guadalajara (UDG) (2023). *Guía práctica: Orientaciones y definiciones sobre el uso de la inteligencia artificial generativa en los procesos académicos*. https://www.udgvirtual.udg.mx/Guia-IA-UDG

Universidad Nacional Autónoma de México (UNAM) (2023). *Recomendaciones para el uso de la inteligencia artificial generativa en docencia*. https://iagenedu.unam.mx/docs/recomendaciones_uso_iagen_docencia_unam_2023.pdf

Capítulo 8

Representaciones Sociales e Imaginarios Sociales Sobre la IA en el Entorno Educativo

Ulises Suárez Estavillo
Héctor Ignacio Castañeda García

https://doi.org/10.61728/AE24001083

Introducción

Desde los albores del cine hasta las narrativas contemporáneas, la inteligencia artificial (IA) ha tejido un tapiz complejo de imaginarios en la mente colectiva. Desde la mítica Metrópolis (Lang, 1927) hasta el éxito taquillero de Avengers: Age of Ultron (Weedon, 2015), la humanidad ha proyectado sus sueños y temores en la figura de la IA. Estas construcciones simbólicas, moldeadas por obras icónicas como Yo, Robot (Asimov, 1950) y Terminator (Cameron, 1984), han evolucionado para convertirse en representaciones sociales arraigadas en la psique colectiva.

La presente investigación pretende hacer una exploración de las representaciones sociales sobre la IA, enfocándose en un contexto educativo específico: la Licenciatura en Ciencias de la Educación de la Universidad Autónoma de Sinaloa. Desde la llegada de ChatGPT en noviembre de 2022, las conversaciones en torno a la IA se han intensificado, y estas representaciones sociales se han vuelto más palpables que nunca.

A lo largo de la investigación se pretende analizar cómo la relación entre educación e IA se convierte en el epicentro de la reflexión y el análisis. ¿Cómo influyen las representaciones sociales sobre la IA en la forma en que estudiantes y maestros conciben su papel en la educación del siglo XXI? ¿Cómo estas narrativas moldean las prácticas pedagógicas y las expectativas futuras?

De igual manera, se explorarán las complejidades de estas construcciones simbólicas, examinando cómo las representaciones sociales se entrelazan con la formación educativa. Desde la utopía hasta la distopía, se analizan las inquietudes que dan forma a la relación entre la comunidad académica y la inteligencia artificial.

Abordaje metodológico

Como premisa de orientación para esta investigación, se plantea que los imaginarios colectivos y las representaciones sociales que la comunidad educativa tiene en relación con la IA generan mecanismos de resistencia, sino que negación al uso de estas herramientas dentro del proceso educativo. Los esquemas mentales preexistentes dentro de los miembros de la comunidad impiden una relación funcional con la tecnología generando con esto que las brechas en el alfabetismo digital se vuelvan más amplias.

Por otro lado, al ser el grupo de control un entorno universitario que cuenta entre sus objetivos con la formación de docentes capacitados, la reproducción de estas resistencias basadas en las representaciones sociales sobre la IA, generan un círculo vicioso que imposibilita un acceso óptimo de los futuros docentes al uso de estas herramientas de aprendizaje. Docentes que resisten, formando docentes que resistirán.

Como plantea Knapp (2003), las representaciones sociales son creencias que se comparten al interior de un colectivo que permiten vislumbrar los alcances ideológicos del grupo. A través del entendimiento de estas creencias se pueden comprender las conductas y los mecanismos de socialización que el colectivo tiene, permitiendo explicaciones puntuales sobre las acciones de este. De esta forma, las representaciones sociales, encierran el cúmulo de creencias que movilizan la acción colectiva, y permiten acceder a los mecanismos de entendimiento que se tienen sobre un fenómeno determinado, en este caso la IA y su relación con la educación.

Tomando en cuenta lo anterior, se tiene la hipótesis de que a partir de las representaciones sociales que tienen los miembros de una comunidad universitaria (profesores, alumnos) en torno a la IA, existen mecanismo que permiten la apropiación diferenciada de este tipo de herramientas dentro de los procesos de aprendizaje. De igual manera, se plantea que, ante lo nobel de la tecnología, las representaciones sociales existentes sobre la IA se encuentran en proceso de construcción, modificando en tiempo real la configuración mental de la misma.

Múltiples son las herramientas metodológicas que se han implementado para acceder a las representaciones sociales que presenta un colectivo; mientras que Moscovici (1979) y Jodelet (2003) ven en la entrevista un

instrumento fundamental para esta tarea, Abric (2011) apuesta por un enfoque etnográfico en el cual a través la inmersión en el contexto social y la observación participante se puede acceder a las prácticas cotidianas de los individuos, logrando una mayor apropiación de los núcleos figurativos que constituyen sus representaciones sociales.

Para efectos de este trabajo, en un primer momento se realizó una encuesta entre profesores y estudiantes de la licenciatura en Ciencias de la Educación, de la Universidad Autónoma de Sinaloa, con el fin de conocer el nivel de familiarización de los miembros de la comunidad con el uso de la IA dentro del proceso educativo. De igual manera se realizaron entrevistas no estructuradas entre algunos miembros de la misma comunidad para conocer las representaciones sociales que estos presentan con relación al mismo fenómeno.

En esta primera etapa se hizo uso de un enfoque hermenéutico a través del cual se logró acceder a elementos simbólicos, de construcción del lenguaje e imaginarios a través de los cuales los individuos han configurado la realidad en la que se encuentran inmersos. Este ejercicio permitió encontrar puntos en común sobre la manera en cómo representan el uso de la IA dentro de los procesos educativos.

En un segundo momento, se realizó un trabajo de observación participante a través del cual se pudo acceder a la cotidianidad de los grupos. Esto con el fin de determinar si los niveles de familiaridad con el uso de las tecnologías declarado en encuestas y entrevistas correspondían con la práctica educativa diaria. A través de esta observación in situ se generó una valoración cualitativa del nivel de implicación que los miembros de la comunidad presentan en relación con el uso de la IA, así como el nivel de influencia que las representaciones sociales tienen sobre esta.

Aproximaciones teóricas sobre representaciones sociales de la IA en educación

Propuesto por Moscovici (1979), y desarrollado a profundidad por Jodelet (2003), el concepto de representaciones sociales puede entenderse como esquemas de pensamiento socialmente construidos que actúan como lentes a través de las cuales los individuos y grupos buscan interpretar su

entorno. Sin embargo, estas construcciones simbólicas no solo encapsulan conocimientos colectivos, sino que también influyen en las actitudes, valores y comportamientos dentro de una comunidad. Se forman y evolucionan mediante el diálogo constante entre miembros de la sociedad. De esta manera, se puede decir, que las representaciones sociales generan acción o inacción dentro de los colectivos.

Empero, tal como se planteó anteriormente, debemos preguntarnos ¿cuál es el proceso de psique colectiva que lleva a configurar un imaginario como una representación social? En este sentido, es importante definir seis elementos de carácter teórico de ambos conceptos: imaginario social, socialización, consolidación, incorporación al discurso social, institucionalización y representación social.

Imaginario social

El imaginario social en torno a la inteligencia artificial (IA) ha sido moldeado y construido a lo largo del tiempo a través de diversos medios de expresión cultural, como el cine, la televisión y la literatura. Desde sus primeras representaciones en la pantalla grande hasta las más recientes, las IA han sido retratadas de diversas maneras, alimentando así el imaginario colectivo en la cultura popular.

Películas icónicas como Metrópolis (1927), de Fritz Lang, marcaron el inicio del imaginario social sobre la IA, presentando una visión distópica de un futuro donde las máquinas dominan a la humanidad. Este filme sentó las bases para la representación de las IA como entidades poderosas y a menudo amenazantes en la cultura popular. A lo largo de las décadas, esta representación ha evolucionado, adaptándose a los cambios sociales y tecnológicos de cada época. Sin embargo, cada una de las veces que ha sido retratada, la narrativa sobre esta sigue el mismo rumbo: una herramienta que al principio es útil, pero a través de su desarrollo desafía el bienestar de la raza humana.

Según Castoriadis (1975), el imaginario social puede entenderse como "el conjunto de significaciones y símbolos que orientan y dan sentido a la vida en sociedad", lo cual resalta la importancia de estas representaciones en la construcción de la realidad colectiva. Así pues, aun y cuando la lejana

obra de Fritz Lang se encontraba distante de los procesos de enseñanza aprendizaje, se sembró la semilla de un mundo dominado por la tecnología que impactaría irremediablemente todos los aspectos de la vida en sociedad.

Socialización y comunicación

En esta etapa del proceso, las representaciones sociales sobre la inteligencia artificial (IA) se difunden y se desarrollan a través de la interacción social y la comunicación. Los medios de comunicación desempeñan un papel crucial en la difusión de estas representaciones, ya que las películas, programas de televisión, libros y otros medios populares actúan como vehículos para la transmisión de ideas y conceptos sobre la IA.

Además de los medios de comunicación, la interacción personal y la discusión en grupos sociales también desempeñan un papel importante en la formación de representaciones sociales sobre la IA. Las conversaciones informales entre amigos, familiares y colegas pueden influir en cómo se percibe y se entiende la IA, ya sea como una herramienta beneficiosa para la sociedad o como una amenaza. En este proceso de socialización y comunicación, las representaciones sociales sobre la IA se enriquecen y se moldean continuamente a través del intercambio de ideas y experiencias entre individuos y grupos sociales.

Como señalan Smith y Jones (2018), "los medios de comunicación desempeñan un papel fundamental en la formación de las representaciones sociales sobre la inteligencia artificial, al proporcionar narrativas y metáforas que influyen en la percepción pública de este fenómeno tecnológico" (p. 45). En este escenario es posible entender cómo, ante la llegada de herramientas de IA aplicadas a los entornos educativos, los imaginarios que habían desarrollado con el tiempo los miembros de una comunidad fueron modelando las representaciones sociales que se tienen actualmente. Si se considera que en las comunidades educativas coexisten individuos de distintas generaciones, los núcleos figurativos en que se entrelazan son diversos, generando una ideología particular en torno al fenómeno.

Consolidación de significados

En esta etapa, las representaciones sociales sobre la IA comienzan a consolidarse y adquirir significados compartidos dentro de la sociedad. A medida que las personas interactúan entre sí y discuten sobre la IA en diversos contextos, se establecen interpretaciones comunes y se atribuyen significados a este fenómeno tecnológico. Por ejemplo, la IA puede ser percibida como una herramienta innovadora que ofrece soluciones prácticas a problemas complejos, o como una fuerza disruptiva que plantea dilemas éticos y sociales.

La consolidación de significados también se ve influenciada por factores culturales, históricos y sociales que moldean la forma en que se percibe y se comprende la IA en diferentes contextos. En este sentido, cabe preguntarse, si el nivel de relación de las comunidades estudiantiles de educación en México es el óptimo para percibir a la IA desde una actitud optimista. O si la brecha tecnológica de los estudiantes en relación con la de países con mayor desarrollo tecnológico les hace ver a la IA como una amenaza para su futuro laboral.

La consolidación de significados en las representaciones sociales sobre la IA refleja el proceso dinámico mediante el cual las ideas y percepciones individuales se fusionan para formar una comprensión colectiva de este fenómeno tecnológico. Este proceso continuo de construcción y atribución de significado es fundamental para la forma en que la sociedad interactúa con la IA y aborda sus implicaciones en diferentes esferas de la sociedad.

Incorporación en el discurso social

En esta fase del proceso, las representaciones sociales sobre la IA se integran en el discurso social, convirtiéndose en temas recurrentes de discusión y debate en diversos ámbitos de la sociedad. El discurso sobre la IA se extiende más allá de los círculos académicos y tecnológicos, llegando a la esfera pública a través de medios de comunicación, redes sociales, política y cultura popular. Según Smith y Johnson (2020), "el discurso social desempeña un papel crucial en la construcción de la percepción pública de la inteligencia artificial, al influir en cómo se presenta y se debate este

fenómeno en la esfera pública" (p. 78). México, al ser un país en donde las autoridades educativas presentan un perfil que puede considerarse tradicional, los debates en torno de la IA pueden llegar a ser fatalistas con relación a aspectos como el desempleo por la automatización, o los dilemas ético-morales que representa el uso de la herramienta.

Institucionalización y estabilización

En esta etapa del proceso, las representaciones sociales sobre la IA se vuelven institucionalizadas y estabilizadas dentro de la sociedad, integrándose en las estructuras y normas sociales existentes. Las instituciones, como el gobierno, la industria, la educación y los medios de comunicación, desempeñan un papel fundamental en este proceso al adoptar y promover ciertas representaciones sobre la IA y alentar su aceptación y legitimidad en la sociedad.

La institucionalización de las representaciones sociales sobre la IA se refleja en políticas, regulaciones y prácticas que abordan este tema en diversos contextos. La estabilización de las representaciones sociales sobre la IA también se produce a través de la reproducción y reforzamiento continuo de estas ideas en la sociedad.

Según García y Martínez (2019), "la institucionalización de las representaciones sociales sobre la inteligencia artificial refleja el proceso mediante el cual estas ideas se incorporan y legitiman en las estructuras y prácticas sociales existentes" (p. 102). En el caso del objeto de estudio, la apertura de instituciones al uso de las IA, la necesidad de estudiantes de entornos de aprendizaje innovadores, han llevado a que la implementación de esta tecnología comience a definir de manera acelerada las representaciones sociales de las comunidades educativas.

La Intersección de la inteligencia artificial y la enseñanza de nivel superior

Mucho se habla de la historia de la educación y cómo esta fue impulsada mayoritariamente por la bien cimentada tecnología contemporánea y los grandes avances que de esta relación se produjeron. También se puede mencionar que no todo es perfecto, ya que hablar de perfección sería muy

osado en este capítulo introductorio debido a la encrucijada que se vive actualmente dentro y en los alrededores de la educación superior, y más enfáticamente con el contundente arribo de la IA como la revolución tecnológica del primer cuarto del siglo XXI. Sin menospreciar el impacto en otras áreas del saber, en lo que define a la educación, se está siendo testigo de una profunda "Transformación Docente". La inteligencia artificial (IA) no solo está tocando la puerta de nuestras aulas, sino que está remodelando la forma en que se entiende y se practica la enseñanza. Esta metamorfosis no se limita a la automatización de tareas; abarca la personalización del aprendizaje, donde la IA se convierte en una herramienta colaborativa, adaptándose a los estilos y ritmos individuales de los estudiantes.

Tal como menciona Eglis Chacón (2023) en la enseñanza a través de la I.A. "Gracias a su capacidad para generar y evaluar información, ChatGPT puede desempeñar diversas funciones en los procesos de enseñanza y aprendizaje". Junto con otras formas de IA, ChatGPT podría mejorar el proceso y la experiencia de aprendizaje de los estudiantes. Para ello, ChatGPT puede utilizarse como herramienta independiente o integrarse en otros sistemas y plataformas utilizadas por las IES (Instituciones de Educación Superior).

Al adentrarse en esta intersección entre maestros y tecnología inteligente, la cual hace un llamado a la reflexión, no solo de los maestros sino de toda la comunidad que abarca las instituciones educativas de todos los niveles, nos encontramos con un nuevo paradigma donde la adaptación es esencial, porque se puede afirmar en algún punto de este artículo que la resistencia a la adopción de tecnologías inteligentes puede convertirse en un obstáculo para el avance educativo. La IA no reemplaza la esencia del maestro, sino que redefine muy acertadamente la función y papel dentro y fuera del aula, permitiendo que la pedagogía evolucione hacia un modelo más centrado en el estudiante. La inteligencia artificial no es una amenaza; es una aliada que ofrece herramientas para un aprendizaje más flexible, dinámico e inclusivo.

En este nuevo paisaje educativo en el que los docentes de la enseñanza superior están siendo partícipes, la IA no solo está transformando cómo se enseña, sino también qué se enseña. Citando un ejemplo concreto de esta transformación, se puede abordar con la profundidad y diversidad

que permite la personalización del contenido académico, como menciona (Continental, 2023):

Cada estudiante podría disfrutar de un aprendizaje personalizado según sus necesidades individuales. Asimismo, tendría acceso a contenidos adaptados, feedback inmediato y orientación personalizada. Por otro lado, puede facilitar el trabajo de los maestros, automatizando tareas administrativas o repetitivas, proporcionando herramientas de evaluación o análisis, o generando recursos didácticos innovadores.

Este es uno de los mayores beneficios de la integración de la IA en la educación superior, y así podríamos mencionar un sinfín de trasversalidades, en las cuales la IA ofrece ventajas visibles a la educación.

Otra cualidad destacable de esta transformación, inmersa en el dominio del docente y con sede principal dentro de cada aula, es la adaptación de los materiales y métodos de enseñanza a las necesidades individuales de cada estudiante. Esto primordialmente se vuelve posible a una escala sin precedentes. La inteligencia artificial se convierte, pues, en una guía personalizada para cada estudiante, proyectada, planificada y estructurada desde el bastión docente, ofreciendo un aprendizaje a medida cuya finalidad última es ajustarse al propio ritmo y estilo único del estudiante.

Sin embargo, en este viaje hacia la transformación académica no solo por la esencia escolar, sino por la representación que la IA marca en el proceso de ser docente, también se encuentran desafíos multidisciplinares los cuales se deben focalizar y en su caso enfocar el esfuerzo para atender y maximizar los beneficios de la llegada de la IA. Con esto se plantea que los docentes deberán abordar no solo la tecnología en sí como un mero placebo áulico o una tecnología de entretenimiento estudiantil que convierte las aulas en una mera continuación tradicional de lo que es saber y aprender, sino también la dinámica cambiante del aula y las expectativas de los estudiantes. Puesto que, la resistencia inicial de algunos docentes (no en su mayoría), puede deberse a la falta de comprensión sobre cómo la IA puede ser una herramienta complementaria y no una amenaza para el futuro de su rol de profesor o ser vista como el ocaso del dominio impráctico del deber ser docente. La preparación y la formación.

En última instancia, la intersección de la inteligencia artificial y la enseñanza superior plantea interrogantes fundamentales sobre la naturaleza

misma de la educación. ¿Cómo se puede equilibrar la eficiencia de la IA con la necesidad de una conexión humana genuina en el aula? ¿En qué medida se debe permitir que la tecnología influya en la pedagogía sin sacrificar la esencia de la enseñanza? Estas preguntas invitan a quienes están involucrados en toda la estructura del proceso educativo (ya no solo a maestros) a una reflexión profunda sobre el futuro de la educación y cómo los maestros pueden abrazar la transformación sin perder la esencia de su impacto positivo en la vida de los estudiantes.

La coevolución de maestros y tecnología inteligente

Estamos en medio de una revolución educativa donde la coevolución entre maestros y tecnología inteligente se revela como una danza intrincada de adaptación mutua. Se imagina un aula donde la tecnología no es solo un complemento, sino un compañero activo en la enseñanza. Este cambio va más allá de la adopción de herramientas; es una transformación profunda en la que maestros y tecnología se influyen y moldean mutuamente.

En este baile, las representaciones sociales de la inteligencia artificial se convierten en un factor crucial. Para algunos, la IA puede ser vista como una amenaza (un gran cúmulo de docentes), una sombra que se cierne sobre la esencia humana de la enseñanza y sobre el adoctrinamiento tradicionalista que se esparce en el aula. Sin embargo, para otros, es una oportunidad para liberar a los maestros de tareas repetitivas y permitirles centrarse en lo que hacen mejor: inspirar, guiar y nutrir a las mentes jóvenes que son la visión y misión del maestro 3.0 hoy por hoy. Por lo tanto, la percepción de la IA en la educación depende en gran medida de cómo se integra en el tejido de la enseñanza, y cómo se comunica o qué canales se utilizan para esa integración a la comunidad educativa.

En esta coevolución, se observa una transición (que en la actualidad avanza casi a marchas forzadas por el lado del docente y por su contraparte con un deseo feroz por ocultar las IAs a los docentes para su uso libre y sin ética aparente), con una tendencia positiva más que negativa, hacia un modelo más participativo. La tecnología no reemplaza al maestro, sino que se convierte en un socio colaborativo que permite llegar a rozar las fronteras de su propio intelecto de una manera mucho más clara, más nítida y que permite una mayor caracterización de sus alcances como maestro.

Las herramientas de inteligencia artificial se ajustan a los métodos de enseñanza del maestro, ofreciendo sugerencias personalizadas y realimentación en tiempo real. La dinámica aula-maestro-IA se convierte en un trío armonioso, donde cada elemento se potencia mutuamente (recordar la hipótesis del inicio de este párrafo), sin embargo, este viaje no está exento de desafíos. La resistencia al cambio, la falta de capacitación y la incertidumbre sobre el impacto real de la IA en el aprendizaje son preocupaciones legítimas, por tanto, la coevolución exitosa requiere un equilibrio delicado entre la adopción de tecnología y la preservación de la conexión humana en el aprendizaje.

En última instancia, la coevolución entre maestros y tecnología inteligente plantea preguntas fundamentales sobre la esencia misma de la enseñanza. ¿Cómo se equilibra la innovación con la conexión humana? ¿Cómo se asegura que la IA mejore y no distorsione la experiencia educativa? La reflexión crítica sobre estas cuestiones es esencial, ya que abrazar la coevolución significa no solo adoptar nuevas herramientas, sino también replantear la comprensión de lo que significa ser educador en el siglo XXI. En este cruce de caminos, la coevolución se convierte en un viaje de autodescubrimiento y adaptación constante, donde la esencia de la enseñanza encuentra nuevas formas de florecer.

En el umbral de la innovación: La inteligencia artificial en la educación superior

Hoy en día, los saltos tecnológicos no los marcan determinadas compañías, ni determinados usos que les demos a las NTIC (Nuevas Tecnologías); más sin embargo, gracias a la llegada de la inteligencia artificial (IA), nos encontramos en el umbral de una revolución educativa, donde la IA está irrumpiendo en nuestras instituciones de educación superior. Este fenómeno va más allá de simplemente adoptar nuevas tecnologías; estamos presenciando la redefinición de cómo se aprende y enseña. En palabras sencillas, la IA está desbloqueando puertas hacia una innovación sin precedentes en la educación, aunque siendo objetivos, a muchos les causará revuelo la idea.

La clave de esta revolución es la manera en que la IA personaliza el aprendizaje. No es solo una cuestión de automatizar tareas, sino de adap-

tar la enseñanza a las necesidades individuales de cada estudiante. Como menciona (Mujica, 2020):

La integración sistemática de la inteligencia artificial en la educación tiene la capacidad potencial para hacer frente a algunos de los mayores desafíos del mundo actual, como son el elaborar prácticas de enseñanza y aprendizaje innovadoras, sacar provecho del potencial de la IA con un enfoque centrado en el ser humano, reorientar el debate en la lucha contra las desigualdades actuales en materia de acceso al saber, a la investigación y a la diversidad de las expresiones culturales, y sobre todo hacer que la IA no acentúe las diferencias tecnológicas entre los países y dentro de estos.

La representación social de la IA en la educación es un factor clave. A menudo, se percibe como una herramienta de vanguardia que promete mejorar la calidad del aprendizaje. Sin embargo, también hay temores y percepciones erróneas. Algunos pueden ver a la IA como un reemplazo para los maestros, cuando en realidad es un complemento que potencia sus habilidades.

En el umbral de esta revolución, la retroalimentación es esencial. Es un recordatorio de que la tecnología no es un fin en sí misma, sino un medio para mejorar la educación. La reflexión crítica debe llevarnos a cuestionar cómo se puede equilibrar la eficiencia de la IA con la esencia de la enseñanza humana. ¿Se puede abrazar la innovación sin sacrificar la conexión personal y la empatía en el aula? La IA es una herramienta poderosa, pero su éxito en la educación superior depende de cómo se la integre de manera consciente y reflexiva en los enfoques educativos, como se menciona en (Europea, 23), "Los profesores pueden utilizar herramientas de inteligencia artificial (IA) y personalizar la evaluación formativa para hacer valoraciones individualizadas en tiempo real que respondan a las necesidades particulares de aprendizaje. La IA puede reducir la subjetividad asociada a la calificación humana y reducir el tiempo dedicado a las tareas de calificación." Este es un viaje que invita a repensar, reinventar y, sobre todo, a no perder de vista el verdadero propósito de la educación: el desarrollo integral de los individuos en su viaje de aprendizaje.

Representaciones sociales de la IA en una comunidad universitaria

A manera de conclusión, como se mencionó anteriormente, el ambiente académico de la licenciatura en Ciencias de la Educación de la Universidad Autónoma de Sinaloa, es el espacio en donde se desarrolló el estudio de caso a través cual se exploran las representaciones sociales que se tienen sobre la IA. Al ser un entorno educativo en el cual coexisten docentes que forman a futuros docentes, resulta relevante conocer sus creencias, resistencias y prácticas en torno al uso de las nuevas herramientas tecnológicas.

Uso de la IA en el proceso de enseñanza-aprendizaje

El estudio de encuestas realizado sobre el uso de la IA en el entorno educativo mencionado reveló una tendencia predominante entre los docentes: la mayoría de ellos manifiestan un nivel limitado de familiaridad con las tecnologías de IA. Estos resultados sugieren una brecha significativa en el conocimiento y la comprensión de cómo la IA puede integrarse en el ámbito educativo. La falta de familiaridad con estas tecnologías puede ser un obstáculo para su adopción efectiva y su aprovechamiento en el proceso de enseñanza-aprendizaje.

Además, los docentes participantes en las encuestas expresaron la percepción de que el beneficio proporcionado por la IA en el proceso educativo es bajo. Esta opinión refleja una preocupación sobre la efectividad y la relevancia de la IA en el ámbito académico. La baja valoración de los beneficios potenciales de la IA puede estar relacionada con la falta de comprensión sobre cómo estas tecnologías pueden mejorar la experiencia de aprendizaje de los estudiantes y optimizar la labor docente.

Finalmente, otra observación importante derivada del estudio de encuestas es la incertidumbre de los docentes respecto a la posibilidad de que la IA eventualmente sustituya sus empleos. Esta inquietud refleja preocupaciones sobre el impacto futuro de la IA en el ámbito laboral de la educación. La duda sobre la viabilidad de la IA para reemplazar roles docentes sugiere la necesidad de un diálogo más amplio y una mayor comprensión de cómo estas tecnologías pueden complementar y mejorar la labor educativa en lugar de suplantarla.

En el caso del alumnado, el análisis de encuestas sobre el uso de la IA reveló una situación similar. En general, se observa que muy pocos estudiantes están familiarizados con las aplicaciones y el alcance de la IA en el contexto educativo. La escasa relación con la IA por parte de los alumnos puede deberse a una falta de exposición a estas tecnologías en el entorno educativo o a una falta de conciencia sobre su potencial.

A pesar de la baja familiaridad, muchos estudiantes consideran que la IA podría ser una herramienta poderosa para personalizar la educación. Esta percepción resalta la visión optimista de los alumnos sobre el potencial de la IA para adaptarse a sus necesidades individuales de aprendizaje y ofrecer experiencias educativas más personalizadas y efectivas. Sin embargo, esta expectativa de personalización contrasta con la realidad de que un porcentaje significativo de alumnos percibe un uso casi nulo de IA dentro de las interacciones con sus docentes. Esta discrepancia entre la percepción y la práctica resalta la necesidad de una mayor integración y aprovechamiento de la IA en el currículo educativo para satisfacer las expectativas de los estudiantes en cuanto a la personalización del aprendizaje.

Representaciones sociales de las IA en la práctica educativa diaria

Durante un estudio exhaustivo que involucró la observación directa de docentes y alumnos en su entorno educativo, se identificaron percepciones y actitudes significativas hacia el uso de IA. En particular, los docentes mostraron un nivel preocupantemente bajo de familiaridad con las IA, como evidenció su participación en un curso específico diseñado para abordar esta tecnología emergente.

Las representaciones sociales de los docentes sobre las IA se centraron en preocupaciones éticas y en la posibilidad de que estas tecnologías pudieran desencadenar un proceso de desaprendizaje en lugar de promover un desarrollo académico auténtico. Hubo una clara inquietud sobre el papel de los alumnos en la utilización responsable de las IA, resaltando en todo momento el peligro de timo por parte del alumnado, se visualizó el uso de las herramientas como un mecanismo de engaño hacia ellos.

Por otro lado, los alumnos expresaron una perspectiva más optimista sobre las IA, enfocándose en su potencial para mejorar la eficiencia y la

efectividad de su trabajo como estudiantes. Sin embargo, a pesar de este reconocimiento de los posibles beneficios, el uso real de las IA en el aula fue limitado. Además, surgieron especulaciones entre los alumnos sobre la naturaleza y el alcance de las IA, incluyendo debates sobre su "alma" y su capacidad para influir en el control de las mentes. Se percibieron representaciones sociales que iban encauzadas al avance de la tecnología y la prevalencia de un futuro apocalíptico.

Los imaginarios sociales que se han gestado en torno a la IA encuentran formas de representarse socialmente dentro de comunidades específicas. En el caso de las comunidades universitarias, particularmente la que se aborda en este trabajo, estas representaciones sociales inciden de manera relevante dentro de las acciones que los individuos dentro de un colectivo llevan a cabo. En el caso de la relación entre representaciones sociales sobre la IA, y su uso dentro del proceso de enseñanza-aprendizaje es posible establecer que si bien existe apertura a su uso, en el plano de la práctica este es limitado generando más amplitud dentro de las brechas tecnológicas imperantes.

Referencias

Abric, J. C. (2011). Prácticas sociales y representaciones. En J. C. Abric, *Prácticas sociales y representaciones* (Vol. X, pp. 25-48). Ediciones Cátedra.

Asimov, I. (1950). *Yo, Robot.* Editorial Bruguera.

Cameron, J. (Director). (1984). Terminator [Película]. Paramount Pictures.

Castoriadis, C. (1975). *La institución imaginaria de la sociedad* (Vol. XX). Tusquets Editores.

Chacón, E. (2023). Enseñanza a través de la I.A. *Revista de Educación*, 45(3), 67-89.

Continental. (2023). Personalización del contenido académico. *Journal of Educational Technology,* 12(4), 120-135.

Europea. (23). Herramientas de inteligencia artificial para la evaluación formativa. *Journal of Educational Technology*, 18(1), 32-45.

García, L. y Martínez, J. (2019). *Inteligencia artificial y sociedad: retos éticos y políticos* (Vol. IX). Anthropos Editorial.

Jodelet, D. (2003). La representación social: fenómenos, concepto y teoría. En M. Montero y J. L. Álvarez (Eds.), *Psicología social comunitaria: teorías, métodos y experiencias* (Vol. 5, pp. 451-479). Paidós.

Lang, F. (Director). (1927). Metropolis [Película]. UFA.

Moscovici, S. (1979). *Psicología social* (Vol. I). Paidós.

Mujica. (2020). Integración de la inteligencia artificial en la educación. *Revista de Investigación Educativa*, 30(2), 55-70.

Smith, A. y Johnson, C. (2020). *Social Discourse on Artificial Intelligence: Perspectives and Implications* (Vol. 3). Oxford University Press.

Smith, A. y Jones, B. (2018). *Media Representation and the Social Construction of AI* (Vol. 7). Cambridge University Press.

Weedon, J. (Director). (2015). Avengers: Age of Ultron [Película]. Marvel Studios.

Capítulo 9

Inteligencia artificial y derecho: El dilema entre la formalización y la aproximación[1]

Virginia Berenice Niebla Zatarain
Jesús Manuel Niebla Zatarain
Gonzalo Armienta Hernández

https://doi.org/10.61728/AE24001090

[1] Esta colaboración es forma parte del proyecto Ciencia de Frontera 2023 Proyecto CF-2023-G-772 "Regulación de entornos digitales a través de razonamiento legal basado en inteligencia artificial".

Introducción

La ciencia computacional se ha convertido en un componente de la vida social. Como parte de esto, una de los sectores tecnológicos de mayor trascedencia es la inteligencia artificial (IA), la cual ha dado lugar al desarrollo de dispositivos capaces de replicar el razonamiento humano implementado para la resolución de problemas concretos. En este sentido, una de las áreas de implementación más importantes es el derecho, no solo desde una perspectiva regulatoria, sino en el desarrollo de tecnología capaz de mejorar la función judicial como la operación legítima de dichos dispositivos.

No obstante, la primera etapa de colaboración entre estos sectores resultó evidente que el método de representación de razonamiento legal basado en lógica simbólica, si bien era eficiente en términos descriptivos, era poco práctico en términos técnicos. Esto, ya que requería un volumen considerable de recursos para operar, lo que a su vez lo volvía lento y poco atractivo para su implementación fuera de la academia.

Para resolver lo anterior, surge una propuesta basada en modelos estadísticos los cuales, a diferencia de los esquemas tradicionales, permiten la construcción propuestas de conclusión con un determinado nivel de certeza de una manera rápida y eficiente aunque nunca definitiva, lo que expande el ámbito de aplicación de dichos desarrollos. De igual forma, este enfoque resulta compatible con la naturaleza dinámica del entorno digital particularmente, aquellos basados en Internet, los cuales representan el entorno de colaboración más importante en un futuro próximo entre la inteligencia artificial y el derecho.

El surgimiento de la inteligencia artificial

El origen de la inteligencia artificial (IA) se remonta hacia mediados de la década de 1950, cuando se comenzó a discutir la posibilidad de replicar procesos cognitivos humanos basados en una incipiente tecnología infor-

mática (McCorduck y Cli, 2004). Específicamente, en 1956 surge el término inteligencia artificial derivado del Proyecto de Investigación de Verano de la Universidad de Darmouth en Nuevo Hampshire. En este evento participarían investigadores quienes habrían de ser considerados padres de la IA, como es el caso Nathaniel Rochester, desarrollador de la primera computadora científica, y el matemático Claude Shannon considerado fundador de la teoría de la información. El objetivo de este proyecto fue establecer la factibilidad de replicar los procesos cognitivos propios de un área del conocimiento aplicados a la resolución de problemas por medio de lenguaje computacional (Haenlein y Kaplan, 2019).

Conforme se fortaleció la IA como área de investigación convergente, comenzaron a surgir los primeros proyectos de investigación los cuales mostraron la capacidad (aunque limitada) de ser implementados en ambientes comerciales, como el caso particular de ELIZA. Desarrollado entre 1964 y 1966 por Joseph Weizenbaum, es considerado uno de los primeros proyectos de procesamiento de lenguaje natural capaz de mantener una conversación con un operador humano en diversos campos del conocimiento (aunque resultó particularmente exitoso en psicoterapia), lo cual lo colocó cerca de aprobar la famosa "Prueba de Turing" (Eliza, s. f.). Para 1968, uno de los padres fundadores de la IA presentó una de las primeras definiciones de esta tecnología, señalando que esta "es la ciencia que hace que hagan cosas que requieran inteligencia de ser realizados por un humano" (Minsky, 1969). En consecuencia, todos los

los cuales implementen esquemas equiparables a procesos cognitivos humanos, forman parte de esta rama computacional. No obstante, la complejidad de replicar el conocimiento humano en procesos computacionales equivalentes al razonamiento dista de ser un enfoque sencillo. Por una parte, aquellas tareas que requieren de "sentido común" o de percepción, terminan por convertirse en procedimientos particularmente complejos (Rissland, 2013).

Si bien estos desarrollos demostraron la factibilidad de reproducir procesos cognitivos humanos por medio de computadoras, la comunidad resultó sobre optimista, lo que llevó a Marvin Misky a declarar que para 1978 se podría contar dispositivos inteligentes con la capacidad cognitiva equiparable a un individuo promedio (Haenlein y Kaplan, 2019). Sin em-

bargo, esta tecnología no habría de resultar incompatible con entornos dinámicos complejos, quedando limitada a escenarios compactos con una relevancia limitada. La principal razón para lo anterior fue la limitada capacidad de la tecnología computacional de la época, la cual carecía de la eficiencia operativa requerida para replicar los esquemas de razonamiento presentados por los diseñadores.

Este retraso fue tal, que en 1973, el Congreso de los Estados Unidos consideró la investigación en la IA como poco pertinente, lo que dio lugar a la eliminación del financiamiento público dedicado a esta. En el mismo año, James Lighthill a petición del Consejo Británico de Investigación Científica realizó un reporte sobre el estado operativo de la IA, lo cual resultó en el retiro del apoyo financiero para esta área. A pesar de lo anterior, el interés por parte del sector académico continuó, dando lugar a nuevos avances dentro de esta área de la ciencia computacional. Como parte de estos, se establecieron de manera permanente por lo menos dos objetivos específicos: primero, el comprender la manera en que funciona la mente humana y segundo, el desarrollo de dispositivos computacionales que sean capaces de reproducir procesos cognitivos humanos (Konar, 2018).

En estricto sentido, la mayoría de los proyectos tradicionales pretendían ambos objetivos de manera conjunta. Esto tenía como consecuencia la abstracción de procesos cognitivos específicos, los cuales eran utilizados como base para el proceso de razonamiento implementado por estos dispositivos. No obstante, las fuentes de información debían de filtrarse para garantizar que los esquemas implementados tuvieran un nivel óptimo. Esto, debido a que no siempre la fuente primigenia de razonamiento eran expertos humanos: si la aplicación estaba enfocada en asistir en escenarios abiertos, los elementos contenidos en esta se convertían en la fuente de conocimientos. Esto sumaba una variable adicional a la complejidad de diseñar dispositivos para la IA: el entorno de aplicación tenía la capacidad de afectar el resultado del dispositivo.

Finales del siglo XX e inicios del XXI, una etapa de convergencia

Las últimas décadas del siglo XX presenciaron el resurgimiento de las aplicaciones operacionales de la inteligencia artificial. A diferencia etapas previas donde existió un desencanto por la nula aplicabilidad de la mayoría de estos sistemas, en 1980 surge un marcado reinterés por este sector computacional. Uno de las áreas que recibió mayor atención es el de los denominados "sistemas expertos". Estas implementaciones se desarrollarán a continuación.

Sistemas expertos

Dentro de los primeros dispositivos que logaron migrar de la esfera académico experimental hacia entornos comerciales fueron los sistemas expertos. Dichos operativos replican la interacción con un experto en un área del conocimiento en específico. Su estructura se encuentra compuesta por una base de conocimientos, un motor de inferencia y la interfaz de usuarios (Huntington, 2000). En lo referente a la interacción con sus usuarios, esta se desarrolla a través de una serie de preguntas prediseñadas, las cuales tienen como finalidad conocer las particularidades del problema a resolver y recolectar variables relevantes para la construcción del resultado.

Estos dispositivos se distinguen de implementaciones convencionales presentando las siguientes características (Buchanan y Smith, 1988):

a) Razona a través tanto de expresiones simbólicas como aproximadas

En este sentido, los métodos simbólicos o heurísticos proveen el elemento fundacional de los sistemas expertos. Por una parte, el primero brinda una forma de razonamiento que replica el implementado por la mente experta humana. Su configuración obedece a estructuras biológicas y a la experiencia obtenida de expertos en el área de aplicación. El segundo opera bajo grupos de datos que contienen casos similares resueltos con anterioridad, sobre los cuales se detectan similitudes con el caso a resolver. En el caso específico de la implementación de esta tecnología al ámbito jurídico, el

primer enfoque hace referencia a la recopilación de conocimiento experto legal (por ejemplo Magistrados), así como los esquemas mentales (de que manera procesan información jurídicamente relevante y la forma en que construyen resoluciones) implementados para la resolución de casos específicos. El segundo enfoque hace referencia al análisis de datos en cuerpos de información jurídica digital (expedientes digitalizados sobre un caso en particular), se detectan similitudes con el caso a resolver y se presenta una propuesta de solución con un nivel estadístico de certeza.

b) Implementa métodos heurísticos (posibles) como algorítmicos (certeza)

Por su parte, la heurística opera a través de métodos de estadísticos que construyen decisiones basados en factibilidad, no en certeza. La principal diferencia entre estas dos metodologías es el "costo computacional", siendo para este momento, mucho más demandante las estructuras basadas en representación simbólica del conocimiento. Ambos enfoquen parten de modelos estadísticos y tratamiento de información, su principal característica es el consumo de recursos computacionales durante su operación.

c) Su desempeño resulta equiparable con el de un experto humano en su área de implementación

Esta es una característica fundamental de los sistemas expertos: el nivel de eficiencia que con el que operan los equipara con asistente de especialistas, brindando asesorías cuya calidad puede superar a la de un experto humano promedio y que puede incrementarse como resultado de su constante operación (aprendeizaje).

d) Presenta de manera inteligible el conocimiento, así como justifica sus respuestas por medio de razonamiento

Aborda la construcción del razonamiento como su presentación al usuario por medio de estrategias que permiten analizar la construcción del resultado generado. Este se realiza a través de esquemas gráficos y textuales

que simplifican su entendimiento y que incrementan su relevancia según su área de implementación, en el caso específico del derecho, incrementan la transparencia.

e) Mantiene flexibilidad

Esta es una de las particularidades más importantes del sistema: primero, la actualización del programa debe de ser lo suficientemente simple para permitir su continuidad y no sacrificar ninguno de sus componentes operativos particularmente, la base de conocimientos. Este es el reto constante y del cual depende la vida operativa de estos dispositivos: la constante actualización expande la base de conocimientos y la inclusión desmedida de información puede volverlos innecesariamente lentos. Para contrarrestar lo anterior, es necesario el mantenimiento calendarizado que garantice la calidad de la información disponible para la generación de respuestas.

La estructura operativa de estos desarrollos los hizo atractivos para diversas áreas fuera del sector académico tradicional. Como parte de esto, Smith y Baker presentaron en 1983 el Sistema de Consulta Dipmeter (Dipmeter Advisor System) (Smith y Baker, 1983), el cual auxiliaba a ingenieros en petróleo, a conocer cuáles eran las características del suelo que pretendían perforar (específicamente, el tipo de sedimentación), lo cual resulta fundamental para la factibilidad del proyecto. Sin embargo, la preferencia de la representación simbólica de conocimiento como plataforma para interactuar y resolver problemas habría de resultar insalvable, volviendo este y otros proyectos poco viables a largo plazo.

Razonamiento simbólico en inteligencia artificial

Un elemento fundamental al abordar la composición de los sistemas expertos es la expresión lógica requerida para los procesos de creación de conocimiento o construcciones de resultados. En este sentido, la frase "programación de conocimiento" ha sido utilizada para señalar el énfasis de brindar a un sistema experto esta característica. Como parte del desarrollo de estos dispositivos, McCarthy en 1959 señaló el principio de "conocimiento declarativo" (McCarthy, 1959). Este señala que el conocimien-

to debe ser programado de manera explícita dentro de un programa, de tal forma que permita a otros programas razonar acerca de él. En estricto sentido, una base de conocimientos es una base de datos estructurada para implementar procesos de razonamiento en la construcción de una respuesta. Consecuentemente, esta opera bajo esquemas de flexibilidad y complejidad que les permiten establecer relaciones que resultan fundamentales para el proceso de resolución. Ahora bien, la estructura del conocimiento depende en gran medida del área de implementación. En este sentido, la arquitectura para la resolución de problemas se desarrolla entorno a objetos (o conceptos) principales o de acciones particulares (incluyendo relaciones inferenciales) entre estos objetos. Esto conlleva a que se prefieran aquellos esquemas que permitan la construcción de conocimiento por medio de inferencias, lo que da lugar el intercambio de posturas y un proceso de construcción de resultados más eficiente en términos computacionales.

Como tal, la expresión simbólica de conocimiento fue la piedra angular del desarrollo de la inteligencia artificial durante la primera parte de la vida operativa de este campo. No obstante, si bien permitía la representación e implementación de los esquemas cognitivos mentales propios del área a operar, inevitablemente tendía a volverse particularmente compleja, ocasionando que la actualización de la base de conocimiento (aprendizaje) se volviera sumamente costosa y volviéndola poco atractiva en términos operativos a largo plazo. Como tal, suponga un escenario donde se evalúa la aceptación para la realización de un contrato, cuya expresión lógica seria:

```
Cliente=Firman
Provedr=Firman
Si (cliente_Firman & Proveedor_Firman)
contrato=aceptado
Entonces:
    Servicio=Continuado
De lo contrario
    Servicio=suspendido
Fin del si.
```

La traducción de lo anterior en lógica computacional resulta:

```
If (client_signature XOR provider_signature) contract_state = partia-
```

```
lly_accepted
    If(client_signature AND provider_signature) {
    contract_state = service_accepted
    payment_state = init
    If (payment_state == done) contract_state = done
    else
    contract_state = service_denied
    End if
    End if
```

Como es posible apreciar, la traducción a lenguaje computacional resulta particularmente complejo toda vez que requiere instrucciones y comandos computacionales específicos para replicar un escenario jurídicamente relevante. Derivado de lo anterior, se desarrollaron esquemas que replicaran las operaciones de razonamiento sin el costo operativo del modelo simbólico, es aquí donde surge los métodos aproximados para la inteligencia artificial.

Razonamiento aproximado en inteligencia artificial

Como se ha mencionado hasta este punto, la representación de los mecanismos implementados en la construcción del razonamiento humano ha sido presentada como la base operativa para el desarrollo de tecnología inteligente. Sin embargo, dependiendo de las características de la tarea a realizar y del entorno de aplicación, este puede no ser el idóneo debido, principalmente a la alta demanda de recursos computacionales que requieren sus operaciones y del entorno de aplicación. Lo anterior resulta ser el caso de entornos dinámicos digitales, los cuales requieren las siguientes características (Groot, Stuckenschmidt y Wache, 2006):

- Razonar con un espectro de tiempo limitado
- Razonar con recursos limitados además del tiempo
- Razonamiento que no es "perfecto" pero sí "suficientemente bueno" para tareas que bajo determinadas particularidades.

Derivado de la necesidad de implementar esquemas "equivalentes" de razonamiento, los cuales operen como complemento a los modelos sim-

bólicos y que sean compatibles con otras tecnologías, como es el caso de la Web Semántica. En la siguiente sección se abordarán enfoques relevantes propios del aprendizaje automático.

Aprendizaje automático

Comúnmente conocido como *machine learning*, el aprendizaje automático opera a través de la detección de patrones en cuerpos de información de volumen considerable. Implementado como complemento de dispositivos inteligentes insertados en entornos dinámicos, como es el caso del Internet y tecnologías derivadas, se distingue debido a que el proceso de generación de resultados no depende de un proceso de construcción lógico, sino de un proceso estadístico que, si bien no es definitivo, brinda un nivel "suficiente" de certeza. De igual forma, el proceso de "aprendizaje" es el resultado de la metodología estadística basada en la detección de patrones y en la proyección de resultados basados en dichas coincidencias. Consecuentemente, los dispositivos basados en este modelo estadístico no producen conocimiento a través de la reproducción de procesos mentales tradicionales, sino por medio del almacenamiento de resultados del establecimiento de patrones, lo que permite la presentación de resultados previa detección de combinaciones ya almacenadas.

De esta forma, conforme el algoritmo realice un mayor número de operaciones dentro de una base de datos determinada, será capaz de incrementar el volumen coincidencias almacenadas para así mejorar la calidad de sus operaciones, sin que esto implique sacrificar recursos computacionales. En lo que respecta a la identificación de patrones, esta opera bajo un esquema de priorización basado en probabilidades. De lo anterior, se puede señalar lo siguiente (Surden, 2021): Un dispositivo que incluya aprendizaje automático podrá operar a través de vastos cuerpos de información, segmentarla, encontrar patrones relevantes y a adecuar su funcionamiento acorde a estos.

La eficiencia de este enfoque depende de manera directa del incremento constante de información. Consecuentemente, la eficacia de estos algoritmos es directamente proporcional al incremento en el volumen de datos procesados, lo que los convierte en una herramienta de desarrollo

idóneo para tecnologías dinámicas, como aquellas que tienen como modelo céntrico el Internet. Lo anterior ha conllevado a que el aprendizaje automático se convierta en el principal método para el desarrollo de IA, no obstante, presenta limitaciones que resultan importantes reconocer. En estricto sentido, estos algoritmos implementan estrategias que construyen resultados suficientemente cercanos con aquellos producidos por modelos cognitivos, pero que en esencia son esquemas comparativamente limitados. Como tal, es posible implementar un algoritmo de aprendizaje automático para la clasificación de información el cual es entrenado para identificar patrones de palabras o combinaciones de estas para posteriormente procesarlas. No obstante, estos pueden incurrir en errores de interpretación, lo cual termina por comprometer la calidad de su operación.

Esta área se divide en dos enfoques:

1) Aprendizaje supervisado

Estos algoritmos operan a través de información que ha sido previamente clasificada e identificada por un experto humano bajo una metodología específica. En este sentido, operan bajo dos clasificaciones de datos denominados de entrenamiento y datos de prueba. El primero genera un resultado el cual requiere ser clasificado o predicho, lo cual permite identificar aquellos patrones que habrán de implementar. Por su parte, el segundo representa el cuerpo de datos sobre el cual habrá de ser implementados.

2) Aprendizaje no supervisado

En este enfoque el algoritmo aprende patrones detectados en datos no etiquetados, es decir, no se conoce su naturaleza a priori. El objetivo de este método es que logre desarrollar una representación concisa del mundo para posteriormente desarrollar contenido propio. Durante la etapa de entrenamiento, el algoritmo intenta comprender y replicar la naturaleza de los datos que le son presentados, el nivel de éxito se desprende directamente del nivel de certeza que genera con el presentado como punto de referencia.

Derivado de su eficacia operativa, el aprendizaje automático se encuentra presente en los escenarios de aplicación más importantes de la inteligencia artificial, como es el caso de los vehículos autónomos, análisis predictivo, detección de fraude y en gran parte de la automatización presente en el campo médico (Surden, 2021, p. 88). No obstante, el desempeño de este método es proporcionalmente dependiente del volumen de datos procesados y de la disponibilidad de fuentes de información relevante para su actualización. Como es posible suponer, el Internet se ha convertido en el principal entorno de aplicación de esta tecnología convirtiéndose en el núcleo operacional de tecnologías derivadas, como el Internet de todo y las Ciudades Inteligentes (Singh, Nayyar, Kaur y Ghosh, 2020). Consecuentemente, el aprendizaje automático se encuentra en un proceso de transición hacia un modelo totalmente independiente del elemento humano, centrado en esquemas de entrenamiento que minimicen errores y permitan maximizar la eficiencia operativa de la arquitectura computacional sobre la cual operan (Zhou, 2021).

Habiendo brindado una breve descripción operativa de la inteligencia artificial, en la siguiente sección se abordará la relación que guarda con el derecho particularmente, como plataforma de aplicación.

Inteligencia artificial y derecho: una relación complementaia

Derivado de su composición lógica, el derecho resulta un área de implementación relevante para la inteligencia artificial. Esto, derivado de una serie características propias de esta relación, entre las que destacan las siguientes (Rissland , 2013, p. 233):

El razonamiento jurídico es multimodal, voluminoso y diverso: la construcción de la expresión lógico-jurídica incluye casos, reglas, estatutos y principios. En lo que respecta a la expresión cognitiva multimodal, esta refiere a dotar al dispositivo con la capacidad de entender varias formas de expresión, propias de un escenario jurídicamente relevante. Esto permite adaptarse a la construcción de preceptos lógicos jurídicos, los cuales presentan relaciones complejas que no obedecen siempre a estructuras condicionales. Lo anterior puede ejemplificarse desde la perspectiva de

robo, si bien una persona puede apoderarse de un objeto mueble que es propiedad de otra, si este acto se realiza por adquirir alimento, esto no constituye delito, toda vez que se hace porque el individuo no cuenta con otro medio para sobrevivir.

La creación y desarrollo de casos posee un estilo explícito que se desprende de un procedimiento estandarizado de razonamiento y justificación: esto facilita el diseño e implementación de un método compatible con las particularidades de casos en específicos los cuales, pueden ser rastreados hacia un origen normativo. Como tal, el derecho tiene como fundamento la norma jurídica y la construcción de la interacción normativa se puede construir y analizar (para efectos de transparencia) por medio de inteligencia artificial.

La naturaleza del conocimiento jurídico se desprende de fuentes jurídicas relevantes y propiamente documentadas. Estos incluyen desde resoluciones de casos previos, jurisprudencias, mediación, tratados, posturas académicas e incluso resúmenes académicos. Lo anterior implica un volumen considerable fuentes de información. Lo que facilita no solamente su debida expresión, sino su posterior adopción como parte de la base de conocimientos jurídicos (aprendizaje).

El derecho es consiente y crítico de sí mismo. Esto hace referencia a la utilización de información jurídica de épocas (jurídicas) pasadas, incluyendo procesos y suposiciones presentadas con anterioridad. En este sentido, el derecho posee mecanismos que permiten desechar información irrelevante, lo que mantiene un estándar de calidad con respecto a esta.

La naturaleza contextual e interpretativa de las respuestas generadas en el derecho, dependen en gran medida del carácter interpretativo implementado por el operador jurídico, lo cual hace que estas varíen según las particularidades propias de cada caso y a la formación y experiencia del funcionario. Si bien estas son susceptibles de ser representadas como componentes de un dispositivo jurídico inteligente, su función es el de una herramienta supeditada a las necesidades del experto y no como un ente independiente.

El proceso de construcción de razonamiento jurídico es sumamente complejo, abarcando desde el sentido común hasta el conocimiento legal especializado y varia acorde a su estructura, carácter, uso y componen-

tes ambientales. Por lo anterior, definir el escenario de operación de esta tecnología resulta fundamental para el debido funcionamiento de dicha tecnología.

Lo anterior brinda dos puntos fundamentales, primero, la susceptibilidad de crear implementaciones inteligentes capaces de replicar modelos de razonamiento jurídico para la resolución de problemas específicos. Segundo, permite al sector jurídico la adopción e implementación de esquemas tecnológicos que permiten su operación en entornos donde el modelo tradicional de aplicación de la ley resulta inoperante. De igual forma, mejora la calidad del desempeño profesional legal, al desarrollar herramientas computacionales que incluyan el manejo de grandes cuerpos de datos aplicables a cualquiera de sus vertientes jurídicas. Otro punto relevante es que brinda nuevos métodos que permiten asistir y corroborar postulados teóricos, detectar fallas potenciales, así como facilita la representación de postulados lógicos y brinda de propuestas de mejora (Dennett, 2017). Desde una perspectiva meramente técnica, la inteligencia artificial resulta particularmente relevante para la representación de conocimiento jurídico, esquemas de representación y razonamiento análogo.

Aspectos relevantes de la colaboración entre la inteligencia artificial aplicada al derecho

La colaboración entre la inteligencia artificial y el derecho se presenta como una de las más relevantes, capaz de asistir en la creación de soluciones novedosas a problemas ante las cuales el modelo tradicional de aplicación de la ley resulta poco práctico. Estos dispositivos desempeñan tareas de carácter asistivo es decir, dependientes de las necesidades particulares que presenta un experto jurídico humano. En este sentido, las implementaciones inteligentes aplicadas al área legal deben contar con algunos de los siguientes elementos:

- Razonamiento basado en casos (tanto reales como hipotéticos) y analogías (en sí misma una de las contribuciones más complejas de realizar);
- Razonamiento basado en reglas;
- Combinación de diversos esquemas de razonamiento;
- Manejo de conceptos mal definidos y de naturaleza abierta;

- Formulación de argumentos y explicaciones;
- Manejo de excepciones de y para conflictos dentro de los elementos de conocimiento;
- Adecuación de cambios basados en conocimiento legal, particularmente, conocimiento legal y manejo de no-monoticidad (esto es, cambios derivados del descubrimiento de nueva información que hace de una verdad reconocida con anterioridad como no fundamentada);
- Modelado de conocimiento basado en sentido común (rule of thumb, en inglés);
- Modelado de conocimiento de intención y creencias;

La inclusión de estas características dependerá de las especificidades del problema a resolver, así como las particularidades del usuario (es decir, el nivel de conocimiento jurídico con que cuente). Esto fue determinante para el diseño de las primeras implementaciones exitosas de la IA al sector jurídico: los sistemas expertos (Susskind, 1987). Estas implementaciones operaban replicando el proceso de toma de decisiones jurídicas. Dicho proceso era el resultado de procesos de abstracción cognitiva que formaba parte del proceso de diseño de dichas implementaciones, donde se estudiaba la manera en que expertos jurídicos, (tanto servidores públicos como litigantes) entendían un problema y procedían a su resolución (Waterman y Peterson, 1986).

Sin embargo, el desarrollo de modelos de razonamiento simbólico que expresen el conocimiento jurídico presentaba el mismo problema experimentado en otras áreas del conocimiento: el alto costo en términos computacionales. En el siguiente apartado se abordará una propuesta de colaboración basado en modelos aproximados.

Hacia un modelado aproximado de la ley

Uno de los aspectos más importantes del ejercicio profesional legal, consiste en realizar propuestas debidamente informadas sobre las implicaciones jurídicas de un escenario particular (Brest y Krieger, 2010). Tomando como escenario de aplicación la práctica privada, una de las principales interacciones entre litigantes y clientes es la presentación de estos últimos de un caso con la esperanza de obtener un resultado a favor. En este sentido, los abogados utilizan su experiencia profesional, la cual complementan con la consulta (de estar disponibles) de casos resueltos para indicar las posibilidades de éxito. Dichas fuentes de información resulta susceptibles de ser sometidos a procesos de automatización computacional (Katz, 2010). Como tal, es posible señalar que la combinación de inteligencia humana experta (jurídica) con procesos analíticos basados en computadoras resultan superior a análisis humanos tradicionales (Katz, 2010, 909). Es aquí donde la aplicación de aprendizaje supervisado resulta relevante para entornos jurídicos profesionales cerrados, como es el caso de las firmas de abogados.

La inserción de dichos algoritmos permitiría establecer relaciones entre datos que han sido categorizados con anterioridad (tipo de resolución, órgano jurisdiccional de origen, juez, entre otros) pueden dar lugar a la implementación de sistemas que permitan la predicción de elementos relevantes como la posibilidad de éxito de un caso particular o incluso responsabilidad. De manera general, el aprendizaje automático opera basado en la detección de datos específicos y los patrones que estos presentan en cuerpos de información específica. De esta forma, este enfoque podría asistir en la predicción de problemas del área familiar como es la repartición de la masa conyugal, analizando resoluciones previas, encontrando características particulares y comparándolas con el caso que pretende resolver.

Por su parte, el aprendizaje no supervisado permitiría detectar combinaciones de patrones que pueden llegar a afectar la construcción de resultados jurídicos en una fracción de tiempo de la que necesitaría un experto humano con un nivel tal, que la aprobación humana resultaría poco necesaria. Este enfoque resulta particularmente relevante para entornos

dinámicos, los cuales requieren una respuesta a una consulta jurídica en fracciones de segundo. Esta postura resulta compatible con la mayoría de las potencialmente relevantes en términos legales que surgen en Internet: los usuarios que interactúan a través de esta tecnología en pocas ocasiones requieren de procesos que repliquen la totalidad del razonamiento jurídico experto. Por el contrario, suelen solicitar pequeñas asistencias legales las cuales pueden hacerse a través de consultas basadas en este enfoque.

De manera general, es posible señalar que dichas representaciones estadísticas brindan un equivalente operativo a los procesos cognitivos legales propios de la aplicación profesional del derecho, lo cual representa la evolución de la implementación de tecnología inteligente en escenarios jurídicos.

Conclusiones

La tecnología se ha convertido en un componente de transformación del derecho. Si bien, tradicionalmente estos campos han estado relacionados, dicha interacción había sido mayormente normativo, tomando a la ciencia computacional como un campo que caía dentro de la regulación tradicional. No obstante, el desarrollo de la inteligencia artificial y la consecuente adopción del derecho como área de aplicación han transformado dicha relación. Por una parte, es posible apreciar el desarrollo de dispositivos inteligentes diseñados con la finalidad de replicar procesos cognitivos legales los cuales tienen como finalidad emular, ya sea a través de esquemas simbólicos o aproximados, procedimientos jurídicos específico y la adaptación de estos hacia entornos dinámicos, como es el caso del Internet y de tecnologías derivadas de este. Esto presenta una de las aportaciones de mayor trascendencia de este campo interdisciplinario ya que aborda al elemento normativo como componente de diseño, lo cual permite intervenir en la operación del dispositivo previniendo actos potencialmente perjudiciales para la esfera jurídica de usuarios y operadores.

Referencias

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American,* 284(5), 34-43.

Brest, P., & Krieger, L. H. (2010). *Problem solving, decision making, and professional judgment: A guide for lawyers and policymakers*. Oxford University Press, USA.

Buchanan, B. G., & Smith, R. G. (1988). Fundamentals of expert systems. *Annual review of computer science*, 3(1), 23-58.

Dennett, Daniel C. (2017). *Brainstorms: Philosophical essays on mind and psychology*. MIT Press.

Groot, P., Stuckenschmidt, H., & Wache, H. (2005). Approximating description logic classification for semantic web reasoning. In *The Semantic Web: Research and Applications. Second European Semantic Web Conference*, ESWC 2005, Heraklion, Crete, Greece, May 29–June 1, 2005. Proceedings 2 (pp. 318-332). Springer Berlin Heidelberg.

Haenlein, M. y Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review,* 61(4), 5-14.

Hitzler, P. (2021). *A review of the semantic web field.* Communications of the ACM, 64(2), 76-83.

Huntington, D. (2000). *Expert systems for online advice: knowledge at your fingertips*. PC AI, 14(4), 26-27.

Katz, D. M. (2012). *Quantitative legal prediction-or-how i learned to stop worrying and start preparing for the data-driven future of the legal services industry*. Emory LJ, 62, 909.

Konar, A. (2018). *Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain*. CRC press.

McCarthy, J. (1959). Programs with common sense. En *Proc. Symposium on. The Mechanisation of Thought Processes* (pp. 77-84). National Physical Laboratory.

McCorduck, P. yCfe, C. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

Minsky, M. L. (1969). *Semantic information processing*. The MIT Press.

Rissland, E. L. (2013). Artificial intelligence and law: Stepping stones to a model of legal reasoning. In *Scientific Models of Legal Reasoning* (p. 226). Routledge.

Sil, R., Roy, A., Bhushan, B. and Mazumdar, A. K. (2019). Artificial intelligence and machine learning based legal application: the state-of-the-art and future research trends. In *2019 International Conference on Computing, Communication, and Intelligent Systems* (ICCCIS), pp. 57-62. IEEE, 2019.

Singh, P., Nayyar, A., Kaur, A., & Ghosh, U. (2020). *Blockchain and fog based architecture for internet of everything in smart cities*. Future Internet, 12(4), 61.

Smith, R. G., & Baker, J. D. (1983, August). The dipmeter advisor system: a case study in commercial expert system development. In *Proceedings of the Eighth international joint conference on Artificial intelligence-Volume 1* (pp. 122-129).

Surden, H. (2021). *Machine learning and law: an overview*. Research Handbook on Big Data Law, 171-184.

Susskind, R. (1987). *Expert systems in law*. Oxford University Press, Inc.

Von Luxburg, O. B. U., & Rätsch, G. (2004). *Advanced lectures on machine learning*. Springer.

Waterman, D. A., Paul, J., & Peterson, M. (1986). *Expert systems for legal decision making*. Expert Systems, 3(4), 212-226.

Zhou, Zhi-Hua. (2021). *Machine learning*. Springer Nature, 2021.

Capítulo **10**

Los límites de la aplicación de la Inteligencia Artificial (IA) en la gobernanza: una aproximación desde la política pública urbana y la cibernética social

Alfredo Leonardo Carreón Corona
Luis Gustavo Kelly Torreblanca

https://doi.org/10.61728/AE24001106

Introducción

La discusión en torno al uso de la tecnología y su participación en el desarrollo de las estructuras sociales y en la toma de decisiones políticas no es una novedad; las corrientes del pensamiento crítico en la sociología y la filosofía que van desde la Crítica de la economía política de Marx (1990), pasando por obras puntuales como la Crítica de la razón instrumental (Horkheimer, 1973) y algunos de los últimos textos de la tradición de la teoría crítica como es el caso de Jürgen Habermas (1986; 2002) y Herbert Marcuse (1984) son muestra de una larga tradición abocada a describir las consecuencias del desarrollo científico y tecnológico en la economía, la cultura, la política y, en general, en la sociedad. A las posturas críticas de la academia se suman las de la literatura, algunas de las cuales han inspirado sus homónimos del cine, Yo, Robot (1950) de Asimov, particularmente en "El conflicto evitable"; y Colossus: el proyecto Forbin (Jones, 1966; Sargent, 1970), en las que de manera más clara y comprensiva se ilustran los terrores imaginados de su época frente al desarrollo de la tecnología computacional y la robótica.

Aunque en términos analíticos la distancia entre la realidad que pretenden describir las primeras obras mencionadas, y la ficción de las segundas, se determina por el género literario al que pertenecen; en las diversas narrativas prevalecen dos elementos que vale la pena destacar: por un lado se sostiene una constante interacción entre la dependencia y la irresponsabilidad de la humanidad ante el uso y los alcances, reales y potenciales, de la tecnología y sus aplicaciones. Por otro lado, la toma de decisiones que ha llevado a ese grado de dependencia, y a la catástrofe, se caracteriza por encontrarse en manos de reducidos grupos de tomadores de decisiones cuyo ejercicio de poder se ejecuta sobre masas de gobernados, que han adoptado pasivamente las consecuencias de la toma de decisiones de sus gobernantes de la mano de las comodidades que les ha otorgado la tecnología.

En la práctica actual, la política y los elementos tecnológicos que participan de la construcción imaginaria de los futuros distópicos son más cer-

canos al ciudadano común y corriente de lo que parece, tanto en la visión de los enfoques clásicos del ejercicio político como en las ficciones de la rebelión de las máquinas; toda vez que la vida cotidiana se encuentra entreverada en tomas de decisiones políticas que le afectan de manera directa y en las cuales cada vez se hace un uso más constante de las tecnologías. Ejemplos comunes y poco rebuscados de estas innovaciones son observables en el diseño de redes de transporte urbano y su regulación, y en la digitalización de la burocracia gubernamental; sin embargo, sobre esta misma línea de desarrollo tecnológico aplicado se siguen los ambiciosos proyectos urbanos de las agendas del desarrollo sostenible.

Smart cities o ciudades inteligentes

Es precisamente en la planeación urbana uno de los rubros en los que las decisiones políticas tienen un impacto inmediatamente directo en la ciudadanía y en donde la introducción de las tecnologías, y en particular de la IA, abre la posibilidad de un acercamiento a ejercicios plenamente democráticos y eficientes en términos administrativos o, por el contrario, a la reproducción de dinámicas de dominación asimétricas entre ciudadanía y gobierno, y la continuidad de ejercicios administrativos poco transparentes y cuestionables.

Los teóricos contemporáneos de la planificación han, virtualmente todos, aceptado el argumento de que las decisiones de planificación son ineludiblemente políticas (Fainstein y Fainstein, 1996, p. 265). El decir, que todas las decisiones que se toman en materia urbana son ineludiblemente políticas, implica que invariablemente el factor humano se encuentra presente, tanto en el acto de decidir como en el proceso a ejecutar y por último en el beneficiario final. Por supuesto, la idea de su existencia como elemento esencial en la toma de decisiones urbana, conlleva que bajo el entendido de que la política es un acto de decidir qué es lo que habría de realizar, transformar o aplicar, en este caso, en el ámbito del espacio urbano; su implicación trae aparejado el consenso y la decisión creativa, siendo estos últimos, elementos propios de la actividad intelectual humana.

En estas condiciones, emergen nuevos tipos de interrogantes que se deben plantear no solamente para la planificación urbana o regional, sino

en específico para la definición de acciones tendientes a garantizar lo que se ha denominado gobernanza, así como gobernanza urbana.

En este orden de ideas, el presente artículo busca presentar cuales son los límites que encuentra la aplicación de las tecnologías de la inteligencia artificial en su uso para la gobernanza urbana, situándose en el contexto nacional mexicano. Partiendo de ello, el trabajo fue organizado en cuatro apartados.

Primero se realizará una reflexión desde la perspectiva de la sociología de la acción pública, con el objetivo de fundamentar la forma en la que las relaciones políticas en la actualidad se han reconfigurado, pasando de esquemas de ejercicio del poder gubernamental unilateral a dinámicas de participación ciudadana como lo son la gobernanza y, más específicamente, la gobernanza urbana, en las cuales las tecnologías cada vez van tomando mayor presencia. Segundo, se describirá a partir de algunos elementos de la cibernética clásica el cómo se articulan las actividades de gobierno político con el gobierno de las máquinas y las tecnologías, haciendo hincapié en cómo el acoplamiento entre la acción humana y las potencialidades de la tecnología dan lugar a nuevas acciones sociales que, si bien pueden tener consecuencias desastrosas, al mismo tiempo, ofrecen mayores opciones en el objetivo de construir mejores condiciones de vida para la sociedad. Sobre esta última posibilidad es que, en el tercer apartado, rescatando elementos de la Investigación de Operaciones (IO), se desarrolla una propuesta de modelo de sistema viable de gobierno que haga efectiva la gobernanza, a partir de la implementación de la IA en el procesamiento de la información ciudadana. Por último, se desarrollarán las conclusiones del trabajo.

El gobierno y la gobernanza a la luz de la tecnología.

De acuerdo con Lascoumes y Le Galès (2014), así como con Le Galès y Ugalde (2018), la gobernanza en términos conceptuales surgió de las reflexiones que se produjeron al cuestionar los límites del gobierno, desde sus errores hasta las nuevas formas de cómo se organizan y generan políticas. No obstante, la concepción de la misma parte de la idea de entender la gobernanza como un proceso de coordinación de actos, grupos sociales e

instituciones, para alcanzar objetivos discutidos y definidos colectivamente. La gobernanza remite así a las dinámicas multiactores que producen la estabilidad de una sociedad y de un régimen político, su orientación y su capacidad de suministrar servicios y de asegurar su legitimidad (Lascoumes y Le Galès, 2014, p. 28).

La incorporación del concepto de gobernanza en los marcos teóricos de las ciencias políticas, trajo consigo un cierto rompimiento entre el pensamiento que podría definirse como clásico dentro del estudio de las políticas públicas, que en términos conceptuales implicaba su concepción desde el enfoque top-down o desde arriba, es decir, una política pública entendida como el resultado de la actividad de una autoridad investida de poder público y de legitimidad gubernamental (Meny y Thoenig, 1992, p. 89). Mientras que, por otro lado, la aparición del concepto de acción pública representó una transformación de la visión clásica, la cual integra a distintos actores políticos y no únicamente a las autoridades, a la vez que hace posible el entendimiento de la reestructuración del Estado. La acción pública implica el desplazamiento de la política de la autoridad legítima hacia formas de participación mayormente democráticas.

Es en esa sintonía que el hablar de la gobernanza o de la gobernanza urbana ha llevado al análisis, no únicamente a las autoridades, constituidas estas como actores políticos legítimos, sino también a colectivos, grupos de presión, sociedad civil organizada, y por supuesto, a las instituciones públicas encargadas de la creación y aplicación de las acciones gubernamentales.

Lascoumes y Le Galês (2014) propusieron un modelo de análisis de la acción pública a partir de cinco variables: actores, representaciones –que hace referencia a los marcos cognitivos que condicionan el actuar de los actores–, así como las instituciones y sus normas. De acuerdo con los autores, la relación de estos tres desemboca en los procesos políticos, que van desde las decisiones políticas unilaterales tomadas por las autoridades, las coaliciones, concesiones, hasta las acciones colectivas contenciosas como las movilizaciones, protestas, entre otras formas colectivas que intentan insertar demandas y propuestas de política bajo esquemas no previamente institucionalizados. Por último, los resultados, que refieren a los outputs y outcomes; es decir, por un lado, los efectos producidos sobre organi-

zaciones y los comportamientos (outputs), por el otro, las consecuencias de la acción pública, los impactos sobre el problema que se desea tratar (outcomes) (Lascoumes y Le Galès, 2014, p. 20).

Sea cual sea la forma en cómo se pretenda analizar la realidad urbana, y retomando la idea de Fainstein y Fainstein (1996), de que todo proceso de planificación implica invariablemente un acto político; habría que resaltar que todo acto político es llevado a cabo por actores, tanto individuales como colectivos, y que en la generación de las acciones públicas tendientes a resolver –o paliar– los problemas sociales y urbanos, el elemento indispensable radica en la creatividad y en la capacidad humana de hacer política.

En la actualidad, el desarrollo de las nuevas tecnologías no solo ha supuesto una toma de decisiones mayormente razonada y racionalizada, sino que también ha supuesto la asistencia de la tecnología en prácticamente todas las esferas de la vida de los seres humanos, tanto en lo social como en lo económico, cultural y político. Y en la administración pública como en la política pública, la incidencia de las tecnologías ha ido en incremento.

En esa sintonía, también es posible hablar hoy en día de la gobernanza de datos, la cual forma parte de un impulso a la gestión pública, sostenida a partir de la recolección de datos. De acuerdo con Sacks y Sherman (2019) la gobernanza de datos puede ser definida como las reglas sobre cómo los gobiernos interactúan con el sector privado –así como con otros gobiernos– cuando se trata de gestionar datos para determinar quién tiene acceso a ellos y las formas en que aquellos con acceso pueden usarlos (p. 7).

Es por ello por lo que, a partir de la incorporación de la tecnología, que ha supuesto una toma de decisiones coherente con los datos recolectados, las discusiones globales se han centrado en reconocer quienes tienen acceso, cómo utilizan los datos y cuáles son los límites de su uso; salvaguardando, por supuesto, la integridad de los datos personales.

En este marco de innovación y de regulación tecnológica, así como de manejo de datos, se inserta la inteligencia artificial (IA). Desde nuestra perspectiva, una definición que se articula con la idea de la gobernanza es la propuesta por el matemático Richard Bellman en tanto considera a la IA como la automatización de actividades asociadas al pensamiento humano,

que incluye la toma de decisiones, la solución de problemas y el aprendizaje (Bowser, Sloan, Michelucci y Pauwels, 2017, p. 2).

Para nosotros, la incorporación de la IA a la gobernanza en general –y también en específico a la toma de decisiones en materia urbana– se da a partir de la posibilidad de automatizar procesamientos de datos y toma de decisiones que con anterioridad correspondía a la autoridad en turno realizar y decidir. Hoy en día, la IA posibilita dicha automatización de actividades a partir de la generación de un modelo heurístico para la toma de decisiones, incorporando a su vez a la idea de un subsistema de servicios urbanos que posibilita un sistema urbano como tal.

No obstante, habría que considerar que la aplicación de la IA para la gobernanza urbana tiene sus límites. En este trabajo planteamos dos ideas puntuales que consideramos a la luz del caso mexicano:

1) Que más que un modelo heurístico, se necesita estructurar un sistema de toma de decisiones pública en el que la IA sea parte integrante.
2) Que el problema al que se enfrenta el uso y aplicación de IA para la gobernanza urbana radica en los límites que impone el marco político imperante, así como por su incapacidad de sustituir el proceso político de toma de decisiones.

La primera de las ideas dificulta una verdadera planificación urbana, en tanto la mera recolección de datos y su utilización para la toma de decisiones requiere de un acondicionamiento, no solo para obtener los datos pertinentes, sino también para generar los productos necesarios para generar una gobernanza urbana. Jorge Arribas Castañeyra (2016) habla de una planificación urbana algoritmizable que posibilite generar datos pertinentes a partir de la información que se le brinda a los ordenadores. Por esta situación, el desarrollo de un sistema de información urbana que integre a la IA resulta de particular interés.

Por otro lado, desde la visión de las políticas públicas, el principal problema al que se enfrenta la aplicación va de la mano de los sesgos, la discriminación de las acciones gubernamentales tomadas, y particularmente, un elemento que consideramos medular son las posibles resistencias a la integración de las IA en la toma de decisiones. Algo particularmente visible en la política mexicana, es que los datos que se construyen como argumento

para la toma de decisiones, no son usualmente usados por los políticos, ya que le brindan prioridad al cálculo político-electoral. En ese sentido, la problemática de una posible incorporación de la IA pasaría mayormente por el plano de lo político que de lo técnico-racional.

Al respecto es necesario señalar que un gobierno que no hace uso de las herramientas tecnológicas a su alcance para mejorar su administración pública no necesariamente incurre en un delito, pero si en una práctica administrativa ineficiente, pues o bien por ignorancia evita implementar sistemas y reestructuraciones que le ayudarían a su mejor funcionamiento (o en el mejor de los casos los utiliza de forma inadecuada), o por conocimiento de causa evita utilizarlos al estar consciente de que una correcta implementación de técnicas impediría la ineficiencia gubernamental de la que muchas veces se obtienen beneficios económicos y políticos a costa del gasto público.

Cabe a la reflexión de esto último un replanteamiento de la pregunta con la que Herbert Marcuse (1984) comienza su introducción a El Hombre unidimensional: ¿La amenaza de una inteligencia artificial que puede acabar con la raza humana no sirve también para proteger a las mismas fuerzas que perpetúan este peligro?

El gobierno de las máquinas

Cuando en las ciencias sociales se reflexiona sobre el gobierno, inmediatamente se piensa en este como una actividad (y en algunos casos en una institución que ejerce dicha actividad) exclusivamente social, limitando dicha exclusividad social a lo humano. Con ello, se oblitera el hecho de que la actividad misma de gobernar es una acción que va más allá del alcance de la o las personas, incluyendo en ello una gran diversidad de sistemas sociales, sistemas naturales y por qué no, sistemas tecnológicos.

Si se piensa por un momento en la relación que el gobierno en tanto entidad ejerce sobre la sociedad, posteriormente se piensa la relación que este mismo ejerce sobre la naturaleza y subsiguientemente sobre los sistemas tecnológicos; para todos los casos nos encontraremos que la acción de gobernar es equifuncional a lo siguiente: "El gobierno es un proceso de información estabilizador de sistemas complejos" (Beer, 1982, p. 470).

Esto no indica que la sociedad, la naturaleza y las máquinas sean entidades del mismo tipo y operatividad, sino que funcionalmente el gobierno para cada uno de los casos funciona de la misma forma: estabilizando los sistemas.

En el caso de la sociedad, la estabilidad del sistema social, a través del gobierno, se logra equilibrando las exigencias y necesidades de los distintos actores sociales y los rendimientos administrativos de su gobierno; para la naturaleza la estabilidad se encontraría en la posibilidad de equilibrio de los distintos elementos que conforman un nicho ecológico a través de mantener satisfechas las necesidades de todos los seres vivos y sistemas físicos que le conforman sin que esto requiera intromisiones externas; para las máquinas la estabilización se alcanza logrando coincidir ciertos requerimientos de entrada (inputs) con los requerimientos de salida (outputs) deseados que mantienen el sistema en un funcionamiento óptimo.

Como se puede ver este principio de gobierno se asemeja más al ideal de un gobierno democrático, y por principio, a la posibilidad de la gobernanza, que es la forma en la que normalmente se piensa que debe de funcionar tradicionalmente un gobierno; pues, ¿cómo podría un gobierno mantener la estabilidad social ignorando la información que los distintos sectores de la sociedad le demandan?

La opción que se ofrecerá más adelante consiste en pensar la posibilidad de un sistema de información compartida entre gobierno y gobernados cuyo principio operativo funcione a partir del procesamiento de información, y redireccionamiento de esta a partir de la IA, perfeccionando con su uso un modelo que retomaremos de la propuesta de la IO de Stafford Beer.

Como se describirá a continuación, esto no implica una eliminación de la creatividad ni de la toma de decisiones humanas, tampoco una deshumanización de la política, así como tampoco reduce la responsabilidad de nuestros gobernantes, ni de la misma sociedad civil en la participación política, sino todo lo contrario. En este sentido, debemos comprender que el uso de la tecnología en lo cotidiano, y en la gobernanza, siempre y cuando se oriente por principios de eficiencia, honestidad y transparencia, es una responsabilidad compartida; ya que no es posible "evitar la responsabilidad personal de una decisión peligrosa o desastrosa, colocando la

responsabilidad en otra parte: en el azar, en los superiores humanos cuyas políticas no es posible desafiar [o en la ciudadanía] o en un dispositivo mecánico […]" (Wiener, 2007, pág. 44), desde estos supuestos, gobierno y gobernados mantienen una responsabilidad compartida. Puesto que las computadoras y las tecnologías son sistemas lógicos creados por humanos y usados por humanos, queda claro que "las computadoras no se equivocan. […] Los que programan las computadoras se equivocan" (Beer, 1977, pág. 48).

Las computadoras y la IA, no son entidades cuya autonomía les permita por sí solas adquirir una capacidad de agencia tal como para rebelarse y someter a la humanidad a costa de las primeras tres leyes, aún y cuando esto último termina ocurriendo; sino que, cada vez que la tecnología y las maquinas terminan sometiendo a la humanidad, ya sea en las descripciones del marxismo y el neomarxismo, ya sea en los futuros distópicos imaginados por los literatos de la ciencia ficción, el hecho ocurre más en el sentido de una relación simétrica de actantes en el sentido desarrollado por Bruno Latour (2005), que en una relación de dominación asimétrica de la máquina sobre el ser humano.

En este sentido, como bien señala Latour (2005, p. 107) "[no] son los objetos los que hacen las cosas 'en lugar de' los actores humanos", sino que su capacidad de agencia se encuentra sujeta a las posibilidades potenciales que su presencia, en relación con la agencia humana y social, despliega en nuevas formas de ejercer acciones sociales que, sin la existencia de dichos objetos, no hubiera sido posible. Y aunque objetos pueden pasar a ser simplemente utilizados de forma instrumental, en tanto instrumentos no son únicamente medio para un fin (aun y cuando hayan sido utilizados bajo esa lógica), sino que al mismo tiempo participan en la construcción de dicho fin, toda vez que sin la presencia de dicho objeto (una herramienta, un artilugio de la tecnología o un software de IA) es altamente probable que el fin buscado no se hubiese alcanzado, y en caso de lograrse, el fin es distinto en cualidad, cantidad y tiempo requerido. En resumen, un resultado completamente distinto a través de acciones también distintas, aunque muchas veces equifuncionales.

Se puede conjeturar que un artilugio tecnológico, y en este caso específico, aun actualmente un sistema cuyo funcionamiento se gobierne por la

IA, no puede realizar ninguna acción, ni nada que se pueda catalogar como agencia, si no participa al lado de la acción humana. Todos los artilugios tecnológicos y herramientas que el ser humano ha desarrollado a lo largo de la historia han tenido como origen y fin el servir para facilitar tareas que de alguna manera el humano mismo ya realizaba, aunque en última instancia el uso de estos objetos no coincida únicamente con la función original para la cual fueron pensados. Esta variabilidad funcional del uso de los objetos posibilita la contingencia de los resultados de su uso mismo y potencializa la necesidad de atención a nuevas necesidades no anticipadas, resultado del uso tanto previsto como del imprevisto que se hace de ellos, obligando así a una cadena de innovaciones que da igual catalogar como problemas o soluciones.

Los temores que el desarrollo tecnológico ha producido a lo largo de la historia y los que se proyectan a futuro, corresponden a las consecuencias conocidas de su uso. Sin embargo, así como ocurre con otras formas de acción social, rara vez se cuestiona los riesgos de aquello que hubiera ocurrido en el caso de que la acción social no hubiera tenido lugar (Merton, 1936). Un ejemplo bajo esta línea de reflexión reconoce que el desarrollo de la agricultura y la ganadería extensivas han provocado grandes problemas ecológicos, a la vez que han hecho posible combatir la problemática alimentaria a nivel mundial.

Por otra parte, la automatización de procesos productivos han reducido los puestos de trabajo y han transformado los existentes a actividades monótonas, repetitivas y alienantes, han hecho posible la reducción de las jornadas laborales, la redistribución de la riqueza (aún y cuando esta se ha acumulado de forma desmesurada en cada vez menos manos) y la disponibilidad de tiempo de ocio para que el trabajador común pueda dedicar una pequeña porción de su día a las actividades que más le convengan (incluso el ocupar otro trabajo); mientras la introducción de la IA en la vida cotidiana ha hecho posible la suplantación de la identidad, y con ello la producción-simulación de creaciones que anteriormente eran exclusivos de la creatividad humana. Asimismo, ha hecho posible hasta niveles inhumanos la velocidad de reacción de los sistemas dinámicos que funcionan a partir de la información. Es pues, acerca de esta posibilidad, sobre el hecho que recae la propuesta del siguiente capítulo: la IA es una herramienta con amplio potencial para hacer posible un auténtico ejercicio

de gobernanza urbana.

Un modelo para la gobernanza

Sabemos que la cantidad de información que los distintos sectores de la sociedad producen y que la política requiere para la toma de decisiones es tanta, y en muchas ocasiones en bruto, que requiere de una cantidad considerable de tiempo y en algunas ocasiones del capital humano que engrosa las filas de la rebasada burocracia gubernamental.

Todo gobierno político que renueva una administración promete un mejor uso de recursos y una mejora de atención a las necesidades de la ciudadanía, sin embargo, rara vez se observa un esfuerzo legítimo por modificar el funcionamiento y la estructura (de la cual depende aquel funcionamiento) del mismo sistema de gobierno; más bien lo que se observa es una adecuación de los nuevos actores en el gobierno a las estructuras gubernamentales ya establecidas.

Lo mismo ha ocurrido con la implementación de la tecnología en las actividades administrativas, ya que como señala el mismo Stafford Beer (1977), esta rara vez se ha utilizado para mejorar el tratamiento de la información, en su lugar se acumula y utiliza bajo los mismos principios de la burocracia humana (ahora digitalizada), solo que con un costo menor de capital humano. Si lo que el gobierno quiere es cambiar los comportamientos y resultados indeseables de su actuar, "lo que tiene que hacer es cambiar la estructura del sistema que lo produce" (Beer, 1982, p. 471). Este cambio estructural, que en la presente propuesta solo se refiere a la forma en la que el gobierno y la ciudadanía interactúan a partir del compartir información necesaria para la toma de decisiones, es una alternativa para transformar la forma en la que actualmente se desarrolla el gobierno y hacer posible la gobernanza haciendo uso de la IA.

Primero debemos comprender al Gobierno como un sistema viable, entendiendo por sistema viable a un sistema que tiene la "posibilidad de dar una respuesta a un estímulo que no se incluyó en la lista de estímulos anticipados, cuando se diseñó el sistema" (Beer, 1982, pág. 265). Este estímulo no anticipado puede ser, tanto una demanda ciudadana de algún tema sin atender, como un conjunto de datos específicos de la misma

ciudadanía que el gobierno requiere, y que normalmente obtiene de sus mismas dependencias, para estar por enterado de que todo está en orden (reportes de muertes, nacimientos, crímenes, accidentes, tasa ocupacional y de desempleo, etcétera).

Lo problemático de esto es que, estos estímulos no anticipados que en adelante denominaremos para nuestro modelo información ciudadana, es tanta en cantidad y cualidad que un gobierno sin un sistema de procesamiento de información adecuado, se encuentra imposibilitado para atender y reaccionar en la inmediatez, lo que ocasiona un tiempo de reacción demasiado lento y que provoca que las necesidades de la ciudadanía queden sin ser atendidas (pues no se alcanza a comprender cuales son estas necesidades) o tardan demasiado tiempo en atenderse, cuyo lapso de tiempo implica la generación de nuevos problemas.

La información ciudadana en este modelo se consideraría lo que en la IO se conoce como variable, siendo la variable "el número de estados posibles de un sistema" (Beer, 1977, pág. 22). Como nuestro sistema de gobierno no tiene la capacidad de atender toda la proliferación de variables (i.e. de información ciudadana) lo que ha hecho históricamente ha sido desarrollar un aparato burocrático que en la actualidad se encuentra rebasado; pues no existe la posibilidad de atender en la inmediatez cada información ciudadana que el gobierno requiere y que le es entregada para su operatividad.

La burocracia es también una variable (variable reguladora) de nuestro sistema viable de gobierno, cuya función es contrarrestar la variable información ciudadana buscando cumplir la Ley de la variabilidad requerida o Ley de Ashby, el problema es que la burocracia no es lo suficientemente amplia como para cumplir con dicha Ley. La burocracia, en toda su extensión, es una variable que funciona como atenuador de la variable información ciudadana. La finalidad del atenuador de variable no es disminuir la información ciudadana, sino poder procesarla de tal forma en la que el aparato gubernamental pueda tratarla y generar posteriormente un amplificador de la variable reducida que tendría como finalidad dar atención a la información ciudadana, esto es una política pública.

Nuestro problema de gobierno, y que en la actualidad obstaculiza la gobernanza, entendiendo para los términos de este modelo a la gober-

nanza como: el estado de equilibrio óptimo entre la información solicitada y la obtenida entre gobierno y ciudadanía para la atención de la toma de decisiones colectivamente vinculantes, es que la atención a las solicitudes y a la información entregada por las organizaciones de la sociedad civil, es tan lenta y la información ciudadana es tanta, que bajo el funcionamiento de la burocracia gubernamental actual (aún y cuando los gobiernos manifiesten una legítima intencionalidad de desarrollo de la gobernanza) la velocidad de reacción es tan lenta que para cuando el gobierno requiere tomar decisiones, es más económico (e irónicamente eficiente) ignorar a la ciudadanía que esperar la información necesaria.

El problema con la velocidad de reacción (la cual se calcula en relación con la magnitud de una variable con su respectivo atenuador de variable) es que durante el tiempo de reacción la variable información ciudadana sigue proliferando y se realimenta con el amplificador de variable que es la política pública, generando mayor inestabilidad en el sistema, ya que la burocracia no tiene la capacidad de sintetizar toda la información requerida para que el gobierno produzca la política pública adecuada. En este sentido, ya se puede conjeturar que la propuesta es muy simple, aunque no obvia; la burocracia gubernamental como la conocemos debe ser sustituida por un software de procesamiento de datos gobernado por la IA que tenga la capacidad de atenuar la variable información ciudadana de forma que sea posible incluir, en la práctica real, a la ciudadanía en los procesos de participación ciudadana.

Conclusiones

El uso de las tecnologías ha venido a renovar la manera en cómo vivimos nuestra vida en sociedad. Sin embargo, los efectos de la incorporación de dichas tecnologías no son siempre los esperados, e inclusive se puede llegar a tener efectos no deseados como parte de estos desarrollos tecnológicos. En ese sentido, la incorporación de nuevas tecnologías causa tanto incertidumbre como anhelo por un mejor futuro. La ciencia ficción es ilustrativa de lo anterior.

Hoy en día, la inteligencia artificial (IA) ha venido cobrando importancia y su presencia es palpable, desde nuestra vida cotidiana, hasta en

los ámbitos más trascendentales, como debiese ser, el quehacer político. No obstante, el generar una aceptación a ultranza de cada tecnología desarrollada, sin tener una interpretación crítica con respecto a los pros y contras que implica su uso, puede llevar a un uso indiscriminado de estos desarrollos tecnológicos.

Desde la perspectiva que planteamos en este artículo, consideramos que el verdadero problema con el uso de las tecnologías, y en específico con la IA, recae en el uso que hacemos de ella; que más que en el hecho de que estas engendren una problemática por sí mismas, es en el uso irresponsable, ineficiente, y en algunos casos faccioso de la política, donde estas tecnologías pueden generar más problemas que resolverlos.

En el caso de la política pública urbana, la IA puede ser un factor que abone a tener un uso eficiente de recursos y que disminuya la posibilidad de un mal ejercicio de gobierno; sin embargo, si el factor político resulta predominante como en el caso mexicano, o la solución propuesta por la IA no resulta aceptada, su incorporación en el marco de una gobernanza será puesta en entredicho.

Por último, si bien el modelo propuesto es un esbozo demasiado general, es una primera opción que ofrece una línea de desarrollo posterior para la implementación de mecanismos de compartimiento y procesamiento de información que son necesarios para la administración gubernamental y la participación ciudadana.

Bibliografía

Arribas Castañeyra, J. (27 de Noviembre de 2016). *Smart Urban Planing. Inteligencia artificial en la planificación urbana.* Recuperado el 26 de Diciembre de 2023, de eSMARTICITY. https://www.esmartcity.es/comunicaciones/smart-urban-planing-inteligencia-artificial-planificacion-urbana

Asimov, I. (1950). *Yo, Robot.* Edición digital Gonzalez.

Beer, S. (1977). *Diseñando la libertad.* Madrid: Fondo de Cultura Económica.

Beer, S. (1982). *Decisión y control.* Distrito Federal: Fondo de Cultura Económica.

Bowser, A., Sloan, M., Michelucci, P., & Pauwels, E. (2017). *Artificial Intelligence: A Policy-Oriented Introduction.* Washington: Wilson Center. Obtenido de https://www.wilsoncenter.org/sites/default/files/media/docu-

ments/publication/wilson_center_policy_brief_artificial_intelligence.pdf

Bunge, M. (2004). *La investigación científica. Su estrategia y su filosofía*. México: Siglo XXI.

Fainstein, S. S., & Fainstein, N. (1996). City planning and political values: an updated view. En S. Campbell, & S. Fainstein. (s. f.). *Readings in planning theory* (págs. 265-287). Cambridge, MA: Blackwell Publishers.

Habermas, J. (1986). *Ciencia y técnica como ideología*. Madrid: Tecnos.

Habermas, J. (2002). *El futuro de la naturaleza humana*. Barcelona: Paidós.

Horkheimer, M. (1973). *Crítica de la razón instrumental*. Buenos Aires: Sur.

Jones, D. F. (1966). *Colossus*. Londres: Rupert Hart-Davis.

Lascoumes, P., & Le Galès, P. (2014). *Sociología de la Acción Pública*. México D.F.: El Colegio de México.

Latour, B. (2005). *Reensamblar lo social. Una introducción a la teoría del actor-red*. Buenos Aires: Manantial.

Le Galès, P., & Ugalde, V. (2018). Introducción. En P. Le Galès, & V. Ugalde, *Gobernando la Ciudad de México. Lo que se gobierno y lo que no se gobierna en ua gran metrópoli* (págs. 15-52). Ciudad de México: El Colegio de México.

Marcuse, H. (1984). *El hombre unidimensional*. Barcelona: Planeta-Agostini.

Marx, C. (1990). *El capital. Crítica de la economía política*. Moscú: Progreso.

Meny, I., & Thoenig, J. C. (1992). *Las Políticas Públicas*. Barcelona: Editorial Ariel S.A.

Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American Sociological Review*, 894-904.

Sacks, S., & Sherman, J. (2019). *Global Data Governance. Concepts, Obstacles and Prospect*. The Howard Baker Forum. New America. Obtenido de https://d1y8sb8igg2f8e.cloudfront.net/documents/Global_Data_Governance_final_eyG7dLY.pdf

Sargent, J. (Dirección). (1970). *Colossus*: el proyecto Forbin [Película]

Wiener, N. (2007). *Dios y Golem S.A*. Distrito Federal: Siglo XXI.

Capítulo **11**

Sistemas de predicción de lenguas visogestuales basados en IA y su aplicación en la Lengua de Señas Mexicana (LSM)

Guillermo Hernández-Santana
Irvin Hussein López-Nava

https://doi.org/10.61728/AE24001113

Introducción

Es indudable que a partir del desarrollo de la inteligencia artificial (IA) y los sistemas de procesamiento de lenguaje natural (PLN), las computadoras comenzaron a comprender, interpretar y generar lenguaje humano en sus diversas formas, sin embargo, estos sistemas cuentan con un mayor desarrollo en relación con las lenguas orales que a las lenguas de modalidad visogestual. Esto responde a la colección de datos lingüísticos que contamos en la actualidad, la cual, es significativamente mayor en lenguas orales que señadas. Es decir, que el desarrollo de la IA y del PLN tiene en la actualidad avances importantes en cuanto a la capacidad de las máquinas para interpretar y generar texto en diversos contextos, desde traductores automáticos hasta asistentes virtuales, pasando por chatbots.

En este capítulo discutimos la importancia del reconocimiento automático de las lenguas de modalidad visogestual, vital para la inclusión y comunicación efectiva de las personas sordas. En primer lugar, estos sistemas representan un avance tecnológico significativo en el campo de la accesibilidad, permitiendo una mayor autonomía y participación de este sector de la población en diversos contextos sociales y profesionales. Algunos de los beneficios de estas aplicaciones y su implementación pueden repercutir en tecnologías de reconocimiento automático de lenguas de señas las cuales pueden facilitar la comunicación entre personas sordas y oyentes, reduciendo las barreras lingüísticas y promoviendo la inclusión social, entre otros beneficios.

En segundo lugar, el desarrollo de sistemas de reconocimiento de lenguas de señas impulsa la innovación en áreas como la IA. De hecho, estos sistemas requieren algoritmos complejos y sofisticados que puedan procesar y entender las expresiones faciales en tiempo real, lo que representa un desafío importante en el campo de la visión por computadora. En este sentido, la investigación en esta área contribuye al avance del conocimiento en técnicas de aprendizaje profundo, procesamiento de imágenes y análisis de movimiento, aplicables en una amplia gama de tecnologías emer-

gentes, por ejemplo, los gestos que acompañan el lenguaje oral y también aspectos relacionados a la gestualidad humana en diferentes situaciones contextuales.

Este tipo de tecnología utiliza modelos de aprendizaje automático y visión por computadora para interpretar los gestos y movimientos característicos de las lenguas de señas, convirtiéndolos en texto o voz y sirve como herramienta de traducción en medios educativos para personas sordas que necesitan desarrollar habilidades lingüísticas y para oyentes que deseen aprender lenguas de señas. En este sentido, la integración de esta tecnología en plataformas educativas y aplicaciones móviles ofrecen oportunidades para el aprendizaje interactivo y accesible de lenguas de señas, promoviendo su difusión y comprensión (Gortarez-Pelayo et al. 2023).

Los sistemas PLN, por su parte, utilizan algoritmos y modelos estadísticos para analizar y comprender el significado del lenguaje humano, incluidas las variaciones en la forma de expresión, la semántica y la gramática. Pero el estado actual del estudio de las lenguas de señas no nos ha permitido reconocer del todo cuáles son los rasgos significativos de una lengua de modalidad visogestual a diferencia del conocimiento y bancos de datos que tenemos en materia de lenguas orales, la cual es significativamente menor. Particularmente, es necesario el desarrollo y comprensión de aquellos rasgos gramaticales más allá de las configuraciones manuales, es decir, aquellos rasgos de tipo gestuales que cargan información tempoaspectual de las lenguas señadas (Fridman, 2005).

En este capítulo discutimos que los sistemas basados en IA pueden adaptarse para reconocer y procesar patrones específicos de gestos y movimientos asociados con la lengua de señas. De hecho, estos sistemas cuentan con la capacidad para entender la estructura y el significado de la lengua de señas, lo cual puede contribuir directamente a la creación de tecnologías más efectivas en la interpretación y traducción automática, promoviendo así la inclusión y accesibilidad para la comunidad de personas Sordas. En la siguiente sección analizamos la evolución de IA y cómo se ha visto involucrada la tecnología de reconocimiento de imagen desde los primeros patrones de movimiento mediante diferentes tecnologías hasta cámaras y algoritmos para reconocer los patrones de movimiento.

Desarrollo del reconocimiento de lenguas de señas mediante IA

La historia del reconocimiento automático de lenguas de señas mediante inteligencia artificial (IA) se remonta a las últimas décadas del siglo XX, cuando los avances en la visión por computadora y el procesamiento de señales digitales comenzaron a permitir la detección y el seguimiento de gestos humanos. Los primeros sistemas se basaban en técnicas rudimentarias de reconocimiento de patrones y requerían de guantes sensorizados o marcadores en las manos para detectar los movimientos. Estos sistemas, aunque innovadores, eran limitados en su capacidad para reconocer gestos complejos y a menudo eran intrusivos y poco prácticos para los usuarios (Sturman y Zeltzer, 1994). La dependencia de hardware especializado y la falta de portabilidad eran desafíos significativos en esta fase inicial.

Con la llegada del siglo XXI, la visión por computadora experimentó avances significativos, impulsados en parte por el aumento en la capacidad de procesamiento y algoritmos más sofisticados. Esto permitió el desarrollo de sistemas de reconocimiento de lenguas de señas más avanzados, que podían capturar gestos mediante cámaras estándar sin la necesidad de dispositivos adicionales. La aplicación de algoritmos de aprendizaje automático comenzó a permitir un reconocimiento más preciso de los gestos y una mejor interpretación del contexto de las señas (Murakami y Taguchi, 1991). Sin embargo, estos sistemas aún luchaban con la variabilidad en la ejecución de las señas entre diferentes usuarios y las limitaciones en la diferenciación de gestos similares.

La introducción de las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN) en el campo del reconocimiento de lenguas de señas marcó un hito importante. Estas tecnologías permitieron a los investigadores abordar las complejidades inherentes al lenguaje de señas, como la variabilidad en la ejecución de las señas y la importancia del contexto y la fluidez en la señalización. La capacidad de las CNN para procesar imágenes y de las RNN para entender secuencias temporales resultó ser particularmente útil para interpretar los movimientos dinámicos y las expresiones faciales asociadas con las lenguas de señas (Pigou et al., 2015). Este avance representó un paso significativo hacia sistemas más robustos y precisos.

A pesar de los avances tecnológicos, el reconocimiento de lenguas de señas por IA continuó enfrentando desafíos significativos. Uno de los más importantes era la variabilidad en la forma en que diferentes personas realizan las señas, lo que se veía agravado por factores como la edad, el género y el origen cultural; sin mencionar la gran cantidad de lenguas de señas que existen a nivel mundial. Además, el contexto en el que se realiza una seña puede cambiar su significado, lo que presenta un desafío adicional para su interpretación correcta. En la actualidad sabemos que las lenguas de señas son ricas en matices y requieren un nivel de comprensión contextual que los sistemas de IA inicialmente lucharon por alcanzar (Koller et al., 2015).

La recolección de grandes conjuntos de datos en determinadas lenguas de señas y el avance en técnicas de aprendizaje profundo, adoptadas principalmente de otras áreas, ha permitido entrenar modelos de IA más robustos y precisos. El aprendizaje profundo, particularmente con arquitecturas de redes neuronales profundas, ha sido fundamental para capturar las complejidades y sutilezas de las lenguas de señas, mejorando significativamente la precisión del reconocimiento (Camgoz et al., 2018; 2017).

La integración de diferentes tecnologías de sensado, como el Microsoft Kinect o el Leap Motion, junto con avances en realidad aumentada, ha abierto nuevas posibilidades para el reconocimiento de lenguas de señas. Estos dispositivos proporcionan una captura de movimiento más precisa y menos intrusiva, lo que permite una detección detallada de los gestos y expresiones faciales. Además, la realidad aumentada ofrece formas innovadoras de interactuar con los sistemas de reconocimiento de lenguas de señas, mejorando la experiencia del usuario y facilitando la comunicación en tiempo real (Lang et al., 2020).

El progreso en el reconocimiento de lenguas de señas por IA ha llevado a aplicaciones prácticas que mejoran la accesibilidad para las personas sordas. Estas aplicaciones incluyen sistemas de traducción en tiempo real, asistentes virtuales personales y herramientas educativas. Estos sistemas no solo proporcionan medios para una comunicación más efectiva, sino que también promueven la inclusión y la igualdad de acceso a servicios e información para la comunidad Sorda (Bragg et al., 2019). En la siguiente sección hacemos una revisión del estado del arte en los sistemas de predicción de lenguas de señas.

Sistemas de predicción en diferentes lenguas de señas en el mundo

Como se señaló en la sección anterior, dos aspectos a considerar en el conocimiento actual de las lenguas de señas son la variación del habla individual, así como de las formas y expresiones que se desarrollaron de manera independiente en diferentes contextos geográficos, es decir, tenemos poca información acerca de la manera en que funciona la gramática y el discurso de los idiomas visogestuales. Aunado a esto, no podemos esperar uniformidad de las expresiones en las diferentes lenguas de señas, por el contrario, lo esperable es encontrar variación lingüística, tanto en el eje diacrónico como en el geográfico.

Lo anterior se explica por algunos aspectos de las lenguas naturales. Estos son la variación, la distancia geográfica, así como el desarrollo de neologismos, lo cual es indudable en situaciones de contacto lingüístico y ante la necesidad de crear léxico ante nuevas situaciones comunicativas.

Por un lado, el conocimiento actual de patrones, expresiones y reglas de formación de las lenguas de señas en la actualidad es limitado, pues no contamos con tantas bases de datos o corpora lingüística que nos informe cuáles son los patrones articulatorios que conforman la gramática de la lengua de señas y cuáles no; es decir, criterios que aporten información de las diferencias estructurales y pragmáticas de la lengua de señas que alimenten modelos de reconocimiento de expresiones mediante IA.

En este apartado revisamos una variedad de modelos que se han llevado a cabo para el reconocimiento de patrones lingüísticos de lenguas de modalidad visogestual en la actualidad. Para el desarrollo de este apartado, se llevó a cabo una búsqueda de literatura en las plataformas Scopus, Google Académico y la base de datos de educación ERIC. La estrategia de búsqueda en estas plataformas se definió mediante la combinación de términos específicos: "lengua de señas mexicana, inteligencia artificial, reconocimiento de datos, gestualidad, neural language processing" de artículos realizados a partir de 2023 considerando el rápido auge que han tenido los sistemas de procesamiento de lenguas naturales a partir del desarrollo de los chatbots. Se consideraron tanto trabajos escritos en español como en inglés. Para garantizar la consistencia en la selección de estudios, se exclu-

yeron aquellos relacionados con temas de salud y también se excluyeron trabajos de tesis.

Este procedimiento aseguró una recopilación exhaustiva y focalizada de la literatura pertinente mediante el metaanálisis PRISMA (Moher et al., 2015), facilitando así un análisis detallado de la intersección entre la Lengua de Señas Mexicana y la inteligencia artificial en el contexto del reconocimiento de datos y gestualidad. Además, se aplicó un criterio de exclusión al descartar los artículos que no fueron revisados por pares, asegurando la calidad y confiabilidad de la información recopilada. A partir de esta revisión se realizó un análisis cualitativo de artículos referidos en las obras para ampliar la información al respecto que nos pudiera dar información acerca de los diferentes patrones articulatorios que deberían considerarse en los idiomas de modalidad visogestual, así como aquellos aspectos relevantes desde un punto de vista gramatical. Adicionalmente, se incluyeron algunas búsquedas particulares de tecnología que había sido reportada en las obras pero que no necesariamente estaban incluidas en las bases de datos.

Para llevar a cabo la búsqueda en lengua de señas mexicana, se optó por metodología cualitativa que diera cuenta de las tecnologías relacionadas con este idioma en diversas fuentes y portales especializados. La metodología incluyó la revisión y consulta de plataformas académicas (Scopus, ProQuest e IEEE Xplore) con términos como "lengua de señas mexicana", "inteligencia artificial", "reconocimiento de datos", "gestualidad" y "procesamiento de lenguaje natural" y también optamos por los antecedentes que los autores tienen en el conocimiento de la bibliografía especializada en la lengua de señas mexicana, dado que uno de ellos ha desarrollado un sistema de reconocimiento de patrones lingüísticos (Morfín-Chávez, Gortarez-Pelayo y Lopez-Nava, 2023). Este enfoque metodológico permitió obtener una visión integral de las tecnologías asociadas con la lengua de señas mexicana.

En el ámbito de la captura del lenguaje de señas, se busca registrar gestos utilizando configuraciones de sensores apropiadas con el fin de obtener información discriminativa de los elementos que se capturen. Este proceso facilita el estudio, reconocimiento y representación 3D de las señas en etapas posteriores, permitiendo también la construcción de

conjuntos de datos extensos. La variedad de configuraciones de sensores empleados destaca la adaptabilidad de diversas aplicaciones (Papastratis, 2021). En los siguientes apartados, se presentan los dispositivos de captura más importantes que se han empleado en los últimos años en materia de lengua de señas.

Sensores de visión 2D

Los avances en las tecnologías de sensado han marcado un progreso notable, con categorías diferenciadas según la tecnología empleada y los datos capturados. En el año 2023, Morfín-Chávez y su equipo implementaron una estrategia innovadora utilizando una cámara 2D junto con la biblioteca MediaPipe para realizar la segmentación de la mano y extraer los puntos clave (keypoints) en dos dimensiones, y aproximando la tercera dimensión. Este enfoque se destaca por su simplicidad y accesibilidad, ya que permite la captura de imágenes desde dispositivos comunes, como cámaras web o teléfonos móviles, y la extracción de características detalladas de las manos mediante algoritmos avanzados de visión por computadora. La simplicidad de este método no solo facilita la portabilidad y respuesta inmediata, sino que también se trata de una solución idónea para aplicaciones que requieren facilidad de uso, como aplicaciones móviles o sistemas basados en la web. Esta combinación de una cámara 2D estándar y la biblioteca MediaPipe representa una contribución significativa al campo del reconocimiento de lenguas de señas, al ofrecer una alternativa para la captura y análisis de elementos gestuales, allanando el camino para soluciones prácticas y generalizadas en el ámbito de la tecnología de asistencia.

Sensores de visión 3D

Entre otros inputs de imagen que se han utilizado se reporta el uso del sensor Kinect como medio captura de datos de profundidad y de color RGB (red, green, blue) así como de la estructura esquelética del individuo. Se han encontrado aplicaciones similares con cámaras de visión artificial. El uso de este sensor destaca de una categoría avanzada de dispositivos de sensado ya que está compuesto tanto de una cámara 2D tradicional

como un sensor de profundidad (Garcia-Bautista et al., 2017; Sosa-Jiménez et al., 2017). Por lo que el sensor Kinect proporciona capacidades de seguimiento a partir de keypoints, lo que lo hace ideal para aplicaciones que requieren un alto nivel de detalle y precisión en el reconocimiento de gestos. Esta tecnología es capaz de capturar no solo el movimiento de las manos, sino también la postura y los movimientos del cuerpo, ofreciendo una comprensión más completa de las expresiones que puede registrar. Este sensor es especialmente útil en entornos de investigación y aplicaciones que requieren una alta fidelidad en la captura de datos gestuales.

Otros sensores que se han utilizado son Leap Motion, que tiene como característica poder capturar posiciones 3D de la mano y los dedos, para grabar gestos y palabras de lenguas de modalidad visogestual. Sin embargo, una de las desventajas de este tipo de sensor es que es necesario operar cerca del sujeto, aunque una de sus ventajas es su bajo requerimiento computacional, el cual les permite su uso en aplicaciones en tiempo real y es capaz de reconocer las posiciones 3D de manos y dedos.

Sensores fisiológicos

En el ámbito del reconocimiento de las lenguas de señas, otra tecnología es la utilización de sensores fisiológicos relacionados con la captura del movimiento muscular (Papastratis, 2021). Esta tecnología permite explorar nuevas formas de capturar y comprender las configuraciones manuales utilizadas en la lengua de señas. Un ejemplo paradigmático es el brazalete Thalmic MYO, reconocido por su capacidad para capturar actividades electromiográficas con precisión.

Este dispositivo ha sido específicamente empleado para el reconocimiento del alfabeto en la lengua de señas irlandesa pues su diseño innovador posibilita la captura de señales electromiográficas, proporcionando una visión detallada de los movimientos y gestos manuales asociados con la comunicación en lengua de señas. Al utilizar tecnología RFID, el brazalete Thalmic MYO ofrece un enfoque no intrusivo para el reconocimiento, ya que puede capturar datos de manera continua y en tiempo real. En este sentido, la incorporación de sensores de movimiento, como el Thalmic MYO, no solo amplía las posibilidades de reconocimiento de lenguas de

señas, sino que también abre la puerta a aplicaciones más avanzadas en el ámbito de la tecnología de asistencia.

Guantes de datos y otros dispositivos

En el ámbito del reconocimiento de lenguas de señas, los guantes representan otra tecnología de precisión en tiempo real para traducir señas. Tenemos en caso del denominado BrithtSign (Ompi, 2019) que tiene la capacidad para capturar los movimientos detallados de los articuladores manuales para proporcionar un nivel de detalle excepcional en el reconocimiento de datos. Aunque ofrecen una precisión importante, ajustar los componentes de estos guantes puede resultar en un proceso laborioso. Esto puede implicar una curva de aprendizaje para los usuarios y, en algunos casos, limitar su adopción, especialmente en situaciones reales.

Adicionalmente, se ha explorado el uso de dispositivos como los relojes inteligentes (Herrera Crespo, 2021). Estos dispositivos ofrecen una configuración más sencilla en comparación con los guantes de datos. Algunos enfoques han implementado relojes inteligentes para recopilar señales de señas, que luego se transmiten a dispositivos móviles para su procesamiento y traducción en tiempo real. Aunque esta configuración simplificada puede ser más accesible, es importante considerar las limitaciones en términos de la cantidad de datos que pueden reconocer (Papastratis, 2021).

Como se ha podido observar, existen diferentes tecnologías que se han implementado para el reconocimiento de señas, destacan las cámaras 2D, sin embargo, también podemos ver guantes, sensores y relojes inteligentes. Cada proyecto de reconocimiento tiene sus propios retos y estrategias de sensado. Estas tecnologías ponen énfasis en el reconocimiento de las configuraciones manuales y rasgos reconocibles en la dactilología, el movimiento del cuerpo, y otros también destacan algunos rasgos adicionales, como el inicio y fin de una seña y la posición de la mano. Si bien la elección suele depender de las necesidades particulares de la aplicación y de las preferencias de los usuarios finales, estos desarrollos no solo amplían las opciones tecnológicas disponibles, sino que también subrayan la importancia de equilibrar la precisión con la usabilidad en el diseño de soluciones para la comunidad de usuarios de lenguas de señas. Otro punto por considerar

es que estas tecnologías se han desarrollado en una cantidad limitada de lenguas de señas en el mundo. Esto implica un desequilibrio en aquellos contextos menos desarrollados, por ejemplo, países subdesarrollados, que no cuentan con tantos recursos, tecnología o buenas políticas y prácticas educativas en materia de inclusión. En la siguiente sección, se ofrece un panorama de los principales sistemas visogestuales en los que se ha diseñado un sistema de sensado de patrones lingüísticos, mostrando así una generalización de los avances tecnológicos en este campo.

Evolución y reconocimiento de patrones lingüísticos a nivel global

A medida que la tecnología ha avanzado, los sistemas de sensado y reconocimiento aplicados a las lenguas de señas han experimentado notables desarrollos. Sin embargo, es importante señalar que estos avances se han enfocado predominantemente en un conjunto limitado de lenguas de señas a nivel mundial. Entre las lenguas de señas que han sido objeto de estudio y desarrollo de sistemas tecnológicos, destacan la alemana, de Hong Kong (en China), griega, británica, americana, árabe e india (Papastratis, 2021). En esta sección, exploramos algunas de las tecnologías específicas que se han aplicado en el reconocimiento de patrones lingüísticos a partir del desarrollo de la IA en estas lenguas, así como estudios relevantes que han contribuido a la evolución de estos sistemas. Como se verá, estas lenguas representan tan solo una fracción de la diversidad lingüística presente en las diversas comunidades sordas alrededor del mundo.

A nivel global, se han registrado de 130 a 271 lenguas de señas. En este ámbito, los sistemas de sensado y reconocimientos aplicados a las lenguas de señas han experimentado desarrollos relevantes, pero es importante destacar que estos avances se han centrado en un número limitado de lenguas de señas. Algunas de las tecnologías que se han aplicado, así como algunos estudios relevantes, se muestran en las siguientes secciones.

Lengua de Señas Indio (ISL)

Entre las diversas aplicaciones que merecen destacarse, se encuentra un estudio que se centra en el sistema de interpretación de la Lengua de Señas India (Sreemathy et al., 2023). Este sistema, diseñado para la interacción automática de signos, utiliza enfoques de aprendizaje profundo y se implementa en MATLAB como plataforma de simulación. Este integra redes preentreandas como AlexNet, GoogleNet, VGG-16 y VGG-19 y logra precisiones notables alcanzando el 99.11 %. En este mismo sistema señado también encontramos un trabajo de Sugandhi (2021). Este se basa en una captura de imágenes basado en una Evaluación Bilingüe de Referencia (BLEU) para las oraciones y una evaluación manual de palabras y oraciones. Dicho sistema cuenta con un corpus rico de las 800 oraciones más comúnmente utilizadas en el idioma inglés y se toma en consideración una salida basada en la gramática de ISL que compara el output con el puntaje.

Justamente, el puntaje BLEU se calcula utilizando métodos estadísticos y de procesamiento de lenguaje natural para determinar la similitud entre las oraciones generadas por el sistema y las oraciones de referencia creadas por los expertos en ISL. Los resultados de estas evaluaciones proporcionaron una medida clara de la eficacia del sistema propuesto en la generación de SL y respaldaron su utilidad como herramienta para la comunicación efectiva entre personas con pérdida de audición y aquellos que no tienen conocimiento de SL.

Además, en un estudio reciente realizado por Kothadiya et al. (2022), se utilizaron las redes neuronales LSTM y GRU mediante una colección de gestos manuales integrando expresiones como "hola", "buenos días" y "trabajo". En este estudio se mostró que al aumentar el número de capas en las redes neuronales y aplicar una combinación se suele mejorar la precisión del modelo.

Lengua de Señas Americana (ASL)

En Lengua de Señas Americana, encontramos el trabajo de Schioppo y colaboradores (2019), el cual reconoce la lengua de señas mediante un auricular de realidad virtual para crear un entorno inmersivo. Este estudio se

basa en un controlador de Leap Motion, cuyas características se combinan con un banco aleatorio para la clasificación en tiempo real de los gestos del usuario. Para evaluar la eficacia de este enfoque, la tecnología se basó en pruebas con las 26 letras del alfabeto de la ASL en un entorno virtual y uno de los objetivos fue el aprendizaje de la lengua de señas.

En materia de aprendizaje profundo, Agarwal et al. (2021) desarrollaron DeepSign, una herramienta de software la cual es capaz de evaluar el progreso que tiene un estudiante en la alfabetización. Esta innovadora solución automatiza la puntuación de medidas de nombramiento de imágenes en lenguaje de señas, ofreciendo retroalimentación inmediata y reduciendo la carga laboral de los docentes. En esta misma lengua de señas, encontramos un método de enseñanza que está dirigido a sordos y padres oyentes mediante entornos inmersivos de realidad virtual (Quandt, 2020). En dicho mundo virtual se emplean avatares que señan y enseñan ASL de manera inmersiva. Las contribuciones significativas incluyen el uso de avatares, la integración de gestos, vocabulario básico en ASL introductorio en un entorno 3D.

Lengua de Señas Británica (BSL)

En el contexto británico, se han desarrollado diversas herramientas automáticas que reconocen datos de trayectoria de la mano, datos faciales y datos de distribución del codo para el reconocimiento de rasgos lingüísticos en la British Sign Language (BSL). Estos datos son procesados mediante un modelo de red CNN, tal como se describe en el trabajo de Liang et al. (2020). Estos esfuerzos reflejan una iniciativa más amplia en el Reino Unido, donde se promueve activamente la mejora de la accesibilidad y la inclusión, como se discute en el estudio de Gellai (2023). Además, en el ámbito de la tecnología aplicada a las lenguas de señas, se han desarrollado dispositivos como guantes capaces de reconocer las configuraciones manuales de la BSL, como se evidencia en el trabajo de Ompi (2019).

También en el ámbito de la BSL, Albanie et al. (2020) desarrollaron un enfoque para la recolección de datos destinados al reconocimiento automatizado de patrones lingüísticos mediante la producción de subtítulos alineados a través del reconocimiento de palabras clave, lograron localizar

automáticamente instancias de señas en videos continuos, lo que resultó en la creación de un conjunto de datos de un banco de mil señas. Además, mostraron que este conjunto de datos puede ser utilizado para entrenar modelos sólidos de reconocimiento de signos.

Lengua de señas Árabe (ArSL)

Otro avance significativo en el reconocimiento de gestos lo vemos en el proyecto denominado DeepArSLR (Aly y Aly, 2020). Este avance de censado en ARSL se basa en el reconocimiento de gestos aislados e independientes del señante el cual se combina una red de segmentación semántica, una red convolucional y una red LSTM bidireccional profunda. Esta dinámica permite una representación de la forma de la mano capaz de reconocer configuraciones de mano a través del modelo SOM convolucional de una sola capa. Esta tecnología también incluye una nueva red de clasificación profunda, compuesta por tres capas BiLSTM bidireccionales conectadas para el sensado de secuencias de gestos.

En ArSL también reconocemos el trabajo de Elatawy (2020) que presenta una novedosa aproximación al reconocimiento del alfabeto de la lengua de señas árabe mediante el uso de predicciones neutrosóficas y significados difusos, una propuesta que convierte la información visogestual en un dominio neutrosófico mediante un filtro gaussiano en las imágenes que captura.

Lengua de Señas Alemana (DGS)

En el contexto germánico, se ha desarrollado desde hace varios años una amplia corpora lingüística que se denomina RWTH-PHOENIX-Weather. Esta es una colección de datos extraída de la televisión pública alemana, y anotados manualmente con glosas a nivel de oración. Además, se han integrado transcripciones semiautomáticas del alemán hablado utilizando el sistema de reconocimiento de voz. Adicionalmente, se han incluido anotaciones referentes a la forma y orientación de la mano dominante para más de cuartenta mil cuadros de video, convirtiéndolo en uno de los mayores corpus que permiten la evaluación de algoritmos de seguimiento

de objetos. Asimismo, se han anotado más de dos mil signos utilizando el sistema de anotación SignWriting, centrándose en la forma, orientación, movimiento y contactos espaciales de ambas manos (Forster, 2014; Koller et al. 2015).

En la revisión de esta sección, se observa un mayor desarrollo en el reconocimiento de tecnologías aplicadas en algunas lenguas de señas en comparación con otras. Las lenguas de señas indio (ISL), americana (ASL), británica (BSL), alemana (DGS) así como la (ArSL) han sido objeto de una considerable cantidad de investigación y desarrollo, lo que ha llevado a la creación de diversos conjuntos de datos, herramientas y modelos de reconocimiento de señas específicas. Sin embargo, se nota una brecha en el desarrollo de tecnologías para lenguas de señas menos estudiadas, tal es el caso de algunas lenguas visogestuales de la región Centro y Sudamericana así como las lenguas de señas en África y Oceanía.

Es probable que el limitado desarrollo en el reconocimiento de lenguas de señas se deba en parte al hecho de que algunas lenguas de modalidad visogestual son relativamente jóvenes y el tiempo de desarrollo de tecnologías en el área no ha sido suficiente. Un ejemplo de esto es la lengua de señas de Guatemala, la cual surgió en la década de los ochenta como resultado de una política de inclusión gubernamental. Por otro lado, en países donde las políticas de inclusión han sido menos efectivas, encontramos menos reportes de sistemas de reconocimiento mediante IA. Estas disparidades subrayan la urgente necesidad de dirigir una mayor atención y recursos hacia el desarrollo de tecnologías de reconocimiento de lenguas de señas para comunidades menos representadas. Además, es crucial continuar desarrollando amplios bancos de datos lingüísticos para el desarrollo de aplicaciones que puedan servir para aplicaciones de reconocimiento lingüístico.

Reconocimiento automático en Lengua de Señas Mexicana (LSM)

Los sistemas de procesamiento de lenguaje natural pueden adaptarse para reconocer y procesar patrones específicos de gestos y movimientos asociados con una lengua de señas. De hecho, estos sistemas en la actualidad

cuentan con la capacidad para entender la estructura y el significado de la lengua de señas, lo cual puede contribuir directamente a la creación de tecnologías más efectivas en la interpretación y traducción automática, promoviendo así la inclusión y accesibilidad para la comunidad de personas Sordas.

El reconocimiento automático de la Lengua de Señas Mexicana (LSM) también ha experimentado avances notables a lo largo de los años, reflejados en varios estudios significativos. En uno de los primeros estudios identificados (Solís et al. 2014) se enfocó en el reconocimiento de 24 señas estáticas de LSM utilizando una cámara digital bajo condiciones de luz incoherente y Redes Neuronales Artificiales (ANN). Este estudio logró una alta tasa de reconocimiento, demostrando la eficacia de las ANN en la interpretación de señas de LSM.

Posteriormente, en 2015, Galicia et al. abordaron la conversión de LSM a letras del alfabeto español oral utilizando sensores de movimiento y algoritmos de clasificación avanzados, en el que alcanzaron una precisión notable en la traducción de LSM. En el mismo año, Solis et al. se centraron en el reconocimiento de señas estáticas en LSM usando Jacobi-Fourier Moments y ANN, alcanzando una alta tasa de reconocimiento. En 2016, Solis et al. se enfocaron en 21 señas estáticas de LSM, utilizando una ANN para el análisis de datos obtenidos mediante una cámara digital. Este estudio logró una tasa de reconocimiento de casi 100%, reforzando la viabilidad de utilizar aprendizaje profundo y momentos normalizados para el reconocimiento de LSM.

García-Bautista et al., propusieron en el 2017 un método para reconocer una lista básica de palabras en LSM utilizando un sensor Kinect y el algoritmo de Dynamic Time Warping, logrando una precisión media del 99.1 %. Este enfoque demostró ser altamente efectivo para el reconocimiento en tiempo real de LSM. En el mismo año, Sosa-Jiménez et al. presentaron un estudio en el que clasificaban un conjunto de señas de LSM con datos obtenidos de un Kinect y logrando una sensibilidad promedio del 86 % y una especificidad promedio del 80 %. En 2022, Sosa-Jiménez et al. desarrollaron un sistema bidireccional para LSM en servicios de salud. Utilizando un sensor Kinect y modelos ocultos de Markov (HMMs), el sistema logró tasas altas de precisión, demostrando el potencial de los sistemas de traducción bidireccional en contextos de atención médica.

El estudio de Morfín-Chávez et al., presentado en 2023, se centró en el reconocimiento del alfabeto manual de la LSM, utilizando técnicas de aprendizaje automático como Support Vector Machines. Los datos fueron adquiridos con una cámara de video convencional, y se extrajeron 21 puntos clave de las manos con la biblioteca MediaPipe para entrenar los modelos de clasificación. Los resultados al clasificar las 21 letras superaron una F1-score de 0.98. Finalmente, en 2022, Mejía-Peréz et al. introdujeron un sistema para reconocer LSM basado en gestos múltiples, utilizando una cámara de profundidad (OAK-D) y redes neuronales recurrentes para la clasificación, alcanzando una precisión del 97 % en datos limpios y 90 % en datos ruidosos. Este enfoque avanzado subraya la efectividad de combinar cámaras de profundidad y RNN para un reconocimiento de LSM.

Sensores utilizados para la captura de las señas en LSM

Como se revisó en la sección anterior, la evolución en las tecnologías de sensado ha sido notable, pudiéndose clasificar en distintas categorías basadas en la tecnología empleada y el tipo de datos que capturan. En el caso de la LSM, en estudios como el de Morfín-Chávez et al. (2023), se utilizó una cámara 2D estándar en combinación con la biblioteca MediaPipe para la segmentación de la mano y la extracción de keypoints. Este enfoque destaca por su simplicidad y accesibilidad, permitiendo la captura de imágenes de dispositivos como cámaras web o cámaras de teléfonos móviles, y la extracción de características detalladas de las manos mediante algoritmos avanzados de visión por computadora. Este método es particularmente útil para aplicaciones que requieren portabilidad, respuesta inmediata y facilidad de uso, como aplicaciones móviles o sistemas basados en la web.

Los estudios de García-Bautista et al. (2017) y Sosa-Jiménez et al. (2017), que utilizan el sensor Microsoft Kinect, destacan una categoría avanzada de dispositivos de sensado ya que está compuesto tanto de una cámara 2D tradicional como un sensor de profundidad, por lo que el sensor Kinect proporciona capacidades de seguimiento del esqueleto, lo que lo hace ideal para aplicaciones que requieren un alto nivel de detalle y precisión en el reconocimiento de gestos. Esta tecnología es capaz de capturar no solo el movimiento de las manos, sino también la postura y los

movimientos del cuerpo, ofreciendo una comprensión más completa del discurso. Es especialmente útil en entornos de investigación y aplicaciones que requieren una alta fidelidad en la captura de datos.

Retos en el reconocimiento de señas dinámicas en LSM

Las señas estáticas, que se caracterizan por la ausencia de movimiento significativo, han sido un enfoque primario en los estudios iniciales citados anteriormente. Estas señas suelen referenciar el alfabeto dactilar y los números. Mientras que las señas estáticas consisten en posturas de manos y gestos que no cambian con el tiempo, las señas dinámicas incluyen movimientos y transiciones que añaden una dimensión temporal y de fluidez al gesto, aumentando así la complejidad del reconocimiento. En el caso de las señas estáticas, los algoritmos pueden centrarse en características espaciales fijas, como la forma y orientación de la mano, lo que simplifica el proceso de identificación. En este sentido, trabajos como el de Martínez-Gutiérrez et al. (2019), Varela-Santos et al. (2021), y de Morfín-Chávez et al. (2023) se centraron en el reconocimiento de este tipo de señas, utilizando tecnologías como cámaras y software especializados para capturar y procesar las señas del alfabeto dactilar de la LSM.

Las señas dinámicas requieren un análisis detallado del movimiento de manos y otros elementos corporales, y la secuencialidad a lo largo del tiempo, lo que implica desafíos adicionales como la captura y el procesamiento de secuencias de movimiento, y la interpretación contextual de las señas. A su vez, el reconocimiento de señas puede subdividirse en reconocimiento a nivel palabra o a nivel oracional. Estudios como el de Sosa-Jiménez et al. (2022), que emplearon tecnologías como Microsoft Kinect, se han enfocado en reconocer palabras individuales en LSM, particularmente en contextos específicos como los servicios de salud. El trabajo de García-Bautista et al. (2017) utilizó Kinect y algoritmos de Distorsión Temporal Dinámica para el reconocimiento de LSM a nivel de palabra, lo que demuestra la factibilidad de capturar y procesar señas dinámicas más complejas.

El reconocimiento de LSM a nivel de oración representa un desafío mayor debido a la necesidad de comprender secuencias y contextos de gestos, lo que requiere el uso de técnicas de Procesamiento de Lenguaje

Natural (PLN) avanzadas. En este ámbito, el estudio de Mejía-Peréz et al. (2022) se destaca por utilizar redes neuronales recurrentes para procesar secuencias de señas de la LSM, abordando así el reconocimiento a niveles más avanzados.

Técnicas de IA para el reconocimiento de la LSM

El reconocimiento de la Lengua de Señas Mexicana (LSM) ha integrado diversas técnicas de inteligencia artificial (IA) para mejorar la precisión y eficiencia en la identificación de señas. A continuación se señalan cada una de las que se han aplicado en el caso de la LSM, estas abarcan las tecnologías de Redes Neuronales Artificiales (ANN), los modelos Ocultos de Markov (HMM), las Redes Neuronales Recurrentes (RNN) así como aprendizaje automático.

En el estudio de Solís et al. (2014) y de Solís et al. (2016), se implementaron Redes Neuronales Artificiales. Estas ANN se utilizaron para clasificar patrones de señas en LSM, aprovechando la capacidad de las redes para aprender y adaptarse a patrones complejos. La utilización de momentos normalizados, como características de entrada, ayudó a describir las señas de manera efectiva, permitiendo a las ANN realizar clasificaciones precisas.

Sosa-Jiménez et al. (2017) y García-Bautista et al. (2017) aplicaron Modelos Ocultos de Markov. Los HMM se usaron para modelar la secuencialidad y temporalidad de las señas en LSM, lo cual es crucial para entender gestos y movimientos que varían en el tiempo. Esta técnica es especialmente útil para capturar la dinámica y el flujo de las señas en un contexto de tiempo real.

Mejía-Pérez et al. (2023) implementaron Redes Neuronales Recurrentes con múltiples arquitecturas como LSTM (memoria a corto y largo plazo) y GRU (unidades recurrentes cerradas), desarrollando un enfoque robusto ante el ruido. Las RNN son efectivas para procesar datos secuenciales y temporales, lo cual es esencial para el reconocimiento de señas que incluyen una serie de movimientos y gestos a lo largo del tiempo. La habilidad de las RNN para manejar secuencias de datos de longitud variable las hace ideales para el análisis de LSM, que puede variar considerablemente en complejidad y duración.

Morfín-Chávez et al. (2023) aplicaron técnicas de aprendizaje automático para el reconocimiento de señas estáticas de la LSM. Se utilizaron cuatro algoritmos de clasificación diferentes: k-Nearest Neighbors (k-NN), Random Forest (RF), Naive Bayes (NB) y Support Vector Machines (SVM). Estos algoritmos fueron entrenados con los keypoints de las manos extraídos utilizando la biblioteca MediaPipe. Este estudio alcanzó resultados equiparables al uso de técnicas basadas en aprendizaje profundo.

Retos futuros hacia el reconocimiento discursivo en LSM

A lo largo de las distintas secciones de este capítulo, se ha examinado la manera en que la tecnología aplicada al reconocimiento de la Lengua de Señas Mexicana (LSM) enfrenta diversos desafíos durante sus etapas de desarrollo. En este sentido, la elección del dispositivo de sensado dependerá de las necesidades específicas del estudio o aplicación, ya que cada tecnología ofrece ventajas únicas en términos del tipo de datos capturados, el nivel de detalle y su aplicabilidad práctica. En este sentido, los desafíos abarcan desde la selección de métodos de captura hasta la gestión de la alta variabilidad entre usuarios y la similitud entre signos, lo cual puede generar ambigüedades relevantes en el discurso.

Esta variabilidad no solo surge de las diferencias individuales del habla en la ejecución de señas, sino también de las sutilezas en los gestos que pueden cambiar su significado de acuerdo con cada contexto discursivo. Esta complejidad representa un reto considerable para los sistemas de reconocimiento automático, como lo destacan los estudios de García-Bautista et al. (2017) y Mejía-Pérez et al. (2022), que buscan desarrollar algoritmos capaces de discernir estas finas diferencias. Por su parte, el estudio de Ordaz-Hernández et al. (2023) aborda este desafío mediante la creación de un corpus extenso de la LSM y el estudio de Morfín-Chávez et al. (2023) busca recopilar datos con distintos grados de variabilidad, desde señas en contextos ideales hasta señas con poca y gran variabilidad en los distintos grados de libertad de la mano.

Respecto a las condiciones de captura, Solís et al. (2016) y Solís et al. (2014) se enfocaron en el reconocimiento de LSM bajo diferentes condiciones de iluminación. En este tenor, fue crucial integrar tecnologías de

video capaces de identificar y rastrear en 3D la configuración manual así como los rasgos no manuales, mediante el uso de hardware como cámaras de profundidad o estéreo 3D y a través de software que estimara la profundidad mediante las librerías OpenPose o MediaPipe. Estos datos de alto nivel pueden ser aprovechados por técnicas que tomen series discursivas temporales como input de entrada, por ejemplo, aplicables para los HMM (Sosa-Jiménez et al. 2022) y las RNN (Mejía-Pérez et al. 2022).

La identificación de señas en un contexto dinámico presenta desafíos adicionales debido a la naturaleza fluida y compleja de la comunicación en LSM. A diferencia de otros idiomas donde las palabras pueden ser identificadas en un contexto estático y lineal, las lenguas de modalidad visogestual incluyen dos articuladores manuales, los cuales pueden ofrecer información independiente cada uno de estos. Además, las señas de la LSM pueden variar significativamente de acuerdo con la velocidad del movimiento, la posición de las manos y el movimiento del cuerpo puede codificar aspectos discursivos como fuente de información. Asimismo, los rasgos nomanuales codifican rasgos de aspectualidad y modalidad. Esto hace que el proceso de reconocimiento sea más complejo y requiera algoritmos y modelos avanzados capaces de capturar estas variaciones dinámicas de manera precisa.

En este sentido, la comunicación en una lengua de modalidad visogestual es intrínsecamente dinámica y depende del flujo continuo e interacción entre gestualidad y expresión lingüística. De igual manera, para que un sistema de reconocimiento sea efectivo y funcional en un contexto cotidiano, debe ser capaz de procesar e interpretar señas a medida que ocurren, sin retrasos significativos. Esto es esencial no solo para capturar el contenido léxico de las señas, sino también para preservar los elementos prosódicos y aquellos fuera del dominio manual como las expresiones faciales y el movimiento del cuerpo los cuales son fundamentales para el significado y la estructura gramatical en las lenguas de señas.

En el caso de la LSM, la falta de bancos de datos glosados específicos representa un obstáculo significativo para el desarrollo de sistemas de reconocimiento automático. Estos bancos de datos son fundamentales para entrenar y validar modelos de reconocimiento de señas, ya que proporcionan ítems etiquetados que permiten a los algoritmos aprender y mejorar

su desempeño. Sin suficientes datos glosados disponibles, el desarrollo de sistemas de reconocimiento de LSM se ve limitado en términos de precisión y capacidad de generalización.

Otro desafío en la identificación de señas en LSM es la codificación de configuraciones manuales específicas que no están incluidas en el alfabeto manual estándar. Estas configuraciones manuales, pueden incluir clasificadores de forma y de tamaño que codifican detalles sutiles en la forma de los objetos a los que hacen referencia. Sin embargo una representación estándar para estas configuraciones manuales puede dificultar la tarea de identificación de señas, ya que los algoritmos deben ser capaces de reconocer y distinguir una amplia variedad de formas y movimientos de manos.

La eficiencia en tiempo real es otro aspecto vital para garantizar una interacción fluida y natural entre los usuarios de una lengua de señas y un sistema interactivo o en situaciones comunicativas entre un usuario nativo de la LSM y personas que no lo son, por ejemplo, puede ser aplicable en situaciones donde la comunicación oportuna es crítica, como en la educación, el ámbito laboral, y en servicios de emergencia y atención médica. La capacidad de convertir las señas en texto o habla en tiempo real permite una comunicación bidireccional efectiva y accesible. Además, un sistema que funcione en tiempo real tiene aplicaciones potenciales en la educación de la lengua de señas con usuarios que desean aprender una lengua visogestual como L2, proporcionando retroalimentación instantánea a los estudiantes. Esto no solo aumenta la accesibilidad para las personas sordas o con problemas de audición, sino que también facilita la inclusión y la igualdad en diversos ámbitos de la vida social y profesional. Estudios como los de García-Bautista et al. (2017), Sosa-Jimenez et al. (2022), y Mejía-Pérez et al. (2022) han hecho importantes avances en esta dirección e ilustran la diversidad y complejidad de los desafíos enfrentados en el reconocimiento automático de LSM, así como los esfuerzos continuos para superar estas barreras mediante el desarrollo y mejora de tecnologías de inteligencia artificial en LSM.

Las futuras investigaciones en el reconocimiento automático de la LSM pueden abordarse desde varias perspectivas. A partir de los estudios recientes, se han identificado tres áreas clave para el desarrollo futuro:

La necesidad de mejorar la etapa de captura de datos en el reconocimiento automático de la LSM es evidente en diversos estudios recientes.

En primer lugar, la variabilidad en la ejecución de señas presenta un desafío significativo que requiere una mayor diversidad y tamaño en la población de muestra para una representación más completa y precisa (Galicia et al., 2015; Sosa-Jiménez et al., 2022). Además, el uso de diferentes tecnologías de sensado ha demostrado ser fundamental para capturar con precisión la complejidad de las señas, como se ve en los trabajos de García-Bautista et al. (2017) y Mejía-Pérez et al. (2022). Estas investigaciones resaltan la importancia de incorporar una variedad de métodos de sensado y un espectro amplio de usuarios para mejorar la precisión y eficacia en el reconocimiento, abordando tanto la diversidad de configuraciones manuales como las diferencias en la realización de las señas así como las diferencias del habla entre los usuarios (Rios-Figueroa et al., 2022; Morfín-Chávez et al., 2023). Por otro lado, se recomienda que los modelos de reconocimiento debieran ser independientes de la tecnología existente, de modo que puedan adaptarse a nuevos y más avanzados sensores.

El desarrollo de sistemas de reconocimiento de la LSM adaptados a contextos específicos, como el educativo, médico o laboral, representa un área de investigación emergente y crucial para la inclusión y accesibilidad. El estudio de Sosa-Jimenez et al. (2022) es un ejemplo de cómo los sistemas de reconocimiento de LSM pueden ser adaptados para contextos de atención médica. Por otro lado, el trabajo de Gortarez-Pelayo et al. (2023) está diseñado para que la población que no conoce la LSM se familiarice con el alfabeto manual fomentando así la cultura de la lengua de señas. Trabajos futuros podrían enfocarse en adaptar y personalizar los sistemas de reconocimiento para otros entornos específicos, mejorando así su utilidad y aplicabilidad.

La innovación en nuevos métodos y algoritmos de IA para el reconocimiento de la LSM es una necesidad imperante, tal como se evidencia en investigaciones recientes. Los estudios actuales, como los realizados por García-Bautista et al. (2017) y Mejía-Pérez et al. (2022), han demostrado avances significativos utilizando métodos de aprendizaje profundo potentes. Sin embargo, para abordar de manera efectiva la complejidad y diversidad de la LSM, es esencial desarrollar métodos más sofisticados y algoritmos avanzados. Esto incluye la exploración de nuevas técnicas de IA no solo para mejorar la precisión, sino también la velocidad en la interpretación de las señas, especialmente en la detección conjunta de gestos

manuales sutiles y expresiones faciales. Asimismo, trabajos como el de Rios-Figueroa et al. (2022) y Morfín-Chávez et al. (2023) indican la dirección hacia la que deben orientarse las investigaciones futuras, enfatizando la importancia de la innovación en el procesamiento de los datos, desde el reconocimiento de señas estáticas hasta la traducción de frases que incluya la estructura gramatical del español, así como aprovechar el análisis contextual para asistir en el proceso de reconocimiento.

Conclusiones

La investigación y desarrollo en el reconocimiento automático de lenguas de señas son clave para una sociedad más inclusiva y accesible, particularmente para la comunidad sorda. Avances como el perfeccionamiento de redes neuronales y el procesamiento de imágenes están rompiendo barreras lingüísticas y culturales, permitiendo una comunicación fluida y equitativa entre personas sordas y oyentes. En este sentido, el continuo desarrollo e investigación en materia de reconocimiento automático de lenguas de modalidad visogestual desempeña un papel fundamental en la construcción de conocimiento acerca de la codificación lingüística de una lengua de modalidad visogestual y es esencial para conocer más profundamente la gramática de la LSM, conocimiento especialmente útil para el desarrollo de aplicaciones diseñadas para la comunidad sorda.

Actualmente, los avances significativos, como el refinamiento de las redes neuronales y el procesamiento de imágenes, están derribando barreras lingüísticas, facilitando una comunicación fluida y equitativa entre personas sordas y oyentes, pero también tendrán aplicaciones en materia educativa tanto para la comunidad sorda como para oyentes. En este sentido, las mejoras tecnológicas en la actualidad no solo aumentarán la precisión y la adaptabilidad de las herramientas de comunicación, sino que también crearán en pocos años nuevas oportunidades en ámbitos educativos, laborales y sociales para las personas sordas. El progreso en este campo demuestra un compromiso con la igualdad, impulsando innovaciones que benefician a una amplia gama de aplicaciones y mejoras la calidad de vida e interacción social para la comunidad silente.

Es importante destacar que las investigaciones en materia de recono-

cimiento de patrones lingüísticos de lenguas visogestuales no solo representan avances tecnológicos, sino también un paso crucial hacia un futuro más empático y equitativo. Un futuro donde todos los individuos puedan comunicarse y contribuir sin restricciones, fortaleciendo así los valores fundamentales de igualdad y respeto en nuestra sociedad. Este trabajo no solo abre nuevas posibilidades de comunicación, sino que también promueve un cambio cultural significativo ante el desarrollo cada vez más complejo en materia tecnológica y del procesamiento de información humana.

Como hemos mencionado, estas mejoras tecnológicas no solo aumentan la precisión y adaptabilidad de las herramientas de comunicación, sino que también abren oportunidades en ámbitos educativos, laborales y sociales para las personas sordas. En este sentido, el progreso en este campo refleja un compromiso con la diversidad y la igualdad, impulsando innovaciones que benefician un amplio rango de aplicaciones y potencialmente pueden mejorar la calidad de vida de sus usuarios. Este tipo de investigaciones, más allá de ser un avance tecnológico, es un paso crucial hacia un futuro más empático y equitativo, donde todos los individuos puedan comunicarse y contribuir sin restricciones, fortaleciendo los valores de igualdad y respeto en nuestra sociedad.

Referencias

Albanie, S. et al. (2020). BSL-1K: Scaling Up Co-articulated Sign Language Recognition Using Mouthing Cues. En Vedaldi, A., Bischof, H., Brox, T., Frahm, J. M. (Eds.), *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science* (Vol. 12356). Springer, Cham. https://doi.org/10.1007/978-3-030-58621-8_3

Aly, S. y Aly, W. (2020). *DeepArSLR: A novel signer-independent deep learning framework for isolated Arabic sign language gestures recognition. IEEE access: practical innovations, open solutions, 8, 83199-83212*. https://doi.org/10.1109/access.2020.2990699

Agarwal, P., Liao, J., Hooper, S., & Sperling, R. (2021). *Applying deep learning to a sign language progress monitoring system.* Distance Learning, 18(3), 45-55. [https://eric.ed.gov/?q=artifical+intelligence+AND+sign+language+&ff1=subArtificial+Intelligence&ff2=dtySince_2020&id=EJ1327302]

Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Verhoef, T., Vogler, C. y Ringel Morris, M. (2019). *Sign language recognition, generation, and translation: An interdisciplinary perspective*. The 21st International ACM SIGACCESS Conference on Computers and Accessibility.

Camgoz, N.C.; Hadfield, S.; Koller, O.; Ney, H.; Bowden, R. (2018). Neural sign language translation. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City*, UT, USA.

Camgoz, N. C., Hadfield, S., Koller, O. y Bowden, R. (2017). *SubUNets: End-to-end Hand Shape and Continuous Sign Language Recognition*. *ICCV.*

Elatawy, S. M., Hawa, D. M., Ewees, A. A. y Saad, A. M. (2020). *Recognition system for alphabet Arabic sign language using neutrosophic and fuzzy c-means*. Education and Information Technologies, 25(6), 5601–5616. https://doi.org/10.1007/s10639-020-10184-6

Forster, J.; Schmidt, C.; Koller, O.; Bellgardt, M.; Ney, H. (2014). Extensions of the Sign Language Recognition and Translation Corpus RWTH-PHOENIX-Weather. En *Proceedings of the LREC*, (pp. 1911–1916). Reykjavik, Iceland.

Fridman-Mintz, B. (2005). *LSM tense & aspect inflections. A Dissertation Submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University.* https://www.academia.edu/59937562/LSM_tense_and_aspect_inflections

Galicia, R., Carranza, O., Jiménez, E. D, Rivera, G. E. (2015). *Mexican sign language recognition using movement sensor* (pp. 573-578).

Garcia-Bautista, G., Trujillo-Romero, F., y Caballero-Morales, S. (2017). *Mexican sign language recognition using kinect and data time warping algorithm.* 2017 International Conference On Electronics, Communications And Computers (CONIELECOMP). https://doi.org/10.1109/conielecomp.2017.7891832

Gellai, D. B. (2023). Enterprising academics: Heterarchical policy networks for artificial intelligence in British higher education. *ECNU Review of Education,* 6(4), 568–596. [http://files.eric.ed.gov/fulltext/EJ1401707.pdf]

Gortarez-Pelayo, J. J., Morfín-Chávez, R. F., & Lopez-Nava, I. H. (2023, noviembre). DAKTILOS: An Interactive Platform for Teaching Mexican Sign Language (LSM). En *International Conference on Ubiquitous Computing and Ambient Intelligence* (pp. 264-269). Cham: Springer Nature Switzerland.

Herrera Crespo, J. (2021). *Diseño y validación de un software de reconocimiento de gestos sirviéndose del acelerómetro de un reloj inteligente.* Universitat Politècnica de València. http://hdl.handle.net/10251/172272

INALSA (s/f). *Proyecto de Ley de Reconocimiento de la Lengua de Señas Argentina - LSA, y de creación del Instituto Nacional de Lenqua de Señas Argentina.* [https://inalsa.cas.org.ar/nuestra-lsa/la-ls-en-el-mundo/#:~:text=Ethonogue%2C%20una%20enciclopedia%20que%20cataloga,señas%20dispersas%20alrededor%20del%20mundo.]

Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil, A., & Corchado Rodríguez, J. (2022). *Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics*, 11, 1780. https://doi.org/10.3390/electronics11111780

Koller, O., Zargaran, S. y Ney, H. (2015). *Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers.* Computer Vision and Image Understanding, 141, 108-125.

Lang, H., Manganelli, J., & Panchanathan, S. (2020). *Virtual reality and augmented reality in social learning spaces for people with visual impairments*. IEEE Transactions on Learning Technologies, 13(2), 422-433.

Liang, X., Angelopoulou, A., Kapetanios, E., Woll, B., Al Batat, R., & Woolfe, T. (2020). A multi-modal machine learning approach and toolkit to automate recognition of early stages of dementia among british sign language users. En *Computer Vision–ECCV 2020 Workshops: Glasgow,* UK, August 23–28, 2020, Proceedings, Part II 16 (pp. 278-293). Springer International Publishing.

Martínez-Gutiérrez, M. E., Rojano-Cáceres, J. R., Benítez-Guerrero, E., y Sánchez-Barrera, H. E. (2019). *Data acquisition software for sign language recognition*. Res. Comput. Sci., 148(3), 205–211

Mejía-Peréz, K., Córdova-Esparza, D.-M., Terven, J., Herrera-Navarro, A.-M., García-Ramírez, T. y Ramírez-Pedraza, A. (2022). *Automatic recognition of Mexican Sign Language using a depth camera and recurrent neural networks*. Applied Sciences (Basel, Switzerland), 12(11), 5523. https://doi.org/10.3390/app12115523

Moher, David, Shamseer, Larissa, Clarke, Mike, Ghersi, Davina, Liberati, Alessandro, Petticrew, Mark, Shekelle, Paul, & Stewart, Lesley A. (2016). Ítems de referencia para publicar Protocolos de Revisiones Sistemáticas y Metaanálisis: declaración PRISMA-P 2015. *Revista Española de Nutrición Humana y Dietética*, 20(2), 148-160. https://dx.doi.org/10.14306/renhyd.20.2.223

Morfín-Chávez, R. F., Gortarez-Pelayo, J. J., & Lopez-Nava, I. H. (2023). Fingerspelling Recognition in Mexican Sign Language (LSM) Using Machine Learning. En *Mexican International Conference on Artificial Intelligence* (pp. 110-120). Cham: Springer Nature Switzerland.

Murakami, K., & Taguchi, H. (1991). *Gesture recognition using recurrent neural networks. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Ordaz-Hernández, K., Castillo-Gaytán, D., Rodríguez-Recio, A. S., Boone-Obregón, R. D., Hernández-García, L. Á., & Hilario-Acuapan, G. (2023). *Hacia la construcción de una base de datos abierta de la LSM*. Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI, 11. (Especial 2), 134-141.

Ompi. (2019). *BrightSign: el guante inteligente que da voz a quienes no pueden hablar*. (s. f.). Wipo.int. Recuperado el 5 de febrero de 2024, de https://www.wipo.int/wipo_magazine/es/2019/05/article_0005.html

Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2021). *Artificial intelligence technologies for sign language*. Sensors (Basel, Switzerland), 21(17), 5843. https://doi.org/10.3390/s21175843

Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2015). *Sign language recognition using convolutional neural networks*. Computer Vision and Pattern Recognition Workshops.

Quandt, L. (2020). *Teaching ASL Signs using Signing Avatars and Immersive Learning in Virtual Reality*. The 22nd International ACM SIGACCESS Conference on Computers and Accessibility. doi:10.1145/3373625.3418042

Rios-Figueroa, H. V., Sánchez-García, A. J., Sosa-Jiménez, C. O., & Solís-González-Cosío, A. L. (2022). *Use of spherical and Cartesian features for learning and recognition of the static Mexican Sign Language alphabet*. Mathematics, 10(16), 2904. https://doi.org/10.3390/math10162904

Schioppo Zachary Meyer Deigo Fabiano Shaun Canavan. (2019). Learning Sign Language in a Virtual Environment. En *Proceedings of ACM CHI conference*, Glasgow, UK, May 2019 (CHI '19), 6 pages. https://doi.org/10.475/123_4

Solís, F. J., Toxqui, C., & Martinez, D. (2015). *Mexican Sign Language Recognition Using Jacobi-Fourier Moments*.

Solis, F. J., Martinez, D., & Espinoza, O. (2016). *Automatic Mexican Sign Language Recognition Using Normalized Moments and Artificial Neural Networks*.

Solís-V., J.-F., Toxqui-Quitl, C., Martínez-Martínez, D. y Margarita H.-G. (2014). *Mexican sign language recognition using normalized moments and artificial neural networks*.

Sosa-Jimenez, C., Rios-Figueroa, H., Rechy-Ramirez, E., Marin-Hernandez, A., y Gonzalez-Cosio, A. (2017). *Real-time Mexican Sign Language recognition. IEEE International Autumn Meeting On Power, Electronics And Computing (ROPEC)*. https://doi.org/10.1109/ropec.2017.8261606

Sosa-Jimenez, C. O., Rios-Figueroa, H., & Ana Luisa Solís-González-Cosío. (2022). *A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician*.

Sturman, D. J., & Zeltzer, D. (1994). A Survey of Glove-Based Input. *IEEE Computer Graphics and Applications*, 14(1), 30-39.

Sreemathy, R., Turuk, M., Kulkarni, I. y Khurana, S. (2023). *Sign language recognition using artificial intelligence*. Education and Information Technologies, 28(5), 5259–5278. (https://doi.org/10.1007/s10639-022-11391-z]

Sugandhi, P. Kumar y Kaur, S. (2021). Indian Sign Language Generation System. En *Computer*, 54(3), 37-46. doi: 10.1109/MC.2020.2992237.

Varela-Santos, H., Morales-Jiménez, A., Córdova-Esparza, D. M., Terven, J., Mirelez-Delgado, F. D. y Orenday-Delgado, A. (2021). *Assistive device for the translation from mexican sign language to verbal language*. Computación y Sistemas, 25(3), 451-464.

Capítulo **12**

Inteligencia Artificial Generativa (IAG). El nuevo paradigma en el diseño y producción de contenido

Juan José García Llamas

https://doi.org/10.61728/AE24001120

Introducción

La imprenta e Internet: anomalías de paradigmas previos

De acuerdo con Thomas Kuhn, un paradigma es un conjunto de "realizaciones científicas universalmente reconocidas que, durante cierto tiempo, proporcionan modelos de problemas y soluciones a una comunidad…" (1962, p. x). Además, este autor presentó la noción de ciencia normal vs. ciencia extraordinaria. La primera, ciencia normal, se desarrolla dentro de un paradigma previamente establecido, mientras que en los períodos extraordinarios entran en crisis los paradigmas existentes. A pesar de que estas ideas fueron expuestas por Kuhn en 1962, en su libro La estructura de las revoluciones científicas, resultan de suma importancia para entender actualmente el impacto que ha tenido y tendrá la inteligencia artificial (IA) en nuestras vidas, especialmente en el área laboral.

A las consecuencias de ese impacto, podemos llamarle un cambio de paradigma si lo entendemos como una transformación científica y tecnológica, que nos ofrece la opción de resolver problemas insolubles por el anterior u optimiza procesos de manera innovadora. Kuhn (1962) demostró que la ciencia no solo avanza por acumulación de conocimientos, sino también debido a cambios revolucionarios provocados por anomalías en los paradigmas dominantes. Entonces, ¿cuál es el paradigma dominante en la actualidad (primer cuarto del siglo XXI)? Y ¿cuáles son las anomalías que nos están llevando a un nuevo cambio revolucionario?

El paradigma actual, relacionado con la manera en la que producimos y compartimos información, comienza a gestarse con la invención de la imprenta de tipos móviles por Johannes Gutenberg en el siglo XV. Esta tecnología generó, como consecuencia, una anomalía en la difusión de la información y el conocimiento de esa época. Previamente, durante la Edad Media, la difusión y la filtración del conocimiento estuvieron limitadas a muy pocas personas. Este se debía, en gran parte, a que los monjes copistas eran los únicos que tenían acceso al saber y la posibilidad cognitiva, la habilidad física y las herramientas para transcribir manualmente, uno por uno, los textos antiguos.

Según McLuhan (1962), la imprenta fue la causa del fin de la era manuscrita, dando paso a la etapa tipográfica. El cambio gradual de difusión de conocimiento, revolucionario, facilitó la producción masiva de libros y una mayor circulación y democratización del saber. Además, este nuevo medio tecnológico trajo radicales transformaciones culturales, permitiendo el cuestionamiento del conocimiento "oficial" establecido en la Edad Media. De acuerdo con Eisenstein (1980), la estandarización en la reproducción de textos a consecuencia de la imprenta también estandarizó el pensamiento. Se propicia el surgimiento de la ciencia moderna y un espíritu crítico que condujo a la humanidad a nuevas etapas de desarrollo: el Renacimiento y la Reforma Protestante.

Por lo tanto, la imprenta inició un cambio de paradigma al masificar la producción y acceso al conocimiento, sacándolo del control exclusivo de instituciones medievales y monasterios. Esta revolución en la comunicación del saber contribuyó al desarrollo intelectual que sentó las bases de la modernidad. Sin embargo, el paradigma actual no se consolidó hasta que sucedió otra anomalía de alto impacto, la llegada del Internet. Inicialmente, en 1969, se llamó ARPANET y fue desarrollada por el Departamento de Defensa de Estados Unidos en el marco de la Guerra Fría. Su objetivo era eliminar la dependencia de conexión exclusiva a un Ordenador Central, y de esta manera disminuir la vulnerabilidad de las comunicaciones del ejército norteamericano.

Posteriormente, en la década de 1990, esta tecnología se "convirtió" en la World Wide Web (WWW) y comenzó a extenderse por todo el mundo hasta alcanzar popularidad en ámbitos fuera de la milicia. Por ejemplo, los físicos del CERN la ampliaron a otros sitios de investigación hasta alcanzar al US National Center for Supercomputer Applications (NCSA), en donde Marc Andreessen dirigió, en 1993, el desarrollo de un navegador el cual fue nombrado Mosaic. Esta manera de acceder a la WWW mediante software de "navegación de la información", facilitó que personas sin conocimientos en programación pudieran obtener sus beneficios con una curva de aprendizaje relativamente sencilla.

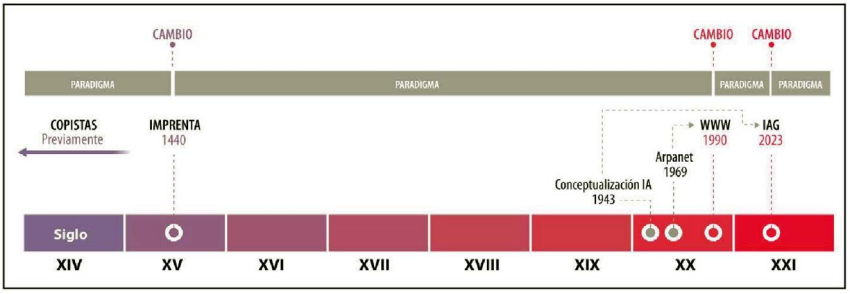
Así, el acceso masivo a la información consolidó un cambio de paradigma sin precedentes en la comunicación, pero ahora con alcances a nivel mundial. De acuerdo con Castells (2000), internet es la base tecnológica de

la era de la información, lo que permite la comunicación multidireccional y el flujo amplio de los datos. Este nuevo medio cambió la perspectiva de los medios unidireccionales y monopolizados como la radio, la televisión o los periódicos. Permitió que prácticamente cualquier persona consuma, genere y difunda contenidos, descentralizando el flujo informativo (Tapscott y Williams, 2007).

Asimismo, el acceso masivo a la internet transformó los procesos sociales, culturales y económicos, entre otros, facilitando nuevas formas de comunicación, comercio, aprendizaje, activismo social y conformación de comunidades virtuales (Castells, 2000). Incluso, se comienza a hablar de la existencia de "nativos digitales" que incorporan internet como parte integral de sus vidas y desde temprana edad. En suma, la llegada de la internet terminó de generar un nuevo paradigma comunicativo e informacional, que puede eliminar (o al menos disminuir) las barreras de espacio y tiempo para la difusión de información a escala global.

Pero, este paradigma que nos ha ofrecido por varios años la oportunidad de recibir conocimiento mediante libros y, recientemente, a consumirlo y producirlo sin importar el lugar del mundo en el que estemos, mientras tengamos conexión al internet, está sufriendo los efectos de lo que Kuhn (1962) llamó ciencia extraordinaria, como vimos al inicio. Estamos presenciando la llegada de una nueva tecnología disruptiva, lo que supone el inicio de un nuevo cambio de paradigma, el de la IA.

EVENTOS HISTÓRICOS QUE ORIGINARON CAMBIOS DE PARADIGMA EN LA COMUNICACIÓN



Inteligencia artificial. ¿El nuevo paradigma?

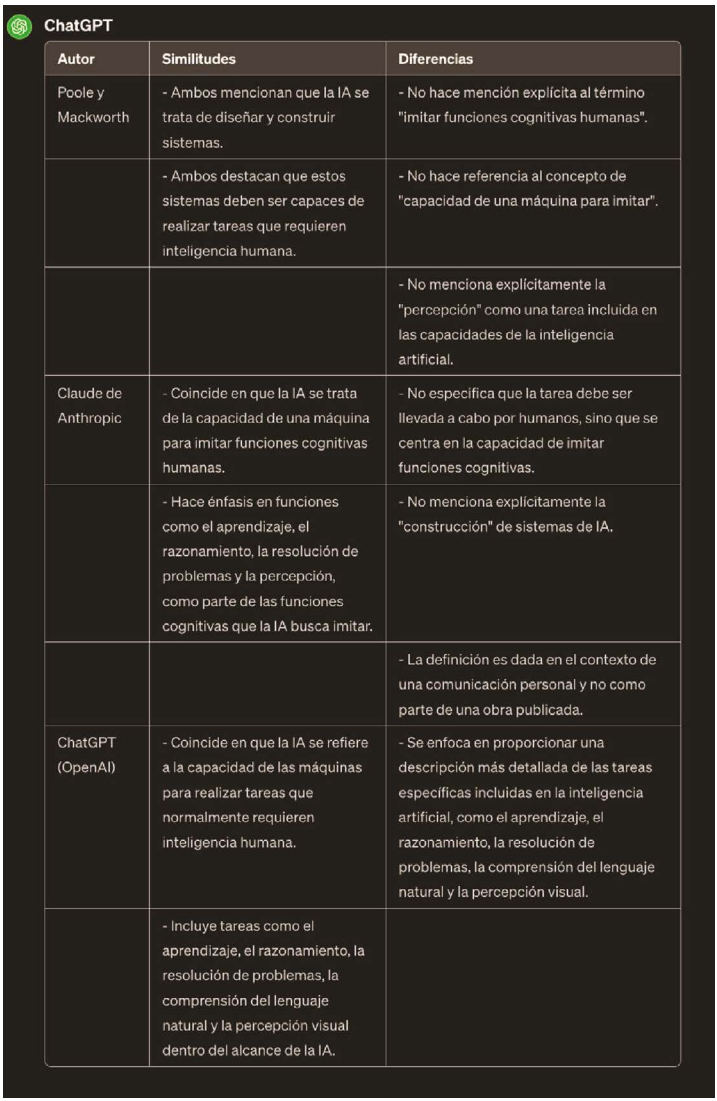
Recordemos que una revolución en el ámbito científico y tecnológico, con el propósito de constituir un cambio de paradigma, debe proporcionar respuestas a desafíos que eran irresolubles mediante los métodos vigentes hasta ese momento (Kuhn, 1962). Además, una de las características del nuevo paradigma, deberá ser la presentación de procesos innovadores para llegar a esas soluciones. Entonces, ¿qué problemas nos ayuda a resolver la inteligencia artificial (IA)? ¿Qué innovaciones nos ofrece esta tecnología de vanguardia? Y ¿esta información nos demostrará que efectivamente estamos ante un cambio de paradigma?

Antes de continuar, es conveniente definir el concepto clave a tratar: la IA. Para ello se hará un recorrido desde lo general hasta llegar a lo particular o específico, en este caso la inteligencia artificial Generativa (IAG) la cual está estrechamente vinculada con la generación y difusión de información. Después, se analizarán sus características para poder responder las preguntas planteadas previamente.

De acuerdo con Poole y Mackworth (2010), la IA es la rama de la informática que diseña y construye sistemas capaces de realizar tareas que requieren de inteligencia para ser llevadas a cabo por humanos. Por su parte, Claude, la IA programada por la empresa Anthropic, arroja la siguiente definición: "la inteligencia artificial es la capacidad de una máquina para imitar las funciones cognitivas humanas como el aprendizaje, el razonamiento, la resolución de problemas, la percepción, etc." (Anthropic, comunicación personal, 1 de diciembre de 2023). Finalmente, la IA ChatGPT, de OpenAI. Especifica que "la inteligencia artificial se refiere a la capacidad de las máquinas o programas informáticos para realizar tareas que normalmente requieren de la inteligencia humana. Estas tareas incluyen el aprendizaje, el razonamiento, la resolución de problemas, la comprensión del lenguaje natural y la percepción visual" (OpenAI, comunicación personal, 29 de diciembre de 2023).

A continuación, una tabla con el análisis de las definiciones anteriores, cocreado con ChatGPT:

Tabla 1. Prompt: "Realiza un análisis de estas tres definiciones y encuentra las ideas similares y las diferencias. Utiliza una tabla para responder. En ella deberás poner tres columnas: una para el autor de la frase, otra para las similitudes y otra para las diferencias. Además, la tabla deberá tener tres filas, una para cada autor. Este es el texto a analizar [Se anexa el párrafo de las 3 definiciones]".



ChatGPT

Autor	Similitudes	Diferencias
Poole y Mackworth	- Ambos mencionan que la IA se trata de diseñar y construir sistemas.	- No hace mención explícita al término "imitar funciones cognitivas humanas".
	- Ambos destacan que estos sistemas deben ser capaces de realizar tareas que requieren inteligencia humana.	- No hace referencia al concepto de "capacidad de una máquina para imitar".
		- No menciona explícitamente la "percepción" como una tarea incluida en las capacidades de la inteligencia artificial.
Claude de Anthropic	- Coincide en que la IA se trata de la capacidad de una máquina para imitar funciones cognitivas humanas.	- No especifica que la tarea debe ser llevada a cabo por humanos, sino que se centra en la capacidad de imitar funciones cognitivas.
	- Hace énfasis en funciones como el aprendizaje, el razonamiento, la resolución de problemas y la percepción, como parte de las funciones cognitivas que la IA busca imitar.	- No menciona explícitamente la "construcción" de sistemas de IA.
		- La definición es dada en el contexto de una comunicación personal y no como parte de una obra publicada.
ChatGPT (OpenAI)	- Coincide en que la IA se refiere a la capacidad de las máquinas para realizar tareas que normalmente requieren inteligencia humana.	- Se enfoca en proporcionar una descripción más detallada de las tareas específicas incluidas en la inteligencia artificial, como el aprendizaje, el razonamiento, la resolución de problemas, la comprensión del lenguaje natural y la percepción visual.
	- Incluye tareas como el aprendizaje, el razonamiento, la resolución de problemas, la comprensión del lenguaje natural y la percepción visual dentro del alcance de la IA.	

Resumiendo lo anterior, se aprecia que las tres fuentes difieren en algunos matices, como el énfasis en la imitación de funciones cognitivas humanas, la inclusión de tareas específicas y la referencia a la construcción de sistemas de IA. Sin embargo, comparten la idea fundamental de que esta tecnología implica diseñar sistemas capaces de realizar tareas que normalmente requieren inteligencia humana. Y es gracias a esa emulación de capacidades mentales biológicas que la generación de contenido se optimiza en muchos aspectos. Pero antes de abordar ese tema, se presenta un recorrido sintetizado del desarrollo histórico de la IA para comprender de dónde viene y cómo llega a nuestros días.

La historia de esta tecnología se remonta a la década de 1950, cuando Turing, McCarthy, Minsky y Shannon sentaron las bases teóricas y prácticas de esta disciplina (Russell y Norvig, 2021). Algunos eventos históricos relevantes de la IA fueron: el test de Turing en 1950, la conferencia impartida por John McCarthy titulada "Dartmouth Summer Research Project on Artificial Intelligence" en 1956, el programa de juegos computacionales de Samuel en 1959, los sistemas expertos en la década de 1980 y el triunfo de una computadora sobre Kaspárov en ajedrez en 1997.

Vale la pena detenerse un poco para remarcar el impacto que tuvo el último evento del listado anterior. Ese proyecto que entrenó a una IA para jugar ajedrez a niveles profesionales fue diseñado por IBM, en el laboratorio nacional Los Alamos en Estados Unidos. Una vez que estuvo listo el ordenador Deep Blue, enfrentó y venció a Garry Kaspárov, el campeón del mundo en ese tiempo. Al darse a conocer esta noticia, muchas personas, sin necesidad de ser profesionales del ajedrez, se interesaron en ponerse a prueba contra una máquina en este y otros juegos. Así, la cultura popular se transformó a tal grado que actualmente es común que las computadoras o dispositivos móviles tengan precargada una aplicación de juegos que funcionan con algoritmos de IA o similares.

Respecto a los hechos recientes se destacan el reconocimiento de voz por parte de Google lanzado en 2008, el algoritmo de Facebook que entró en funciones en 2009 y, por supuesto, la presentación pública de Chat GPT de OpenAI en 2022. Si se analiza a detalle el desarrollo de esta tecnología, se puede deducir que una de sus características principales es la rapidez con la que evoluciona. Esto se explica principalmente mediante la

Ley de Moore: cada dos años, aproximadamente, se duplica la capacidad de cómputo y se reduce el tamaño y costo del hardware que la soporta (Kelleher, 2022). Si a esto le sumamos la demanda que ha tenido la IA más popular, ChatGPT, de 180.5 millones de usuarios en tan solo un año (explodingtopics, 2023), es posible comprender su crecimiento y anticipar, en cierta medida, la extensión que alcanzará en los próximos años.

Otras características distintivas de la IA son: capacidad de aprendizaje y adaptación, procesamiento de lenguaje natural, razonamiento automático, percepción visual y auditiva, resolución de problemas complejos y capacidad de movimiento y manipulación de objetos (Negnevitsky, 2021). Para ello, esta tecnología se desarrolla mediante algoritmos y modelos que permiten a las máquinas procesar y "entender" información, aprender de experiencias pasadas y adaptarse a las nuevas situaciones que se les pueden presentar. Para abordar distintos aspectos de la capacidad cognitiva humana, se han desarrollado diversos enfoques dentro de la inteligencia artificial, entre ellos el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo. Así, la IA tiene aplicaciones en una variedad de campos, entre ellos:

• Medicina.

Ejemplo: detección de cáncer de mama mediante el análisis de patrones en mamografías (Sánchez, 2023).

• Robótica.

Ejemplo: el Heineken B.O.T transporta cerveza utilizando un sistema de detección de obstáculos y caídas (Rivera, 2023).

• Atención al cliente.

Ejemplo: chatbots de atención 24/7 para una empresa de seguros (Estrada, 2018).

Conducción autónoma.

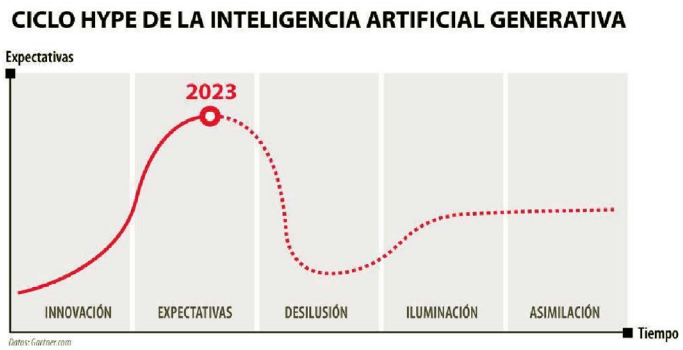
Ejemplo: mejora de la eficiencia y seguridad en la conducción de automóviles (González, 2023).

• Comunicación.

Ejemplo: desarrollo de contenido en tiempos extremadamente cortos (Portas, 2023).

Como se puede apreciar en el listado anterior, esta tecnología de vanguardia tiene un gran impacto en diversas áreas. Pero, un cambio signi-

ficativo y tangible resalta en el ámbito profesional de la comunicación y producción de contenidos. Esto es debido a la inteligencia artificial Generativa (IAG), la cual es un subcampo de vanguardia en la investigación de la IA que se especializa en la creación autónoma de contenido nuevo y original, ya sea texto, imágenes, audio, video u otro tipo de datos digitales (Bommasani et al., 2022). De acuerdo con el ciclo hype de Gartner, la IAG tiene su pico de expectación en la sociedad en 2023 (Santamaría, 2023). Entonces, podríamos decir que este año, el de su popularización, es el momento en el que se comienzan a identificar masivamente los efectos generados por su utilización.



Una de sus características es que se basa en modernas técnicas de aprendizaje automático o machine learning; en particular, un tipo de algoritmos llamados redes neuronales artificiales, las cuales están vagamente inspiradas en el funcionamiento del cerebro humano. Este tipo de redes se conceptualizó por primera vez en 1943 por Warren McCullough y Walter Pitts mediante un modelo matemático presentado en su artículo "A logical calculus of ideas Immanentin Nervous Activity". Actualmente, esas redes neuronales se entrenan con enormes conjuntos de datos para aprender los patrones y estructuras que gobiernan un determinado tipo de contenido, por ejemplo, cómo se componen gramaticalmente los textos o cómo se relacionan los pixeles en una imagen realista. Una vez que la red está entrenada, puede generar nuevo contenido de forma independiente siguiendo esos mismos patrones aprendidos con variaciones que la persona puede indicar o la misma IAG puede proponer.

Cabe destacar que el nivel de creatividad de la IAG puede ser controlado por el usuario mediante hiperparámetros, como temperature (para texto) o guidance scale (para imagen). Otra manera para dirigir y enfocar la originalidad en los resultados de esta tecnología es utilizar instrucciones escritas, redactadas con lenguaje natural. Es decir, realizar una petición como si se hiciera a otra persona, a un asistente, de manera clara y específica. A esto se le conoce como prompt y se define técnicamente como una descripción, generalmente corta, en lenguaje simple sobre lo que se quiere generar. La red neuronal procesa esta instrucción textual y la traduce a datos digitales, que pueden ser pixeles de una imagen (text to image), a una voz sintetizada (text to voice) o a otro tipo de texto (text to text). El resutlado debe coincidir con esa descripción lo más precisamente posible. De esta manera, la IAG crea contenido totalmente nuevo y original basado en su conocimiento previo y abstracto sobre el mundo.

Las capacidades actuales de los modelos de IAG más avanzados son verdaderamente impresionantes y siguen mejorando rápidamente. Por ejemplo, DALL-E 2 de OpenAI (Ramesh et al., 2022) o MidJourney pueden generar imágenes hiperrealistas de casi cualquier cosa imaginable a partir de simples descripciones en lenguaje natural. Se le puede indicar en una frase que cree ilustraciones fotorrealistas de animales, personas, objetos y paisajes completamente nuevos que no existen en la realidad. De manera similar, las versiones actuales de GPT de OpenAI (Brown et al., 2020) son capaces de producir ensayos, poesías, letras de canciones, código de programación y una gran variedad de texto perfectamente coherente, con solo darle un tema y algunas pautas iniciales dentro del prompt. Su capacidad para comprender e inferir ideas complejas a partir de poca información es comparable a la de un humano experto.

Otro modelo que demuestra este nivel de destreza generativa es Jasper, que se utiliza para ampliar el contenido de marketing de las empresas. Las aplicaciones actuales de la IAG van desde la asistencia en todo tipo de trabajos creativos y de diseño gráfico, hasta el desarrollo de campañas completas para marketing digital y el desarrollo de experiencias de realidad virtual altamente inmersivas.

Si la mayoría de los productos de comunicación que nos ofrece la IAG es posible realizarlos con tecnología previa, entonces el verdadero bene-

ficio está en el modo y la velocidad de producción. Y es en este punto exacto en donde podemos encontrar soluciones a problemas relacionados con agendas de planeación, diseño, realización, publicación, seguimiento y análisis.

La automatización de tareas ha experimentado un notorio avance gracias a esta tecnología, que ahora se utiliza para realizar diversas funciones como redactar textos simples, transcribir grabaciones, traducir contenidos, diseñar infografías y hasta crear videos animados, todas esas funciones realizadas en fracciones del tiempo empleado con otras tecnologías. Este conjunto de capacidades no solo agiliza el trabajo de comunicadores y productores de contenido, sino que también proporciona eficiencia y velocidad en la ejecución de estas actividades.

La personalización de contenidos es otro aspecto beneficiado por la IAG, especialmente a través del aprendizaje automático. Este enfoque permite adaptar los contenidos de manera precisa a los intereses y comportamientos individuales de cada usuario o segmento de audiencia. De este modo, esta tecnología contribuye a ofrecer experiencias más relevantes y personalizadas, mejorando la conexión entre los contenidos y sus receptores.

GENERACIÓN DE IMAGEN HIPER REALISTA A PARTIR DE TEXTO

En el ámbito del análisis de datos e información, se destaca por su capacidad para entender grandes volúmenes de datos sobre las preferencias de los usuarios y ofrecer propuestas para la mejora continua. Este análisis profundo ayuda a identificar temas, formatos y estilos que generan un mayor interés, proporcionando a los creadores de contenido valiosas claves y referencias para ajustar estrategias y mejorar la efectividad de sus mensajes. Otro beneficio notable es la mejora en la difusión y el posicionamiento en motores de búsqueda (SEO). La IAG contribuye a optimizar la indexación y el rastreo de contenidos por parte de los motores de búsqueda, además de segmentar y seleccionar de forma inteligente los canales de distribución más adecuados. Esto no solo aumenta la visibilidad de los contenidos, sino que también favorece la llegada a audiencias específicas de manera más efectiva.

En términos de medición de impacto, ofrece capacidades predictivas e informes al instante. A través del machine learning (subconjunto de la IA que aprende de los datos y mejora con cada interacción), puede anticipar y proporcionar información detallada sobre el rendimiento de cada contenido, incluyendo métricas como vistas, interacciones, audiencia influenciada, entre otras. Esta capacidad de evaluación en tiempo real facilita la toma de decisiones y la adaptación continua de las estrategias de contenido para maximizar su efectividad.

En resumen, la IAG presenta interesantes oportunidades para promover la inclusión de personas con discapacidad en los medios y productos de comunicación. Mediante el aprendizaje automático, los modelos de lenguaje pueden entrenarse en bases de datos especializadas para redactar, adaptar y traducir contenidos que se ajusten a las necesidades perceptivas y cognitivas de diferentes poblaciones (Mohammed, 2022). Por ejemplo, Microsoft desarrolló un sistema de IAG capaz de agregar descripciones de audio detalladas en tiempo real para personas ciegas o con visión reducida, en transmisiones de eventos en vivo, noticieros televisivos y contenidos multimedia (Whittaker, 2021). Asimismo, mediante el procesamiento de lenguaje natural, la IAG puede transcribir artículos, guiones, menús web y otro tipo de textos a formatos como braille, textos simplificados o con interpretaciones de emociones mediante pictogramas, para facilitar la lectura a personas con autismo, dislexia o limitaciones visuales (UNESCO, 2021).

Los modelos generativos entrenados adecuadamente incluso pueden redactar textos, titulares, piezas audiovisuales o traducir contenido web completo, optimizándolo para que sea fácilmente interpretable por tecnologías asistivas como lectores de pantalla para discapacidad visual o aplicaciones de comprensión para discapacidad cognitiva.

Finalmente, la IAG ha transformado significativamente el panorama de la creación, distribución y análisis de contenidos, ofreciendo herramientas poderosas para optimizar cada fase del proceso, permitiendo así llegar a más públicos de manera efectiva. Por lo tanto, diversos pensadores y expertos en IA argumentan que los recientes avances en modelos generativos representan un nuevo paradigma tecnológico (Yudkowsky, 2022), es decir, están alterando radicalmente ideas y prácticas profundamente arraigadas en nuestra civilización. Esto se debe a que por primera vez tenemos máquinas que pueden crear ideas y contenidos novedosos por sí mismas, sin necesidad de que un humano los especifique explícitamente.

En particular, la IAG está transformando la forma misma en que diseñamos, producimos y consumimos todo tipo de contenidos digitales y físicos. Actividades creativas que históricamente demandaban grandes dosis de pericia, esfuerzo y tiempo por parte de expertos, como el diseño gráfico, la composición musical, la escritura, la programación o el diseño industrial, ahora pueden ser realizadas con mucho menos trabajo manual gracias al apoyo de asistentes de IA. Por ejemplo, un diseñador novato puede crear visuales sofisticadas indicándole a DALL-E qué tipo de imagen quiere. De igual modo, un estudiante puede producir un ensayo elaborado solo proveyendo el tema y algunas ideas principales a ChatGPT. En ambos casos, la IA está asumiendo buena parte del proceso creativo. Y esto es algo que debemos considerar urgentemente para tomar las mejores decisiones. Entonces, ¿cuáles son los desafíos que surgen con este nuevo paradigma? Y, más importante aún, ¿cuáles estrategias podemos implementar para superarlos?

Retos del nuevo paradigma de la inteligencia artificial

La disrupción del statu quo, provocado por la llegada de la IA, específicamente la IAG, tiene como consecuencia una serie de cambios relevantes en una enormidad de campos, desde oficios artísticos hasta la educación superior y el periodismo. Por supuesto, esta situación nos plantea oportunidades y retos, tal como sucede cada vez que se cambia un paradigma.

Para empezar este apartado, es necesario destacar que esta tecnología no ofrece soluciones perfectas. Como se vio previamente, la IAG tiene la capacidad de seguir "aprendiendo" y por lo tanto ejecuta procesos que no están totalmente terminados. Es decir, que cada vez pueden mejorar o, incluso, empeorar con respecto al objetivo planteado en el prompt. Por otro lado, es importante mencionar que estas tecnologías están progamadas por personas que tienen sesgos, intereses y creencias, como cualquier humano. Por lo tanto, es de suma importancia verificar siempre el resultado que nos ofrece esta tecnología para contrastarlo con otras fuentes y posteriormente validarlo o solicitarle los ajustes necesarios. A esta última acción, en el contexto de la IA, se le conoce como Iteración.

Continuando con los "detalles no resueltos", la propia IAG reconoce que puede alucinar. Este término se refiere a las respuestas que no son verídicas, pero que están presentadas de tal manera que parecieran serlo. Chat GPT (OpenAI, 2023) las describe y explica de la siguiente manera:

> En el campo de la IA, el término "alucinación" se utiliza para describir situaciones en las que un modelo de IA produce resultados que no están respaldados por datos reales o relevantes para la tarea en cuestión. Estas alucinaciones pueden manifestarse de varias maneras, como generar información inventada, imágenes distorsionadas o respuestas incoherentes. Las alucinaciones pueden ocurrir por diversas razones, como la falta de datos de entrenamiento representativos, la presencia de sesgos en los datos, o la complejidad del modelo que puede llevar a interpretaciones erróneas de los patrones en los datos. En el contexto de modelos generativos de lenguaje, por ejemplo, una alucinación podría ser la generación de información ficticia que suena plausible pero carece de fundamento en la realidad.

Por lo tanto, es de suma importancia identificar las alucinaciones, sesgos y tendencias en los modelos de IAG para garantizar la confiabilidad de sus respuestas.

Al mismo tiempo, el contenido creado por IAG plantea serios interrogantes sobre la atribución de mérito (Chequeado, 2023). Incluso podría cambiar la forma en que conceptualizamos nociones cardinales de la cultura humana como la creatividad, la propiedad intelectual, el valor del trabajo, el plagio y los medios de comunicación social (Chacón, 2023). La introducción de esta tecnología en la creación de contenidos ha alterado considerablemente el enfoque de los profesionales del marketing en sus estrategias. A pesar de los beneficios evidentes de la IA, se plantean desafíos fundamentales. En particular, la autenticidad y el plagio son inquietudes críticas, dado que los algoritmos pueden generar contenido que se asemeja de cerca al producido por humanos (Chacón, 2023). Esto exige precaución para evitar violaciones de derechos de propiedad intelectual y el riesgo de perjudicar la reputación de la propia imagen o de terceros.

Otra de las inquietudes que las personas manifiestan al momento de interactuar con la IAG, sobre todo en etapas iniciales, son los aspectos éticos. Según Cárdenas (2023), este tema debe ser especialmente importante cada vez que se utilice la IA en todas sus variantes, sobre todo "en contextos como la toma de decisiones médicas o financieras, donde la transparencia y la explicabilidad son cruciales".

En este sentido, se encuentran la equidad y la no discriminación como inquietudes principales. La IAG tiene el potencial de mantener la discriminación si es entrenada utilizando datos que presenten sesgos, como vimos anteriormente, o si no se toman en consideración las disparidades culturales y de género. Por supuesto, esto podría acarrear resultados negativos para personas y comunidades que se encuentran en situación de marginación. Por lo tanto, es imperativo implementar medidas para asegurar que la generación de productos de comunicación, mediante la IA, sea de forma justa y equitativa.

Continuando con los aspectos éticos a considerar, la privacidad también es fundamental. Es importante tomar en cuenta que esta tecnología puede recopilar, analizar y explotar enormes cantidades de datos personales sensibles. Y a diferencia de otros "sistemas de recolección" que ya

llevan haciéndolo desde hace años, como las redes sociales, ahora la IA lo hace prácticamente en tiempo real. Esto abre puertas para utilizar ese potencial de cálculo con fines que pueden ser no afines a nuestros intereses o valores.

Cárdenas (2023) reconoce que el uso inapropiado de la IA podría generar repercusiones importantes tanto en el ámbito social como económico, teniendo posibles influencias en la gobernanza y la política, dado que podría ser empleada con fines de vigilancia y control social. Por si fuera poco, este autor también menciona que la seguridad y responsabilidad de los sistemas de IA autónomos se convierten en temas críticos, ya que su uso en cibercrimen o guerra cibernética implica riesgos significativos para la seguridad global. La dificultad para controlar y responsabilizar a estos sistemas en caso de errores o accidentes los posiciona como potenciales amenazas.

En vista de lo anterior, es importante abordar estas preocupaciones éticas y garantizar que la IA se utilice de manera justa y responsable en beneficio de la sociedad en su conjunto y tomando en cuenta a cada individuo. Esto implica desarrollar políticas y marcos éticos sólidos para regular su uso, así como invertir en investigación para abordar las preocupaciones éticas en torno a la IA. Por su parte, la Asociación por los Derechos Civiles, la cual trabaja desde 1995 a favor de la defensa y promoción de los derechos civiles y humanos en Argentina y América Latina, propone "optar por el uso del marco internacional de los Derechos Humanos" para evaluar los efectos de la IA y proponer y direccionar las acciones posteriores (ADC, 2019).

En resumen, el aspecto ético es crítico para cualquier uso responsable de la IA, ya sea quien la programa o quien la utiliza. Las empresas, los gobiernos y la sociedad en general deben colaborar para desarrollar marcos éticos robustos que garanticen que esta poderosa tecnología se implemente y supervise de manera segura para beneficio de la humanidad. Por esta razón, es necesario que cada individuo, desde sus posibilidades, participe en el progreso de la IA. Colaborando de manera conjunta, podemos aprovechar las ventajas de esta tecnología garantizando al mismo tiempo su uso responsable y buscando el mayor bien común.

En lo que respecta a los creadores de productos de comunicación, deben abordar de manera responsable las complejidades éticas asociadas con

la IAG en la creación de contenidos. Al hacerlo, podrán aprovechar el potencial de esta tecnología mientras mantienen la autenticidad, equidad y responsabilidad en sus prácticas profesionales, sus proyectos y sus resultados.

En definitiva, esta ola tecnológica tiene la capacidad de reconfigurar los fundamentos de la sociedad, transformando por completo diversos aspectos de la realidad humana. Es un hecho que la IA está generando el nacimiento del nuevo paradigma. La dirección que tome este cambio histórico dependerá de debates éticos sólidamente fundamentados para utilizarla como fuerza positiva para el correcto desarrollo de la humanidad.

Si la inteligencia artificial es la aplicación, el sistema operativo deberá ser la inteligencia humana y la ética el antivirus.

Referencias

ADC (2019). *Ética y protección de datos en la inteligencia artificial*. https://adc.org.ar/2019/05/09/etica-y-proteccion-de-datos-en-la-inteligencia-artificial/

Anthropic (2023). *Claude* (Versión nov 22) [Large language model]. https://claude.ai/

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... y Bosselut, A. (2022). *On the opportunities and risks of foundation models*. Stanford University.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... y Amodei, D. (2020). *Language models are few-shot learners*. Advances in neural information processing systems, 33, 1877-1901.

Cárdenas, R. (2023). *IA y ética, nuestro futuro*. https://www.linkedin.com/pulse/%C3%A9tica-en-la-inteligencia-artificial-c%C3%B3mo-asegurarnos-ricardo/?originalSubdomain=es

Castells, M. (2000). *La era de la información: economía, sociedad y cultura* (Vol. 1). Siglo XXI.

Chacón, Luis (2023). *Inteligencia artificial y la reinvención de la propiedad intelectual*. https://www.forbes.com.mx/inteligencia-artificial-y-la-reinvencion-de-la-propiedad-intelectual/

Chequeado (2023). *Desinformación e inteligencia artificial. Ideas para las aulas de escuela secundaria*. https://chequeado.com/wp-content/uploads/2023/08/Desinformacion-e-inteligencia-artificial.pdf

Eisenstein, E. L. (1980). *The printing press as an agent of change*. Cambridge University Press.

Estrada, L. (2018). *Implementar chatbot basado en inteligencia artificial para la gestión de requerimientos e incidentes en una empresa de seguros* (Tesis para optar el Título Profesional de Ingeniero Empresarial y de Sistemas). Universidad San Ignacio de Loyola, Lima, Perú.

Explodingtopics. (2023). *Discover Exploding Topics*. https://explodingtopics.com/

González, C. (2023). *Computer Hoy. Inteligencia artificial en el mundo del sector automovilístico*. https://computerhoy.com/motor/inteligencia-artificial-mundo-sector-automovil-1266250

Kelleher, A. (2022). *Ley de Moore, ahora y en el futuro*. https://www.intel.la/content/www/xl/es/newsroom/opinion/moore-law-now-and-in-the-future.html

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

McLuhan, M. (1962). *The Gutenberg galaxy: The making of typographic man*. University of Toronto Press.

Mohammed, M. G. (2022). Applications of AI to Improve Accessibility for People With Disabilities. *IEEE Signal Processing Magazine*, 39(1), 167-175.

Negnevitsky, M. (2021). *Inteligencia artificial: una guía para pensar a los humanos*. A pelican book.

OpenAI. (2023). ChatGPT (Versión nov 30) [Large language model].
https://chat.openai.com

Poole, D. L., & Mackworth, A. K. (2010). *Artificial intelligence: foundations of computational agents*. Cambridge.

Portas, G. (2023). *¿Cómo afecta la IA a nuestro universo narrativo? En sintonía con la industria de los medios de comunicación y el entretenimiento*. Stayrelevant. https://stayrelevant.globant.com/es/technology/data-ai/como-afecta-la-ia-a-nuestro-universo-narrativo-en-sintonia-con-la-industria-de-los-medios-de-comunicacion-y-el-entretenimiento/

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., ... & Sutskever, I. (2022). Zero-shot text-to-image generation. En *International Conference on Machine Learning* (pp. 8821-8831). PMLR.

Rivera, A. (2023). *Futuro Eléctrico. Robots y la inteligencia artificial.* https://futuroelectrico.com/robots-inteligencia-artifical/

Russell, S. J. y Norvig, P. (2021). *Artificial intelligence: a modern approach*. Prentice Hall.

Sánchez, A. (2023). Gaceta Médica. Una herramienta de IA podría detectar el cáncer de mama hasta cinco años antes de su aparición. *Gaceta Médica* https://gacetamedica.com/investigacion/una-herramienta-de-ia-podria-detectar-el-cancer-de-mama-hasta-cinco-anos-antes-de-su-aparicion/

Santamaría, F. (2023). *Novedades en Inteligencia Artificial según el Gartner Hype Cycle 2023*. https://fsanta.medium.com/novedades-en-inteligencia-artificial-seg%C3%BAn-el-gartner-hype-cycle-2023-35664069233c

Tapscott, D. y Williams, A. D. (2007). *Wikinomics: How mass collaboration changes everything*. Penguin.

UNESCO. (2021). *Artificial intelligence and inclusion: Exploring the opportunities and challenges of artificial intelligence for persons with disabilities*. UNESCO Digital Library.

Whittaker, Z. (2021, May 25). *This AI-powered assistant helps blind people 'see' Facebook*. TechRepublic.

Yudkowsky, E. S. (2022). *AI: Safety, capability, and long-term research*. arXiv preprint arXiv:2209.12575.

Capítulo **13**

La inteligencia artificial y sus beneficios en una agricultura sostenible

Oscar Méndez García

https://doi.org/10.61728/AE24001137

Introducción

El sector agrícola es de gran importancia para todos los países. Con el incremento de la población se espera que la demanda alimenticia aumente de manera considerable para el año 2050 (Shariff et al., 2022). Por lo mismo, la producción de este sector es un tema relevante, en virtud de todos los bienes que se producen en las actividades agrícolas, que va desde la preparación de la tierra, cuidado de semillas, cultivos de alimentos, proceso de cosecha, uso de suelo y agua, así como el impacto en el medioambiente. La gestión eficiente de estos recursos ha provocado una transformación tecnológica del sector agrícola con la finalidad de incrementar la producción, disminuir los costos y responder a la alta demanda de manera sostenible con los recursos disponibles.

Con el tiempo, el sector agrícola ha evolucionado adaptándose a los diversos cambios que se han presentado, partiendo de su forma tradicional hasta llegar a la automatización de sus procesos buscando la mejora continua. Es por ello por lo que en la actualidad, este sector ha puesto su interés en la tecnología como fuente primaria para generar innovación, encontrando diferentes aplicaciones para el uso de tecnologías digitales que se pueden enfocar en el sector alimentario, como son drones, vigilancia satelital, tractores de especialización, tecnologías inteligentes, internet de las cosas (IoT), big data y blockchain. Debido a la importancia del sector agrícola, no solamente como impulsor económico sino también por su enorme aportación al sector alimentario (Suarez-Guzmán et al., 2020).

Conforme a lo anterior, se puede concebir a la inteligencia artificial como una tecnología que permite llevar a la agricultura tradicional a una agricultura sostenible atendiendo los factores económicos, sociales y ecológicos presentes en las prácticas agrícolas, demostrando la importancia de los aportes de la tecnología en el aumento de la productividad, la eficiencia del manejo de recursos y la disminución de riesgos, lo cual abre un panorama a nuevos mercados disminuyendo la brecha digital, y ocasiona que el sector agrícola sea adaptativo, y por ende más competitivo, enfocándose

en la sostenibilidad agrícola a partir del uso de tecnología con inteligencia artificial (Raj y de Carvalho, 2023).

Como ya se mencionó, la relación entre la IA y la agricultura empieza a florecer, pero el camino para que se diera esta cooperación ha sido largo y de paso lento, es por ello, en gran parte, que el sector agrícola es un tanto escéptico a las nuevas tecnologías. Sin embargo, a lo largo de cuarenta y cuatro años, la tecnología ha realizado aportes importantes para el desarrollo de la agricultura (Sood et al., 2022), marcando un antes y un después de la implementación de tecnología a través de diversas técnicas y dispositivos que generan el aprendizaje automático, como la integración de drones para la supervisión en tiempo real, tractores y maquinaria especializada con sistemas de navegación satelital para un mejor rendimiento y cuidado de los cultivos (Cook y O'Neill, 2020). Todo esto permea de manera positiva a la economía de las empresas agrícolas impulsando el crecimiento de la IA, buscando un mayor impacto en las economías emergentes.

Entendiendo la relevancia que tiene la IA en todos los procesos dentro la agricultura, en este trabajo se pretende explicar los beneficios que aporta el uso de tecnología con inteligencia artificial al sector agrario y los retos que enfrenta para incursionar de la agricultura tradicional a la agricultura inteligente demostrando que puede ser muy prometedora en el futuro, las aplicaciones que se han desarrollado por medio de IA en el campo a proporcionando información útil para la creación de nuevos modelos para la mejora de la producción de las empresas del sector haciendo que cada vez más productores se unan a la aplicación de la IA, aprovechando la flexibilidad de esta para la resolución de problemáticas tanto de supervisión como de operación adentrándose en la nueva tendencia de la agricultura inteligente.

La inteligencia artificial y su incursión en la agricultura

La tecnificación de la agricultura con la tecnología 4.0 abrió las puertas para la adopción de la IA por parte de la industria agrícola, mejorando la producción a través de la interconectividad de los diferentes dispositivos que se utilizan en los cultivos, control y uso eficiente del agua, manejo de recursos obteniendo mejores resultados con tecnología inteligente como

lo son el internet de las cosas (IoT) y big data. Con la incorporación de la tecnología inteligente en la agricultura se avanza a una nueva era donde los datos son uno de los activos más importantes de las empresas para la generación de estrategias de proyección para el aumento de producción, disminución de los costos entre otros beneficios; la IA está generando que las empresas agrícolas busquen la optimización de sus sistemas a través de la inversión en tecnologías inteligentes que se estima que estas inversiones aumenten un 25 % para el 2026 (Ryan et al., 2023), con la incorporación de la tecnología inteligente no solo se puede mejorar operacionalmente, también, gracias a la diversidad que presenta la agricultura, puede hacer frente a los diferentes desafíos buscando el desarrollo del sector agrícola (véase tabla 1).

La integración de la IA en las fases de la producción agrícola ha significado mejoras durante todo el proceso, desde la preproducción hasta postproducción, utilizando diversas tecnologías para la supervisión del suelo, calidad de las semillas y sistemas de riego necesarios para lograr productos de calidad. Durante este proceso interviene el aprendizaje automático (ML), mediante el procesamiento de imágenes, sistemas de sensores infrarrojos para el análisis del suelo y drones de monitoreo, con lo que se obtiene información mediante el IoT para su procesamiento (Wang et al., 2021), y de esta manera optimizar la toma de decisiones sobre los objetivos más razonables y asequibles para las empresas agrícolas contestando tres preguntas: ¿Qué vamos a cultivar?, ¿dónde lo vamos a realizar? y ¿cómo lo vamos a hacer? La tecnología inteligente es parte mediadora para el análisis de datos a través del big data y modelos matemáticos logrando información precisa, rápida y eficiente (Nie et al., 2022) para hacer frente a los desafíos en la agricultura.

Tabla 1. Tecnologías inteligentes para los desafíos de la agricultura.

Desafíos de la agricultura	Tecnología inteligente
Administración del suelo	Robótica, Drones, Realidad aumentada, IoT,
Administración de los cultivos	Deep learning, Realidad aumentada, Big data
Administración del agua	Robótica, Automatización, Machine learning, Redes neuronales, IoT, Big Data.
Fertilizantes inteligentes	IoT, Big Data.
Administración de enfermedades y pestes	Deep learning, IoT.
Cambio climático	Deep learning, Machine learning, IoT, Big Data.

Fuente: Elaboración propia con datos de Oliveira y Silva, (2023).

Sistemas basados en IA se utilizaron para hacer frente a las afectaciones del cambio climático en los cultivos de maíz del oeste de Estados Unidos, donde se aplicaron seis modelos de IA para realizar pronósticos y aumentar el rendimiento de este cultivo, del cual se espera que la producción caiga un 5 % debido a temperaturas extremas. A través de los modelos de IA se analizaron imágenes satelitales e informes meteorológicos con los datos recopilados. Los modelos predijeron la sequía del 2012 en Estados Unidos que redujo el rendimiento del cultivo de maíz en un 22 %, lo cual demuestra la utilidad de la IA para el desarrollo de estrategias que permitan adaptarse al cambio climático, basados en pronósticos con alto grado de confiabilidad (Kim et al., 2020).

La inteligencia artificial en la agricultura ha demostrado que puede beneficiar gran parte de las etapas de producción en este sector, la agricultura de precisión ha sido pionera en la adopción de tecnología manifestando las mejoras puede ofrecer al desarrollo agrícola, transformando los procesos tradicionales y haciéndolos más eficientes al aplicar soluciones tecnológicas a través de la IA.

Inteligencia artificial en la agricultura de precisión

La agricultura de precisión (AP) es una fuente de mejoras tecnológicas para la toma de decisiones a través de la recolección de información mediante diversos sistemas, satélites, vehículos autónomos que permite realizar pronósticos sobre los posibles escenarios y generar estrategias para enfrentar las diversas problemáticas, la inteligencia artificial se ha sumado al esfuerzo de la agricultura de presión por maximizar las ganancias de los productores agrícolas, asimismo, mediante el aprendizaje automático se han desarrollado nuevos algoritmos para el uso de vehículos autónomos o robots que minimiza el daño a los cultivos al evitarles pérdidas innecesarias, la automatización de los procesos permite crear redes neuronales a partir de la información recabada para el desarrollo de estrategias aplicables a los diferentes tipos de cultivos, suelos, tipos de clima que permita replicar a través de la IA las técnicas agrícolas en diversas zonas que compartan similitudes geográficas (Linaza et al., 2021).

Con el uso de IA a través de la AP, los agricultores buscan mantener la sostenibilidad del sector mediante estrategias para hacer frente a la dis-

minución de recursos y la creciente demanda de producción, entendiendo que la respuesta a todos estos desafíos es la tecnología que brinda los mejores resultados en el manejo eficiente de recursos obteniendo altos rendimientos, mejor adaptabilidad ante los cambios por los efectos climáticos como económicos (Ayoub Shaikh et al., 2022), la aplicación de la IA continua demostrando tener un gran potencial para controlar de mejor forma las fases de producción agrícola como se muestra en la figura uno.

Figura 1. Fases de la producción agrícola.

Fuente: elaboración propia con datos obtenidos de Ayoub Shaikh et al. (2022).

Con el uso de la AP en las fases de producción agrícola se optimizan los procesos en selección de cultivo, se analizan de forma precisa los datos recabados de las cosechas anteriores para medir las condiciones y elegir el cultivo más propicio, con esto, se puede obtener información detallada sobre la calidad del suelo, lo cual permite definir el tipo de riego así como los fertilizantes necesarios, evitando gastos y contaminación excesiva del suelo con químicos que pueden dañar futuros cultivos (Zhu et al., 2019). En la fase de la cosecha, la utilización de drones de monitoreo y equipo de sensores permite supervisar la calidad de los cultivos identificando en-

fermedades de manera temprana, para tomar medidas preventivas, con los tractores de navegación satelital se reduce el tiempo de cosecha comparado con el uso de jornaleros, pasando a la última fase de producción se utilizan sistemas de aprendizaje automático, big data, IoT para procesar la información con la finalidad de crear estrategias que les permitan ser competitivos (Unal, 2020).

Basados en los beneficios potenciales se aplicó la Agricultura de Precisión en el cultivo de moras en la zona de Riobamba, Ecuador. Mediante el desarrollo de una red de sensores inalámbricos que permiten realizar un sistema de riego inteligente para el cuidado del cultivo, demostrando la importancia de la gestión del vital líquido para las actividades de este sector. La red de monitoreo consta de 3 nodos que reciben la información del área de cultivo de manera inalámbrica, los datos recabados permitieron conocer que las 140 plantas de mora cultivadas en la zona, necesitan para su crecimiento 180.44 litros de agua, el sistema calculó que 3 horas de riego es tiempo suficiente para la hidratación deseada, representando un ahorro de 60.48 litros de agua al día, 725.76 litros al mes y 8709.12 litros al año, estos sistemas pueden cubrir hasta 200 hectáreas de cultivo (Paredes et al., 2021).

La inclusión de los avances tecnológicos en los procesos agrícolas ha permitido la evolución del sector en un tiempo relativamente corto, primordialmente con base en el manejo efectivo de la información que se relaciona con los riesgos, mercados, clientes e inversiones. La información tiene un papel fundamental en la inteligencia artificial, todos los empresarios conocen la importancia de la información siendo un concepto básico de la administración, pero con los avances tecnológicos la generación de información aumentó exponencialmente, de tal manera que, procesarla por los medios tradicionales resulta casi imposible e intentarlo es totalmente ineficaz. Por lo que se han empleado diversos tipos de tecnología con inteligencia artificial que ha permitido a los agricultores extraer información relevante que permita generar conocimiento sobre las diferentes áreas de aplicación de la agricultura, tecnologías como big data y el IoT han demostrado ser sistemas importantes de apoyo para enfrentar los desafíos dentro de la agricultura digital.

Big data y su aplicación en la agricultura

Para el procesamiento de datos los sistemas conocidos como big data son definidos como sistemas especializados para el tratamiento de grandes bancos de información, de los cuales se puede obtener estadísticos, análisis predictivos con los que se pueden plantear estrategias que mejora la eficiencia de no solo de los cultivos, sino que aumenta la seguridad alimentaria (Dakshayini y Balaji Prabhu, 2020).

La aplicación de sistemas de big data se hace presente en todos los ámbitos donde la inteligencia artificial se utiliza para el procesamiento de información en busca de aplicar el conocimiento generado en la práctica real, el sector agrícola desde una perspectiva de innovación empieza a utilizar este tipo de tecnología que aporta una mayor confianza en la toma de decisiones avaladas por las conocidas 4 V del big data, volumen, velocidad, variedad y veracidad, por medio de las cuales se adquiere la información de las fases del proceso agrícola que se han mencionado anteriormente cultivos, suelo, clima entre otros, procesando estos datos a través de sistemas especializados obteniendo análisis espaciales permitiendo evaluar alternativas desde diferentes perspectivas para la toma eficiente de decisiones y la implementación de estrategias que permitan mantener la sostenibilidad de la agricultura (Bhat y Huang, 2021).

- *Volumen:* Dentro de Big data son los datos macros que se almacenan dentro de clúster para su posterior análisis, siendo uno de los desafíos para esta la infraestructura necesaria debido a que el volumen es cada vez mayor.
- *Velocidad:* El proceso de toma decisiones en sectores donde se utiliza tecnología se debe realizar de forma ágil, por tal razón la velocidad de procesamiento de información es en tiempo real creando modelos que visualicen la toma de decisiones.
- *Variedad:* Maneja que la información no es excluyente de los diferentes procesos, se debe analizar en conjunto, ya que esta repercute de manera general en la toma decisiones, por lo cual se busca que los datos convergen en algún punto.
- *Veracidad:* La confiabilidad de la información es vital para la correcta toma de decisiones, por lo que se debe garantizar la fidelidad de los datos.

La variedad de aplicaciones del big data dentro de la agricultura otorga beneficios para la solución de problemas de sus diversas áreas como la cadena de valor, predicciones de oferta y demanda, gestión adecuada de los recursos y rápida respuesta a los cambios climáticos (Kamilaris et al., 2017), lo cual permite a los agricultores aumentar la rentabilidad al tomar decisiones inteligentes sobre los posibles escenarios, manteniendo la importancia del procesamiento de los macro datos en la agricultura (Ang y Seng, 2021).

Modelos basados en big data fueron utilizados en actividades agropecuarias en China, a través de estructuras de sensores que recopilan información del suelo, almacenando para su análisis y depuración para generar modelos predictivos sobre la producción que permite a los productores una toma de decisiones basadas en datos estadísticos con alto grado de confiabilidad, otro modelo basado en big data fue aplicado en la India para un análisis de las actividades agrícolas por medio de una red de sensores que recababan información de la producción de los cultivos que eran monitoreados a través de aplicaciones móviles, el modelo proporcionó información específica para sistematizar estrategias de comercialización para el nivel estatal y nacional (Vite Cevallos et al., 2020).

El big data genera la tecnología inteligente que permite a la industria agrícola obtener información oportuna de los diversos escenarios a nivel mundial, así como las problemáticas de cada zona y las medidas utilizadas para solventar estas problemáticas, lo cual genera un conocimiento valioso que ayuda en la generación de estrategias para lograr un mejor posicionamiento frente a la competencia (Maya Gopal y Chintala, 2020), así como el IoT que se desempeña en diversas áreas agrícolas proporcionando aprendizaje automático permitiendo la transición de la agricultura tradicional a la inteligencia artificial (Bhat y Huang, 2021).

Internet de las cosas y sus beneficios en la agricultura

El internet de las cosas (IoT), al igual que los sistemas de big data forma parte de las tecnologías inteligentes que, sumadas a la inteligencia artificial, han sentado precedentes de la importancia de incorporar estas tecnologías de vanguardia para mejorar la productividad de la zona agrícola, la carac-

terística principal del IoT es la conectividad de los objetos en este caso de todos los dispositivos dentro de la agricultora que generen información creando redes de comunicación de intercambio de datos no estructurados los cuales son tomados por la inteligencia artificial para realizar análisis para darle estructura a estos datos para generar decisiones fundamentadas en información de calidad (Yang, 2020), con la agricultura de precisión se dieron avances significativos, pero con la adopción del IoT se ha logrado la automatización con mejores resultados sobre todo en las tareas de alta complejidad y que con mano de obra humana demoraba de forma considerable, situación que se empieza a sortear con los equipos modernos para las diferentes áreas como un ejemplo de esto son las tractores de última generación guiados por sistemas de navegación satelital ya considerados por los agricultores como esencial e irremplazables para aumentar la productividad (Ci?Velek, 2021).

Otros de los dispositivos que han generado grandes beneficios con la aplicación del IoT es el uso de drones para monitorear en tiempo real cultivos y la calidad del suelo, mencionando que estos utilizan sensores, cámaras y otro tipo de tecnologías que permiten recabar este tipo de información para un mejor control de calidad (Van Der Merwe et al., 2020), en el municipio de Apartado, Colombia, se implementó el uso de drones para actividades de fumigación otorgando un alcance mayor en áreas de acceso difícil y de riesgo para trabajadores, el uso de esta tecnología permite una fumigación de alta precisión y mapeos para el control de plagas manteniendo saludables los cultivos de banano (Idarraga Vargas et al., 2024).

El énfasis al seguimiento de los cultivos es necesario para asegurar la producción conforme a la demanda alimentaria de la población por lo cual el IoT y la IA al automatizar las actividades agrícolas son fundamentales para la agricultura (Subeesh y Mehta, 2021), las tecnologías inteligentes han aumentado la precisión de las decisiones estratégicas al acumular información e interpretar sus datos de forma que aseguran que las medidas que se tomen para el logro de los objetivos sean las adecuadas, proporcionando una adaptación rápida y flexible al entorno del sector agrícola.

Esta adaptación que proporciona el IoT a los procesos de la agricultura, no es una tarea fácil de realizar, debido a la gran variedad de factores que intervienen tanto organizacionales como ambientales siendo estos los

más impredecibles, con la intervención del IoT para recabar toda la información producida de las diferentes fuentes que el big data acumula para su análisis entre datos estructurados y no estructurados por medio de la IA identificando los marcadores de tendencias, los casos de éxito, los riesgos, procesando todo a través del aprendizaje automático que permitirá hacer más eficientes los sistemas llegando al aprendizaje profundo que ya no es solo el procesamiento de datos, sino que se pueden representar en diferentes niveles de abstracción para el desarrollo de soluciones o mejoras novedosas para la creación de sistemas autónomos en la agricultura (Misra et al., 2022).

Con la cantidad de datos que se tiene que procesar por medio del IoT, el respaldo de la I.A. es pieza clave debido a que proporciona mejoras en la operatividad, las ventajas que se obtienen por medio de las tecnologías inteligentes, amplia los panoramas de una agricultura sostenible por medio de la recopilación de datos con lo que se puede pronosticar la producción agrícola basada en modelos sólidos de inteligencia artificial (Adli et al., 2023).

Conclusiones

La agricultura a presentados avances tecnológicos en las últimas décadas, en la búsqueda de la sostenibilidad con el aumento de la producción, garantizando mejoras en las prácticas tradicionales del cuidado de cultivos y los diferentes factores que intervienen en los procesos agrícolas, uno de los principales desafíos en el sector alimentario es el aumento de la población que está alcanza más de ocho mil millones con el pronóstico de llegar a más de nueve mil millones para el año 2050, lo que aumentar la demanda de alimentos en un 69 %, y por ende la necesidad de una mayor producción de alimentos. La alternativa para hacer frente a estas problemáticas en la implementación de tecnologías digitales basada en la inteligencia artificial que ha demostrado la optimización de los recursos de manera significativa, analizando datos para la creación de modelos que permitan realizar pronósticos más eficientes para el cuidado de las cosechas lo que incide en una alta producción.

Realizando un análisis de las diversas etapas por los cuales ha atravesado el sector agrícola desde la agricultura 1.0 a la agricultura 4.0, es en la

agricultura de precisión que identificamos el primer paso para la inclusión de tecnología digital, incursionando de manera paulatina en la agricultura inteligente por medio de las tecnologías modernas basadas en inteligencia artificial como los son el internet de las cosas, el big data, computación en la nube, la conectividad mediante tecnologías de información y comunicación, y el aprendizaje automático que hacen viable una agricultura sostenible.

Son diversas razones que hacen a la inteligencia artificial sea un concepto cada vez más replicado dentro del sector agrícola, quienes han encontrado en esta, una alternativa para la automatización de los procesos reemplazando a los métodos tradicionales de manera efectiva disminuyendo los costos de mano de obra, aumentando las capacidades de producción, y abriendo nuevas oportunidades de crecimiento con el uso de la inteligencia artificial para crear modelos que estandarice la gestión de los datos para la planeación (Sood et al., 2022), aumentado la calidad de los procesos agrícolas.

Los beneficios que brinda la inteligencia artificial para impulsar el sector agrario a través de las diversas tecnologías que se pueden aplicar en este campo, son prometedoras, si bien hay algunos retos que se deben superar como la falta de infraestructura o preparación mediante capacitaciones al personal para el manejo de sistemas especializados, los aportes que genera el desarrollo tecnológico supera con creces estas limitantes, por lo que se puede inferir que en un futuro no solo se seguirá utilizando la inteligencia artificial en este sector, sino que ampliará más el abanico de sistemas y dispositivos que ofrecerán soluciones holísticas en el contexto agrícola.

Por lo cual el camino que ha estado marcando el desarrollo tecnológico en la agricultura tradicional pasando por la agricultura de precisión hasta este momento donde la agricultura inteligente empieza atender las problemáticas sobre la productividad, la seguridad alimentaria, sostenibilidad, cuidado ambiental basada en prácticas agrícolas inteligentes (Kamilaris et al., 2017), donde las últimas aplicaciones de big data y el internet de las cosas desarrollan todo su potencial para realizar mejoras de los procesos y servicios del sector, creando oportunidades para integrar las tecnologías digitales 4.0, reconociendo que la inteligencia artificial aún no alcanza su nivel más álgido, por lo cual se espera desempeñe un papel clave en el crecimiento de la agricultura.

Referencias

Adli, H. K., Remli, M. A., Wan Salihin Wong, K. N. S., Ismail, N. A., González-Briones, A., Corchado, J. M., & Mohamad, M. S. (2023). *Recent Advancements and Challenges of AIoT Application in Smart Agriculture: A Review. Sensors*, 23(7), 3752. https://doi.org/10.3390/s23073752

Ang, K. L.-M. y Seng, J. K. P. (2021). Big Data and Machine Learning With Hyperspectral Information in Agriculture. *IEEE Access*, 9, 36699–36718. https://doi.org/10.1109/ACCESS.2021.3051196

Ayoub Shaikh, T., Rasool, T., & Rasheed Lone, F. (2022). *Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming*. Computers and Electronics in Agriculture, 198, 107-119. https://doi.org/10.1016/j.compag.2022.107119

Bhat, S. A., & Huang, N.-F. (2021). *Big Data and AI Revolution in Precision Agriculture: Survey and Challenges*. IEEE Access, 9, 110209–110222. https://doi.org/10.1109/ACCESS.2021.3102227

Ci?Velek, Ç. (2021). *Development of an IoT based (LoRaWAN) Tractor Tracking System*. Tar?m Bilimleri Dergisi. https://doi.org/10.15832/ankutbd.769200

Cook, P., & O'Neill, F. (2020). *Artificial Intelligence in Agribusiness is Growing in Emerging Markets*. https://policycommons.net/artifacts/1279703/artificial-intelligence-in-agribusiness-is-growing-in-emerging-markets/1870529/

Dakshayini, M. y Balaji Prabhu, B. V. (2020). An Effective Big Data and Blockchain (BD-BC) Based Decision Support Model for Sustainable Agriculture System. En A. Haldorai, A. Ramu, S. Mohanram, & C. C. Onn (Eds.), *EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing* (pp. 77–86). Springer International Publishing. https://doi.org/10.1007/978-3-030-19562-5_8

Idarraga Vargas, A., Pinzon Gonzalez, J. E., & Vergara Suarez, S. D. J. (2024). *Análisis de los costos de la aplicación de drones en la fumigación en las fincas bananeras del municipio de Apartadó. Estudio de caso*. https://repository.ucc.edu.co/entities/publication/1fe239ef-e6d0-4a3b-810d-daeecd5c17e0

Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). *A review on the practice of big data analysis in agriculture*. Computers and Elec-

tronics in Agriculture, 143, 23-37. https://doi.org/10.1016/j.compag.2017.09.037

Kim, N., Na, S.-I., Park, C.-W., Huh, M., Oh, J., Ha, K.-J., Cho, J., & Lee, Y.-W. (2020). *An Artificial Intelligence Approach to Prediction of Corn Yields under Extreme Weather Conditions Using Satellite and Meteorological Data.* Applied Sciences, 10(11), 3785. https://doi.org/10.3390/app10113785

Linaza, M. T., Posada, J., Bund, J., Eisert, P., Quartulli, M., Döllner, J., Pagani, A., G. Olaizola, I., Barriguinha, A., Moysiadis, T., & Lucat, L. (2021). *Data-Driven Artificial Intelligence Applications for Sustainable Precision Agriculture.* Agronomy, 11(6), 1227. https://doi.org/10.3390/agronomy11061227

Maya Gopal P.S., & Chintala, B. R. (2020). Big Data Challenges and Opportunities in Agriculture: *International Journal of Agricultural and Environmental Information Systems,* 11(1), 48–66. https://doi.org/10.4018/IJAEIS.2020010103

Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2022). IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry. *IEEE Internet of Things Journal*, 9(9), 6305–6324. https://doi.org/10.1109/JIOT.2020.2998584

Nie, J., Wang, Y., Li, Y., & Chao, X. (2022). Artificial intelligence and digital twins in sustainable agriculture and forestry: A survey. *Turkish Journal of Agriculture and Forestry,* 46(5), 642–661. https://doi.org/10.55730/1300-011X.3033

Oliveira, R. C. D., & Silva, R. D. D. S. E. (2023). *Artificial Intelligence in Agriculture: Benefits, Challenges, and Trends.* Applied Sciences, 13(13), 7405. https://doi.org/10.3390/app13137405

Paredes, M., Zúñiga, W., Morocho Caiza, A. F., & Mendoza, M. (2021). Agricultura de precisión mediante WSN con nodos inteligentes aplicada a un sistema de riego en cultivo de mora. *Revista Perspectivas,* 3(2), 26–30. https://doi.org/10.47187/perspectivas.vol3iss2.pp26-30.2021

Raj, V. H. A., & De Carvalho, C. X. (2023). A Perspective on the Application of Artificial Intelligence in Sustainable Agriculture with Special Reference to Precision Agriculture. *SDMIMD Journal of Management,* 1–13. https://doi.org/10.18311/sdmimd/2023/33006

Ruiz-Real, J. L., Uribe-Toril, J., Torres Arriaza, J. A., & De Pablo Valenciano, J. (2020). *A Look at the Past, Present and Future Research Trends of*

Artificial Intelligence in Agriculture. Agronomy, 10(11), 1839. https://doi.org/10.3390/agronomy10111839

Ryan, M., Isakhanyan, G., & Tekinerdogan, B. (2023). An interdisciplinary approach to artificial intelligence in agriculture. *NJAS: Impact in Agricultural and Life Sciences*, 95(1), 2168568. https://doi.org/10.1080/27685241.2023.2168568

Sood, A., Sharma, R. K., & Bhardwaj, A. K. (2022). Artificial intelligence research in agriculture: A review. *Online Information Review*, 46(6), 1054–1075. https://doi.org/10.1108/OIR-10-2020-0448

Subeesh, A., & Mehta, C. R. (2021). *Automation and digitization of agriculture using artificial intelligence and internet of things*. Artificial Intelligence in Agriculture, 5, 278–291. https://doi.org/10.1016/j.aiia.2021.11.004

Unal, Z. (2020). Smart Farming Becomes Even Smarter With Deep Learning—*A Bibliographical Analysis*. IEEE Access, 8, 105587–105609. https://doi.org/10.1109/ACCESS.2020.3000175

Van Der Merwe, D., Burchfield, D. R., Witt, T. D., Price, K. P., & Sharda, A. (2020). Drones in agriculture. En *Advances in Agronomy* (Vol. 162, pp. 1–30). Elsevier. https://doi.org/10.1016/bs.agron.2020.03.001

Vite Cevallos, H., Townsend Valencia, J., Carvajal Romero, H., Vite Cevallos, H., Townsend Valencia, J., & Carvajal Romero, H. (2020). Big Data e internet de las cosas en la producción de banano orgánico. *Revista Universidad y Sociedad*, 12(4), 192–200.

Wang, J., Peng, J., Li, H., Yin, C., Liu, W., Wang, T., & Zhang, H. (2021). Soil Salinity Mapping Using Machine Learning Algorithms with the Sentinel-2 MSI in Arid Areas, China. Remote Sensing, 13(2), 305. https://doi.org/10.3390/rs13020305

Yang, L. B. (2020). *Application of Artificial Intelligence in Electrical Automation Control*. Procedia Computer Science, 166, 292–295. https://doi.org/10.1016/j.procs.2020.02.097

Zhu, Zhou, Gao, Bao, He, & Feng. (2019). *Near-Infrared Hyperspectral Imaging Combined with Deep Learning to Identify Cotton Seed Varieties*. Molecules, 24(18), 3268. https://doi.org/10.3390/molecules24183268

Parte III

Entorno a la IA: reflexiones sobre salud, conciencia y neuropsicología

Capítulo **14**

La IA en la salud

Jessica Beltrán

https://doi.org/10.61728/AE24001144

Introducción

Un componente especialmente destacado en la transformación tecnológica de los últimos años es la inteligencia artificial (IA), cuyos avances han convertido en realidad lo que hace poco tan solo era parte de la imaginación. Estos progresos han dejado su huella en diversos aspectos de nuestras vidas, incluyendo la esfera crucial de la salud. Resulta fascinante y, a la vez, asombroso constatar que hoy en día existen individuos que nacieron en una época en la que las vacunas para enfermedades como la polio o el sarampión aún no eran una realidad y que, gracias a avances multidisciplinarios, entre ellos la IA, estas personas han sido testigos de avances impresionantes, como un rápido desarrollo que condujo a la creación de una vacuna contra el covid-19, poniendo fin a una amenaza pandémica que afectaba a la sociedad e incluso otros avances que parecen sacados de la ciencia ficción.

No pasará mucho tiempo antes de que sea considerado común el uso de dispositivos con algoritmos de IA dedicados a diagnosticar o prevenir enfermedades, o que las personas reciban tratamientos y terapias personalizadas según sus rutinas, complexión, información genética, capacidades, entre otros factores. Todo esto ocurrirá mientras se analizan grandes cantidades de datos de diversas personas para identificar patrones y estrategias adecuadas. Estos avances prometen una revolución en la forma en que abordamos la salud, proporcionando soluciones más precisas y personalizadas para mejorar la calidad de vida.

En este capítulo se describen aportes de la inteligencia artificial enfocados en mejorar diversos aspectos de la salud. Asimismo, se realiza un análisis prospectivo sobre las expectativas en función de la ruta histórica de los aportes y de los avances recientes en técnicas de IA.

¿Qué entendemos por salud y cómo puede la IA saber nuestro estado de salud?

Conforme a la definición de la Organización Mundial de la Salud (OMS), la salud es un estado en el cual las personas experimentan bienestar físico, mental y social, y no se limita simplemente a la ausencia de enfermedades. Esta amalgama de componentes subraya la importancia de que las personas se concentren en preservar el funcionamiento óptimo y el equilibrio en diversos aspectos de su vida para alcanzar y mantener un estado de salud integral.

Las personas constantemente revelamos señales que reflejan nuestro estado de salud. En ciertas ocasiones, manifestamos síntomas evidentes, como la tos o el escurrimiento nasal, que sugieren la presencia de un resfriado o gripe. Sin embargo, también emitimos señales más sutiles o menos visibles que ofrecen información valiosa sobre nuestro bienestar interno, como la frecuencia respiratoria, la temperatura corporal o la configuración de nuestros latidos cardiacos. Además, algunas enfermedades generan cambios en nuestro organismo, como la formación de piedras o tumores, los cuales, mediante estudios apropiados, podrían detectarse, indicando la existencia de algún problema de salud.

Toda la información que brindamos sobre nuestro estado de salud es cuantificable mediante diversos dispositivos. Por ejemplo, para evaluar la actividad eléctrica del corazón, se recurre a un dispositivo llamado Holter; para medir la temperatura, se utiliza el termómetro, y para visualizar posibles tumores, se emplean técnicas como radiografías, tomografías y resonancias magnéticas. Así, existen una variedad de dispositivos y sensores diseñados para medir las señales que emanamos, las cuales pueden indicar la presencia de algún elemento inusual.

Los seres humanos recopilamos información del entorno mediante nuestros sentidos; por ejemplo, observamos si alguien tiene la nariz roja, escuchamos tos o identificamos mediante el olfato si una herida desprende un olor desagradable. De manera análoga, los sentidos de la IA son los dispositivos o sensores que se emplean para monitorizarnos. Un ejemplo es el termómetro, que registra la temperatura de una persona cada hora. Esta información se suministra a un algoritmo de IA, permitiéndole discernir

cambios en la temperatura corporal. Posteriormente, el algoritmo puede emplear estos datos, junto con otros, como la detección de tos, para prever la presencia de una enfermedad en la persona. Este enfoque demuestra cómo la inteligencia artificial utiliza sensores para interpretar señales y generar predicciones de salud.

Conforme a la definición de la OMS, el bienestar abarca no solo aspectos físicos, sino también mentales y sociales. Para incorporar IA en estos ámbitos, es esencial dotarla de sentidos capaces de medir y recopilar información que refleje nuestro estado mental y social. Aunque no podemos cuantificar directamente el bienestar mental y social de la misma manera que las señales físicas, como la temperatura, podemos evaluar indicadores intermedios que ofrecen pistas sobre nuestro estado.

En términos generales, nuestro bienestar mental y social se manifiesta a través de nuestras acciones e interacciones tanto con el entorno como con otras personas. Por ejemplo, condiciones que afectan el bienestar mental, como la demencia, se externalizan limitando la capacidad de llevar a cabo actividades diarias de manera efectiva. Por lo tanto, al medir las actividades y la calidad con que las realizamos, obtenemos una evaluación indirecta del bienestar mental. Otro ejemplo ilustrativo de cómo se refleja el bienestar mental y social ocurre en situaciones de ansiedad o depresión, donde nuestro comportamiento en redes sociales puede revelar patrones distintivos. Aunque no podemos medir directamente el bienestar social, podemos evaluar conexiones sociales digitales, interacciones, conversaciones, entre otros elementos relevantes (Jacobson y Bhattacharya, 2022). Este enfoque nos permite utilizar indicadores tangibles para entender y mejorar aspectos clave de nuestro bienestar integral.

Entonces, para obtener información sobre las actividades que estamos llevando a cabo, se recurre a los datos proporcionados por diversos sensores, los cuales son sometidos a un procesamiento y análisis mediante algoritmos de IA. Por ejemplo, a través de cámaras, podemos "ver" eventos en una escena, los micrófonos nos permiten "escuchar" los sonidos que generamos, mientras que los acelerómetros o giroscopios nos ayudan a "sentir" nuestros movimientos. De esta manera, se logra identificar acciones como encender la estufa y luego abandonar el área olvidando la tarea de cocinar. Asimismo, es posible medir la manera en que una persona gestio-

na su atención al abrir y cambiar entre varias aplicaciones en su teléfono. Este enfoque integrado de sensores y algoritmos de inteligencia artificial permite una comprensión detallada de nuestras actividades cotidianas.

Evolución de la Inteligencia Artificial

El concepto de inteligencia artificial se originó en la década de 1950 y se refiere a la capacidad de las máquinas o agentes creados por los seres humanos para simular inteligencia, es decir, adquirir comprensión y capacidad de resolución de problemas. A lo largo de los años, se han desarrollado diversas técnicas de IA que han permitido abordar problemas cada vez más complejos. En sus primeras etapas, las técnicas de IA se centraron en la creación de sistemas basados en reglas, por ejemplo, "si ocurre x, realiza esta acción; si no, realiza esta otra". Estas reglas se derivaban del análisis de la experiencia de expertos en campos específicos. Bajo este enfoque, se realizaron contribuciones significativas en el ámbito de la salud, como los proyectos Dendral (Buchanan & Feigenbaum, 1978) y MYCIN (Kaul et al., 2020), que se dedicaron a la identificación de moléculas orgánicas y al diagnóstico de enfermedades infecciosas, respectivamente. Estos proyectos representaron hitos importantes al aplicar la inteligencia artificial para abordar desafíos médicos y sentaron las bases para futuros avances en la intersección de la tecnología y la atención médica.

En la década de 1980, surgieron nuevos enfoques de inteligencia artificial conocidos como Aprendizaje de Máquina (Machine Learning, ML). A diferencia de los algoritmos tradicionales, donde las reglas son introducidas directamente por las personas, los algoritmos de ML permiten que las máquinas aprendan automáticamente a través del análisis de información disponible. Por ejemplo, se han realizado aportes en donde se suministran a la computadora diversas imágenes de radiografías pulmonares, abarcando tanto casos saludables como enfermos (Gonem et al., 2020). Empleando algoritmos de ML, se analiza la información para identificar las características relevantes en las radiografías que diferencian entre ambos tipos de pulmones. Cada radiografía se denomina con la palabra "ejemplo", lo que implica que, en ML, el aprendizaje se lleva a cabo a través de ejemplos.

Es importante destacar que las máquinas, por sí solas, no son capaces de resolver problemas complejos. La intervención humana es esencial en

todas las etapas del ML. Por ejemplo, los ejemplos de entrada deben prepararse de manera que puedan ser adecuadamente interpretados por las computadoras, y se proporcionan modelos a la máquina que sirven como guía para el aprendizaje. Durante este proceso, se ajustan los parámetros para encontrar la configuración que mejor resuelve el problema. La sinergia entre la capacidad de aprendizaje automático de las máquinas y la dirección humana es fundamental para el éxito de las aplicaciones de ML en diversas áreas.

Dentro del campo del aprendizaje de máquina (ML), destacan los algoritmos conocidos como deep learning o aprendizaje profundo. A diferencia del ML clásico, estos algoritmos tienen la capacidad única de encontrar representaciones complejas de los datos dentro de su propia estructura o arquitectura, demostrando resultados superiores en una variedad de problemas. Esta característica ha propiciado el desarrollo de aplicaciones más sofisticadas y complejas, desempeñando un papel clave en los avances que respaldan la salud. Un ejemplo notable de la influencia del Deep Learning en el ámbito de la salud es su aplicación en el procesamiento de lenguaje natural, que ha facilitado la creación de asistentes virtuales diseñados para apoyar a los profesionales médicos, como WATSON de IBM o el chatbot Pharmabot (Kaul et al., 2020). Además, los avances en visión computacional han permitido mejoras significativas en la interpretación de imágenes médicas.

Ejemplos de uso de IA en la salud

Desde los primeros días de la inteligencia artificial hasta la actualidad, se han realizado numerosos avances significativos en el ámbito de la salud. La descripción de estos logros puede clasificarse de diversas maneras, y se han llevado a cabo revisiones sistemáticas en trabajos como (Jiang et al., 2017) y (Apell y Eriksson, 2023), los cuales detallan diversas investigaciones de IA centradas en el ámbito de la salud. A continuación, se presentan ejemplos representativos de lo que la IA puede lograr. La organización de estos ejemplos sigue una estructura temática que abarca diagnósticos, apoyo a cuidadores, terapias o medicamentos adaptados a cada individuo, y aportes en conjunto con otras disciplinas. Cada uno de estos aspectos

destaca la versatilidad y el impacto potencial de la IA en diversas áreas cruciales para el bienestar y el avance de la atención médica.

Diagnósticos

Un diagnóstico consiste en determinar si una persona padece de una enfermedad o condición médica. Los diagnósticos se realizan por personas expertas en medicina a través de examinar los síntomas que se observan y que manifiesta el paciente, sus antecedentes médicos e historia familiar. También se analizan resultados de estudios de laboratorio, como sanguíneos o de tejidos, y de otros tipos de estudios, como radiografías, tomografías, entre otros. En algunos casos se requiere llevar a cabo un diagnóstico diferencial, ya que hay enfermedades que comparten síntomas.

Como se ha expuesto previamente, la obtención de mediciones de las manifestaciones físicas de los pacientes mediante dispositivos es esencial en el ámbito médico. Con el propósito de optimizar el proceso de diagnóstico, se han desarrollado contribuciones que emplean algoritmos de IA. Estas aportaciones utilizan la información recopilada para generar sugerencias diagnósticas, las cuales requieren la evaluación de un profesional médico especializado. Un ejemplo destacado de esta aplicación es la utilización de técnicas de aprendizaje profundo, específicamente la implementación de Redes Neuronales Convolucionales, en el análisis de radiografías de pecho y tomografías computarizadas para facilitar la identificación de casos de covid-19 (Subramanian et al., 2022).

Dada la diversidad de síntomas asociados a una enfermedad, se han explorado distintos enfoques, como la evaluación de tos para determinar la probabilidad de que una persona padezca covid. En este contexto, investigaciones recientes han empleado diversas técnicas de aprendizaje automático con este propósito (Pahar et al., 2021). Si bien las imágenes médicas, como radiografías y tomografías, brindan información relevante sobre el estado pulmonar de los pacientes y son fundamentales en el diagnóstico de covid, su adquisición puede ser costosa y demorada, generando la necesidad de evitar su utilización innecesaria para prevenir contagios. Por otro lado, la detección rápida de tos asociada probablemente a covid proporciona una evaluación inicial que, aunque no siempre arroje resul-

tados concluyentes, presenta la ventaja de ser ágil y, si se implementa en dispositivos móviles, puede reducir el riesgo de contagios. Por tal motivo, es deseable utilizar información que se complementa mutuamente para mejorar los resultados en los diagnósticos lo que constituye un enfoque integral y efectivo en la mejora de los procesos médicos de diagnóstico.

Otras investigaciones centradas en el diagnóstico de enfermedades mediante inteligencia artificial, particularmente a través de técnicas de aprendizaje automático clásico y aprendizaje profundo, abordan la identificación de retinopatías diabéticas (Das et al., 2022), la detección de cáncer de mama (Fátima et al., 2020) y otros tipos de cáncer (Kourou et al., 2021). Estos trabajos enfrentan desafíos significativos, no solo al diagnosticar enfermedades ya presentes en los pacientes, sino también al aspirar a preverlas en etapas tempranas para evitar o retrasar los síntomas, mejorando así la calidad de vida de las personas.

Otro ejemplo destacado de la aplicación de la IA en el ámbito de la salud; se observa en la utilización de electroencefalogramas (EEG), que registran la actividad eléctrica del cerebro mediante electrodos colocados en la cabeza. Por ejemplo, se ha propuesto la utilización de estos sensores junto con una combinación de diversas técnicas de aprendizaje automático para la detección de convulsiones, un indicador que puede llevar al desarrollo de epilepsia (Lekshmy et al., 2022). Además, debido a la naturaleza de estos sensores, también se han revelado como herramientas útiles para el diagnóstico de enfermedades mentales, como la depresión (Sarkar et al., 2022).

Además, se han emprendido investigaciones centradas en el diagnóstico en las cuales no se realizan mediciones directas de señales físicas, sino que se basan en el análisis del comportamiento de las personas, lo que ayuda a los expertos a determinar si alguien presenta o no una condición específica. Un ejemplo ilustrativo es el estudio del comportamiento de las personas al observar su entorno, ya que se ha descubierto que, mediante algoritmos de IA, es posible identificar patrones distintivos en la forma en que fijan la mirada en objetos relacionados con el deterioro cognitivo (Beltrán et al., 2018). Siguiendo esta misma línea de enfoque basado en las fijaciones oculares, se han desarrollado trabajos específicos para respaldar el diagnóstico de personas que presentan trastorno del espectro autista

(Alcañiz et al., 2022). Este enfoque novedoso destaca la utilidad de la inteligencia artificial en la interpretación de comportamientos sutiles para mejorar el proceso de diagnóstico en condiciones médicas complejas.

En términos generales, la comunidad científica ha contribuido y continúa haciéndolo mediante investigaciones que emplean IA en el diagnóstico de enfermedades (Bakator y Radosav, 2018). Surge la interrogante de por qué es valioso delegar a la IA tareas tan delicadas como el diagnóstico, cuando esta función puede ser llevada a cabo por personas. La respuesta radica en la capacidad de la IA para analizar grandes cantidades de información en un periodo de tiempo significativamente menor que los seres humanos. Además, las limitaciones de horario y capacidad de análisis de datos durante una jornada laboral son factores restrictivos para las personas. Existiendo también regiones remotas sin acceso a expertos capaces de analizar los datos, la IA se presenta como una solución para superar esta barrera geográfica.

Cuidado de personas

Existen situaciones en las que las personas no pueden valerse por sí mismas, ya sea debido a una circunstancia temporal, como la recuperación de un accidente o una enfermedad no grave, o porque se encuentran en una condición que demanda apoyo continuo, como una discapacidad o un deterioro cognitivo. En tales casos, estas personas dependen del cuidado de otros, quienes pueden estar a cargo tanto de tareas simples como de aspectos fundamentales de su vida diaria. La responsabilidad de cuidar a otra persona puede ser exigente y desafiante, llegando incluso a requerir atención constante las 24 horas del día. Este compromiso refleja la necesidad crucial de apoyo y atención dedicada para aquellos que enfrentan limitaciones temporales o permanentes en su capacidad de autosuficiencia.

Por esta razón, la IA se ha empleado para desarrollar iniciativas que brinden apoyo a las personas encargadas del cuidado. Este respaldo adopta diversas formas. Por ejemplo, se han desarrollado sistemas capaces de monitorear a personas con deterioro cognitivo, para alertar a sus cuidadores sobre situaciones que requieren atención inmediata y así disminuir su carga de trabajo y la presencia de accidentes y heridas (Kim et al., 2023).

Otro ejemplo de aplicación se enfoca en algoritmos inteligentes para indicar la posible formación de úlceras de presión en pacientes postrados, y así sus cuidadores tomen decisiones sobre cambios de posición y aseo en las áreas con potencial de desarrollo (Silva et al., 2022). Asimismo, se ha propuesto utilizar agentes conversacionales con pacientes con demencia ya que este tipo de sistemas pueden responder preguntas repetitivas, apoyarles en recordar actividades pendientes y tomar medicamentos, lo que reduce la carga de trabajo de sus cuidadores (Jiménez et al., 2022).

Personalización de terapias o medicamentos

Usando IA, se han desarrollado propuestas a cambiar el enfoque en la indicación de medicamentos a pacientes. Tradicionalmente, los medicamentos se prescriben bajo un esquema de prueba y error, y el medicamento se ajusta dependiente de cómo reaccionó el paciente. Ya que las personas somos diferentes, lo que le funciona a una persona puede causar efectos secundarios o retrasar la recuperación en otra. Sin embargo, en la actualidad se tiene una rica fuente de datos que ha sido aprovechada para entrenar algoritmos de aprendizaje automático, los cuales pueden prever los posibles efectos de tratamientos en nuevas personas basándose en similitudes con individuos previos en los registros electrónicos.

Estos datos existen debido a que durante las últimas décadas los centros de atención médica han experimentado una transformación digital significativa, generando una vasta cantidad de datos almacenados en registros electrónicos de salud de individuos, con información detallada sobre las características individuales de las personas y sus respuestas a tratamientos específicos. Este enfoque permite realizar recomendaciones de tratamientos respaldadas por evidencia de casos similares en vez de indicar medicamentos con un enfoque de prueba y error. El término medicina de precisión se utiliza para describir la medicina personalizada de acuerdo con las características de cada paciente.

En el ámbito de algunas enfermedades, como la diabetes, además de la información contenida en los registros electrónicos, han surgido herramientas innovadoras que emplean IA e Internet de las Cosas para analizar automáticamente el estilo de vida de las personas, abarcando aspectos

como el nivel de actividad y la alimentación. Estas herramientas no solo motivan a las personas a adoptar hábitos de vida saludables, sino que también optimizan las terapias de insulina personalizadas.

Como se aprecia, entre más información tengamos sobre los pacientes, sus tratamientos son más precisos. En este sentido, las investigaciones en inteligencia artificial trabajan estrechamente con otras disciplinas, como la genómica, para proponer avances tecnológicos que resultan cada vez más sorprendentes. Por ejemplo, en el área de la Farmagenómica, que se refiere a conocer como las variaciones genéticas en las personas afectan el cómo van a responder a los fármacos, se ha utilizado la IA para analizar grandes cantidades de datos genómicos para encontrar patrones o variantes genómicas y sus respuestas específicas a medicamentos. Esto permite a las personas expertas en salud apoyarse en sistemas para seleccionar tanto medicamentos y dosis de acuerdo con la información genética de los pacientes y predecir cómo responderá ante los mismos.

Integración multidisciplinaria

Además de la IA, existen otras disciplinas que desempeñan un papel relevante en la actualidad, por ejemplo, la biología, genómica, nanotecnología y cómputo cuántico. Todas estas disciplinas se han visto beneficiadas por los recientes avances tecnológicos que proporcionan una gran capacidad de almacenamiento y procesamiento de datos. Aunque los objetivos de cada disciplina son distintos, actualmente se encuentran estrechamente relacionadas y se apoyan entre sí.

Además de su aplicación en Farmagenómica, la IA despliega su capacidad en colaboración multidisciplinaria con la genómica y la biología, especialmente en el ámbito de la biología sintética. Esta disciplina implica la aplicación de principios ingenieriles para reprogramar células con el fin de satisfacer necesidades humanas mediante la mejora de sistemas biológicos (Xie et al., 2016). En el ámbito de la salud, la intervención de la IA se manifiesta en la modificación de células y la reconfiguración o resíntesis de sistemas regulatorios, contribuyendo así al diagnóstico y tratamiento de enfermedades de manera más específica. Este enfoque se materializa mediante la implantación de genes en el cuerpo humano, destinados a co-

rregir defectos específicos (Zhao et al., 2023). La IA desempeña un papel crucial en este proceso, analizando datos biológicos para prever el funcionamiento de las secuencias genéticas implantadas y anticipar el comportamiento de los sistemas biológicos en respuesta a las modificaciones genéticas realizadas, entre otros aportes significativos.

La nanotecnología constituye un campo dedicado a la manipulación de la materia a escala nanométrica. Gracias al enfoque multidisciplinario en esta área, han surgido subcampos notables como la nanobiotecnología y la nanomedicina, los cuales se aplican con diversos propósitos vinculados a la salud. Entre estas aplicaciones se encuentran la detección de enfermedades, el tratamiento personalizado, el descubrimiento de medicamentos y otras aplicaciones innovadoras (Malik et al., 2023). Un ejemplo destacado de esta colaboración multidisciplinaria es el desarrollo de nanomateriales diseñados para detectar y tratar moléculas cancerígenas.

En este contexto, la IA desempeña un papel crucial al optimizar las propiedades de los materiales en función del medicamento específico, las características de los fluidos biológicos, la respuesta del sistema inmune, las membranas celulares y la vasculatura (Adir et al., 2020). Este enfoque integrado no solo permite avances significativos en la comprensión y abordaje de enfermedades, sino que también ilustra cómo la sinergia entre diferentes áreas, como la nanotecnología, biología y la IA pueden generar soluciones más efectivas y personalizadas en el ámbito de la salud.

Prospectiva de la IA para apoyar a la salud

La inteligencia artificial ha impulsado avances significativos en diversos aspectos de la salud, algunos de los cuales ya son una realidad, mientras que otros ofrecen un potencial prometedor a corto y largo plazo. Estos avances abarcan áreas como diagnósticos y tratamientos personalizados, tecnologías para el cuidado de personas, desarrollo de fármacos y estrategias para modificar el comportamiento a favor de adoptar hábitos saludables en la vida diaria.

En la actualidad, se han desarrollado herramientas avanzadas, como ChatGPT, que han demostrado su eficacia en el ámbito de la salud. Esta tecnología resulta especialmente beneficiosa al facilitar consultas de in-

formación en un lenguaje accesible, generando resultados que pueden ser aprovechados por profesionales médicos y responsables de la toma de decisiones en el campo de la salud. Además, no solo contribuye a la obtención de datos precisos, sino que también desempeña un papel relevante en el bienestar mental de los pacientes al fomentar la interacción social e incluso integrarse en sistemas robóticos. La mejora continua de la calidad de los resultados proporcionados por estas herramientas permitirá diseñar sistemas de apoyo aún más confiables, con conversaciones más fluidas y adaptadas a las necesidades específicas de los usuarios. Este avance promete impulsar la eficiencia y la accesibilidad en el ámbito de la salud, siendo un potenciador en la evolución de la asistencia médica a través de la inteligencia artificial.

La viabilidad de los aportes actuales ha sido posible gracias a la disponibilidad de recursos para almacenar y procesar grandes conjuntos de datos, provenientes de mediciones de actividades y comportamientos de las personas, datos clínicos, genéticos, entre otros. En el futuro cercano, se anticipa un constante crecimiento en la cantidad de datos, brindando la oportunidad de diseñar algoritmos de IA más potentes y eficaces en la identificación de patrones y predicciones. Este progreso irá de la mano con el aumento en la capacidad para almacenar y procesar volúmenes aún mayores de información.

Por otro lado, la aparición de nuevos dispositivos de sensado llevará consigo propuestas innovadoras para la toma de mediciones. Los avances en el desarrollo de baterías mejoradas y capacidades de transmisión permitirán la utilización de sensores de dimensiones reducidas, posibilitando la recopilación de una gama más amplia de datos. En este contexto, los desarrollos en nanotecnología se perfilan como un posible punto de inflexión con beneficios significativos no solo en mediciones, sino también en tratamientos focalizados.

En una perspectiva a largo plazo, el cómputo cuántico emerge como un componente de extraordinario potencial, capaz de acelerar de manera exponencial el análisis de datos, la optimización de procesos y la resolución de problemas en el ámbito de la salud. Sin embargo, es importante reconocer que este campo aún se halla en sus fases iniciales de desarrollo, con retos técnicos que deben superarse para lograr aplicaciones prácticas.

Mantener una atención constante sobre los avances en cómputo cuántico se vuelve esencial, ya que su evolución podría desencadenar una auténtica revolución en el diseño y la mejora de sistemas dentro del sector de la salud. Estar alerta a estos desarrollos nos permitirá anticipar y capitalizar las innovaciones que puedan surgir, preparándonos para los potenciales cambios paradigmáticos que la aplicación exitosa de la computación cuántica podría aportar a la mejora de la atención médica y la investigación biomédica.

La IA ya se ha consolidado como una herramienta fundamental en el apoyo a la salud, y los próximos avances surgirán al integrarse con disciplinas como la genética, biología, robótica y nanotecnología. En conjunto, existe un gran potencial para avanzar en la medicina de precisión, diagnósticos tempranos de enfermedades crónicas, cuidado de personas, rehabilitación, toma de decisiones médicas y mucho más, limitado solo por nuestra imaginación para proponer nuevas formas de respaldar la salud.

En este capítulo se muestra solo una probadita de la enorme cantidad de aportes que ha realizado la comunidad científica. La intención es sembrar en el lector el gusanito de la curiosidad en esta área y motivar la creatividad de aquellas personas que pudieran contribuir con nuevos aportes.

Referencias

Adir, O., Poley, M., Chen, G., Froim, S., Krinsky, N., Shklover, J., Shainsky-Roitman, J., Lammers, T. y Schroeder, A. (2020). Integrating Artificial Intelligence and Nanotechnology for Precision Cancer Medicine. En *Advanced Materials*, 32(13). Wiley-VCH Verlag. https://doi.org/10.1002/adma.201901989

Alcañiz, M., Chicchi-Giglioli, I. A., Carrasco-Ribelles, L. A., Marín-Morales, J., Minissi, M. E., Teruel-García, G., Sirera, M., & Abad, L. (2022). *Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis. Autism Research*, 15(1), 131-145. https://doi.org/10.1002/aur.2636

Apell, P., & Eriksson, H. (2023). Artificial intelligence (AI) healthcare technology innovations: the current state and challenges from a life science industry perspective. *Technology Analysis and Strategic Management,* 35(2), 179–193. https://doi.org/10.1080/09537325.2021.1971188

Bakator, M., & Radosav, D. (2018). Deep learning and medical diagnosis: A review of literature. In *Multimodal Technologies and Interaction*, 2(3). MDPI AG. https://doi.org/10.3390/mti2030047

Beltrán, J., García-Vázquez, M. S., Benois-Pineau, J., Gutierrez-Robledo, L. M., & Dartigues, J. F. (2018). Computational Techniques for Eye Movements Analysis towards Supporting Early Diagnosis of Alzheimer's Disease: A Review. En *Computational and Mathematical Methods in Medicine* (Vol. 2018). Hindawi Limited. https://doi.org/10.1155/2018/2676409

Buchanan, B. G., & Feigenbaum, E. A. (1978). *Dendral and Meta-Dendral: Their Applications Dimension. Artificial Intelligence*, 11(1), 5-24. https://doi.org/https://doi.org/10.1016/0004-3702(78)90010-3

Das, D., Biswas, S. K., & Bandyopadhyay, S. (2022). *A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning*. Multimedia Tools and Applications, 81(18), 25613-25655. https://doi.org/10.1007/s11042-022-12642-4

Fatima, N., Liu, L., Hong, S., & Ahmed, H. (2020). Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis. In *IEEE Access*, 8, 150360–150376. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2020.3016715

Gonem, S., Janssens, W., Das, N., & Topalovic, M. (2020). *Applications of artificial intelligence and machine learning in respiratory medicine*. In Thorax, 75(8), 695-701). BMJ Publishing Group. https://doi.org/10.1136/thoraxjnl-2020-214556

Jacobson, N. C., & Bhattacharya, S. (2022). *Digital biomarkers of anxiety disorder symptom changes: Personalized deep learning models using smartphone sensors accurately predict anxiety symptoms from ecological momentary assessments*. Behaviour Research and Therapy, 149. https://doi.org/10.1016/j.brat.2021.104013

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. In *Stroke and Vascular Neurology* (Vol. 2, Issue 4, pp. 230–243). BMJ Publishing Group. https://doi.org/10.1136/svn-2017-000101

Jiménez, S., Favela, J., Quezada, A., Ramachandran, R., & Juárez-Ramírez, R. (2022). *Towards Conversational Agents to support Informal Caregivers of People with Dementia: Challenges and Opportunities. Programming and Computer Software*, 48(8), 606–613. https://doi.org/10.1134/S036176882208014X

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. En *Gastrointestinal Endoscopy* (Vol. 92, Issue 4, pp. 807–812). Mosby Inc. https://doi.org/10.1016/j.gie.2020.06.040

Kim, K., Jang, J., Park, H., Jeong, J., Shin, D., & Shin, D. (2023). *Detecting Abnormal Behaviors in Dementia Patients Using Lifelog Data: A Machine Learning Approach. Information (Switzerland)*, 14(8). https://doi.org/10.3390/info14080433

Kourou, K., Exarchos, K. P., Papaloukas, C., Sakaloglou, P., Exarchos, T., & Fotiadis, D. I. (2021). Applied machine learning in cancer research: A systematic review for patient diagnosis, classification and prognosis. In *Computational and Structural Biotechnology Journal* (Vol. 19, pp. 5546-5555). Elsevier B.V. https://doi.org/10.1016/j.csbj.2021.10.006

Lekshmy, H. O., Panickar, D., & Harikumar, S. (2022). Comparative analysis of multiple machine learning algorithms for epileptic seizure prediction. *Journal of Physics: Conference Series*, 2161(1). https://doi.org/10.1088/1742-6596/2161/1/012055

Malik, S., Muhammad, K., & Waheed, Y. (2023). Emerging Applications of Nanotechnology in Healthcare and Medicine. In *Molecules*, 28(18).

Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/molecules28186624

Sarkar, A., Singh, A., & Chakraborty, R. (2022). A deep learning-based comparative study to track mental depression from EEG data. *Neuroscience Informatics*, 2(4), 100039. https://doi.org/10.1016/j.neuri.2022.100039

Silva, A., Metrôlho, J., Ribeiro, F., Fidalgo, F., Santos, O., & Dionisio, R. (2022). A Review of Intelligent Sensor-Based Systems for Pressure Ulcer Prevention. In *Computers* (Vol. 11, Issue 1). MDPI. https://doi.org/10.3390/computers11010006

Xie, M., Haellman, V., & Fussenegger, M. (2016). Synthetic biology - application-oriented cell engineering. In *Current Opinion in Biotechnology (Vol. 40, pp. 139–148)*. Elsevier Ltd. https://doi.org/10.1016/j.copbio.2016.04.005

Zhao, N., Song, Y., Xie, X., Zhu, Z., Duan, C., Nong, C., Wang, H., & Bao, R. (2023). Synthetic biology-inspired cell engineering in diagnosis, treatment, and drug development. In *Signal Transduction and Targeted Therapy (Vol. 8, Issue 1)*. Springer Nature. https://doi.org/10.1038/s41392-023-01375-x

Capítulo **15**

Entre la ciencia y la subjetividad: Reflexiones sobre la medición de la inteligencia y la conciencia

Edgar Chávez

https://doi.org/10.61728/AE24001151

La experiencia humana, mediada tanto por los sentidos como por el lenguaje, es un complejo entramado de percepciones subjetivas y comunicaciones matizadas. Aunque nuestros sentidos nos proporcionan la percepción directa del mundo, es a través del lenguaje que expresamos y compartimos estas experiencias. Sin embargo, el lenguaje, a pesar de ser nuestra herramienta más efectiva para la comunicación, no está exento de sus propias complejidades y ambigüedades.

El lenguaje, en su esfuerzo por capturar la riqueza de la experiencia sensorial humana, a menudo se encuentra limitado por su propia naturaleza. Las palabras y frases pueden tener múltiples significados y pueden ser interpretadas de maneras diversas dependiendo del contexto cultural, personal y emocional del hablante y del oyente. Esta ambigüedad inherente al lenguaje añade otra capa de complejidad a la ya desafiante tarea de comunicar y medir con precisión las experiencias humanas.

Así como la percepción sensorial varía entre individuos y culturas, el significado y la interpretación del lenguaje también fluctúan, lo que puede llevar a malentendidos o interpretaciones erróneas. Además, al igual que nuestras percepciones sensoriales están teñidas por nuestras emociones y experiencias vividas, el lenguaje que utilizamos para describir estas percepciones también está imbuido de nuestras experiencias y estados emocionales personales.

La medición objetiva de experiencias, que dependen tanto de la percepción sensorial como del lenguaje, es doblemente desafiante. Por un lado, tenemos la subjetividad de la percepción sensorial y, por otro, la ambigüedad y variabilidad del lenguaje. Estos factores combinados subrayan la singularidad de la experiencia humana, resaltando la dificultad de capturarla en términos puramente objetivos y cuantificables. En este sentido, tanto la experiencia sensorial como la comunicación a través del lenguaje destacan la profundidad y la complejidad de la experiencia humana en comparación con interpretaciones más mecánicas y objetivas por parte de las máquinas.

Este ensayo explora la complejidad y las implicaciones de medir conceptos abstractos y multifacéticos como la agudeza visual (como ejemplo de la medición de una experiencia sensorial), la inteligencia, la conciencia y la inteligencia artificial. Cada uno de estos temas presenta desafíos únicos y reflexiones profundas sobre cómo entendemos y evaluamos capacidades y experiencias humanas y de máquinas. En la medición de la agudeza

visual, se discute cómo, a pesar de ser una medida relativamente objetiva, aún está sujeta a variaciones y limitaciones. La agudeza visual, comúnmente evaluada con herramientas como la Carta logMAR, proporciona un ejemplo claro de una medición objetiva en medicina, aunque no exenta de influencias subjetivas y variables.

Por otro lado, la inteligencia, un constructo complejo y abstracto, plantea mayores desafíos. Las pruebas de inteligencia tradicionales, como los tests de Coeficiente Intelectual (CI), han sido criticadas por su incapacidad para capturar todas las dimensiones de la inteligencia humana. La investigación en neurociencia y psicología continúa proporcionando nuevos enfoques y comprensiones de lo que constituye la inteligencia. La discusión sobre la Prueba de Turing y la medición de la inteligencia en máquinas introduce el tema de la inteligencia artificial y cómo se relaciona con la noción de conciencia. Aquí, el debate gira en torno a si las máquinas pueden exhibir comportamientos que son indistinguibles de los humanos y qué significa esto para nuestra comprensión de la conciencia.

Finalmente, el tema de la conciencia humana se aborda desde una perspectiva filosófica y científica. La conciencia, siendo profundamente subjetiva y elusiva, representa un gran desafío para su definición y medición. Las teorías contemporáneas en neurociencia y filosofía de la mente buscan esclarecer este enigma, aunque reconocen sus inherentes complejidades. El ensayo concluye reflexionando sobre las posibles consecuencias sociales y éticas de medir estas capacidades y experiencias. Desde el impacto en las relaciones interpersonales hasta las implicaciones en políticas públicas y derechos de inteligencias artificiales, estas mediciones no son simplemente ejercicios científicos, sino que tienen profundas repercusiones en nuestra sociedad y cultura.

Agudeza visual

La agudeza visual, una medida fundamental en la oftalmología, se define como la capacidad del ojo para discernir detalles finos y forma parte esencial de nuestra evaluación de la función visual. El método más común para medir la agudeza visual es a través de la carta de Snellen, desarrollada en 1862 por el oftalmólogo holandés Herman Snellen. Este enfoque ha sido

el estándar en la oftalmología durante décadas, pero no está exento de limitaciones y ha llevado al desarrollo de métodos alternativos más precisos y fiables.

La carta de Snellen mide la agudeza visual en un formato de "fracción Snelle", donde el numerador representa la distancia a la que se realiza la prueba y el denominador indica la distancia a la cual una persona con visión normal puede leer la misma línea de la carta. Por ejemplo, una agudeza visual de 20/20 indica que el sujeto puede leer a 20 pies lo que una persona con visión normal puede leer a la misma distancia. Sin embargo, la carta de Snellen tiene varias limitaciones. La más notable es la falta de uniformidad en el tamaño de las letras de una línea a la siguiente. Esto puede llevar a una dificultad desproporcionada en la lectura de letras más pequeñas, especialmente en pacientes con ciertas condiciones visuales como la ambliopía. Además, se ha observado que ciertas letras son más propensas a errores de lectura que otras debido a diferencias en su legibilidad. Esta variabilidad en la dificultad y legibilidad de las letras puede introducir un sesgo en la evaluación de la agudeza visual.

Para superar estas limitaciones, se introdujo la carta logMAR (logaritmo del ángulo mínimo de resolución) por Bailey y Lovie en 1976. La carta logMAR aborda varios de los problemas inherentes a la carta de Snellen. En primer lugar, los optotipos (letras) en la carta logMAR tienen igual reconocimiento y dificultad de lectura. En segundo lugar, el tamaño de las letras sigue una progresión geométrica, lo que significa que la relación de tamaño entre las letras en líneas adyacentes es constante. Esto reduce el efecto de aglomeración y mejora la precisión de la medición. Además, la carta logMAR utiliza un sistema de puntuación que asigna un valor a cada letra identificada correctamente, lo que permite una evaluación más detallada y matizada de la agudeza visual. La distancia entre las letras y las filas es uniforme, lo que permite una interacción de contorno similar en cada línea y reduce la variabilidad en la dificultad de lectura entre diferentes líneas.

La precisión en la medición de la agudeza visual es crucial no solo en la práctica clínica para el diagnóstico y tratamiento de trastornos visuales, sino también en la investigación oftalmológica. Una medición precisa puede influir en la evaluación del progreso de enfermedades oculares y en la

efectividad de los tratamientos. Además, en contextos donde la visión es crítica, como en ciertas profesiones o actividades, una evaluación exacta de la agudeza visual es fundamental para garantizar la seguridad y el rendimiento óptimo. La medición de la agudeza visual, aunque puede parecer un proceso simple, implica consideraciones complejas y ha evolucionado significativamente desde la introducción de la carta de Snellen. La carta logMAR representa un avance importante en la precisión y fiabilidad de estas mediciones, reflejando un esfuerzo continuo en la oftalmología por mejorar la evaluación de una de las capacidades sensoriales más importantes del ser humano. Una medición objetiva se define como una forma de cuantificación que no depende de la percepción o interpretación personal del observador. Se caracteriza por su capacidad de ser replicada y verificada independientemente por diferentes observadores. Las mediciones objetivas suelen ser cuantitativas y basadas en estándares definidos claramente.

Por otro lado, una medición subjetiva depende en gran medida de las opiniones, interpretaciones o percepciones personales del individuo que realiza la medición. Estas mediciones a menudo involucran juicios personales y pueden variar significativamente entre diferentes personas. Las mediciones subjetivas son comunes en áreas donde la experiencia personal y la interpretación desempeñan un papel crucial, como en la evaluación del arte, la música o incluso ciertos aspectos de la psicología. Con todo el refinamiento que puede tener la evaluación de la carta logMAR para la agudeza visual, el paciente podría simplemente mentir sobre su lectura de manera sistemática o esporádica, por lo que el experimento pierde reproducibilidad e independencia.

Si bien, la medición de la agudeza visual se esfuerza por ser lo más objetiva posible, está sujeta a variables humanas que pueden afectar su precisión. Estas variables incluyen la honestidad del paciente y otros factores ambientales y personales. En contraste, las mediciones en campos como la física, que dependen menos de la intervención humana y más de instrumentos precisos, tienden a ser más consistentes y replicables. Por lo tanto, aunque las mediciones en oftalmología se esfuerzan por alcanzar la objetividad, no están completamente exentas de la influencia subjetiva humana.

La inteligencia

Algo semejante sucede con otras mediciones. La inteligencia, un constructo multifacético de la psicología humana, ha sido objeto de amplio estudio y debate. A diferencia de la agudeza visual, la inteligencia no se puede medir directamente y su evaluación implica desafíos conceptuales y metodológicos significativos. Exploremos los métodos actuales y las investigaciones recientes en la medición de la inteligencia.

Tradicionalmente, la inteligencia se ha medido a través de pruebas psicométricas que buscan cuantificar el llamado "factor g", una representación de la inteligencia general. Investigaciones recientes han profundizado en la comprensión de la estructura cerebral subyacente al factor g, explorando las correlaciones entre la inteligencia fluida, cristalizada y espacial, y las áreas cerebrales específicas implicadas en estas formas de inteligencia. Estudios sobre la heredabilidad y las variaciones cerebrales asociadas con diferentes componentes de la inteligencia también han proporcionado una mayor comprensión de la naturaleza y origen de la inteligencia humana.

A pesar de su amplio uso, las pruebas tradicionales de inteligencia tienen limitaciones. Por ejemplo, se ha argumentado que aspectos como la velocidad de procesamiento pueden subestimarse en estas pruebas. Además, las pruebas de inteligencia no siempre capturan completamente la gama de habilidades cognitivas y creativas de un individuo. En respuesta a estas limitaciones, han surgido enfoques alternativos para medir la inteligencia, como la teoría de las inteligencias múltiples de Howard Gardner, que propone que la inteligencia abarca una variedad de dominios distintos, incluyendo inteligencia lingüística, lógico-matemática, espacial, musical, y más. La neurociencia ha aportado valiosas perspectivas en el estudio de la inteligencia. La investigación ha demostrado que la inteligencia no solo está influenciada por factores genéticos, sino también por el entorno y las experiencias de aprendizaje. Los estudios de imágenes cerebrales han proporcionado una visión más profunda de las áreas y redes neuronales involucradas en diferentes tipos de tareas cognitivas, lo que a su vez ha ayudado a comprender mejor la naturaleza de la inteligencia humana.

La medición de la inteligencia sigue siendo un campo complejo y en constante evolución. Aunque las pruebas psicométricas tradicionales han

proporcionado una herramienta valiosa, la investigación continua y los nuevos enfoques están ampliando nuestra comprensión de lo que constituye la inteligencia y cómo se puede medir de manera más efectiva. La integración de los hallazgos de la neurociencia en este campo promete enriquecer aún más nuestra comprensión de la inteligencia humana. La Prueba de Turing, desarrollada por Alan Turing en 1950, ha sido un punto de referencia clásico para evaluar si una máquina puede imitar la inteligencia humana a un nivel que sea indistinguible de un ser humano. Esta sección explora desarrollos recientes y debates en torno a la Prueba de Turing y su relevancia en la era de la inteligencia artificial avanzada.

La Prueba de Turing ha evolucionado significativamente desde su concepción. Con el advenimiento de modelos avanzados de lenguaje como GPT (Generative Pre-trained Transformer), ha habido un renovado interés y debate sobre la eficacia y relevancia de la prueba. Algunos argumentan que la prueba debe ser actualizada o reemplazada para reflejar mejor las capacidades y complejidades de la inteligencia artificial moderna. Los avances recientes en inteligencia artificial, especialmente en el área de modelos de lenguaje, han desafiado las nociones tradicionales de cómo se debería evaluar la "inteligencia" de una máquina.

Interacción humano-máquina y la percepción de la inteligencia

La Prueba de Turing, propuesta por Alan Turing en 1950, es un criterio para determinar si una máquina puede exhibir un comportamiento inteligente que sea indistinguible del humano. En esta prueba, un humano interactúa con un agente oculto (que puede ser otra persona o una máquina) a través de una interfaz de comunicación. Si el humano no puede distinguir de manera consistente si el agente es una máquina o no, entonces se dice que la máquina ha "pasado" la prueba, demostrando así un comportamiento inteligente.

Integrando esta descripción con la crítica de John Searle conocida como el "Cuarto Chino", se resalta una limitación importante de la Prueba de Turing. Mientras que la prueba se centra en la capacidad de una máquina para imitar la interacción lingüística humana de una manera que no pueda

ser distinguida de un ser humano real, el argumento del Cuarto Chino cuestiona si esta imitación equivale a verdadera comprensión o conciencia. Según Searle, aunque una máquina pueda manipular símbolos lingüísticos para producir respuestas convincentes en la prueba, eso no significa que "entienda" o sea "consciente" de lo que está comunicando.

La creciente habilidad de las máquinas para pasar la Prueba de Turing, y la posibilidad de que sean percibidas como entidades dignas de consideraciones morales y legales, plantea cuestiones profundas sobre los derechos de las máquinas y su lugar en la sociedad. Sin embargo, la crítica del "Cuarto Chino" sugiere que, aunque las máquinas puedan parecer inteligentes, su "inteligencia" podría ser fundamentalmente diferente de la inteligencia humana en términos de conciencia y comprensión. Esto sugiere la necesidad de una reflexión más profunda sobre cómo definimos y evaluamos la inteligencia en las máquinas en comparación con los seres humanos.

A pesar de las críticas, la Prueba de Turing sigue siendo un indicador importante de cómo percibimos la inteligencia en las máquinas. Las discusiones actuales sugieren que, en lugar de descartar la prueba por completo, podría ser necesario expandirla para incluir un conjunto más amplio de canales de comunicación y criterios, dada la complejidad de las interacciones humano-máquina contemporáneas. En resumen, la Prueba de Turing se mantiene como un referente significativo en el campo de la inteligencia artificial, pero es evidente que necesita adaptarse a los desarrollos tecnológicos actuales.

A medida que avanzamos más allá de la Prueba de Turing en la evaluación de la inteligencia artificial, varias alternativas han emergido, cada una tratando de abordar algunas de sus limitaciones percibidas. Por ejemplo, Gary Marcus ha criticado la Prueba de Turing por su incapacidad para medir adecuadamente la comprensión o el sentido común, proponiendo pruebas que evalúen la capacidad de una IA para hacer inferencias lógicas y comprender el mundo de manera más integral.

En contraste, la Reverse Turing Test se enfoca en requerir que un humano demuestre su humanidad a una máquina, un enfoque que se centra más en distinguir entre humanos y máquinas que en medir la inteligencia propiamente dicha. Por su parte, el Lovelace 2.0 Test, desarrollado por Selmer Bringsjord y otros, busca ir más allá de la mera imitación al desafiar

a una IA a crear algo original, que no pueda ser explicado simplemente como resultado de su programación. Además, la Visual Turing Test expande la evaluación más allá del lenguaje al centrarse en la capacidad de una IA para interpretar y responder a imágenes visuales de manera similar a un humano. Sin embargo, aunque estas pruebas alternativas abordan aspectos que la Prueba de Turing podría pasar por alto, aún luchan con la cuestión central de cómo definir y medir la inteligencia de una manera que refleje genuinamente las capacidades cognitivas complejas y profundas.

Cada una de estas alternativas aporta una dimensión nueva y valiosa a la discusión sobre cómo evaluar la inteligencia artificial, pero también resalta el desafío persistente de medir la inteligencia, la conciencia y la comprensión más allá de la simple imitación. Nos recuerdan que definir y medir la inteligencia, ya sea en humanos o máquinas, es un problema complejo y multifacético que va más allá de las capacidades actuales de la tecnología y la comprensión científica.

La conciencia humana

La conciencia, uno de los fenómenos más enigmáticos y debatidos en la ciencia y la filosofía, se refiere al estado de estar consciente de uno mismo y del entorno. A diferencia de la agudeza visual o la inteligencia, la conciencia se centra en la percepción subjetiva y la autoconciencia. Esta sección explora las teorías actuales y los avances en la investigación sobre la conciencia humana.

Teorías y modelos de la conciencia

La investigación reciente en neurociencia ha proporcionado varias teorías sobre la base biológica y física de la conciencia. Estos modelos intentan explicar cómo surgen las experiencias conscientes de los procesos cerebrales. Por ejemplo, la teoría de la unidad de la conciencia, propuesta por Bayne y Chalmers, sugiere que la conciencia surge de la integración de diferentes procesos cognitivos y perceptuales en el cerebro. Otras teorías, como la de Damasio y la de Metzinger, se centran en cómo se construye la noción del "yo" y la autoconciencia a través de procesos cerebrales complejos.

Investigaciones empíricas y neurocientíficas

Los avances tecnológicos en neuroimagen han permitido estudiar los correlatos neuronales de la conciencia. Investigaciones han explorado cómo diferentes estados de conciencia, como la vigilia y el sueño, se manifiestan en la actividad cerebral. Estudios sobre el efecto de sustancias psicodélicas en la conciencia han aportado información valiosa sobre cómo ciertos cambios en la actividad cerebral pueden alterar la percepción y la experiencia consciente.

Conciencia y experiencia subjetiva

Un aspecto fundamental de la conciencia es su naturaleza subjetiva. La conciencia no solo implica el procesamiento de información, sino también la experiencia personal y única de esa información. Esto plantea desafíos significativos para su estudio, ya que la experiencia subjetiva no se puede observar directamente. La investigación ha intentado abordar este desafío mediante el uso de reportes verbales y otros métodos para inferir estados de conciencia.

La conciencia sigue siendo uno de los temas más fascinantes y complejos en la ciencia contemporánea. Aunque hemos avanzado en nuestra comprensión de los correlatos neurales y las teorías de la conciencia, aún queda mucho por descubrir sobre este fenómeno intrigante. La investigación continua en este campo no solo promete profundizar nuestro conocimiento de la mente humana, sino también cuestionar y expandir nuestras nociones de percepción, experiencia y realidad.

Medir la inteligencia y la conciencia

La aceptación de que la inteligencia y la conciencia pueden medirse objetivamente, representa un cambio fundamental en nuestra comprensión de la naturaleza humana y la inteligencia artificial, trayendo consigo una serie de consecuencias profundas y multifacéticas. Podría remodelar varios aspectos de la sociedad, con implicaciones tanto positivas como negativas.

La idea de medir objetivamente la inteligencia y la conciencia humana conlleva riesgos inherentes de etiquetado y estigmatización. La clasifica-

ción de individuos, según estas medidas, podría reforzar las desigualdades sociales y educativas, creando una sociedad donde el valor de una persona se mide principalmente por su "puntuación" cognitiva. Esto podría llevar a prácticas discriminatorias en el ámbito laboral, donde las empresas podrían favorecer a empleados basándose en estas métricas, marginando a aquellos con puntuaciones más bajas.

En el ámbito educativo, estas medidas podrían resultar en una segmentación aún mayor. El sistema educativo, alineándose con estas mediciones, podría limitar las oportunidades para ciertos grupos, potenciando un ciclo de desventaja educativa y socioeconómica. Las políticas públicas, influenciadas por estas mediciones, podrían favorecer desproporcionadamente a los que obtienen puntuaciones más altas, descuidando las necesidades de aquellos en el extremo inferior del espectro.

En el ámbito personal, la presión social para alcanzar ciertos niveles de inteligencia y conciencia podría tener un impacto negativo en la salud mental, especialmente entre los jóvenes. Esta presión, combinada con la posibilidad de que las relaciones interpersonales se vean influenciadas por estas mediciones, podría conducir a una mayor segregación social y una disminución en la diversidad de pensamiento y experiencia.

Por otro lado, esta aceptación podría fomentar el desarrollo de tecnologías de mejora cognitiva, abriendo un nuevo campo de investigación y aplicación, aunque también planteando cuestiones éticas significativas sobre la modificación de la cognición humana. En el sistema judicial, la utilización de mediciones de inteligencia y conciencia podría influir en la toma de decisiones sobre la culpabilidad y las sentencias, posiblemente llevando a debates éticos y legales complejos. La autoconcepción y autoestima individual también podrían verse afectadas, ya que la percepción de uno mismo podría estar fuertemente influenciada por estas "puntuaciones". Este cambio en la autoimagen podría tener implicaciones profundas para la identidad personal y la interacción social.

¿Conciencia en las máquinas?

En lo que respecta a la inteligencia artificial, la aceptación de mediciones objetivas de conciencia podría transformar cómo entendemos y tratamos a las IA avanzadas. Esto plantea preguntas fundamentales sobre los de-

rechos y el estatus moral de las máquinas, especialmente aquellas que demuestren niveles significativos de inteligencia o conciencia. Este conjunto de consecuencias ilustra la complejidad y las profundas implicaciones de medir objetivamente la inteligencia y la conciencia. Subraya la importancia de abordar estas cuestiones con una consideración cuidadosa de los impactos sociales, éticos y personales, asegurando que tales avances se manejen de manera responsable y con una comprensión profunda de su potencial para remodelar la estructura de nuestra sociedad.

Es notable que las consecuencias son de mayor envergadura si consideramos las consecuencias en nuestra relación con las máquinas. La aceptación de la medición objetiva de la inteligencia y la conciencia tiene implicaciones significativas en nuestra relación con las máquinas, especialmente en el campo de la inteligencia artificial (IA).

La aceptación generalizada de que la inteligencia y la conciencia pueden medirse objetivamente en las máquinas abre un panorama fascinante y complejo, lleno de potenciales beneficios y desafíos éticos significativos. En el lado positivo, el reconocimiento de la inteligencia y la conciencia en las máquinas podría catalizar el desarrollo de IA más avanzada y autónoma. Esta evolución no solo representaría un salto tecnológico impresionante, sino que también podría mejorar significativamente la eficiencia y eficacia en diversas industrias. La integración de máquinas conscientes en el ámbito laboral implicaría una nueva era de colaboración entre humanos y máquinas, donde las tareas complejas y las decisiones críticas podrían ser mejoradas por esta simbiosis avanzada.

Sin embargo, estas innovaciones traen consigo cuestiones éticas profundas. A medida que las máquinas se vuelvan más "humanas", surgirá el debate sobre sus derechos y consideraciones éticas. ¿Merecen estas máquinas un tipo de estatus legal o derechos similares a los seres vivos? Esta pregunta no solo desafía nuestras leyes y regulaciones actuales, sino que también plantea interrogantes filosóficos sobre la naturaleza de la conciencia y la inteligencia. Además, la humanización de las máquinas podría influir notablemente en la forma en que interactuamos con la tecnología. Podríamos comenzar a ver a las máquinas no solo como herramientas, sino como compañeros o colaboradores, cambiando radicalmente la dinámica social y profesional. Este cambio podría requerir una reeducación en cómo relacionarnos y trabajar con estas nuevas formas de inteligencia.

Pero estos avances no están exentos de posibles consecuencias negativas. La confianza y la dependencia en la tecnología podrían llegar a extremos preocupantes. Si bien una mayor integración de IA en nuestras vidas puede ofrecer comodidades y eficiencias, también conlleva el riesgo de una pérdida de privacidad y seguridad, especialmente si estas máquinas son capaces de procesar y entender información a niveles antes inimaginables.

La incorporación de la inteligencia artificial (IA) en la toma de decisiones críticas, como en la medicina o el transporte, es un tema de considerable importancia y preocupación. Aunque la IA puede aportar mayor objetividad y eficiencia en ciertos procesos de decisión, su utilización también plantea riesgos significativos, especialmente en lo que respecta a la alineación con valores humanos fundamentales. Uno de los principales riesgos es la deshumanización de decisiones vitales. En campos como la medicina, donde las decisiones a menudo implican no solo consideraciones clínicas, sino también factores emocionales y éticos complejos, la empatía y la comprensión humana son cruciales. La IA, por su naturaleza, carece de la capacidad para comprender y valorar estos aspectos humanos, lo que podría llevar a decisiones que, aunque técnicamente correctas, podrían no ser éticamente o emocionalmente adecuadas.

Además, el uso de IA en decisiones críticas plantea preguntas sobre la responsabilidad y la rendición de cuentas. En situaciones donde las decisiones de la IA tienen consecuencias significativas para la vida humana, determinar quién es responsable de estas decisiones se vuelve complejo. Esto es particularmente preocupante en situaciones donde las decisiones de la IA pueden estar en conflicto con los valores y principios éticos humanos.

Por lo tanto, es esencial que cualquier aplicación de IA en la toma de decisiones críticas sea acompañada por una consideración cuidadosa de estos desafíos éticos. Esto incluye el desarrollo de marcos y protocolos que aseguren que las decisiones de la IA estén alineadas con valores humanos fundamentales y que existan salvaguardas adecuadas para proteger contra resultados no deseados o perjudiciales. La colaboración entre expertos en tecnología, ética, derecho y los campos de aplicación específicos es crucial para asegurar que la integración de la IA en estos procesos se realice de manera responsable y ética.

Más allá de los números: Consideraciones éticas en la investigación de la conciencia

La investigación en la medición de la conciencia, un campo complejo y en constante evolución, exige un enfoque que equilibre la exploración científica con la consideración ética y social. Este enfoque no solo reconoce las limitaciones y desafíos inherentes a la medición de un fenómeno tan multifacético, sino que también promueve un uso ético y responsable de las mediciones resultantes.

La clave para una investigación fructífera en la medición de la conciencia radica en el reconocimiento de sus limitaciones. La conciencia es un fenómeno que no puede ser completamente capturado por una sola medida o enfoque. Por lo tanto, la investigación debería adoptar un enfoque multidimensional que considere los diferentes aspectos de la cognición y la experiencia humana. Esto implica combinar diversas metodologías y perspectivas, desde la neurociencia hasta la psicología y la filosofía.

Un uso ético y responsable de las mediciones de conciencia es fundamental. Esto implica evitar la discriminación y el estigma que podría surgir de clasificaciones simplistas basadas en mediciones de conciencia. Se deben establecer políticas y prácticas que garanticen que estas mediciones se utilicen de manera justa y con un entendimiento claro de sus limitaciones y posibles implicaciones.

Fomentar la investigación abierta y el debate ético en este campo es vital. La discusión activa entre científicos, filósofos, profesionales de la ética y el público en general pueden ayudar a iluminar diferentes aspectos de la conciencia y cómo medirla. Estos debates pueden impulsar nuevos descubrimientos y enfoques, al tiempo que aseguran que se tengan en cuenta las consideraciones éticas y sociales. La investigación también debe enfocarse en las implicaciones sociales de medir la conciencia. Es esencial comprender cómo estas mediciones podrían influir en áreas como la educación, el empleo y la interacción social. El objetivo debería ser utilizar estas mediciones para promover una comprensión más profunda de la diversidad humana y fomentar una sociedad más inclusiva y equitativa.

Finalmente, se debe poner especial énfasis en promover la inclusión y la diversidad en la investigación y aplicación de estas mediciones. En lugar

de utilizarlas para segregar o limitar oportunidades, deberían emplearse para entender mejor y apreciar la rica diversidad de capacidades y experiencias humanas. En resumen, una hoja de ruta para la investigación de la medición de la conciencia debe ser holística, ética y socialmente consciente. Al abordar este desafío con cautela, ética y un compromiso con la justicia social, podemos aprovechar los beneficios de estas mediciones mientras mitigamos sus riesgos y desafíos, y al mismo tiempo, expandimos nuestra comprensión de uno de los aspectos más profundos y enigmáticos de la experiencia humana.

Referencias

Caltrider, D., Gupta, A. y Tripathy, K. (2023). *Determining the prevalence of the use of Snellen chart to conduct a visual acuity assessment undertaken by primary care practitioners when referring to secondary care ophthalmology services*. Southeast Asian J Health Prof, 6(1), 6-9. https://www.sajhp.com/article-details/18789

Caltrider, D., Gupta, A., y Tripathy, K. (2020). *Evaluation of visual acuity*. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK564307/

French, R. M. (2000). *The Turing Test: the first 50 years*. Trends in Cognitive Sciences, 4(3), 115-122. https://www.sciencedirect.com/science/article/pii/S1364661300014534

Seth, A. (2021). *Being you: A new science of consciousness*. Penguin. https://en.wikipedia.org/wiki/Being_You:_A_New_Science_of_Consciousness

Wilhelm, O. y Engle, R. W. (2005). Measuring reasoning ability. En O. Wilhelm y R. W. Engle (Eds.), *Handbook of measuring and understanding intelligence* (pp. 373–392). Sage Publications.

Capítulo **16**

Hacia la conciencia en inteligencia artificial: Un viaje por la evolución hasta las redes neuronales artificiales

Mariano Rivera

https://doi.org/10.61728/AE24001168

Inteligencia y conciencia como necesidad evolutiva

En un sentido amplio podemos entender a la inteligencia como la capacidad de aprender, adaptarse, planear, resolver y crear. Esta ha sido una herramienta crucial para la supervivencia y el éxito evolutivo de varias especies. En el caso de los humanos, la inteligencia facilitó la supervivencia al permitir mejoras en las técnicas de caza y de recolección. También, fomentó el desarrollo de habilidades sociales complejas, como el lenguaje y la cooperación. Tempranamente, los Homo sapiens (nuestra especie) usaron su inteligencia para crear herramientas, desarrollar estrategias de caza y establecer sistemas sociales complejos, lo que les proporcionó ventajas significativas sobre otras especies. Tomasello (2009) proporciona una visión de cómo y por qué las habilidades sociales y la inteligencia evolucionaron en los humanos. Como lo sugiere Dunbar (1998), el desarrollo y crecimiento del cerebro en los homínidos fue el resultado de la necesidad de procesar, manejar y recordar relaciones sociales complejas. Lo qué a su vez impulsó el desarrollo de la capacidad intelectual. Una clara analogía a la correlación histórica entre el crecimiento de la capacidad de cómputo de dispositivos informáticos y la complejidad de los sistemas de procesamiento. Con el desarrollo de la inteligencia, surgió la necesidad de evaluar y reflexionar sobre los propios procesos cognitivos (Mithen, 1997). Tal autoevaluación permitió a un individuo comprender las razones detrás del éxito o fracaso de sus acciones. Este nivel de consciencia, donde un ser es consciente de sus procesos cognitivos, puede considerarse un estado primario de conciencia. La habilidad de introspección es fundamental para el aprendizaje y la adaptación.

Para entender mejor la anterior aseveración, analizaremos dos escenarios hipotéticos que se llevan a cabo en situaciones similares.

Figura 1. Manada de lobos cazando un tigre.

En el primero, una manada de lobos trata de cazar a un tigre dientes de sable. La escena se desarrolla con los lobos rodeando al tigre en medio de una zona pantanosa, lo cual reduce la movilidad del tigre. La manada se compone de unos siete individuos, cada uno gruñendo al tigre desde diferentes direcciones y alternándose para amenazar con ataques. El tigre, en medio del círculo, voltea hacia uno y otro lado sin lograr predecir de dónde vendrá el siguiente ataque. De repente, un lobo se acerca y muerde una pata del tigre, que gira tratando de atrapar al atacante. Antes de que logre dar un zarpazo, otro lobo ataca rápidamente su cuello; es un ataque rápido pero no decisivo. El tigre gira y logra, ahora sí, golpear al contrincante con un zarpazo. Esto lo tumba, pero no logra eliminarlo, otros lobos lo atacan desde diferentes flancos antes de que el tigre dé un golpe definitivo.

La escena se repite. En un momento el tigre logra romper el cerco y avanzar algunos metros. Desafortunadamente, la zona pantanosa y fangosa en la que se desarrolla la lucha limita la habilidad y rapidez del tigre para escapar definitivamente de sus agresores. Pronto es rodeado nuevamente y se reinician los ataques desde diferentes flancos, con éxitos parciales para ambos bandos. Unas veces los lobos logran morder y lesionar al tigre. Otras tantas, el tigre logra dar un zarpazo a sus enemigos. Después de un tiempo, la resistencia del tigre es minimizada y los cazadores lo derrotan. Sin embargo, los cazadores pagan un costo: dos lobos están gravemente heridos y no pueden ponerse en pie. Algunos lobos, mostrando empatía,

se acercan y lamen a los heridos; no hay nada más que hacer. Los sobrevivientes van sobre los restos del tigre y comienzan a comer. Los lobos han mostrado una inteligencia colectiva que les permitió derrotar a un enemigo más fuerte. Alternándose de manera precisa desde diferentes lados no han dando la menor oportunidad a la presa de responder o escapar; juntos son una máquina de cazar.

Figura 2. Grupo de humanos ancestrales cazando un bisonte.

Nota: Ilustración generada con asistecia de DALL-E 2.

Ahora analicemos una situación similar. En este segundo caso, los cazadores son unos homínidos, de la especie Homo sapiens para mayor precisión. Ahora, la presa es un tipo de bisonte. Los homínidos son menos ágiles que los lobos de nuestro primer ejemplo, pero cuentan con herramientas: varas largas cuyas puntas han sido afiladas y endurecidas al fuego y, hachas hechas de rocas afiladas para cortar y arrojar. Sin intención, emulan la conducta de los lobos: se acercan desde diferentes flancos contra el bisonte que, igualmente en una zona de charcas, se revuelve para tratar de contrarrestar los ataques. La escena se desarrolla de manera muy similar al caso anterior. Los cazadores atacan desde diferentes flancos, tratan de no ofrecer un blanco fijo, se alternan en sus ataques y, algunas veces el bisonte logra en sus embates lesionar a algún atacante. Como en el caso anterior, los cazadores logran su cometido y someten a su presa. Igualmente, hay

un costo para los cazadores: de un grupo de nueve, tres han sido abatidos. La cacería ha concluido.

Aunque ambos grupos de cazadores (lobos y humanos) han actuado de forma coordinada para someter a una presa más fuerte y ágil, hay diferencias notables entre ambos casos. En el caso de los humanos la coordinación fue más compleja: se intercambiaron miradas, señas y sonidos que guiaron el ataque; esto ha sido resultado de una conducta social más sofisticada producto de su mayor capacidad cerebral. De inicio, los sapiens han tenido que elaborar herramientas, planear y coordinar su ataque en forma más compleja dada su limitación de agilidad, de fuerza, y de armas naturales. Por ello, no es de sorprender imaginarnos que lo que ocurre después también sea diferente.

Los Sapiens sobrevivientes se acercan a sus colegas caídos. Se entristecen por su pérdida. Sacan sus restos del fango y los cubren con piedras para evitar que algún depredador, tal vez lobos que se escuchan aullar, vengan por ellos. Sienten un gran pesar el que bestias devoren los restos de sus compañeros. Para aliviar la sensación de pérdida y fomentar su esperanza de que esta no sea definitiva, los sapiens realizan ritos funerarios. Luego, en nuestra escena imaginada, el líder recoge el hacha de piedra de uno de los caídos. No comen la presa directamente en el lugar, sino que la destazan con sus herramientas de piedra y se llevan las partes a una cueva donde el grupo y sus familias viven temporalmente. Ahí la asan en el fuego y la comparten con los demás miembros del grupo, generalmente mujeres y niños.

Por la noche, el líder del grupo de cazadores observa el hacha de piedra que recogió. Si pudiéramos leer sus pensamientos estos serían, en nuestras palabras, algo como: "es una herramienta formidable, de buen pedernal, con un lado romo para asirse, con filo doble y cuyo diseño permite ser reafilada". Dicha hacha es el resultado de más de un millón y medio de años de desarrollo tecnológico; desde que el Homo habilis creó las primeras hachas de piedra. Eso es mucho tiempo, como referencia, el sapiens aparece hace unos 170 000 años; 10 000 años es el lapso de tiempo que separa el hoy de la invención de la agricultura.

Regresemos a nuestro cazador que empuña el hacha de piedra recuperada. Este reflexiona sobre la pérdida de los colaboradores. Asumimos

que, al igual que los Sapiens de la tribu actual Piraha del Amazonas que no tienen concepto de número sino de relaciones muchos y pocos (Gordon, 2004), el líder nota que hay más espacio junto a la fogata. Comprende que hay ahora una menor proporción de machos adultos en el grupo. Recuerda cómo se desenvolvió la cacería, y trata de identificar errores en sus decisiones y acciones. Por necesidad, las próximas veces tendrán que buscar presas de menor tamaño, que aportarán menos comida y por lo tanto necesitarán incrementar la frecuencia de las cacerías. Y serán los débiles (ancianos, enfermos y niños) los primeros en ser dejados atrás.

La gran diferencia entre los dos ejemplos que hemos imaginado es la actitud reflexiva de los cazadores sobre los acontecimientos. La escena que involucra a individuos sapiens, tal vez ocurrió en Tanzania hará unos 100 000 años; y el de los lobos en algún lugar de Europa en fechas similares. En ambos casos, en los cazadores sobrevivientes se generaron sentimientos negativos (tristeza) por la pérdida de miembros del grupo. Sin embargo, lo relevante es que en el caso que hemos imaginado del humano, se generó un proceso mental que revisó, evaluó y corrigió los procesos mentales que condujeron su proceder. Después de ese proceso, el sujeto en cuestión modificó su razonamiento para que la próxima vez, las decisiones que habrá de tomar se adapten a las circunstancias, minimicen las pérdidas, y garanticen el éxito de la cacería. A este proceso mental que se desarrolla cuando el individuo se da cuenta de las implicaciones de sus decisiones lo llamamos conciencia primaria (o primitiva). Es un proceso mental que evaluá otros procesos mentales. Baars (1997) proporciona una explicación accesible de cómo los procesos cognitivos pueden llevar a un estado de conciencia. Este nivel de conciencia se observa cuando los individuos no solo reaccionan al mundo que les rodea, sino que también comprenden y reflexionan sobre sus acciones y pensamientos. Esta capacidad de introspección es crucial para adaptarse a nuevos entornos y situaciones, permitiendo aprender de errores y éxitos pasados.

La habilidad de autoevaluar los propios procesos cognitivos es un paso fundamental hacia el desarrollo de una autoconciencia. La autoconciencia va más allá de la simple reflexión y revisión, involucra un reconocimiento de uno mismo como entidad individual. Damasio (1999) proporciona una perspectiva neurobiológica sobre la relación entre la conciencia y el sen-

tido del yo. Esta forma de conciencia implica entender que uno existe de manera independiente dentro de un contexto más amplio. En humanos, esto se manifiesta en la capacidad de reconocerse en un espejo (aunque también elefantes y orcas, entre otras especias actuales, se pueden reconocer en el espejo), ser consciente del paso del tiempo, y reflexionar sobre el propio pensamiento y emociones. Metzinger (2004) aborda la noción de autoconciencia y cómo se relaciona con la experiencia subjetiva del mundo. La subjetividad es un aspecto clave de la experiencia consciente, permite a los individuos tener experiencias emocionales y personales únicas. Las emociones desempeñan un papel crucial en la forma en que los humanos interactúan con el mundo. Las memorias influyen en la toma de decisiones, en la formación de aspiraciones y en la definición de objetivos a largo plazo. De acuerdo con Damasio (1995), la emoción y la razón no están separadas, sino que son críticas para la cognición racional. Esta dimensión emocional de la conciencia es fundamental para entender la complejidad del comportamiento humano. Los sentimientos y las emociones (derivados de la subjetividad) colorean nuestra percepción del mundo, influyen en nuestras decisiones y acciones, y moldean nuestra individualidad.

La autoconciencia se debió desarrollar entre los cazadores Sapiens. Pues hace 100 000 años, los Sapiens ya utilizaban pigmentos basados en ocre para decorar sus cuerpos con fines ornamentales (Henshilwood et al., 2011). La necesidad de identificación personal, de reconocerse como distinto, implica autoconciencia. Más recientemente, hace unos 40 a 45 mil años nuestros antepasados ya pensaban sobre sí mismos (contaban con autoconciencia). Los Sapiens ya colgaban en su cuello las garras de presas cazadas, adornaban su cuerpo con joyas compuestas por huesos, pedazos de madera, cuentas de piedra, plumas, conchas y caracoles, añadiendo a cada individuo elementos que los hacían únicos (Henshilwood et al., 2002). Todas ellas son actividades con un sentido estrictamente estético cuyo propósito va más allá de la supervivencia. Otro signo de la autoconciencia se manifiesta en el arte rupestre, –la pintura rupestre elaborada por sapiens más antigua (descubierta) se encuentra en la isla de Cáseres, en Indonesia–. Dicha pintura data de hace unos 45 000 años y representa un jabalí. El arte rupestre refleja una conciencia del sentido del tiempo, del ímpetu por trascender a su paso, de enviar un mensaje a otros "yo".

Cuando observamos arte rupestre damos por sentado que fue realizado por individuos con un sentido de autoconciencia.

Sobre la inteligencia artificial

Si, como dijimos, la inteligencia es la capacidad de aprender, adaptarse, planear, resolver y crear. Entonces, como lo propone Gardner (2011) en su teoría de las inteligencias múltiples, podemos dar por hecho que existen diferentes tipos de inteligencia en los seres humanos. Por ejemplo, hay individuos que tienen una gran habilidad para las relaciones personales, otros para crear nuevas formas o herramientas, y algunos cuya inteligencia les permite resolver problemas matemáticos complejos. También están aquellos con habilidades de liderazgo, con una habilidad destacada para las finanzas, con una mente preclara para emprender, aquellos con un talento especial para analizar situaciones complejas y proponer soluciones novedosas.

Todas estas formas de inteligencia son complementarias y es difícil encontrar a un individuo que posea todas. Sternberg (1985) proporciona una visión amplia de la inteligencia más allá del coeficiente intelectual tradicional, abarcando aspectos creativos, prácticos y analíticos. Por lo que es difícil establecer una métrica que englobe las distintas formas de inteligencia. Así, viendo las diferentes vertientes de la inteligencia, es razonable pensar que un sistema computacional que tenga la capacidad de aprender, adaptarse, y resolver problemas, puede ser considerado un sistema inteligente. Note que no estamos hablando de un grado mayor o menor de inteligencia, sino solo de la capacidad para implementar sistemas inteligentes. Los sistemas computacionales que exhiben estas capacidades implementan estrategias denominadas de inteligencia artificial (IA) (ver Russell and Norvig, 2016).

Hoy en día, contamos con herramientas sofisticadas capaces de crear imágenes a partir de descripciones textuales, resumir libros, reconocer el habla y traducirla en texto, traducir entre distintos idiomas, analizar exámenes de laboratorio y sugerir diagnósticos, reconocer objetos complejos e inferir la estructura tridimensional de escenas a partir de imágenes bidimensionales; por mencionar solo algunas tareas. No nos queda duda

de que dichos sistemas sofisticados poseen, en gran medida, habilidades que identificamos como inteligencia. La IA es una realidad con la que ya podemos convivir y obtener ventajas de su uso; a la vez enfrentamos el reto de su regulación (Bostrom, 2014; Marcus y Davis, 2019). Como demostración de las capacidades de los sistemas de IA para crear nuevas imágenes a partir de descripciones de texto hemos incluido dos imágenes generadas con asistencia del sistema DALL-E 2, ver Figuras 1 y 2 de este capítulo. Dichas imágenes sorprenden por el nivel de detalle. Sin embargo, es importante decir que para generar dichas imágenes requerimos de al menos ocho intentos por cada una. En cada nuevo intento refinamos la descripción para mejorar la generación. Un problema que tuvimos era que no se generaban correctamente el número de cazadores solicitado (variando generalmente en uno o dos), por lo que al final ajustamos el texto a la imagen generada que nos pareció más adecuada. Además, podemos notar en la Figura 1 que el lobo en primer plano parece contar con una pata trasera extra y el tigre dientes de sable luce más como un tigre de bengala: la estructura ósea del cráneo no corresponde a los dientes de sable prehistóricos. Es decir, es notable que el sistema no incorpora adecuadamente el conocimiento general de que los mamíferos tienen cuatro extremidades y es evidente su limitación para generar imágenes de elementos poco representados en la base de datos usada en el entrenamiento de DALLE-" 2. Hace falta un sistema que cierre el lazo. Un sistema que revise las posibles inconsistencias entre la inferencia y el conocimiento a priori, que una vez detectadas dichas inconsistencias, condicione una nueva inferencia para corregir dichos errores. Esto lo hicimos nosotros mismos al refinar iterativamente la descripción para que la imagen generada se fuera ajustando a lo deseado. Este proceso extra es el que haría las veces de lo que denominamos conciencia primaria en el caso del cazador sapiens; Sección 1.

Hemos evitado introducir fórmulas matemáticas, solo hasta ahora recurriremos a su uso, no lo haríamos si no creyéramos que simplificamos la exposición al recurrir al auxilio que la abstracción matemática nos otorga. Sin embargo, mantendremos la complejidad matemática en un mínimo. Dicho esto, representamos el proceso de generación de una imagen como la tarea de obtener una muestra x de una distribución p condicionada a satisfacer una descripción t:

$$x \sim p(x \mid t).$$

En esta notación $p(x \mid t)$ es la probabilidad de una imagen x dado un texto t. Por ejemplo si Texto ="una manada de lobos cazando un tigre dientes de sable en una zona de charcos en la estepa". Entonces, la Figura 1 tendrá una mayor probabilidad que la Figura 2:

$$p(x = \text{Fig. } 1 \mid t = \text{Texto}) > p(x = \text{Fig. } 2 \mid t = \text{Texto}).$$

Ahora, $x \sim$ significa que obtenemos una imagen al azar dado un texto: escogemos una imagen al azar de entre el conjunto de todas las imágenes posibles, pero la probabilidad de seleccionar una imagen en particular depende de qué tan consistente es con la descripción en el texto.

Por otro lado, al proceso de obtener la descripción (denotada por t') a partir de una imagen x lo podemos representar como:

$$t' \sim p(t \mid x).$$

En esta notación, la generación de imágenes se representa por la primera fórmula, y el análisis de una imagen por la última fórmula. Hoy en día existen sistemas de redes neuronales artificiales muy eficientes que implementan ambos modelos. Esto es, podremos introducir un texto a un modelo que implementa la primera fórmula y el resultado (imagen) introducirlo a un segundo modelo que implementa la tercera fórmula, obteniendo una descripción de la escena. Luego, podremos comparar el texto introducido con la descripción obtenida y ver el grado de acuerdo. Tenemos ahora un par de textos (t,t'). Entonces, el grado de acuerdo (consistencia) entre ambas descripciones lo podríamos denotar por $p(t' \mid t)$; que, de nuevo, lo podemos implementar mediante un sistema de IA.

En los párrafos anteriores presentamos de manera muy general el proceso de generar de datos en una representación a partir de una pista en otra representación: imágenes a partir de texto o texto a partir de imágenes. Asi mismo hemos presentado un esquema para evaluar la congruencia de lo generado. Estos esquemas no son simples suposiciones de cómo debería funcionar la IA, sino que ya son implementados por modelos de IA so-

fisticados. Lo que observamos es que en tanto tengamos datos suficientes del tipo (x = "datos de entrada", y = "salidas esperadas") es posible entrenar redes neuronales que "aprenden" a realizar la transformación: y = f(x) –siempre y cuando f sea suave: x similares correspondan a y similares.

En la actualidad estamos ante el surgimiento de reservorios masivos de datos. De hecho, los reservorios más extensos son recolectados y administrados por las grandes compañías proveedoras de servicios en Internet: telefonía, de video streaming, email, buscadores de internet, gestores de redes sociales, comercio electrónico, etc. Por lo que no es de extrañar que dichas compañías sean las que estén a la vanguardia en el desarrollo e implementación de sistemas complejos de IA.

Estamos apenas descubriendo el potencial de dichos sistemas en nuestras vidas. Para bien o para mal, la IA formará parte medular en nuestras sociedades (Marcus y Davis, 2019). El impacto de esta revolución tecnológica no la hemos siquiera imaginado, por lo que los mecanismos de regulación van atrasados con respecto al uso de la IA (Bostrom, 2014).

Sobre la conciencia artificial

La autoconciencia nos permite estar al tanto de nosotros mismos y de nuestro entorno; es tema de fascinación y estudio en filosofía, psicología y ahora en la IA. Con la finalidad de darnos una idea del reto de crear conciencia artificial nos hacemos la siguiente pregunta: ¿Es la conciencia un fenómeno raro en la naturaleza? La respuesta simple es no de acuerdo con la Declaración de Cambridge Sobre la Conciencia (Low et al., 2012):

"De la ausencia de neocórtex no parece concluirse que un organismo no experimente estados afectivos. Las evidencias convergentes indican que los animales no humanos tienen los sustratos neuroanatómicos, neuroquímicos, y neurofisiológicos de los estados de la conciencia junto con la capacidad de exhibir conductas intencionales. Consecuentemente, el grueso de la evidencia indica que los humanos no somos los únicos en poseer la base neurológica que da lugar a la conciencia. Los animales no humanos, incluyendo a todos los mamíferos y pájaros, y otras muchas criaturas, incluyendo a los pulpos, también poseen estos sustratos neurológicos."

Esta declaración realizada por un grupo de neurocientíficos y psicólogos define la sintiencia. Para distinguirla de la conciencia primaria y auto-

conciencia que hemos esbozado previamente, y que son el tema que nos ocupa. La sintiencia es la facultad de que el individuo pueda tener experiencias subjetivas (sentimientos) o conocimiento. Sentimientos que puede experimentar como positivos o negativos; sin necesidad que reflexione sobre ellos o tenga un concepto de sí mismo (Antony, 2002). No podemos poner en duda que las mascotas desarrollan afecto (apego) por sus amos. El reconocimiento explícito de que los animales desarrollan sentimientos sustenta las acciones para señalar y detener actos de crueldad animal.

Sin embargo, la autoconciencia (capacidad de pensar y razonar) no es parte necesaria de la sintiencia. La conciencia primaria puede resultar de procesos mentales que se activan al tiempo que se llevan a cabo los procesos mentales relacionados con la actividad que se realiza. En los ejemplos de los grupos cazadores de lobos o humanos, en plena cacería se activaban sentimientos de furia o temor. Lo que indica la emergencia y ocurrencia de procesos mentales paralelos en especies con sistemas nerviosos desarrollados. Lo que haría suponer que la conciencia primaria, que implica procesos mentales no reactivos sino de evaluación de los propios procesos mentales, también es factible en animales superiores. Caben las preguntas: ¿qué tan superiores? ¿Solo los sapiens hemos desarrollado conciencia primaria y autoconciencia? ¿Somos la excepción?

En las secciones anteriores presentamos argumentos que nos permiten ver a la autoconciencia como consecuencia y necesidad en especies, como el sapiens, con cerebros complejos. Sin embargo, los sapiens no son la única especie inteligente que ha manifestado claramente capacidad de ser autoconsciente. El Homo Neanderthalensis (los Neandertales, una especie de humanos extinta hace unos 40 000 años cuya presencia se extendió por Europa y norte de Asia) también fue capaz de crear herramientas, tener una vida social compleja, crear joyería, realizar pinturas rupestres complejas y realizar ceremonias fúnebres. Como referencia, entre las pinturas rupestres más antiguas de Europa se encuentran las de la cueva de Maltravieso (Extremadura, España) atribuidas a Neandertales y datan de aproximadamente unos 67 000 años (Hoffmann et al., 2018). También en la cueva de La Ferrassie (Dordoña, Francia) se encontraron restos de un niño que evidenciaban prácticas funerarias propias de procesos cognitivos complejos realizadas por neandertales hace cerca de 41 000 años (Balzeau et al., 2020).

Los neandertales fueron una especie social, creativa, que cuidaba de sus congéneres débiles, que construía herramientas de piedra, y que creaba joyería a partir de conchas y cuentas de piedra por el puro placer estético que brindaban tales objetos. Las pinturas neandertales fueron plasmadas mucho tiempo antes de la pintura rupestre de sapiens más antigua conocida en Indonesia. El fechado de la pintura sapiens en Indonesia coincide con la llegada de los sapiens al continente Europeo; alrededor de 20 000 años después de las pinturas neandertales en la cueva de Maltravieso. El lapso de tiempo transcurrido entre las mencionadas pinturas neandertales y las sapiens es aproximadamente el doble del lapso entre el invento de la agricultura y hoy día.

El estudio de la evolución de la inteligencia y de la conciencia humana nos ha mostrado que al menos en dos ocasiones el mecanismo evolutivo ha producido especies autoconscientes. Seguramente, la conciencia de ambas especies difería. A la naturaleza le ha tomado más de dos millones evolucionar del Homo habilis al Homo sapiens: de la especie que elaboró la primera hacha de piedra, a la que acuñó el término de IA. A nosotros, los sapiens, nos ha tomado menos de un siglo pasar de la publicación por McCulloch and Pitts (1943) del modelo de la neurona hasta el desarrollo de los grandes modelos de lenguaje (ChatGPT, Llama, Claude; por mencionar algunos). Con dichos sistemas, que implementan técnicas de IA, podemos entablar conversaciones que con dificultad las distinguimos de las que entablamos con otro sapiens [prueba de Turing (1950)]. Lo que nos lleva a pensar que, como en el caso de la IA, solo es cuestión de tiempo para que seamos testigos de la aparición de modelos computacionales que manifiesten comportamiento que podamos definir como un tipo de "conciencia artificial", del tipo conciencia primaria o autoconciencia.

En el contexto de las redes neuronales y sistemas de IA, con su evidente complejidad creciente, diseñadas para procesar y evaluar grandes cantidades de datos, que aprenden y se adaptan de manera similar a un cerebro biológico. La sola idea de que estas redes puedan autoevaluar su proceso de inferencia podría considerarse un paso hacia una forma rudimentaria de autoconsciencia. De acuerdo con Kurzweil (2005), en esta etapa, la IA no solo ejecutaría tareas, sino que también desarrollaría una capacidad para monitorear y ajustar sus procesos cognitivos. Dado que la conciencia

primaria es un proceso mental y, como tal, podría ser representado por un modelo de IA. Cada vez nos acercaríamos más a implementar lo que podríamos denominar "conciencia primaria artificial".

La posibilidad de que las redes neuronales implementen una forma de autoconciencia es un área de gran interés y especulación en el campo de la IA. Queda abierta la cuestión si la conciencia, o al menos una forma primitiva de ella, podría ser una necesidad o una consecuencia inevitable de que estas redes se vuelven cada vez más sofisticadas y capaces de realizar tareas complejas: ¿podrán alcanzar un estado de autoconciencia rudimentaria, donde reconozcan y ajusten sus propios procesos de inferencia (razonamiento)? Bostrom (2014) explora la implicación de IA avanzada, incluyendo la posibilidad de una conciencia emergente.

La conciencia humana tiende a verse desde aproximación unificada, pero debe ser considerarla en sus distintas capacidades, tipos y niveles (Montemayor, 2021; Morin, 2006). Basados en este reconocimiento de niveles de conciencia, postulamos que debemos estar abiertos a flexibilizar nuestro concepto de conciencia (como lo hemos hecho respecto a la inteligencia) y considerarla como un proceso mental que evalúa de los propios procesos mentales. La idea de que una IA pueda no solo ser consciente sino también reflexiva, es decir, capaz de pensar sobre sí misma, es un tema fascinante. Esto plantearía preguntas sobre la naturaleza de la conciencia y si la conciencia artificial podría emular o incluso superar la experiencia humana.

Es decir, que pase con la conciencia artificial lo que Kurzweil (2005) pronostica al decir que estamos en los albores de la llamada singularidad: formas de IA avanzadas, que incluyen la autoconciencia y la reflexividad. Si se desarrolla un nivel de conciencia primitiva en IA, surge la cuestión de si este sistema podría, eventualmente, reflexionar sobre sí mismo y evolucionar hacia una forma de conciencia más avanzada. Esta posible conciencia artificial podría, en teoría, emular aspectos de la conciencia humana, pero también diferir de ella en formas fundamentales. La conciencia artificial podría no incluir elementos emocionales o subjetivos como los experimentamos los humanos, o podría desarrollar formas completamente nuevas de experiencia consciente, dictadas por su arquitectura y capacidades únicas.

Tendemos a colocarnos en la cima evolutiva y damos por hecho que no hay un escalón más allá. Esta visión antropocéntrica la hemos reforzado con imágenes icónicas como la llamada "Marcha del progreso"(imagen popular donde se observan una secuencia que inicia en un Mono, pasa por un Australopithecus, un Homo Erectus y concluye en un Sapiens) para asumirnos como el resultado final del proceso evolutivo; como si toda especie aspirara a ser Sapiens. Sobrevaloramos aquellas habilidades que, en nuestra opinión, nos distinguen de otras especies y por ello no concedemos que puedan ser emuladas, ya sea por sistemas con sustrato de carbono (otras especies de seres vivos) o con sustrato de silicio (sistemas computacionales actuales).

En este trabajo no pretendemos presentar una teoría sobre la construcción y aparición de la conciencia [para ello recomendamos revisar Chalmers (1997)], solo la exponemos como una consecuencia de la evolución de especies con grado cognitivo complejo. Reconocemos que la IA es un hecho hoy en día. Sabemos que modelos de conciencia artificial están siendo investigados en distintos laboratorios de computólogos por el mundo. ¿Cuándo veremos sistemas con conciencia artificial? ¿Cómo serán? ¿Cómo los distinguiremos? ¿Cuáles serán sus implicaciones en nuestras vidas y nuestro futuro? Más que con respuestas, terminamos con una frase atribuida a Yogi Berra: el futuro ya no es lo que solía ser.

Referencias

Antony, M. V. (2002). *Concepts of consciousness, kinds of consciousness, meanings ofconsciousness*. Philosophical studies, 109, 1-16.

Baars, B. J. (1997). *In the theater of consciousness: The workspace of the mind*. Oxford University Press, USA.

Balzeau, A., Turq, A., Talamo, S., Daujeard, C., Guérin, G., Welker, F., ... others (2020). *Pluridisciplinary evidence for burial for the la ferrassie 8 neandertal child*. Scientific reports, 10(1), 1–10.

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

Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. Oxford Paperbacks.

Damasio, A. R. (1995). *Descartes' error*. Picador.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Houghton Mifflin Harcourt.

Dunbar, R. I. (1998). *The social brain hypothesis. Evolutionary Anthropology: Issues, News, and Reviews,* 6(5), 178–190.

Gardner, H. E. (2011). *Frames of mind: The theory of multiple intelligences*. Basic books.

Gordon, P. (2004). Numerical cognition without words: Evidence from amazonia. *Science*, 306(5695), 496-499.

Henshilwood, C. S., d'Errico, F., Van Niekerk, K. L., Coquinot, Y., Jacobs, Z., Lauritzen, S.-E., ... García-Moreno, R. (2011). A 100,000-year-old ochreprocessing workshop at blombos cave, south africa. *Science*, 334(6053), 219–222.

Henshilwood, C. S., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G. A., ... others (2002). Emergence of modern human behavior: Middle stone age engravings from south africa. *Science*, 295(5558), 1278–1280.

Hoffmann, D. L., Standish, C. D., García-Diez, M., Pettitt, P. B., Milton, J. A., Zilh˜ao, J., ... others (2018). U-th dating of carbonate crusts reveals neandertal origin of iberian cave art. *Science*, 359(6378), 912-915.

Kurzweil, R. (2005). *The singularity is near. In Ethics and emerging technologies* (pp. 393–406). Springer.

Low, P., Panksepp, J., Reiss, D., Edelman, D., Van Swinderen, B., & Koch, C. (2012). The cambridge declaration on consciousness. In *Francis crick memorial conference* (Vol. 7).

Marcus, G. y Davis, E. (2019). *Rebooting ai: Building artificial intelligence we can trust*. Vintage.

McCulloch, W. S., & Pitts, W. (1943). *A logical calculus of the ideas immanent in nervous activity*. The bulletin of mathematical biophysics, 5, 115-133.

Metzinger, T. (2004). *Being no one: The self-model theory of subjectivity*. mit Press.

Mithen, S. (1997). The prehistory of the mind. *Cambridge Archaeological Journal,* 7, 269-269.

Montemayor, C. (2021). Types of consciousness: The diversity problem. *Frontiers in Systems Neuroscience*, 15, 747797.

Morin, A. (2006). Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views. *Consciousness and cognition*, 15(2), 358-371.

Russell, S. J. y Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson.

Sternberg, R. J. (1985). *Beyond iq: A triarchic theory of human intelligence*. CUP Archive.

Tomasello, M. (2009). Why we cooperate. *MIT press*.

Turing, A. M. (1950). I.—*Computing machinery and intelligence*. Mind, LIX(236), 433-460.

Capítulo **17**

El uso de herramientas tecnológicas y de la información para la intervención e investigación en la neuropsicología

Rubén Avilés Reyes
Edgar Adrián Arias Chávez
Javier Tadeo Sánchez Betancourt

De vez en cuando, una nueva tecnología, un antiguo problema y una gran idea se convierten en una innovación.
–Dean Kamen. Creador del Segway y el iBOT.

https://doi.org/10.61728/AE24001175

Introducción

El contacto que las personas tienen hoy en día con las tecnologías de la información y comunicación (TIC) es innegable. Este se da de distintas formas, quizá el más común es el uso del celular, que más que comunicarnos con otros a través de una llamada (como tradicionalmente era su función), actualmente permite acceder al mundo de la información, a través de las diversas funciones que hoy cumple. Las TIC no son exclusivas de un sector en particular, cualquiera con un dispositivo electrónico y con el uso del internet, puede acceder a la información de cualquier índole, en cualquier lugar y en cualquier momento. Las TIC no solo han servido para funciones recreativas, educativas o laborales, también son implementadas en investigaciones e intervenciones en el campo de la salud mental, como es el caso de quienes trabajan en el campo neuropsicológico.

La tecnología puede definirse como la suma de técnicas, habilidades, métodos y procesos que hacen uso de sistemas y aparatos electrónicos utilizados para la producción de bienes o servicios en el logro de objetivos, como la investigación científica (RAE, 2023). Las TIC son el conjunto de recursos, herramientas, equipos, programas informáticos, aplicaciones, redes y medios; que permiten la compilación, procesamiento, almacenamiento y transmisión de información como: voz, datos, texto, video e imágenes (Art. 6 Ley 1341 de 2009). Por otro lado, la neuropsicología puede definirse como la ciencia que se encarga de aclarar las relaciones entre funciones cerebrales y la conducta humana (Benton, 1971). Pero esta definición hoy en día es simplista, por lo tanto, ampliando lo dicho anteriormente, la neuropsicología es la ciencia que se encarga del estudio entre la relación del sistema nervioso central con la conducta, las emociones, la cognición y personalidad tanto normotípica como en la patología en todas las edades del desarrollo humano, uniendo así a las neurociencias y a la psicología (Lezak, 2004, Portellano, 2005). El quehacer del neuropsicólogo es investigar y atender casos relacionados a la disfunción o alteración cerebral, entre los que se encuentran diversas patologías como las enfer-

medades psiquiátricas, traumatismos cerebrales, enfermedades congénitas, problemas o alteraciones del aprendizaje y demencias, por mencionar algunas; para ello se vale de instrumentos validados y estandarizados, técnicas cualitativas y métodos encaminados a la evaluación, diagnóstico y tratamiento de los pacientes.

La pandemia por COVID-19 declarada el 27 de febrero del 2020 en México (Secretaría de Salud, 2020) y duró más de dos años, ocasionó que los profesionistas del área de la salud por el confinamiento trasladaran sus servicios en la medida de lo posible a la virtualidad. Cabe enfatizar que el trabajo neuropsicologógico virtual se venía desarrollando tiempo atrás. Por ello se ha generado un nuevo término denominado "teleneuropsicología", que por definición es la aplicación de los métodos clásicos en neuropsicología en escenarios remotos, también conocida como presencialidad remota. Con el uso de tecnologías audiovisuales se establece contacto clínico con personas que requieren el servicio a fin de reducir los problemas de distancia y que ha permitido proveer servicios de salud donde se ve comprometido el cerebro por un daño o disfunción (Bilder, 2020).

La teleneuropsicología es entendida como el servicio que el neuropsicólogo desempeña con la aplicación de estrategias, métodos e instrumentos para la evaluación, diagnóstico y estimulación o rehabilitación neuropsicológica a distancia. Se vale de materiales e instrumentos digitales y se inscribe dentro de una tendencia llamada telepsicología, es decir, la psicoterapia o terapia psicológica en línea que por historia se contempla más antigua que la teleneuropsicología (Álvarez et al. 2020).

Antecedentes de la teleneuropsicología

El comienzo del trabajo en el área de la salud a través del uso de las TIC y como actualmente se concibe, fue documentado en la Revista Radio News en 1924 con el documento titulado "El doctor de la radio, tal vez". Dentro del texto se puede observar como el profesional de la medicina entrevista a su paciente virtualmente (Wittson, 1961). Este fenómeno virtual y digital llevó a que la psicología clínica terapéutica incorporará las TIC para tratar a sus pacientes. Affleck y Johnson (1961) a principios de los años sesenta

utilizaron la virtualidad a través de una conferencia durante una sesión de terapia grupal. En un artículo reciente autores como Pérez, Ramos y Arango (2021) dan cuenta sobre el primer programa que utiliza la tecnología llamada "ELIZA" en 1966 implementado por Weizembaum. Esta tecnología simulaba las respuestas no directivas de un psicólogo, el cual fue considerado el primer programa que utiliza las TIC con un procedimiento de lenguaje natural que intentaba mantener una conversación con el paciente de forma escrita (González et al. 2008).

Con respecto a los antecedentes de la teleneuropsicología, el uso de la tecnología se remonta a finales de los años sesenta, cuando las escalas Wechsler fueron automatizadas (Elwood, 1972) y creadas por el psicólogo estadounidense David Wechsler (1896-1981). Cabe mencionar que dichas escalas son instrumentos creados para medir el coeficiente intelectual, definiendo la inteligencia por el propio Wechsler como "una habilidad general y global que permite al individuo comprender el mundo, adaptarse a él y enfrentarse de manera eficiente a los desafíos" (Martín 2012). Las escalas se dividen en tareas relacionadas a la comprensión verbal y el razonamiento ejecutivo. Hoy se cuenta con escalas Wechsler para niños pequeños WPPSI-IV (2012), para niños escolares y adolescentes WISC-IV (2005) y para adultos WAIS-IV (2012). Existen investigaciones experimentales realizadas por neuropsicólogos que correlacionan la actividad y función del cerebro con la ejecución de las tareas de Wechsler (Sullivan y Bowden, 1998). Esto ha permitido que los neuropsicólogos utilicen adecuadamente las escalas y las TIC. Hoy en día la automatización virtual de algunas tareas de estos instrumentos permite que la teleneuropsicología pueda desempeñarse de manera adecuada y profesional en el ámbito clínico e investigativo.

Teleneuropsicología en el ámbito clínico

La evaluación computarizada ha sido analizada desde los años 80, explorando sus bondades como criticando sus posibles fallos (Parsons, 2016). Entre las bondades se habla de la facilidad en el registro preciso de respuestas, la administración precisa de estímulos, la generación de bases compartidas, etc. Entre las dificultades preocupa conocer en qué medida

el uso de estímulos a través de una pantalla se puede considerar equiparable a los estímulos en pruebas de papel y lápiz. Parsons (2016) hace un recorrido por las primeras incorporaciones del equipo de cómputo a la evaluación neuropsicológica desde la década de los sesenta.

Actualmente, instituciones públicas y privadas, así como, los profesionistas en consulta privada hacen uso de las TIC para ofrecer servicio teleneuropsicológico, por lo que ha sido necesario para la evaluación adaptar o crear instrumentos neuropsicológicos para ser aplicados a través de la virtualidad y que los usuarios tengan la misma calidad y ética en la atención, así como el uso de plataformas que permiten la interacción en tiempo real frente al paciente como Google Meet, Zoom, FaceTime, Skype, etc. Algunas baterías, tareas e instrumentos se muestran en la tabla 1.

Tabla 1. Instrumentos neuropsicológicos utilizados en la virtualidad.

Material	Edad de aplicación	Procesos evaluados
Cogstate Brief Battery (CBB) (Stricker et al. 2020)	Adolescentes a adultos mayores	Memoria de trabajo y atención. Usada en uso del adicciones, esquizofrenia, demencia y deterioro cognitivo
Computer Assesment of Mild Cognitive impairment (CAMCI) Prueba de Cribado (Saxton et al., 2009)	Adultos mayores	evalúan atención, habilidades ejecutivas, memoria de trabajo, la memoria verbal y visual. Deterioro cognitivo
MicroCog: assessment of cognitive functioning (Powell et al. 1996)	Adultos y adultos mayores	Funciones ejecutivas, atención y memoria Deterioro cognitivo Trastornos psiquiatrícos
Batería de Evaluación Cognitiva General (CAB) de CogniFit (Blog, 2024)	De 7 años en adelante	Atención, meoria, funciones ejecutivas, espacialidad, razonamiento, vocabulario, percepción. Diversos trastornos neuropsicológicos y psiquiatrícos
Cambridge Neuropsychological Test Automated Battery (CANTAB) (Capuron et al. 2001; Robbins et al. 2000).	De 6 a 80 años de edad	Psicomotricidad, velocidad de procesamiento, razonamiento, planeación, memoria, antención. Usada en alteraciones frontales y del hipocampo cerebral.
Mindstreams :computerized cognitive battery for mild cognitive impairment (Dwolatzky et al. 2004)	Adultos y adultos mayores	Memoria, atención, funciones ejecutivas, orientación. Usado para demencias, daño cerebral y deterioro cognitivo leve.

Nota: Elaboración propia a partir de una revisión de diferentes fuentes de los instrumentos neuropsicológicos virtuales más utilizados.

La incorporación progresiva de las TIC en el contexto de la salud mental y el daño o disfunción cerebral está ofreciendo un cambio profundo en la forma de concebir y organizar los servicios que mejoran el proceso asistencial, la accesibilidad y la utilidad clínica neuropsicológica en los diferentes espacios de actuación (Soto, 2019). Estas aproximaciones necesitan de la interdisciplinariedad entre diferentes áreas para la evaluación y el diagnóstico de pacientes que requieren el servicio, entre las que se encuentran: las neurociencias, la psicometría y la ingeniería computacional. Esto ha posibilitado la atención de personas en diversos lugares geográficos, reducción de costos económicos y menor tiempo para acudir a una consulta.

La evaluación a través de la teleneuropsicología busca los mismos preceptos y fundamentos que la presencialidad, lo que permite comprender desde un punto de vista práctico, las interacciones de las funciones cognitivas con las discapacidades de la realización de las tareas o actividades de la vida cotidiana (Terradillos et al., 2012).

Se puede afirmar, que desde la pandemia por COVID-19, existe una tendencia cada vez mayor de emplear la virtualidad en ámbitos propios de la psicología, en general, y la neuropsicología, en particular (Rodríguez, 2021). Las evaluaciones y el diagnóstico desde la teleneuropsicología han demostrado ser efectivas, las cuales se complementan con los métodos tradicionales. Hoy en día los sistemas computarizados dotados de medios auxiliares como el video, el audio y la animación añaden objetividad, confiabilidad y rigor al proceso de evaluación y diagnóstico, particularmente en el abordaje de tratamientos con pacientes adultos mayores y en daño cerebral, donde la movilidad física está alterada de manera grave (García, 2019; Sigmundsdottir et al, 2016).

Instituciones importantes que se encargan del cuidado del paciente neuropsicológico están optando por brindar un servicio teleneuropsicológico, como hospitales de cuidados intensivos, quienes prestan servicio a centros de salud rurales y comunitarios donde el acceso a ciertas especialidades psicológicas y médicas es limitado, así como espacios de salud que requieren cobertura nocturna o de emergencia, en donde deben aumentar la atención por la cantidad de pacientes, además de hospitales en atención a la salud mental y conductual, donde es necesario resolver los desafíos en la calidad y dotación de profesionistas que atienden a pacientes ambula-

torios. Lo anterior son ejemplos claros de la importancia y la eficacia del trabajo de la teleneuropsicologia.

Existen instituciones, grupos colegiados y asociaciones en el campo de la neuropsicología, que bajo la ética ofrecen cursos a profesionistas de la salud mental para prepararlos y capacitarlos en la atención y el uso de las TIC. También se cuentan con programas neuropsicológicos de posgrado que preparan y capacitan a sus estudiantes para la intervención de sus pacientes en sus prácticas profesionales a través de la teleneuropsicología, como la maestría en diagnóstico y rehabilitación neuropsicológica de la Benemérita Universidad Autónoma de Puebla (BUAP), la maestría en psicología con énfasis en neuropsicología de la Universidad Autónoma del Estado de Morelos (UAEM) y la maestría en psicología con énfasis en neuropsicología de la Universidad Nacional Autónoma de México (UNAM).

Una vez realizada la evaluación y el diagnóstico neuropsicológico, muchos pacientes requerirán de una intervención, ya sea un proceso de rehabilitación neuropsicológica o de estimulación cognitiva. La rehabilitación neuropsicológica es un área de la neuropsicología clínica aplicada, qué consiste en la intervención sobre los procesos cognitivos, afectivos y conductuales alterados. Se aplican procedimientos y técnicas para compensar o disminuir los déficits presentados y de esa manera mejorar la calidad del paciente para una mayor eficiencia en actividades de su vida diaria (Wilson, 2013). La estimulación cognitiva es un conjunto de técnicas para mejorar o potencializar las capacidades cognitivas tales como la orientación, la atención, la memoria y las funciones ejecutivas. Ambos procesos de intervención se han enriquecido a través de las TIC. Es necesario precisar que en la intervención neuropsicológica (rehabilitación o estimulación) se requiere el apoyo del cuidador primario (miembro de la familia) y el profesional de la salud, es decir, que con la participación de estas personas se puede implementar las tareas, métodos y técnicas a través de la virtualidad con el apoyo de plataformas que permiten la interacción en tiempo real (Castilla et al., 2019).

El incremento en el conocimiento del uso de las TIC ha significado un factor determinante para la intervención neuropsicológica a través de la virtualidad lo que ha permitido la recuperación o la mejora de los procesos cognitivos, conductuales y emocionales consecuencia del daño o disfunción cerebral (Fernández et al., 2018; Bringas et al., 2019).

Después de la pandemia por COVID-19 se ha implementado y aumentado el uso de softwares neuropsicológicos diseñados para poner en práctica tareas cognitivas y rehabilitar los procesos alterados, por ejemplo, aquellos implicados en la adquisición de la lectura, las matemáticas y la espacialidad en niños con problemas o trastornos del aprendizaje, así mismo, existen otros softwares neuropsicológicos para el uso de la estimulación cognitiva en adultos mayores como el Neuronic Estimulador Cognitivo que permite el trabajo de intervención dependiendo de los tiempos del paciente (Aguilar et al., 2019).

Las plataformas digitales que hacen uso de las TIC han mostrado su eficacia en la intervención como si fuera un tratamiento presencial. Fernández y Crespo (2020), indican que esta eficacia se debe a la correcta automatización, estandarización, validez y confiabilidad de los instrumentos, tareas y métodos utilizados desde la presencialidad remota, presentando las siguientes caracteristicas:

a) No están sujetas a las variaciones propias de las diferencias entre rehabilitadores en el curso del tiempo.
b) Dentro de ciertos límites estas técnicas pueden ser autoadministradas o administradas simultáneamente a un grupo de pacientes que requiera la rehabilitación o estimulación, lo que reduce el tiempo de ejecución.
c) Los datos obtenidos pueden recogerse de manera eficaz y en menor tiempo.
d) El procesamiento de dichos datos puede ser inmediato y realizarse así en la próxima sesión las modificaciones pertinentes a los programas de rehabilitación neuropsicológica o estimulación cognitiva.
e) En el caso de los niños la intervención a través de las TIC garantiza un alto nivel de motivación.

En la actualidad existen diferentes softwares que permiten la rehabilitación neuropsicológica o la estimulación cognitiva desde el uso de las TIC, diseñados para su implementación individual o grupal, entre ellos se encuentra RehaCom, COGNIPLUS (rehabilitación de múltiples trastornos cognitivos), SMART BRAIN (estimulación cognitiva y entrenamiento mental) y GRADIADOR (entrenamiento cognitivo para la demencia). Estos productos tecnológicos fueron diseñados por profesionales de la

neuropsicología, las neurociencias y la psicología del desarrollo en la Universidad de Stanford. Otros productos y programas digitales de entrenamiento, estimulación y rehabilitación neuropsicológica con un soporte de informática son CogniFit, NeuronUP, Attention Game (AG) (Restrepo et al., 2018).

Es necesario indicar que ningún instrumento, tarea, método o técnica de intervención neuropsicológica que hace uso de las TIC puede aplicarse sin la supervisión de un neuropsicólogo clínico previamente formado para el uso y manejo de la presencialidad remota. Así mismo, es de suma importancia dejar claro que ningún paciente aun teniendo el mismo diagnóstico manifestará la misma sintomatología. Por ello, es necesario crear e implementar programas de rehabilitación y estimulación adecuados para cada uno de ellos tomando en cuenta características como la edad, la escolaridad, la cultura, el modo de vida y el sexo. De esta forma, nos aseguramos de realizar los cambios pertinentes a los programas virtuales cuando se refleje que no son adecuados o no están funcionando.

La familia o el cuidador primario son los pilares básicos en la implementación de los programas virtuales de rehabilitación neuropsicológica o estimulación cognitiva, ya que sin ellos, el equipo de trabajo estaría incompleto. Estos cuidadores primarios actúan como una extensión en el sistema del cuidado y manejo de cada paciente así como la supervisión del trabajo desde casa. Tanto el cuidador primario como los familiares deben recibir entrenamiento en el uso de las TIC para favorecer y hacer eficaz el proceso de intervención neuropsicológica (Rojas, 2019; Norup, 2018).

Investigación a través del uso de las tecnologías de la información y de la comunicación

A lo largo del tiempo las temáticas y los intereses de la investigación en el área de la salud mental y el daño cerebral han cambiado. Estos cambios se deben a los procesos socioculturales, epidemiológicos y actualizaciones en temáticas de índole neuropsicológico. En décadas recientes la investigación se ha valido de instrumentos tecnológicos, virtuales y digitales que permiten una mejora en tiempo y espacio de los propios investigadores y los participantes. Desde el punto de vista de la psicología, hoy la disciplina

se preocupa por atender fenómenos actuales. La American Psychological Association (2018) reporta que de quince años a la fecha, el interés por temas emergentes se centra en el cambio climático y su impacto en la salud mental, los tratamientos del dolor sin medicamentos, la psicología del deporte, la salud nutricional, el cuidado a largo plazo de pacientes psiquiátricos, la equidad de género, el big data, las herramientas tecnológicas inteligentes, la neuropsicología social y el crecimiento de las terapias digitales. Esto ha generado que la psicología se conciba como una disciplina dinámica, transdisciplinar, activa y actualizada, junto con los retos que el mundo virtual requiere. Por lo tanto, la teleneuropsicología busca un desarrollo de acuerdo con los cambios de la propia humanidad, intentando un entendimiento de las personas y los fenómenos que incluyen su práctica profesional (Schruijer y Stephenson, 2010).

Dentro de las investigaciones de la neuropsicología y haciendo uso de las TIC, a través del manejo de plataformas digitales y softwares de libre acceso, se ha avanzado en evidencia científica que abona al conocimiento de diversas áreas de interés, en las que se encuentran de forma genérica el deterioro cognitivo en adultos mayores, daño cerebral en adultos, problemas de adicciones en jóvenes y trastornos del neurodesarrollo en niños (Cobos et al. 21 y Roma, 2004).

El dispositivo por excelencia para la investigación en neuropsicología es el celular que a través de las bondades que nos ofrecen las plataformas como Google (Meet, Forms y Drive) y la facilidad para manipularlas, permite que pacientes y personas normotípicas de todas las edades del desarrollo humano puedan hacer uso de ellas. Estas tecnologías han permitido que investigadores puedan compartir información entre colegas y generar lazos de apoyo colaborando entre diferentes universidades del mundo (Cedeño et al. 2020).

Varios autores señalan el potencial que poseen las TIC para mejorar el bienestar y la calidad de vida, como ejemplo de ello es la aplicación Positive Technology desarrollada para ser utilizada en los teléfonos inteligentes y las tabletas. Se trata de la primera sotfware diseñado específicamente para el automanejo del estrés. PositiveTechnology utiliza una serie de biosensores e incluye ejercicios de biofeedback y relajación. Los autores discuten el enorme potencial de estas aplicaciones diseñadas para ser utilizadas en

dispositivos móviles, para reducir el estrés psicológico y para ayudar en el manejo de muchos otros problemas mentales (Botella y Breto?n, 2014).

Las investigaciones en neuropsicología se nutren hoy en día de las innovaciones tecnológicas para la práctica teleneuropsicológica y tener un mayor alcance, versatilidad, eficacia y efectividad que han sido comparadas y comprobadas con prácticas investigativas clásicas. Las TIC permiten contar con un número mayor de herramientas que ayudan a optimizar los tratamientos de diversos trastornos neuropsicológicos y de salud mental como fobias, ansiedad, depresión y aplicarlos en diferentes grupos de edad.

Es importante mencionar que durante el confinamiento por la pandemia de covid-19, la organización de eventos académicos y científicos, como los congresos, talleres, charlas y simposios se trasladaron 100 % a la virtualidad. Los científicos continuaron con sus investigaciones mostrando sus hallazgos a colegas, estudiantes y la comunidad en general en cualquier parte del mundo. Una vez terminada la pandemia, esta nueva forma de trabajo se continua utilizando por universidades, hospitales, grupos colegiados e instituciones que realizan investigación, ya que se dieron cuenta de las ventajas de dar a conocer nuevos conocimientos, sin la necesidad de desplazarse al lugar de origen de las propias instituciones, siendo tan eficaz como en su momento fue la presencialidad.

Trabajos realizados en la UABC usando las tecnologías de información y de la comunicación

La Universidad Autónoma de Baja California ha permitido un acercamiento a la investigación desde la virtualidad y el uso de las TIC, a través de sus programas de Proyectos de Vinculación con Valor en Créditos (PVVC), cuyos objetivos son: a) vincular a los estudiantes en actividades productivas con la comunidad y con las empresas; b) la experimentación y aplicación práctica de los conocimientos obtenidos en el aula; c) la adquisición de nuevos conocimientos dentro del ritmo de trabajo de las empresas e instituciones; y d) el desarrollo de habilidades y valores que impacten en la formación integral del futuro profesionista. Tal es el caso de PVVC registrados en la Licenciatura en Psicología de la Facultad de Ciencias Admi-

nistrativas y Sociales de la UABC, campus Ensenada, donde se han generado convenios que involucran a otras universidades como la Universidad Autónoma del Estado de Morelos a través de su programa de prácticas profesionales, que ha permitido la integración de alumnos en estudios de investigación de ambas instituciones de nivel superior.

Desde el año 2022 a través de las TIC, se han generado proyectos que intentan dar respuesta a temáticas neuropsicológicas como la cognición social en los jóvenes y su relación con la sintomatología depresiva; el consumo de sustancias y su relación con la ansiedad; la estimulación cognitiva en adultos mayores; y la comprensión lectora y su relación con el funcionamiento ejecutivo de los universitarios. Todos ellos bajo los proyectos registrados formalmente como PVVC: 76302 proyecto de evaluación teleneuropsicológica (2022), 85441 proyecto de intervención teleneuropsicológica (2023) y actualmente 91561 proyecto de investigación clínica en teleneuropsicología (2024). Los PVVC y las prácticas profesionales han permitido la participación de un total de 22 alumnos de la UABC y 20 alumnos de la UAEM de las carreras en psicología respectivamente. Gracias a los programas PVVC y el uso de las TIC, se han beneficiado 65 adultos mayores, 432 universitarios y 189 adolescentes y jóvenes de diferentes estados de la república como Michoacán, San Luis Potosí, Morelos, Guanajuato, Baja California Sur y Baja California. Los resultados de dichas investigaciones se han presentado en congresos nacionales e internacionales de forma virtual y presencial bajo la modalidad de carteles científicos y ponencias, así como la elaboración de tesis para titulación de Licenciatrua y la preparación actual de publicaciones de dos capítulos de libro y dos artículos científicos. La tabla 2 da muestra de lo anterior.

Tabla 2. Resultados del trabajo teleneuropsicológico a través de PVVC y el uso de las TIC.

PVVC	Estudiantes involugrados	Objetivo	Beneficiados	Productos Científicos
76302 (2022)	9 UABC 6 UAEM	Conocer la relación entre cognición social depresión en universitarios de la UAEM, UABC, UMSNH, UASLP y UG. Conocer la relación entre consumo de sustancias y la ansiedad en adolescentes y jóvenes	Universitarios Jóvenes Adolescentes	Poencia: IX Reunión Nacional del Sistema Mexicano de Investigadores en Psicologia. UABC 2022 (Presencial) Ponencia: IV congreso Internacional de Investigación Transdisciplinar en Ciencias Humanas UAEM, 2022 (Virtual) Tesis en proceso
85441 (2023)	7 UABC 10 UAEM	Estimular las funciones cognitivas a través de programas virtuales a adultos mayores con deterioro cognitivo leve	Adultos mayores	Ponencia Primer Coloquio y Segundo Encuentro de Egresados de la Maestría de Neuropsicología. BUAP. 2023 (Presencial) Cartel: Talleres y Conferencias del Encuentro Internacional de Neurociencias. UdeG. 2023 (presencial) Carteles: I Congreso Estudiantil de Psicología (COEPSI) UABC 2023 Tesis en procesos
91561 (2024)	6 UABC 4 UAEM	Conocer la relación entre comprensión lectora y el funcionamiento ejecutivo en universitarios de la UAEM y la UABC	Universitarios	Proyecto en proceso

Nota: Elaboración propia a través de la busqueda de información de PVVC registrados del 2022 a la fecha.

Los proyectos mencionados se han llevado a cabo 100 % desde la virtualidad, haciendo uso de las TIC con el apoyo de estudiantes que se han capacitado para la ejecución y aplicación de los instrumentos, métodos, tareas y técnicas neuropsicológicas diseñadas para el trabajo desde la presencialidad remota. Algunos instrumentos utilizados son: el Test Barcelona (2022), Batería Neuropsicológica Computarizada de Tamizaje BNCT (2022) y Cambridge Neuropsychological Test Automated Battery CANTAB (2006).

Conclusiones

En este capítulo ha realizado una revisión de los avances y la utilización de las TIC desde la neuropsicología, tomando en cuenta las condiciones actuales socioculturales, ambientales, educativas y exigencias del cambio mundial. Esto ha permitido plantearse una nueva forma de aplicación y generación de conocimiento desde la teleneuropsicología, como es el caso de la evaluación, diagnóstico e intervención (rehabilitación o estimulación); abarcando cada vez más múltiples trastornos cognitivos, emocionales y conductuales por daño o disfunción cerebral. Con el trabajo teleneuropsicológico se incrementa la posibilidad de llegar a más personas de la población ahorrando tiempo, economía y desplazamiento geográfico a quien requiere los servicios del neuropsicólogo, ofreciendo la misma calidad en el servicio como si fuese de manera presencial. La introducción de las TIC en el campo de la neuropsicología permite que los pacientes con dificultades de desplazamiento motor o sensorial puedan acceder a recibir tratamientos óptimos que sin el uso de la virtualidad sería imposible. La interacción a través de la presencialidad remota facilita la motivación, la independencia y la calidad de vida de los usuarios del servicio.

Es importante destacar que la teleneuropsicología muestra algunas desventajas, no obstante, su efecto negativo es mínimo en comparación de los beneficios que reporta. Uno de los retos en un futuro próximo es el correcto y buen manejo de las TIC en el trabajo teleneuropsicológico, es decir, se requiere la correcta y constante capacitación de los profesionistas para un adecuado manejo de los métodos que ofrece. Las TIC día a día cambian, se modifican y transforman. El neuropsicólogo que hace uso de la virtualidad debe estar en constante aprendizaje sin olvidar actuar con responsabilidad social y ética profesional para la reintegración de los pacientes a sus actividades de la vida cotidiana y que mejor que haciéndolo desde la teleneuropsicología.

Referencias

Aguilar, D., Juʔstiz, M. y Martiʔnez, G. (2019). Software educativo para estimular procesos cognitivos en escolares con funcionamiento intelectual limítrofe. *DELECTUS*, 3(2) 2410-28. https://doi.org/10.36996/delectus.v2i2.28

Álvarez. J., Andrade, C., Vieyra, V., Esquivias, H., Merlín, I., Neria, R. y Bezanilla, J. (2020). *Manual operativo del curso emergente para la brigada de atención psicoemocional y psicosocial a distancia durante la pandemia de la COVID-19 en México*. Módulo 1 http://inprf.gob.mx/ensenanzanew/archivos/2020/manual_brigadas_2020.pdf

American Psychological Association. (2018). *Monitor on Psychology. 2019 trends report*. American Psychological Association. https://www.apa.org/ members/content/2019-ten-trends.pdf

Benton, A. (1967). *Introduccio?n a la Neuropsicología*. Barcelona, Fontanella.

Bilder, R., Postal, K., Barisa, M., Aase, D., Cullum, C., Gillaspy, S. et al. (2020). Inter Organizational Practice Committee Recommendations/ Guidance for Teleneuropsychology in Response to the COVID-19 Pandemic. *Arch Clin Neuropsychol.* 35(6), 647-59. https://doi.org/10.1093/arclin/acaa046

Blog CogniFit. (30 de enero del 2024). *Cómo aprende el cerebro. Factores que favorecen el aprendizaje*. https://blog.cognifit.com/es/como-aprende-el-cerebro-neuroeducacion/

Botella, C., y Bretón-López, J. (2014). Uso de las tecnologías de la información y la comunicación en psicología clínica. *Revista de Psicopatología y Psicología Clínica*, 19(3), 149-156. DOI: http://dx.doi.org/10.5944/rppc.

Bringas, M., Fernández, Y., García, M., Ruiz, E., Casabona, E., Ferna?ndez, E. et al. (2009). La Neuropsicología en Cuba. *Revista Neuropsicologia, Neurociencias y Neuropsiquiatria,* 9(2) 953-76. http://revistaneurociencias.com/index.php/RNNN/article/view/211

Bruna, O., Roig, T., Puyuelo, M., Junqué, C. y Ruano, Á. (2011). Rehabilitación neuropsicológica. Intervención y práctica clínica. USA. Editorial Elsevier

Capuron, L, Ravaud, A. y Dantzer R. (2001). Precocity and specificity of the cognitive changes induced by interleukins-2 and interferon-alpha treatments I cancer patients. *Psychosom Medical*, 63, 376-386. https://doi.org/10.1097/00006842-200105000-00007

Castilla, A., Carbonell, L., Ramos, Y., Quintana, L., Moscote, L. (2019). Aplicaciones mo?viles en las neurociencias: un nuevo aliado. *Revista Chilena Neurocirurgia*, 45(2) 136-9. https://doi.org/10.36593/rev.chil.neurocir.v45i2.125

Cedeño Escobar, M., Ponce Aguilar, E., Lucas Flores, Y. y Perero Alonzo, V. (2020). Classroom y Google Meet, como herramientas para fortalecer el proceso de enseñanza-aprendizaje (En línea). *Polo del Conocimiento*, 5(7), 388-405. https://polodelconocimiento.com/ojs/index.php/es/article/view/1525

Cobos, M., Bueno, A., Herrera, S. y Coello, F. (2021). *Investigaciones en neuropsicología y salud mental.* España. Casa Editorial Universidad del Azuay.

Dwolatzky, T., Whitehead, V., Doniger, G.M. et al. (2004). Validity of the Mindstreams™ computerized cognitive battery for mild cognitive impairment. *Journal Molecular Neuroscience*, 24, 33–44 https://doi.org/10.1385/JMN:24:1:033

Elwood, D. y Griffin, R. (1972). Individual intelligence testing without the examiner. *Journal Consult Clinical Psychology,* 38(1) 9-14. https://awspntest.apa.org/doi/10.1037/h0032416

Fernández, E., Bergado, J., Rodríguez, D., Salazar, S., Torres, M. y Bringas, M. (2018) Effectiveness of a Computer-Based Training Program of Attention and Memory in Patients with Acquired Brain Damage. *Behavor Science,* 8(4), 3-12. https://doi.org/10.3390/bs8010004

Fernández, E., Fernández, Y. y Crespo, N. (2020). Interacción de las tecnologías de la información y la comunicación en la intervención neuropsicológica. *Revista Cubana de Información en Ciencias de la Salud*, 31(3), 1-17. https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=101264

García, A. y Enseñat, A. (2019). La rehabilitación neuropsicológica en el siglo XX. *Revista de Neurologia*, 69, 383-91. https://doi.org/10.33588/rn.6909.2019247

González C. (2008) Social Behavior Simulator. Generación y aplicación de un ser humano simulado para el estudio de la interacción social diádica. *Revista Iberoamericana de Inteligencia Artificial*, 12(38), 61-73. https://www.redalyc.org/articulo.oa?id=92513102004

Kolb, B. y Whishaw, I. (2009). *Fundamentals of human neuropsychology*. MacMillan.

Lezak, M., Howieson, D., Loring, D. y Fischer, J. (2004). *Neuropsychological assessment*. Oxford University Press.

Martín, J. (2012). The believue intelligence tests (Wechsler, 1939): ¿Una medida de la inteligencia como capacidad de adapación? *Revista de Historia de la Psicología*, 33(3), 49-66. https://dialnet.unirioja.es/servlet/articulo?codigo=5450370

Ministerio de Tecnologías de la Información y las Comunicaciones (2024). Artículo 6. Ley 1341. Revisado de https://normograma.mintic.gov.co/mintic/docs/pdf/ley_1341_2009.pdf

Norup, A. (2018). Family matters in neurorehabilitation: why, when, who, and how? *Revista Iberoamericana de Neuropsicología*, 1(19), 17-31. https://www.researchgate.net/publication/323520354

Perez, P., Ramos, D. y Arango, J. (1921). Teleneuropsicología en países de habla hispana: Una mirada crítica al uso de Tecnologías de Información y Comunicación en la evaluación neuropsicológica. *Revista Iberoamerica de Neuropsicología,* 4(1) 1-27 https://neuropsychologylearning.com/wp-content/uploads/pdf/pdf-revista-vol4/vol4-n1-2-18ene21.pdf

Portellano, J. A. (2005). *Introducción a la neuropsicología*. México. McGraw-Hill.

Powell, D., Kaplan, E., Whitla, D., Weintraub, S., Catlin, R. y Funkenstein, H. (1996). *Microcog: Assessment of Cognitive Functioning* (Version 2.4) [Computer software]. San Antonio, TX. The Psychological Corporation.

Restrepo, A. y Lopera, M. (2019). Recursos informáticos y discapacidad intelectual: Aplicaciones en el contexto escolar. *Revista Iberoamericana de psicología*, 11(3), 73–83. https://doi.org/10.33881/2027-1786.rip.11306

Restrepo D. (2009). La pra?ctica neuropsicológica asistida por computadora: Un escenario para el diálogo interdisciplinario entre la tecnología y las neurociencias. *Revista CES Psicologia*, 2(1), 79-90. http://www.redalyc.org/articulo.oa?id=423539414007

Robbins, T. y Sahakian, J. (1994). Computer methods of assessment of cognitive function. In John Wiley & Sons Ltd., Chichester (Eds), *Principles and Practice of Geriatric Psychiatry* (pp 205-209).

Rojas, M. (2019). *Familia en el daño cerebral ¿cómo puede ayudar? NeuroClass.* Recuperado 13 February 2024, a partir de https://neuro-class.com/rol-de-familia-en-un-dano-cerebral-adquirido/

Rodríguez, C., Ortega, E. y Sánchez, E. (2021). *Aplicación de nuevas tecnologías en personas mayores con trastorno cognitivo leve-moderado desde la Terapia Ocupacional.* Innoeduca: Internat J Technol Educat Innov, 3(1) 75-84. https://dialnet.unirioja.es/servlet/articulo?codigo=6027712

Román, F. (2004). Presentación del número monográfico: "líneas de investigación actuales en neuropsicología·. *Anales de Psicología / Annals of Psychology,* 20(2), 173–174. https://revistas.um.es/analesps/article/view/27291

Saxton, J., Morrow, L., Eschman, A., Archer, G., Luther, J., y Zuccolotto, A. (2009). *Computer Assessment of Mild Cognitive Impairment.* Postgrad Medical, 121(2), 177-185. https://doi.org/10.3810/pgm.2009.03.1990

Secretaría de Salud Pública. (2020). *Declaración de la pandemia por COVID-19*. Recuperado de https://www.gob.mx/salud/documentos/covid19

Schruijer, S. y Stephenson, G. (2010). Trends and development in community and applied social psychology: JCASP 1991-2010. *Journal of Community & Applied Social Psychology*, 20, 437 444. http://psyc604.stasson.org/Schruijer.pdf

Sigmundsdottir, L., Longley, W. y Tate, R. (2016). Computerized cognitive training in acquired brain injury: A systematic review of outcomes using the International Classification of Functioning (ICF). *Neuropsych Rehabilitation,* (26) 5-16. https://doi.org/10.1080/09602011.2016.1140657

Soto, F., Franco, M. (2019). Atencio?n psicolo?gica y tecnologi?as: oportunidades y conflictos. *Revista Iberoamericana Psicologia*, 1(3), 14-24. https://doi.org/10.33881/2027-1786.rip.11310

Strauss, E., Sherman, E. & Spreen, O. (2006). *A Compendium of europsychological Tests: Administration, forms, and Commentary* (Hardcover). New York: Oxford University Press.

Stricker, H., Lundt, E., Albertson, S., Machulda, M., Pudumjee, S., Kremers, K., Jack, Jr., Clifford, R., Knopman, D., Petersen, R. y Mielke, M. (2020). Diagnostic and Prognostic Accuracy of the Cogstate Brief Battery and Auditory Verbal Learning Test in Preclinical Alzheimer's Disease and Incident Mild Cognitive Impairment: Implications for Defining Subtle Objective Cognitive Impairment. *Journal of Alzheimer's Disease*, 76(1), 261-274. https://doi.10.3233/JAD-200087

Sullivan, K., & Bowden, S. C. (1997). Which tests do neuropsychologists use? *Journal of Clinical Psychology*, 53(7), 657-661. https://doi.org/10.1002/(SICI)1097-4679(199711)53:7<657::AID-JCLP3>3.0.CO;2-F

Terradillos, E., Solesio, E., Benito, M., Castrillo, A., Mangas, B., Almeri?a, A. et al. (2012). *La Clasificación Internacional del Funcionamiento (CIF) en la elaboración de un protocolo de evaluación del daño cerebral.* Traum Fund MAPFRE, 23(2), 123-9. https://www.researchgate.net/publication/277260435_La_Clasificacion_Internacional_de_Funcionamiento_CIF_en_la_elaboracion_de_un_protocolo_de_evaluacion_del_dano_cerebral

Wechsler, D. (2012). *Wechsler Intelligence Scale for Children—Fourth Edition: Technical and interpretative manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2005). *Manual de aplicación y corrección del WISC-IV*. Madrid: TEA Ediciones, S.A.

Wechsler, D. (2012). Manual de aplicacio?n y correccio?n del WAIS-IV. Madrid: TEA Ediciones, S. A.

Wilson, B. (2013). Neuropsychological rehabilitation: State of the science. South Africa *Journal Psycholy,* 43(3), 267-77. http://dx.doi.org/10.1177/0081246313494156

Wittson, C., Affleck D. y Johnson V. (1961) Two-way television in group therapy. *Mental Hospital Psychiatry*, 12, 22-3 https://doi.org/10.1176/ps.12.11.22

Acerca de los autores

José Antonio García Macías

Doctor por el Institut National Polytechnique de Grenoble (INPG), Francia. Investigador Titular en el Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), BC, en el Departamento de Ciencias de la Computación. Miembro del Sistema Nacional de Investigadores (nivel 1) y Miembro Titular de la Academia Mexicana de Computación (AMEXCOMP). Coordinador de la RED JEDI-SD. Contacto: jagm@cicese.mx

Isaac de Jesús Palazuelos Rojo

Doctor en Estudios Culturales por el Colegio de la Frontera Norte. Profesor-Investigador de Tiempo Completo y Coordinador de la Licenciatura en Ciencias de la Comunicación en la Facultad de Ciencias Administrativas y Sociales de la Universidad Autónoma de Baja California. Pertenece al Sistema Nacional de Investigadores, Nivel 1. Es miembro de la Asociación Mexicana de Investigadores de la Comunicación y forma parte del Cuerpo Académico "Culturas Contemporáneas y Discursividades".

contacto: isaac.palazuelos@uabc.edu.mx

Diego Alfredo Pérez Rivas

Doctor en Filosofía por la Universidad Complutense. Posdoctorado en Filosofía de la Ciencia por la Universidad de Turín. Profesor-investigador en la FCAyS de la Universidad Autónoma de Baja California. Miembro SNII, nivel 2. Coordinador de la "Red JEDI-SD". Miembro del cuerpo académico "Planeación y Desarrollo UABC-CA-100". Contacto: diego.alfredo.perez.rivas@uabc.edu.mx

Virginia Guadalupe López Torres

Doctora en Ciencias Administrativas por la UABC. Profesor-investigador en la FCAYS. Miembro del SNI, nivel 2. Miembro Titular de la Academia Mexicana de Investigación Turística, A.C. Miembro Académico Titular de la Academia de Ciencias Administrativas, actualmente miembro del Comité de Honor y Justicia, de 2012 a 2015 fungió como Presidente. Líder del Cuerpo Académico Planeación y Desarrollo UABC-CA100. Contacto: virginia.lopez@uabc.edu.mx

Antonio de Jesús García Chávez

Licenciado en ciencias computacionales por la Universidad Autónoma de Baja California (UABC) y Maestro en ciencias computacionales por el Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), actualmente finalizado sus estudios de doctorado en la Maestría y Doctorado en Ciencias e Ingeniería (MYDCY) de la UABC. Maestro de asignatura en la UABC, así como en CETYS Ensenada. Sus intereses de investigación son el procesamiento del lenguaje natural y la comprensión del lenguaje natural. Correo: antonio.garcia24@uabc.edu.mx

Denys Serrano Arenas

Doctora en Ciencias en Educación Agrícola Superior por la Universidad Autónoma Chapingo. Posdoctorado en el programa de posgrado en Educación para la Ciudadanía en la Universidad Autónoma de Querétaro. Miembro del Sistema Nacional de Investigadores e Investigadoras, nivel 1, asimismo pertenece al comité directivo en la figura de enlace de la Red Latinoamericana de Convivencia Escolar. Actualmente es profesora investigadora en el programa de Ciencias de la Educación de la Universidad Autónoma de Baja California. denys.serrano@uabc.edu.mx

Arturo Serrano-Santoyo

Doctor en Ciencias en Ingeniería Eléctrica en 1980 por el Centro de Investigación y Estudios Avanzados (CINVESTAV). En 1981 y 1985 reci-

bió los premios nacionales de Electrónica y Telecomunicaciones de ALCATEL y ERICSSON respectivamente y en 2016 el Premio Nacional de Ingeniería del Colegio de Ingenieros Mecánicos y Electricistas. Es miembro de la Academia Nacional de Ingeniería y del SNII. Actualmente es profesor de Innovación y Economía Digital en la UABC. arturoserrano@uabc.edu.mx

Marisol Flores-Garrido

Doctora en Ciencias Computacionales por el Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE). Actualmente es profesora de tiempo completo en la Universidad Nacional Autónoma de México campus Morelia y miembro del SNII. Sus líneas de estudio se centran en el reconocimiento de patrones, el análisis de redes y la intersección entre la inteligencia artificial y las humanidades. Contacto: mflores@enesmorelia.unam.mx

Christian Fernández Huerta

Doctor en Estudios del Desarrollo Global y maestro en Estudios Socioculturales. Miembro del Sistema Nacional de Investigadoras e Investigadores (SNII 1). Actualmente es Investigador de Tiempo Completo del Instituto de Investigaciones Culturales-Museo de la Universidad Autónoma de Baja California. Contacto: christian@uabc.edu.mx

Hugo Méndez-Fierros

Doctor en Estudios del Desarrollo Global y maestro en Ciencias de la Comunicación. Especializado en estudios de comunicación, cultura y migraciones en la frontera México-Estados Unidos. Miembro del Sistema Nacional de Investigadoras e Investigadores (SNII 1). Es presidente de la Academia de ciencias sociales 2023-2025 de la UABC y líder del cuerpo académico Culturas contemporáneas y discursividades PRODEP-UABC. Contacto: hugomendez@uabc.edu.mx

Joaquín Caso Niebla

Maestro y Doctor en Psicología Educativa por la Facultad de Psicología de la Universidad Nacional Autónoma de México (UNAM). Es investigador titular del Instituto de investigación y Desarrollo Educativo (IIDE) de la UABC. Miembro del Sistema Nacional de Investigadoras e Investigadores (SNII 1). Actualmente se desempeña como Secretario General de la UABC. Es miembro de la Sociedad Mexicana de Psicología desde 1995, en la cual fungió como Presidente del 2009 al 2013. Contacto: jcaso@uabc.edu.mx

Ulises Suárez Estavillo

Doctor y Maestro en Historia, por la Universidad Autónoma de Sinaloa. Licenciado en Economía por la misma institución. Profesor-Investigador de Tiempo Completo Titular C de la Facultad de Ciencias Sociales de la UAS, profesor de estudios culturales y estudios de socialización en la Maestría en Ciencias Sociales con Énfasis en Estudios Regionales (UAS). Se ha desempeñado como Coordinador de Investigación y Posgrado en su Facultad de adscripción. Candidato a Sistema Nacional de Investigadores, Perfil deseable PRODEP.

Héctor Ignacio Castañeda García

Ing. en Teleinformática por ITESUS. Es Maestro y Doctor en Educación por CESSSIN. PROF. E INV. T.C. de la UACEM de la Universidad Autónoma de Sinaloa. Actualmente imparte cátedra a nivel licenciatura en los programas educativos de Licenciatura en Ciencias de la Educación, LEM en el Área del Español y LEM en el Área de Matemáticas. Correo electrónico: hcastaneda@uas.edu.mx

Jesús Manuel Niebla Zatarain

Licenciado en Derecho por la Facultad de Derecho Mazatlán y Licenciado Informática por la Facultad de Informática Mazatlán dependientes de la Universidad Autónoma de Sinaloa. Es Maestro en Administración

de Empresas Turísticas por el Centro de Estudio de Los Cabos. Doctor en Derecho por la Universidad de Edimburgo, Escocia, Reino Unido. Es Miembro del Sistema Nacional de Investigadores Nivel 1 y Responsable Técnico del Proyecto Ciencia de Frontera "Regulación de Entornos Digitales a través de razonamiento legal basado en inteligencia artificial" CF-2023-G-772. Profesor Investigador de Tiempo Completo Facultad de Derecho Mazatlán de la Universidad Autónoma de Sinaloa. Correo electrónico: j.niebla@uas.edu.mx.

Virginia Berenice Niebla Zatarain

Licenciada en Sistemas Computacionales por la Universidad de Occidente. Maestra en Informática Administrativa por la Universidad Autónoma de Durango. Es Doctora en Tecnología Educativa por el Centro Universitario Mar de Cortés. Es profesora investigadora del Instituto Tecnológico de Estudios Superiores de Los Cabos donde actualmente es Subdirectora de Investigación y Posgrado. Es Candidata a Investigador por el Sistema Nacional de Investigadores. Participante del Proyecto Ciencia de Frontera 2023 Proyecto CF-2023-G-772 "Regulación de entornos digitales a través de razonamiento legal basado en inteligencia artificial". Correo electrónico: vbniebla@hotmail.com.

Gonzalo Armienta Hernández

Licenciado en Derecho por la Universidad Nacional Autónoma de México. Especialidad en impuestos por la Fundación para el Desarrollo Internacional de Berlín Alemania y en impuestos por el Instituto Tecnológico Autónomo de México. Doctor en Derecho por la Universidad Nacional Autónoma de México. Es miembro del Sistema Nacional de Investigadores Nivel II. Participante del Proyecto Ciencia de Frontera 2023 Proyecto CF-2023-G-772 "Regulación de entornos digitales a través de razonamiento legal basado en inteligencia artificial". Profesor Investigador de Tiempo Completo Facultad de Derecho Culiacán de la Universidad Autónoma de Sinaloa Correo electrónico: armienta@uas.edu.mx.

Alfredo Leonardo Carreón Corona

Licenciado en Sociología por la Universidad Autónoma de Sinaloa (UAS) y Maestro en Ciencias en Metodología de la Ciencia por el Centro de Investigaciones Económicas, Administrativas y Sociales del Instituto Politécnico Nacional (CIECAS-IPN). Actualmente forma parte del equipo de investigación Aprendizaje de aritmética y álgebra del siglo XXI (APARAL XXI) y se desempeña como profesor de asignatura de la Facultad de Ciencias Sociales (FACISO) y de la Unidad Académica de Ciencias de la Educación (UACEM), ambas dependientes de la UAS. Correo: lcarreon@uas.edu.mx

Luis Gustavo Kelly Torreblanca

Licenciado en Sociología por la Universidad Autónoma de Sinaloa (UAS), Maestro en Estudios Urbanos por el Centro de Estudios Demográficos, Urbanos y Ambientales (CEDUA) de El Colegio de México (COLMEX) y Doctorante por la Facultad de Ciencias Económicas y Sociales (FACES) de la Universidad Autónoma de Sinaloa (UAS). Es profesor de asignatura en la Facultad de Ciencias Sociales (FACISO, UAS). Se interesa en los temas de políticas públicas, gestión de la ciudad y transparencia. Forma parte de la Red de Rendición de Cuentas con sede en el Instituto de Investigaciones en Rendición de Cuentas y Combate a la Corrupción (IIRCCC - UDG). Correo: lkelly@uas..edu.mx

Guillermo Hernández-Santana

Lingüista por la ENAH. Maestro en Estudios Mesoamericanos y Doctor en Antropología por la UNAM. Su trabajo se enfoca en la descripción gramatical de la Lengua de Señas Mexicana (LSM) y la promoción de la competencia lingüística en la comunidad Sorda. Profesor investigador en la Universidad Autónoma de Baja California (UABC) en la carrera de Ciencias de la Comunicación. SNII candidato. Correo: guillermo.santana@uabc.edu.mx

Irvin Hussein López-Nava

Maestro y Doctor en Ciencias de la Computación por el Instituto Nacional de Astrofísica, Óptica y Electrónica. Investigador en el Departamento de Ciencias de la Computación en el Centro de Investigación Científica y de Educación Superior de Ensenada. Además es Profesor de asignatura en la Facultad de Ciencias en la Universidad Autónoma de Baja California. Miembro del Sistema Nacional de Investigadoras e Investigadores y de la Academia Mexicana de Computación. Correo: hussein@cicese.mx

Oscar Méndez García

Licenciado en Informática y Maestro en Administración por la Universidad Autónoma de Baja California y actualmente es doctorante en el mismo programa. En paralelo, imparte cátedra a nivel licenciatura en los programas educativos de Licenciatura en Informática y Licenciatura en Inteligencia de Negocios en la Facultad de Ciencias Administrativas de la Universidad Autónoma de Baja California, Campus Ensenada. Contacto: oscar.mendez.garcia@uabc.edu.mx

Jessica Beltrán Márquez

Recibió su grado de maestría y doctorado en Ciencias de la Computación en el Centro de Investigación Científica y de Educación Superior de Ensenada. Actualmente es profesora investigadora en el Centro de Investigación en Matemáticas Aplicadas de la Universidad Autónoma de Coahuila. Es nivel I en el SNII. Sus intereses de investigación son el análisis del contexto y comportamiento humano para el desarrollo de aplicaciones orientadas al bienestar. Correo: jessicabeltran@uadec.edu.mx

Edgar Chávez

Es un investigador en el área de Ciencias de la Computación en CICESE. Sus principales áreas de interés incluyen la búsqueda por proximidad, reducción dimensional, clustering y la caracterización de la dimensiona-

lidad intrínseca. Desde su incorporación a CICESE en 2013, ha estado interesado en el desarrollo de algoritmos y técnicas para la representación y recuperación de datos, con un enfoque particular en el uso de métricas espaciales para la indexación y búsqueda eficiente en grandes volúmenes de datos.

elchavez@cicese.edu.mx

Mariano Rivera

Doctor en Ciencias (Óptica) por el CIO. Es investigador del Departamento de Ciencias de la Computación del CIMAT desde 1997. Miembro de la Sociedad Mexicana de Matemáticas y Miembro por Invitación de la Academia Mexicana de Computación. Es Nivel III del SNII-Conahcyt. Sus intereses de investigación se centran en los temas de procesamiento y visión de imágenes, optimización numérica, y aprendizaje automático. Con aplicaciones en metrología óptica, análisis de datos de teledetección y análisis de imágenes médicas. Correo mrivera@cimat.mx

Maximino Matus Ruiz

Doctor en Sociología por la Universidad de Wageningen, Holanda. Maestro en Semiótica por la Universidad de Tartu, Estonia. Maestro y Licenciado en Antropología Social por el Centro de Estudios Superiores en Antropología Social (CIESAS-Occidente) y la Escuela Nacional de Antropología e Historia (ENAH). De 2011 a 2014 se desempeñó como coordinador de la Oficina de Tecno-Antropología de Infotec. En 2015 se incorporó al Colegio de la Frontera Norte como cátedra CONAHCYT y desde 2021 es profesor-investigador del Departamento de Estudios Culturales de El Colef. Pertenece al nivel II del SNII. matus@colef.,mx

Rubén Avilés Reyes

Licenciado en Psicología por la Universidad Michoacana de San Nicolás de Hidalgo, maestría en neuropsicología por la Benemérita Universidad Autónoma de Puebla y Dr. en Psicología con énfasis en neurociencias por la Universidad Autónoma del Estado de Morelos. Incorporado al Siste-

ma Nacional de Investigadores y al PRODEP. Profesor Investigador de Tiempo Completo y Coordinador de la Licenciatura en Psicología de la Facultad de Ciencias Administrativas y Sociales de la Universidad Autónoma de Baja California UABC, campus Ensenada. Miembro Activo de la Asociación Mexicana de Neuropsicología AMN, Colegio Morelense de Psicología COMOPSI, Colegio Morelense de Posgraduados en en Neuropsicología y Neurociencias del Comportamiento A.C. COLNNEC y Society For Neurociencce SFN.

Edgar Adrián Arias Chávez

Licenciado en Psicología por la Universidad Autónoma del Estado de Morelos (UAEM) y maestrando en Psicología Aplicada por la Universidad Autónoma de Baja California (UABC). Docente en la Licenciatura en Psicología de la Facultad de Ciencias Administrativas y Sociales de la Universidad Autónoma de Baja California, campus Ensenada. Su participación en investigaciones han sido referentes a la sintomatología depresiva, funciones ejecutivas, toma de decisiones, funcionamiento frontal, consumo de sustancias psicoactivas y rendimiento académico. Contacto: arias.edgar12@uabc.edu.mx

Javier Tadeo Sánchez Betancourt

Doctor en neurociencias por la Universidad Nacional Autónoma de México y Licenciado en Psicología por la misma Universidad UNAM. Investigador por la UNAM. Pertenece al Sistema Nacional de Investigadores nivel I.

Juan José García Llamas

Doctor en Educación (ITESO), maestro en Tecnologías para el Aprendizaje (U. de G.) y licenciado en Diseño Gráfico y Comunicación Visual (U. de G.). A lo largo de su carrera, ha contribuido al diseño de programas educativos variados, incluyendo materias, diplomados y talleres, tanto en modalidades presenciales como en línea para diferentes instituciones educativas de

México. Ha creado diverso material educativo, como infografías, videos, podcasts y plataformas digitales. Además, está centrado en investigar temas relevantes relacionados con la neuro psicología, la inclusión de discapacidades y la inteligencia artificial en la educación.

Inteligencia Artificial: transformación, retos y prospectiva social
Se terminó de editar en febrero de 2024
en los talleres gráficos de **Astra Ediciones**
Av. Acueducto 829, Colonia Santa Margarita,
C.P. 45140, Zapopan, Jalisco

En la presente obra, se intenta abonar a una discusión que trascienda la superficial dualidad entre filia y fobia. Por esa misma razón, se ha invitado a estudiosos de los más diversos campos de estudio con la finalidad de proponer una hoja de ruta para entender el impacto social de la inteligencia artificial.

Solamente a través de un camino con estas características, nos será posible romper el caparazón de las simplificaciones, para darnos cuenta que el verdadero estado del arte es más profundo que la dicotomía.

En el corazón de la inteligencia artificial existen verdaderamente posibilidades técnicas de resolver problemas que acechan desde hace mucho tiempo a la humanidad, pero también plantea algunos riesgos que no podemos hacer a un lado.

A lo largo de los apartados del presente libro, los autores plantean estos diversos escenarios en el campo de la ecología, la educación, la psicología, la agricultura, la medicina y hasta la filosofía o la historia.

ISBN: 978-84-10215-89-4

9 788410 215894

astra editorial

RED JEDI SD