

La Inteligencia Artificial y sus Implicaciones

Juan M. Gutiérrez-Nava

Facultad de Ciencias Básicas, Ingeniería y Tecnología
Universidad Autónoma de Tlaxcala
Tlaxcala, México
juanmario.gutierrezn@gmail.com

Rocio Ochoa-Montiel

Facultad de Ciencias Básicas, Ingeniería y Tecnología
Universidad Autónoma de Tlaxcala
Tlaxcala, México
<http://orcid.org/0000-0002-0147-7666>,
ma.rocio.ochoa@gmail.com

Resumen—Actualmente, la dinámica en el cambio en diversos ámbitos a nivel mundial es debida al auge en el desarrollo de la Inteligencia artificial (IA) y su uso a gran escala. Las herramientas, sistemas y aplicaciones basadas en IA se encuentran en todas partes, en su mayoría facilitando diversas tareas. No obstante, el pleno conocimiento de esta disciplina es limitado en algunos sectores, haciendo difícil la creación y el uso consciente de los productos generados por las IAs. En este trabajo se presenta una revisión del impacto, retos e implicaciones de la Inteligencia Artificial para un uso ético y responsable en beneficio de la sociedad.

Index Terms—inteligencia artificial, explicabilidad, inteligencia artificial responsable

I. INTRODUCCIÓN

El creciente avance tecnológico ha transformado radicalmente diversos aspectos de nuestra sociedad, desde la forma en que interactuamos con la tecnología hasta la toma de decisiones en campos tan diversos como la educación, la medicina, la cultura y la agricultura, entre otros. Como consecuencia de este avance y del surgimiento de campos del conocimiento como la Inteligencia Artificial (IA) se ha generado un impacto significativo en nuestra vida cotidiana y en la economía global, planteando preguntas sobre su uso de forma ética y responsable [1], por lo que es fundamental explorar su impacto en diversos entornos en la vida del ser humano.

Aunque no existe una definición universalmente aceptada de la IA, generalmente la Inteligencia artificial se considera como una disciplina dentro de la computación que tiene por objetivo dotar a las máquinas de cierto grado de inteligencia similar a la que exhiben los seres humanos o las diferentes formas de vida animal o vegetal. Si bien la IA ha fomentado la eficiencia en diversos sectores, también ha generado incertidumbre sobre la pérdida de empleos y la desigualdad económica que existe hoy en día [2].

La comunidad internacional está de acuerdo en algo; es necesario desarrollar políticas que mitiguen estos efectos negativos y promuevan una integración responsable de la IA en la sociedad. La regulación de la IA es un tema complejo que abarca aspectos legales y éticos. Es necesario establecer marcos regulatorios que fomenten la innovación en IA mientras se protegen los derechos y la privacidad de las personas [3]. La colaboración internacional es esencial para abordar los desafíos éticos y legales para garantizar que su desarrollo beneficie a la sociedad en su conjunto.

En este trabajo se revisan aspectos que en conjunto permiten generar un criterio respecto al estado actual de la Inteligencia artificial, su influencia en diversos sectores y algunos aspectos a considerar en el desarrollo y uso de tecnología basada en esta disciplina. En las siguientes secciones se presenta el análisis de los rubros considerados en la revisión del impacto, retos e implicaciones de esta disciplina.

II. ANTECEDENTES

El desarrollo de la IA ha evolucionado notablemente desde sus inicios hasta convertirse en una de las tecnologías más disruptivas de la actualidad. La IA como área de estudio, debe sus fundamentos a casi un siglo de investigaciones y desarrollos pluridisciplinarios en distintas ramas de la ciencia, desde la filosofía, la matemática, la psicología, la lingüística, la física e ingeniería [4]. Hoy en día, la IA se utiliza en sectores tan variados como la medicina, la educación, la industria, las finanzas, etc; impulsando cambios significativos en la vida de las personas y en las organizaciones.

II-A. Inteligencia Artificial

En los últimos años, con el desarrollo de la electrónica y áreas afines se han creado dispositivos de menor volumen y mayor desempeño haciendo posible el desarrollo de modelos de IA como las redes neuronales y otros modelos de aprendizaje automático. Esto ha fomentado que la IA se asuma como un sinónimo de “Aprendizaje automático” [5]. El aprendizaje automático y el aprendizaje profundo son dos enfoques comúnmente utilizados en la Inteligencia artificial. En la Figura 1 se muestran un esquema de las principales áreas de aplicación de estos enfoques.

Por otra parte, un inconveniente importante de los modelos de IA es la dificultad de conocer el proceso en la toma de decisiones [7]. La transparencia y explicabilidad, son características fundamentales para garantizar la confianza en los sistemas de Inteligencia artificial, especialmente en aplicaciones críticas como los diagnósticos médicos asistidos por algoritmos de Inteligencia artificial. Otra área particularmente importante es el área legal, donde la evaluación correcta de criterios puede determinar el curso de vida de una o más personas.

En este sentido, la explicabilidad es una característica relevante en el desarrollo de tecnología basada en inteligencia artificial. El programa Explainable Artificial Intelligence (XAI) desarrollado por la Agencia de Proyectos de Investigación

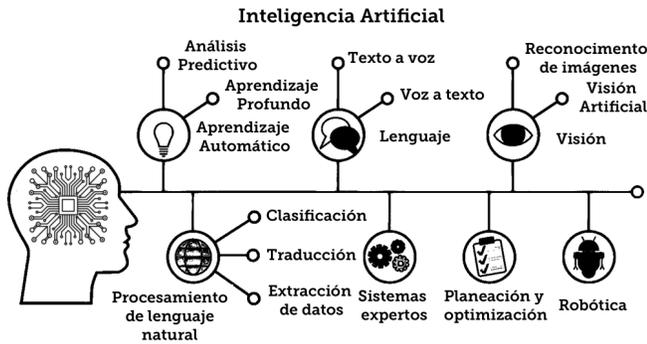


Figura 1: Áreas de aplicación de la Inteligencia Artificial [6].

Avanzados de Defensa (DARPA) busca desarrollar sistemas de IA cuyos modelos y decisiones puedan ser comprendidos y generen más confianza a sus usuarios [8]. Se enfoca en crear modelos de IA explicables, diseñar interfaces de explicación y comprender los requisitos psicológicos para explicaciones efectivas.

XAI se dirige a un usuario final que depende de decisiones o recomendaciones producidas por un sistema de IA, o acciones tomadas por él, y por lo tanto necesita comprender la justificación del sistema como se muestra en la Figura 2. El programa aborda dos áreas clave de problemas operativos: el análisis de datos y la autonomía. En el análisis de datos, se enfoca en clasificar eventos de interés en datos multimedia heterogéneos, mejorando la precisión en la identificación de patrones y eventos relevantes. En el ámbito de la autonomía, trabaja en el desarrollo de políticas de decisión para sistemas autónomos, lo que permite que estas máquinas tomen decisiones efectivas y en tiempo real, adaptándose a diversos entornos y situaciones complejas.

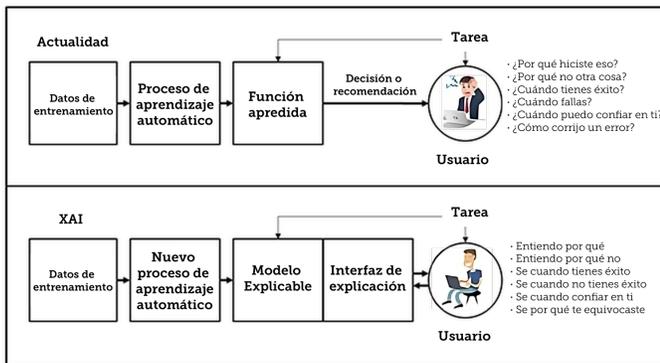


Figura 2: Diagrama de flujo programa XAI de DARPA [9].

II-B. Regulación de la Inteligencia Artificial

Las tecnologías de IA ofrecen beneficios económicos y sociales en sectores como la salud y las finanzas, mejorando la capacidad de predicción y la personalización de servicios a los usuarios. Sin embargo, generan preocupaciones sobre la

violación de derechos fundamentales, como la privacidad y la no discriminación. La Unión Europea (UE) ha impulsado una regulación centrada en el ser humano "para garantizar que la IA cumpla con los valores, principios y normativas de la UE. Para lograr esto ha tomado a una postura legislativa, abogando por normas armonizadas que rijan el desarrollo e implementación de la IA. Esto refleja un compromiso con la adopción responsable de la IA, promoviendo su uso mientras se mitigan los riesgos asociados [10].

El Proyecto de Ley de Inteligencia Artificial (AI Act) es un referente pionero a nivel mundial de la regulación horizontal de sistemas de IA. Contiene elementos cruciales como el diferenciar a los sistemas de AI según el nivel de riesgo como se muestra en la Figura 3, introducir prohibiciones y facilitar el escrutinio social a través de una base de datos pública de los sistemas [11].

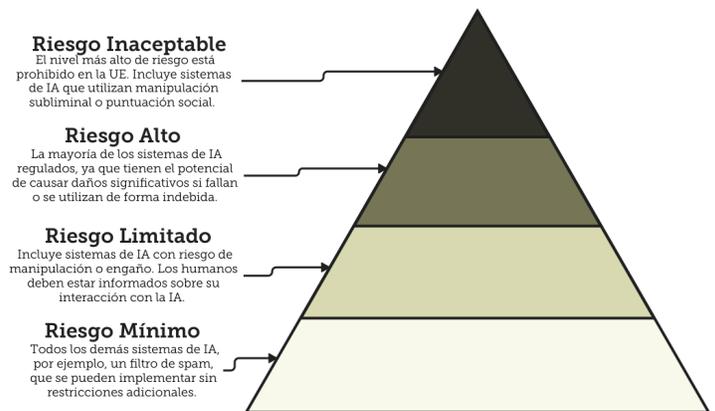


Figura 3: Pirámide de niveles de riesgo de sistemas basados en IA [12].

III. IMPACTO DE LA INTELIGENCIA ARTIFICIAL

La Inteligencia artificial está transformando nuestra sociedad en múltiples niveles. Desde la asistencia en la toma de decisiones en el sector privado hasta la personalización de la atención médica. Su impacto es evidente en la forma en que nos comunicamos, trabajamos y vivimos en la actualidad.

El incremento de aplicaciones basadas en IA en distintas industrias es el resultado del avance tecnológico asociado a sucesos como la Cuarta Revolución Industrial ya que en la actualidad los entornos productivos son más dinámicos y se encuentran cada día más interconectados. Es en este contexto donde la IA ha demostrado tener un impacto significativo ya que es una tecnología clave para realizar análisis predictivos basados en grandes cantidades de datos y ayudar en el proceso de toma de decisiones. Es por ello que el uso de IA como una tecnología clave para transformar procesos se ha extendido más allá del ámbito industrial, teniendo aplicaciones en economía, educación, salud e incluso en la política y gobernanza [13] [14].

A lo largo de la historia, el desarrollo tecnológico ha mejorado y restringido simultáneamente la capacidad de las personas

para tomar decisiones. La IA moderna está desarrollando una estructura donde la información fluye desde el futuro al presente, en lugar de fluir desde el pasado hacia el presente, como era habitual. Mediante sensores, datos y algoritmos, las máquinas logran captar información que sirve como un indicativo acerca de eventos futuros y se emplea para crear servicios y productos de tipo anticipatorio. Estas máquinas serán capaces de predecir nuestros comportamientos y deseos, gracias a un diseño orientado a la anticipación y predicción [15].

La IA ha surgido como una tecnología revolucionaria en el ámbito educativo. En la actualidad, se están investigando varias aplicaciones de la IA en el sector educativo, entre las que se incluyen la personalización del aprendizaje, la evaluación automatizada, la tutoría inteligente y la identificación temprana de dificultades para el aprendizaje en los estudiantes. La IA puede mejorar significativamente al sector educativo al ofrecer novedosos métodos de enseñanza personalizada y retroalimentación instantánea para cada estudiante. No obstante, diversos obstáculos como la falta de datos adecuados, la resistencia al cambio y la limitada interpretabilidad de los sistemas de IA deben superarse para facilitar su implementación en el ámbito educativo. Además, la IA brinda la posibilidad de hacer la educación más accesible para todos los estudiantes, incluyendo a aquellos con discapacidades y aquellos que no tienen acceso a recursos educativos de calidad [16].

III-A. Hacia una Inteligencia Artificial Responsable

El enfoque de una IA responsable es el de garantizar que los sistemas de IA sean desarrollados, implementados y utilizados de manera que se respeten los valores y legislaciones actuales. Esto se logra al considerar la transparencia, la privacidad y la seguridad en todas las etapas del ciclo de vida de un sistema de IA. El objetivo es asegurar que la IA beneficie a la sociedad en su conjunto, minimizando los riesgos potenciales y promoviendo un uso ético y sostenible de esta tecnología. En la Figura 4 se muestran seis rubros básicos para el desarrollo de una IA responsable.

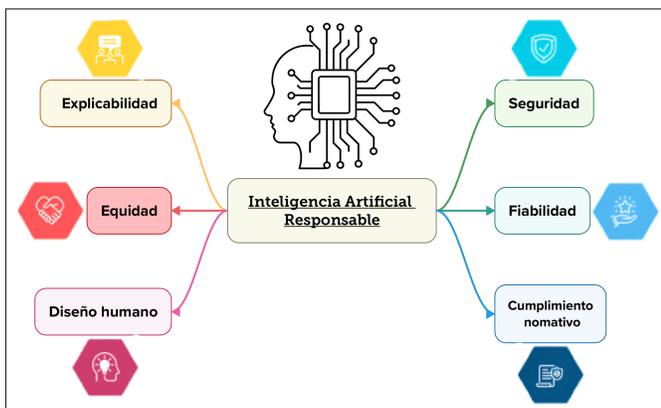


Figura 4: Representación gráfica de una Inteligencia Artificial Responsable [17].

La UNESCO acordó que el desarrollo de la IA debe estar controlado por humanos y centrado en las personas; el despliegue de la IA debe estar al servicio de las personas para mejorar las capacidades humanas y la IA debe diseñarse de manera ética, no discriminatoria, equitativa, transparente y auditable. El impacto de la IA en la sociedad y en los individuos debe ser constantemente monitoreado y evaluado a lo largo de todas las cadenas de valor [18].

La IA plantea serias preguntas del desarrollo de los estados, de las personas, de la economía, entre muchos otros aspectos sumamente relevantes en la actualidad. Los gobiernos tendrán que definir estrategias para aprovechar las oportunidades que ofrece esta tecnología y actuar con rapidez para mantenerse competitivos, estableciendo mecanismos que mitiguen los riesgos asociados al uso e implementación de la IA [2]. Esta competencia está en marcha y algunos países la están liderando mientras otros recién están tomando conciencia de la magnitud del desafío que tienen por delante [19].

IV. APORTACIONES Y TENDENCIAS

En este trabajo se han presentado diversos aspectos para comprender el concepto de IA, así como las implicaciones en su uso en nuestra sociedad. Se ha explorado el impacto que tiene esta disciplina en diversas áreas que van desde la educación hasta la salud (ver Figura 5), y el potencial de revolucionar industrias por completo. Y se han identificado tendencias clave en este campo que se encuentra en constante desarrollo.



Figura 5: Áreas de impacto de la IA en el cuidado de la salud [20].

De igual manera, identificamos la importancia de la IA explicable resaltando la necesidad de que los sistemas y los

nuevos modelos de aprendizaje basados en IA sean capaces de explicar sus procesos en la toma de decisiones, de una manera comprensible para los humanos. Esto con el propósito de mejorar la capacidad de tomar decisiones y resaltando su rol como una de las alternativas a las crecientes preocupaciones éticas que se han derivado de la IA. En cuanto a la regulación de la IA, se recalcó la importancia de que los marcos regulatorios garanticen un uso ético de la tecnología, mediante la implementación de marcos regulatorios que promuevan la transparencia y la responsabilidad en el desarrollo y uso de la IA, un ejemplo claro de estos esfuerzos internacionales es el Artificial Intelligence Act desarrollado por la unión Europea. Una de las tendencias actuales en inteligencia artificial (IA) apunta hacia un mayor enfoque en la IA explicable, donde los sistemas tendrán la capacidad explicar sus decisiones de manera comprensible y se conviertan en una herramienta para los usuarios al mejorar la toma de decisiones. Esto es crucial para generar confianza y aceptación. Además, se espera un crecimiento significativo en el uso de IA en aplicaciones de alto valor como en la atención médica, donde la tecnología puede mejorar la precisión de los diagnósticos y la eficacia de los tratamientos. Otras tendencias que tiene esta tecnología incluyen el desarrollo de aplicaciones de IA en la nube, la automatización de procesos y la personalización de la educación con contenido de alta calidad para estudiantes de cualquier lugar del mundo.

Por otra parte, en la nueva era de la IA que surgió con el desarrollo de arquitecturas de aprendizaje llamadas transformadores y que permitieron el desarrollo de grandes modelos de lenguaje que son la base de aplicaciones como ChatGPT en sus distintas versiones [21]; es notable un cambio de paradigma en el alcance que puede tener el uso no consciente de herramientas que a simple vista proporcionan respuestas similares a las obtenidas por un humano como en el caso de los modelos de Inteligencia artificial generativa cuya característica principal es la capacidad de generar datos nuevos a partir de un conjunto de datos de entrada que pueden ser de tipos distintos a los generados, como el caso de modelo CLIP que puede recibir imágenes o video como entrada, y dar como salida una descripción de dicha imagen o video [22].

V. CONCLUSIONES

La IA está cambiando indudablemente a la sociedad, transformando industrias y revolucionando la forma en que vivimos. Impactando a áreas tan diversas como la medicina, la educación, la industria y el gobierno. Sin embargo, este avance no está libre de desafíos. La necesidad de regular la IA para garantizar su uso ético y responsable es más importante que nunca. Este trabajo se ha enfocado en examinar que entendemos por Inteligencia Artificial, sus implicaciones y la importancia del desarrollo e implementación de sistemas de IA responsables.

Se presentaron detalles sobre el desarrollo de sistemas de IA explicables y su funcionamiento ya que estos tienen la capacidad de informar al usuario como es que el sistema toma una decisión. Ya que es crucial que los sistemas de IA sean

diseñados y utilizados de manera transparente, equitativa y no discriminatoria. Además, se recalcó la importancia de monitorear de cerca el desarrollo de la IA y evaluar constantemente su impacto en la sociedad.

Es de esperarse que la IA siga evolucionando y desempeñe un papel aún más importante en nuestras vidas en el futuro cercano. Por lo tanto, es fundamental saber cómo funcionan estos sistemas y comprender los beneficios y riesgos que conlleva el uso de la IA, para así abordar estos desafíos de manera proactiva y colaborativa, asegurando que la IA beneficie a toda la sociedad y se utilice para promover el desarrollo colectivo.

AGRADECIMIENTOS

Los autores agradecen el apoyo a la Universidad Autónoma de Tlaxcala y a la Red de Inteligencia Computacional Aplicada (RedICA de CONACYT) por el apoyo brindado para el desarrollo de este trabajo.

REFERENCIAS

- [1] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, pp. 82–115, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1566253519308103>
- [2] E. Parliament, "Unesco presenta reporte de evaluación del estado de preparación de inteligencia artificial de México," jul 2024. [Online]. Available: <https://www.unesco.org/es/articles>
- [3] —, "EU AI act: First regulation on artificial intelligence," jun 2024. [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- [4] G. Baños, *El sueño de la Inteligencia Artificial: El proyecto de construir máquinas pensantes: una historia de la IA*. Shackleton Books, 2024.
- [5] J. E. Ricardo, M. Y. L. Vázquez, A. J. P. Palacios, and Y. E. A. Ojeda, "Inteligencia artificial y propiedad intelectual," *Universidad y sociedad*, vol. 13, no. S3, pp. 362–368, 2021.
- [6] SwissCognitive, "What is artificial intelligence?" 2022. [Online]. Available: <https://swisscognitive.ch/2020/11/24/artificial-intelligence-2/>
- [7] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable AI: A review of machine learning interpretability methods," *Entropy*, vol. 23, no. 1, 2021. [Online]. Available: <https://www.mdpi.com/1099-4300/23/1/18>
- [8] D. Gunning and D. Aha, "DARPA's explainable artificial intelligence (XAI) program," *AI magazine*, vol. 40, no. 2, pp. 44–58, 2019.
- [9] J. Ashley, "¿Qué es la IA explicable (XAI)? | Blog de NVIDIA," 2021. [Online]. Available: <https://la.blogs.nvidia.com/blog/que-es-la-ia-explicable/>
- [10] T. Madiaga, "Artificial intelligence act," *European Parliament: European Parliamentary Research Service*, 2021.
- [11] M. Veale and F. Zuiderveen Borgesius, "Demystifying the draft eu artificial intelligence act—analysing the good, the bad, and the unclear elements of the proposed approach," *Computer Law Review International*, vol. 22, no. 4, pp. 97–112, 2021.
- [12] Trail, "EU AI Act: How risk is classified," 2023. [Online]. Available: <https://www.trail-ml.com/blog/eu-ai-act-how-risk-is-classified>
- [13] J. M. T. López, "Tendencias e impacto de la inteligencia artificial en comunicación: cobotización, gig economy, co-creación y gobernanza," *Fonseca, journal of communication*, no. 22, pp. 5–22, 2021.
- [14] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook," *IEEE access*, vol. 8, pp. 220 121–220 139, 2020.
- [15] D. Innerarity, "El impacto de la inteligencia artificial en la democracia," *Revista de las cortes generales*, vol. 109, pp. 87–103, 2020.
- [16] C. S. González-González et al., "El impacto de la inteligencia artificial en la educación: transformación de la forma de enseñar y de aprender," 2023.

- [17] B. Dooley, "Responsible AI: A 6-part blog series by ben dooley," Nov. 2023. [Online]. Available: <https://www.infocepts.ai/blog/responsible-ai-a-6-part-blog-series-by-ben-dooley/>
- [18] I. C. y. I. C. UNESCO, Organización de las Naciones Unidas para la Educación, *International Conference on Artificial Intelligence and Education, Planning Education in the AI Era: Lead the Leap, Beijing, 2019*, 2019. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- [19] M. I. Mejía Jaramillo and J. Torres Páez, "Uso responsable de la inteligencia artificial en el sector público," 2020.
- [20] T. Laurinavicius, "7 life-saving AI use cases in healthcare," Jul. 2024. [Online]. Available: <https://www.v7labs.com/blog/ai-in-healthcare>
- [21] OpenAI, J. Achiam, S. Adler, and et al., "GPT-4 Technical Report," 2024. [Online]. Available: <https://arxiv.org/abs/2303.08774>
- [22] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," 2021. [Online]. Available: <https://arxiv.org/abs/2103.00020>