

Inteligencia Artificial: Desafíos éticos en un mundo en cambio

Por Mg.Pablo Legna – Prof.de ética -UBA

Introducción

¿Quién no ha oído - y hablado - sobre inteligencia artificial (IA) últimamente? En ámbitos académicos o empresariales, en el hogar para hacer la tarea escolar o en la oficina para ayudar en un informe, la IA está omnipresente. La IA se ha posicionado como una de las tecnologías más influyentes del siglo XXI, transformando múltiples aspectos de la vida cotidiana y la estructura social. En este artículo abordaremos una de las múltiples facetas a las que nos enfrenta la IA: la ética. El propósito de este escrito es ofrecer una visión inicial sobre las implicancias éticas y los riesgos asociados al uso de la inteligencia artificial, un tema cuya exploración profunda resulta crucial para comprender y analizar el comportamiento humano a futuro; un futuro que está comenzando a llegar.

En este contexto, la ética se convierte en un elemento crucial para guiar el desarrollo y la implementación de sistemas de IA. La intersección entre IA y ética nos plantea preguntas fundamentales sobre la responsabilidad moral no solo de los desarrolladores, sino también de las instituciones y usuarios que utilizan estas tecnologías.

El avance de la IA presenta desafíos éticos significativos. De acuerdo con el análisis de Maldonado (2024), es esencial reflexionar sobre cómo las decisiones impulsadas por la IA pueden impactar la vida humana, ya que estas tecnologías poseen un "lado oscuro", - tales como los sesgos en los algoritmos -, que puede manifestarse sin una sólida base ética. Esto subraya la necesidad de establecer principios éticos que aseguren que la IA sirva como una herramienta para el beneficio humano y la mejora de la calidad de vida, en lugar de convertirse en una fuente de discriminación o daño.

A medida que se integra la IA en diferentes sectores, como la salud, la educación y la justicia, surgen dilemas éticos que requieren atención especial. La responsabilidad en el uso de IA debe ser una prioridad y no puede delegarse únicamente a los algoritmos; es vital que los seres humanos mantengan el control y la supervisión de los sistemas que influyen en sus vidas. Por lo tanto, la responsabilidad ética y la transparencia son fundamentales para el diseño y la implementación de IA, así como para la confianza pública en estas tecnologías.

La introducción de la inteligencia artificial en diversos ámbitos conlleva la necesidad inmediata de abordar las cuestiones éticas relacionadas, exigiendo un esfuerzo colectivo para garantizar que estas herramientas contribuyan positivamente a la sociedad y respeten los derechos humanos fundamentales.

Evolución

La Inteligencia Artificial (IA) ha sido objeto de estudio y debate desde sus inicios, y su marco teórico se basa en múltiples disciplinas. La IA se define como la capacidad de un sistema para realizar tareas que, tradicionalmente, requerirían inteligencia humana, como el aprendizaje, el razonamiento y la auto-corrección. Esta tecnología ha evolucionado significativamente desde la década de 1950, cuando se establecieron las bases teóricas iniciales que hoy sustentan su desarrollo (Huet, 2023)

Los avances recientes han permitido el surgimiento de diferentes tipos de IA, que se pueden clasificar en IA débil y fuerte. La IA débil, conocida como IA estrecha, está diseñada para tareas específicas, mientras que la IA fuerte busca simular la inteligencia humana de manera más generalizada (Dávila, 2023). Sin embargo, aún no se ha logrado completamente la IA fuerte, lo que evidencia la complejidad y los retos que enfrenta este campo.

Desde el punto de vista ético, es fundamental considerar los efectos de la IA sobre la sociedad. La ética de la IA involucra la reflexión sobre la responsabilidad, la transparencia, la privacidad y la equidad en el uso de estos sistemas, lo que se traduce en la necesidad de marcos normativos que guíen su desarrollo. La UNESCO ha destacado la importancia de establecer principios éticos básicos para asegurar que la IA promueva el bien común, evitando la reproducción de injusticias y discriminaciones ya existentes en la sociedad (Argentina, 2024). Algunas investigaciones enfatizan que la justicia en la IA no solo es un asunto técnico, sino un imperativo moral que debe abordarse en todas las etapas del desarrollo de sistemas inteligentes (WEF, 2025)

Por lo tanto, la intersección de la IA con la ética y la sociedad exige un enfoque multidisciplinario que garantice una implementación responsable y equitativa de esta tecnología, asegurando que sus beneficios sean accesibles para todos, y que se minimicen los riesgos asociados a su uso (ISO, 2025)

Implicaciones éticas de la IA: Cómo influye en nuestra vida cotidiana

La ética proporciona un marco para abordar los desafíos y oportunidades que surgen de la integración de IA en la vida cotidiana. Las implicancias éticas del uso de la inteligencia artificial son diversas y abarcan varios aspectos críticos que necesitan un análisis profundo. Entre las más destacadas se encuentran el sesgo algorítmico, la privacidad de los datos, la responsabilidad y la transparencia, así como el impacto en el empleo.

Sesgo algorítmico: Uno de los riesgos más destacados que implica el uso de la IA es el **sesgo y la equidad en los algoritmos**. Los sistemas de IA tienen el potencial de amplificar los sesgos existentes si se entrenan con datos que reflejan desigualdades sociales. Por ejemplo, se ha evidenciado que herramientas de diagnóstico utilizadas en la salud pueden mostrar un rendimiento dispar según la identidad étnica o la demografía de los pacientes, perpetuando así la discriminación en atención médica. Este fenómeno lleva a un debate ético sobre cómo garantizar que los sistemas sean justos y equitativos. Otro ejemplo, del sesgo algorítmico se presenta en los sistemas de reconocimiento facial, los cuales son menos precisos para personas de piel más oscura, lo que subraya la necesidad de conjuntos de datos más representativos y diversos. Incluso eliminando variables explícitas como raza o género no se soluciona el problema, ya que los algoritmos pueden identificar correlaciones indirectas conocidas como "variables proxy", lo que complica garantizar la justicia en estos sistemas. (WEF, 2025)

Privacidad de los datos: La recopilación masiva de datos personales plantea serias preocupaciones sobre la privacidad. A medida que la IA se vuelve más omnipresente, es esencial que se implementen normativas estrictas para proteger la información personal de posibles abusos. La falta de controles adecuados puede resultar en violaciones a la privacidad y en la desconfianza hacia las tecnologías (Aguerre, 2024) (Matthew, 2025)

Responsabilidad y transparencia: La cuestión de la transparencia es crucial en la interrelación entre ética y IA. Los usuarios deben comprender cómo las decisiones son tomadas por los sistemas automatizados, lo que implica que los desarrolladores deben esforzarse por crear IA que sea explicable. La falta de claridad en los algoritmos puede crear desconfianza, especialmente si los usuarios sienten que sus vidas están influenciadas por decisiones opacas. La falta de claridad en cómo los sistemas de IA toman decisiones plantea dificultades en la rendición de cuentas. Las organizaciones deben esforzarse por desarrollar modelos que sean comprensibles y que permitan a los usuarios cuestionar decisiones automatizadas. Esto incluye la necesidad de auditorías y evaluaciones regulares de los algoritmos para asegurar un funcionamiento ético y transparente (WEF, 2025) (Knowmad modo, s.f)

Impacto en el empleo: La IA está generando transformaciones significativas en el mercado laboral. Si bien crea nuevas oportunidades económicas y mejora la productividad, también representa una amenaza sustancial debido a la automatización y al desplazamiento de trabajadores, particularmente aquellos en actividades rutinarias. La automatización impulsada por la IA tiene el potencial de desplazar trabajos, lo que genera preocupaciones éticas sobre la equidad y la justicia social en el mercado laboral. En este punto, existe una gran variedad en el impacto de la IA según el tipo de trabajo e industria. También hay bastante discrepancia entre distintos informes que estudian la temática. Según un reporte de PWC y el WEF, en el futuro inmediato, la IA solo será capaz de realizar algunas tareas específicas en lugar de trabajos enteros. Aunque aproximadamente el 60% de los empleos incluyen tareas que pueden ser automatizadas, solo el 5% corre el riesgo de ser completamente reemplazado. Según McKinsey, para 2030, el 30% de los trabajos en Estados Unidos podría estar automatizado, mientras que el 60% se vería modificado por herramientas de IA. Goldman Sachs predice que hasta un 50% de los empleos

podrían automatizarse completamente para 2045, impulsados por la IA generativa y la robótica. La firma estima también que cerca de 300 millones de empleos, es decir, el 25% del mercado laboral mundial, podrían ser desplazados. Mas allá de estas cifras, que por naturaleza pueden ser muy volubles, esto subraya la importancia de invertir en la capacitación de trabajadores y la adopción de medidas como la inclusión de redes de seguridad social y programas de *reskilling* permitiendo una transición justa hacia nuevas oportunidades laborales (WEF, 2025) (Carreiro, 2023)

Desafíos y retos actuales

A continuación, se desarrollan los principales desafíos y retos actuales relacionados con la IA. Este análisis resalta temas que tendrán importantes impactos sociales y económicos en los años venideros.

Falta de regulación clara y global

Uno de los mayores desafíos actuales es la ausencia de regulaciones globales claras y consistentes para el desarrollo y uso de la IA. A pesar de intentos como la Ley de IA de la Unión Europea (AI Act), aún enfrentamos barreras significativas para la adopción de enfoques regulatorios homogéneos que permitan garantizar un desarrollo ético y responsable de la tecnología en un contexto internacional. (WEF, 2025). Dado que la tecnología estará siempre avanzando en forma más acelerada que lo que es posible para las regulaciones, debemos considerar la utilidad de impulsar leyes que pronto quedarán obsoletas. Algunos especialistas sostienen que es más práctico utilizar el derecho positivo actual para penalizar efectos ilegales de la utilización de la IA y no buscar regular de antemano algo bastante etéreo.

Dificultades para operacionalizar principios éticos

Aunque se han desarrollado más de 160 conjuntos de principios éticos para la IA, existe un consenso limitado sobre cómo implementarlos efectivamente. Los principios, como la transparencia y la explicabilidad, pueden entrar en conflicto y presentan desafíos adicionales según el contexto geográfico y cultural. Esto requiere la elaboración de guías prácticas claras y de medios formales de rendición de cuentas, así como cooperación internacional para fomentar un desarrollo ético.

Inclusión y diversidad en el desarrollo de IA

La falta de diversidad en los equipos que desarrollan IA es un factor crítico que perpetúa problemas como la discriminación y los sesgos sociales. Los equipos homogéneos tienden a reflejar las limitaciones y cosmovisiones de sus integrantes, lo que obstaculiza la igualdad y la inclusión. Fomentar equipos de desarrollo más diversos no solo puede reducir las desigualdades, sino también impulsar innovaciones que beneficien a una gama más amplia de usuarios y contextos.

Concentración geopolítica y económica

Según el citado informe de PwC, el desarrollo de la IA está concentrado en regiones como América del Norte y China, que capturan el 70% de los beneficios económicos anticipados, mientras que los demás países desarrollados capturan la mayoría del resto entre los que sobresalen Japón, Alemania y Reino Unido. Solo India está comenzando a desarrollar su industria de IA en los últimos tiempos. Según un informe de la ONU, Estados Unidos y China concentraron alrededor del 33% de las publicaciones sobre IA y el 60% de las patentes en 2022. Esta concentración puede exacerbar divisiones globales, generar competencias entre países (incluyendo rivalidades geopolíticas) y marginar a las economías en desarrollo que carecen de los recursos necesarios para competir. Además, la explotación de AI para fines polarizantes, como la creación de *deepfakes* y la manipulación de información, representa una nueva amenaza para la confianza pública y la estabilidad social.

Este análisis sintetiza desafíos altamente complejos, cuya solución requiere un enfoque colaborativo e inclusivo que promueva tanto el progreso tecnológico como el bienestar de la sociedad.

Conclusiones

Las interrelaciones entre la inteligencia artificial (IA) y la ética son complejas y fundamentales en el contexto actual de desarrollo tecnológico. En las conclusiones, es crucial destacar que la aplicación de ética en la IA no solo es deseable, sino esencial para guiar su evolución y uso.

El futuro de la ética en la inteligencia artificial se centra en la necesidad de establecer marcos legales y éticos robustos que guíen el desarrollo de esta tecnología. A medida que la IA se integra en diversas áreas de la vida cotidiana, surge una urgencia por fomentar la diversidad en los equipos de desarrollo, lo que puede mejorar los resultados éticos y la innovación en este campo (Dávila, 2023)

Una consideración crucial es la adaptación de los principios éticos a nuevas realidades, teniendo en cuenta los rápidos avances tecnológicos. Esto implica que la regulación y la ética deben evolucionar en respuesta a los desafíos emergentes, como los sesgos algorítmicos y las implicaciones de la toma de decisiones automatizadas. (Tavasci, 2023)

La cooperación internacional se vuelve indispensable para establecer estándares globales que aseguren el uso ético de la IA. Organizaciones como la UNESCO han emitido recomendaciones que abarcan derechos humanos y la promoción de la dignidad, enfatizando la necesidad de una gobernanza que integre estos principios en las aplicaciones de IA. (López, 2025) (ONU, 2023)

Así, el futuro de la ética en la IA no solo dependerá de avances tecnológicos, sino también de cómo las sociedades decidan estructurar sus marcos legales y éticos para abordar estos nuevos desafíos, permitiendo que la IA sea una herramienta para el bien común y la equidad social. (Meltzer, 2021) (Linares, 2024)

En resumen, el futuro de la ética en la IA es crucial por la necesidad de marcos éticos robustos, un enfoque inclusivo en el desarrollo, y una colaboración global clara. Es esencial que las acciones actuales se alineen con estos principios para garantizar que la IA se utilice de manera justa y beneficiosa para toda la sociedad.

Referencias bibliográficas

1. Maldonado, C. E. (2024). *Inteligencia artificial y ética*. Editorial Desde Abajo. Recuperado de https://www.academia.edu/116687638/Inteligencia_artificial_y_%C3%A9tica
2. Mora Naranjo, B. M. (2023). Ética y responsabilidad en la implementación de la inteligencia artificial en la educación. *Revista Ciencia Latina*, recuperado de <https://ciencialatina.org/index.php/cienciala/article/view/8833/13147>
3. Veron, J. (2022). La ética de la inteligencia artificial. *Revista Bioquímica y Patología Clínica*, 86(3), 20-21. <https://doi.org/10.62073/bypc.v86i3.225>. Recuperado de <https://openwebinars.net/blog/etica-en-la-inteligencia-artificial/>
4. Dávila Morán, R. C., & Agüero Corzo, E. del C. (2023). Desafíos éticos de la inteligencia artificial: implicaciones para la sociedad y la economía. *Conrado*, 19(94), 137-144. Epub 10 de octubre de 2023. Recuperado el 14 de abril de 2025, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442023000500137&lng=es&tlng=es
5. World Economic Forum. (2025). *Artificial Intelligence: Strategic Intelligence Briefing*. Recuperado de <https://www.argentina.gob.ar/la-supervision-humana-ayuda-etica-la-inteligencia-artificial>
6. ISO. (2025). Construir una IA responsable: cómo gestionar el debate ético sobre la IA. Recuperado de <https://www.iso.org/es/inteligencia-artificial/etica-ia-responsable>
7. UNESCO. (n.d.). Recomendación sobre la ética de la inteligencia artificial. Recuperado de <https://www.unesco.org/es/artificial-intelligence/recommendation-ethics>
8. Hanna, M. G., et al. (2025). Consideraciones éticas y de sesgo en la inteligencia artificial y el aprendizaje automático. Recuperado de <https://translate.google.com/translate?u=https://www.sciencedirect.com/science/article/pii/S0893395224002667&hl=es&sl=en&tl=es&client=srp>
9. Knowmad Mood. (n.d.). Inteligencia Artificial y Ética (II): El desafío del sesgo. Recuperado de <https://www.knowmadmood.com/blog/ia-y-etica-desafio-del-sesgo>
10. Aguerre, L. (2024). Ética anticipatoria e inteligencia artificial. *Biblioteca del Congreso de la Nación*. Recuperado de <https://bcn.gob.ar/informes-cientifico-legislativos-sobre-ia/etica-anticipatoria-ia>
11. Carreiro, J. (2023). Ética y sesgo en la inteligencia artificial. Recuperado de <https://www.ironhack.com/es/blog/etica-y-sesgo-en-la-inteligencia-artificial>
12. Linares, J. (2024). Los desafíos éticos de la inteligencia artificial. Recuperado de https://unamglobal.unam.mx/global_revista/los-desafios-eticos-de-la-inteligencia-artificial/
13. López, J. (2025). Importancia de la ética en la inteligencia artificial. Recuperado de <https://www.qualoom.es/inteligencia-artificial-etica-principios-y-consideraciones-en-el-desarrollo-y-uso-de-ia/> y <https://protecciondata.es/principios-eticos-inteligencia-artificial/>
14. Naciones Unidas (ONU). (2023). El debate de la ética de la IA no es tecnológico sino social. Recuperado de <https://news.un.org/es/interview/2023/05/1520557>
15. Brookings Institution. (n.d.). *Strengthening International Cooperation on Artificial Intelligence*. Recuperado de <https://translate.google.com/translate?u=https://www.brookings.edu/articles/strengthening-international-cooperation-on-artificial-intelligence/&hl=es&sl=en&tl=es&client=srp>