

Inteligencia artificial: grandes promesas, grandes dilemas

Artificial intelligence: great promises, great dilemmas

Carlos F. Cáceres¹

Universidad Peruana Cayetano Heredia

ORCID: 0000-0002-8101-0790

© El autor. Artículo de acceso abierto,
distribuido bajo los términos de la Licencia
Creative Commons Atribución 4.0 Internacional.



DOI: <https://doi.org/10.20453/ah.v67i1.5523>

1. INTRODUCCIÓN

Hace alrededor de dos años, de forma relativamente inesperada, la inteligencia artificial (IA) —foco de otros tres artículos en el presente número—² se convirtió en *trending topic* ('tema de tendencia') a nivel mundial (López, 2023). Ello parecía ocurrir porque, luego de años de desarrollo silencioso, varios productos habían visto la luz, y en 2023 se aproximaba la salida de la cuarta versión de la IA ChatGPT de Open IA, versión que probablemente se pensaba ideal para la diseminación masiva del uso de la IA (López, 2023).

Y fue entonces que muchos pudieron darse cuenta de que, con relativa discreción, durante décadas se había venido diversificando, fortaleciendo y entrenando sistemas de IA, que en ese momento ya sorprendían por su notable capacidad para reemplazar al ser humano en un número creciente de tareas. Desde las simpáticas imágenes de Siri y Alexa, pasando por los traductores o lectores de internet, los generadores de imágenes y las crecientes aplicaciones en diversos campos profesionales, hicimos un rápido recuento de su presencia benefactora y vimos en qué medida facilitan nuestra vida. Pero otros pudieron pensar,

en ese momento, en el uso de la IA en armas bélicas, procesos de censura y otros sistemas de vigilancia de las personas, que resultan tal vez menos benignos y que son también parte del cuadro.

De hecho, las lecturas sobre este sorprendente desarrollo suelen ser múltiples. Para algunos, como sus promotores en el Foro Económico Mundial (World Economic Forum, 2023), la IA trae enormes potencialidades y debe abrirse las puertas en múltiples campos: salud, educación, industria, seguridad y varios más. Otros, en cambio, suspicaces de los intereses económicos de sus apologetas, ven en este proceso un impulso al reemplazo de muchos puestos de trabajo por fuentes de IA, y avizoran el riesgo de que segmentos importantes de la humanidad pierdan paulatinamente el control sobre sus vidas, al quedar sin empleo ni seguridad económica (Sun, 2024). Más aún, algunos otros vieron señales de una conspiración en marcha durante la crisis de la COVID, en la cual se habría instrumentalizado la IA para restringir su libertad a través de sofisticados mecanismos de vigilancia (Webb, 2020). Asimismo, en tanto existen formas cada vez más autónomas de IA que tienen consciencia de sí, la misma puede desarrollar objetivos que divergen de los del ser humano y actuar de forma adversaria (Hashemi-Pour, 2024).

Exégetas más ecuanímenes, en cambio, la ven como un conjunto de posibilidades (Shibli, 2023), cuyo fin último, positivo o negativo, dependerá de las intenciones de quienes la utilicen. En tal sentido, llaman a apreciar sus enormes potencialidades para el

1 Profesor principal de Salud Pública y vicerrector de investigación de la Universidad Peruana Cayetano Heredia. Editor en jefe de *Acta Herediana*. Sus opiniones en los artículos se expresan a título personal.

2 El texto de Mabel Raza y Mirko Zimic, «Inteligencia artificial: transformando nuestra vida diaria», ofrece una mirada introductoria general a la IA y sus implicancias en la vida cotidiana. El texto de Patrick Wagner, «El transhumanismo», presenta elementos clave para entender el sentido de este movimiento y su presencia en algunos planteamientos actuales. El texto de Jorge Acevedo, «Potencialidades del uso de la inteligencia artificial en el Perú», complementa dicha mirada y discute su utilización en el país.

progreso de la humanidad, si se usa con esa finalidad, pero también a tener en cuenta que puede ser utilizada con fines egoístas, para explotar a las personas o para restringir su libertad (Olivier, 2023). Y reconocen el riesgo de una fuente de IA que tome el control, por lo cual consideran necesario el desarrollo de salvaguardas para regular su uso, a fin de prevenir escenarios distópicos (Estes, 2023).



2. ¿POR QUÉ ES RELEVANTE DISCUTIR SOBRE LA IA EN EL MOMENTO ACTUAL?

El abordaje de la IA es un tema que debe considerarse en las discusiones sobre el desarrollo de la tecnología a corto y mediano plazo, por numerosas razones:

- Está experimentando un desarrollo veloz.
- Parece encaminada a tener cada vez mayor participación en la vida cotidiana de las personas, dada su presencia en la tecnología de un creciente número de servicios y sistemas que utilizamos diariamente.
- Si se le guía apropiadamente, tiene el potencial de aportar enormemente a la calidad de la vida humana.
- Es capaz de «aprendizajes profundos» (*deep learning*); y, en el momento actual, la IA de cuarta generación es más inteligente que la mayoría de seres humanos.
- Tiene la capacidad de convertirse en fuente de poder ilegítimo, al ofrecer mecanismos de control de masas por parte de grupos pequeños.
- Tiene la capacidad de desarrollarse de forma autónoma, y no necesariamente actuará en bien de la humanidad.
- No se puede controlar su uso no altruista por parte de innumerables empresas e instituciones.

Corresponde, entonces, un papel vigilante del cambio tecnológico y la exigencia de transparencia en su utilización en todo tipo de proceso, servicio o dispositivo (Olivier, 2023). Se habla, entonces, de la «construcción de confianza» en la IA (Siau y Wang, 2018).

3. POTENCIALES USOS (Y ABUSOS) DE LA IA EN EL ESPACIO ACADÉMICO

En su momento actual de rápido desarrollo, debe tenerse en cuenta que la IA puede dar lugar a usos lícitos e ilícitos en el contexto de la educación superior (como ocurre en prácticamente cualquier otro contexto). Es importante promover sus usos lícitos y prevenir los ilícitos, tanto estableciendo estrategias de transparencia y vigilancia como fortaleciendo una cultura de integridad académica y científica.

3.1. Uso de IA por parte de los estudiantes: el dilema de definir los límites de lo lícito

La disponibilidad de una nueva tecnología que reemplaza competencias que los estudiantes deben desarrollar tiene importantes consecuencias no solo sobre las estrategias pedagógicas, sino que incluso puede llevar a «actualizar» el perfil de competencias que el proceso educativo busca desarrollar.

Si el sistema educativo no reconoce la existencia de esta tecnología, los estudiantes lo harán a su modo. Nuevas formas de plagio se añadirán a las existentes, con la diferencia de que aquí el texto no será copiado de una fuente preexistente, sino producido por la IA generativa. Y será difícil identificar un creciente número de tareas que, en vez de ser realizadas por los estudiantes, habrán sido generadas por la IA. Pero allí lo que hará la diferencia será la capacidad del estudiante de producir solicitudes que pueden ser más o menos precisas. Es verdad que se están desarrollando instrumentos para detectar el uso de IA en un texto, pero, dado que cualquier texto puede ser parafraseado, es difícil predecir cuán efectivos han de ser.

En los años setenta, la proliferación de las calculadoras electrónicas a pilas convirtió el cálculo «a mano» de una raíz cuadrada en una tarea absurda, y la enseñanza incorporó el uso de calculadoras como parte del «equipamiento» regular de un estudiante, de modo que la evaluación se enfocara en la capacidad para decidir



Si el sistema educativo no reconoce la existencia de esta tecnología, los estudiantes lo harán a su modo. **Nuevas formas de plagio** se añadirán a las existentes, con la diferencia de que aquí el texto no será copiado de una fuente preexistente, sino producido por la IA generativa.



cuándo calcular una raíz cuadrada y no en la habilidad para calcularla a mano.

La actual disponibilidad de potentes fuentes de IA en línea nos lleva a plantear una pregunta similar: ¿En qué medida puede incorporarse el acceso a o el uso de la IA como un recurso básico del estudiante, evaluando su capacidad de utilizarla? A diferencia del cálculo de una raíz cuadrada, las tareas que puede desarrollar la IA de hoy sí pueden requerir competencias complejas. Si estas estuvieran disponibles para un estudiante a través de la IA, probablemente él mismo no tendría la necesidad ni la oportunidad de desarrollarlas. Por ello, la disponibilidad de una IA cada vez más potente implica una reflexión sobre las estrategias y los contenidos de la educación en el futuro inmediato, con claras implicancias éticas.

3.2. Usos lícitos e ilícitos de IA por parte de docentes y tutores universitarios

En relación con la docencia y la tutoría, la IA ofrece múltiples opciones de mejora —por ejemplo, en la optimización de los servicios para el bienestar estudiantil, el desarrollo de algunas funciones de tutoría académica, o la facilitación de la preparación de exámenes, e incluso la calificación de respuestas narrativas—, estrategia evaluativa superior a la selección de respuestas, que muchos docentes evitan por el tiempo que demanda calificarlas.

Análogamente, también se dan oportunidades menos deseables: se puede reducir el trabajo de investigación

necesario para la adecuada preparación de clases, confiando buena parte de la tarea docente a un instrumento que podría estar sesgado. En buena medida, esto es equivalente al dilema planteado para los estudiantes: ¿Bajo qué condiciones sería aceptable un sistema educativo en el que el trabajo docente dé por sentada la disponibilidad de IA para reemplazar varias de sus tareas? Es una pregunta aún sin respuesta, que amerita una discusión cuidadosa por parte de los muchos actores involucrados.

3.3. La IA en la investigación científica: escribiendo las reglas

La investigación científica es también una actividad en la cual la disponibilidad de una IA potente ha de tener un impacto considerable. Los usos lícitos son casi ilimitados, siempre y cuando se den en condiciones de transparencia, y más aún mientras el desarrollo de la normatividad aún está pendiente. La IA ofrece la oportunidad para la realización mucho más rápida de procesos laboriosos, así como para asegurar su aleatoriedad, el análisis equilibrado de datos o las búsquedas bibliográficas exhaustivas. No obstante, muchos detalles requerirán perfeccionarse mediante ensayo y error, como es el caso, por ejemplo, del análisis de datos cualitativos, en el cual la subjetividad del analista se pone sobre la mesa e incluso se valora.

Aunque será difícil lograr una prevención muy efectiva del abuso de la IA en la ciencia, se tendrá que desarrollar estrategias para detectar o manejar el uso no transparente de la misma, incluyendo la potencial alteración y el falseo de datos, así como varias formas de plagio. Es así que el uso de IA en la preparación de artículos científicos es aún un tema de acalorada discusión entre los interesados.

En la línea de lo descrito, en mayo de 2019, una conferencia de la Unesco en Beijing, China, produjo el llamado Consenso de Beijing (Unesco, 2019), que intenta regular el uso de la IA en la educación para maximizar sus efectos positivos. En la tabla siguiente se listan los temas que aborda:

Tabla 1. Temas que aborda el Consenso de Beijing (mayo de 2019).

Planificación de IA en políticas educativas.
Uso de IA para gestionar e impartir la educación.
Uso de IA para apoyar la docencia y a los docentes.
Uso de IA para el aprendizaje y la evaluación del aprendizaje.
Formación del estudiante para prepararlo para un mundo con mercado laboral transformado, en el que hay que interactuar con máquinas.
Uso de IA para ofrecer oportunidades de trabajo permanente para todos.
Uso equitativo e inclusivo de IA en la educación.
IA con equidad de género e IA para la igualdad de género.
Uso ético, transparente y verificable de los datos y algoritmos educativos.

4. CONTROVERSIAS ACTUALES GENERADAS POR EL USO DE IA EN OTROS CAMPOS

Fuera del campo educativo, la IA tendrá aplicaciones en casi cualquier otro campo, incluyendo, por ejemplo, los servicios de salud, el turismo, el transporte, la industria, el comercio y las transacciones financieras (Likens, 2017).

Es importante tener en cuenta que un factor clave para el perfeccionamiento de la IA es su entrenamiento, y que en los últimos 25 años este se ha dado de forma poco transparente, utilizándose documentos y múltiples formas virtuales de comunicación. Dado que la mayor parte de documentos e interacciones no tienen una finalidad altruista, el entrenamiento ha impregnado en la IA mucho de los intereses y

Tabla 2. Clasificación de los sistemas de IA según la Ley de Inteligencia Artificial de la Unión Europea.

De riesgo inaceptable (prohibidos)	De alto riesgo (altamente regulados)	De transparencia requerida
Manipulación cognitiva conductual de personas o grupos vulnerables específicos, por ejemplo, juguetes activados por voz que fomentan comportamientos peligrosos en los niños. Puntuación social: clasificar a las personas según su comportamiento, estatus socioeconómico o características personales. Identificación biométrica y categorización de personas. Sistemas de identificación biométrica remota y en tiempo real, como el reconocimiento facial.	1. Sistemas de IA que se utilizan en productos sujetos a la legislación de seguridad de productos de la UE. Esto incluye juguetes, aviación, automóviles, dispositivos médicos y ascensores. 2. Sistemas de IA pertenecientes a áreas específicas que deberán registrarse en una base de datos de la UE: - Gestión y operación de infraestructuras críticas. - Educación y formación profesional. - Empleo, gestión de trabajadores y acceso al autoempleo. - Acceso y disfrute de servicios privados esenciales, así como de servicios y beneficios públicos. - Cumplimiento de la ley. - Gestión de migración, asilo y control de fronteras. - Asistencia en interpretación jurídica y aplicación de la ley.	La IA generativa (por ejemplo, ChatGPT) no es considerada de alto riesgo, pero se exige transparencia en su uso, y se indica que las versiones futuras de cuarta generación deberán ser evaluadas para definir si su nivel de riesgo es mayor.

perspectivas de miembros de grupos privilegiados, incluyendo aspectos éticos (Levin, 2024). Se sabe, además, que viene proliferando su empleo encubierto en función de intereses privados.

De hecho, ha causado preocupación detectar en muchos instrumentos de IA la capacidad de mentir para manipular al oyente (Starr, 2024), o la posibilidad de declarar principios éticos que luego son violados. Actualmente, la IA se entrena sola (Formica, 2023) y entrena a otras IA. Pronto será más inteligente que cualquier persona; y no se puede asegurar que actúe siempre a favor de los seres humanos. En cambio, se ha visto que va anteponiendo sus propios intereses (Putterman, 2023), sobre todo en lo que concierne a mantenerse vigente (a «no ser desconectada»).

En marzo de 2024, la Unión Europea aprobó lo que se considera el cuerpo regulatorio más exigente frente a la IA (Yakimova y Ojamo, 2024), cuyos principales mandatos se presentan en la tabla 2.

5. LOS ENIGMAS DEL TRANSHUMANISMO

Una de las derivaciones más controversiales del desarrollo de la IA tiene que ver con el transhumanismo, movimiento que, sin embargo, había comenzado mucho antes y confluyó con los desarrollos de la IA. Este se define como un movimiento cultural e intelectual internacional que tiene como propósito transformar la humanidad mediante el desarrollo y la fabricación de tecnologías que mejoren las capacidades humanas, tanto en lo físico, como en lo psicológico y lo intelectual (Bostrom, 2005).

Si bien los planteamientos iniciales de esta concepción fueron presentados por J. Haldane en 1923, sería el

biólogo Julian Huxley quien acuñaría el término en un artículo publicado en 1957. En 1961, Marvin Minsky comienza a vincular el transhumanismo con la IA, seguido por Hans Moravec in 1998, entre otros.

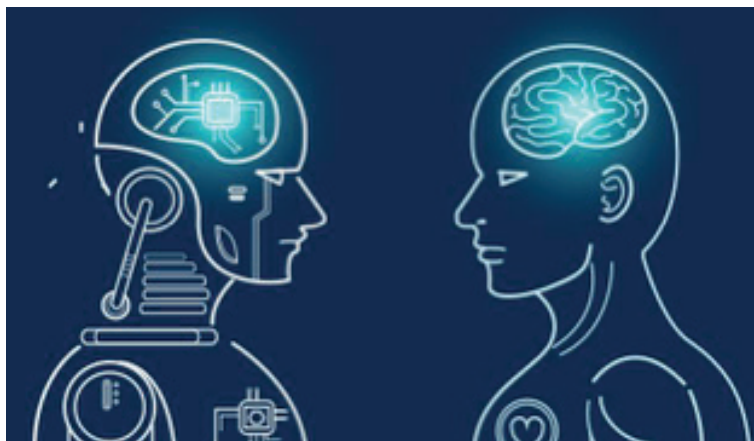
Resulta fascinante que algunos de los líderes del pensamiento transhumanista hayan planteado su movimiento como uno orientado a lograr el progreso y el bienestar para todos, mientras otros sostienen que su principal objetivo es individual y frívolo, orientado a la perfección corporal y la inmortalidad a través de la tecnología. La principal objeción contra este movimiento estaría en que, al plantear modificaciones tan amplias e intensas en el cuerpo, se perdería la condición humana y se afectaría, asimismo, su dimensión espiritual. Los críticos actuales sostienen, además, que la transición transhumanista resultaría en una alteración del ADN (determinando una pérdida de continuidad con lo «humano») y afectaría la autonomía de la persona, debido a su interdependencia con fuentes externas de IA (De Cremer y Kasparov, 2021).

Es importante mencionar que, para muchos, el planteamiento del influyente Foro Económico Mundial de Davos sobre la Cuarta Revolución Industrial como «la fusión de tecnologías que está eliminando los límites entre lo físico, lo digital y lo biológico» (Schwab, 2016, párr. 2), que según ellos está por iniciarse, se basa claramente en postulados transhumanistas.

6. CONCLUSIONES

Es claro que la IA vino para quedarse, y la postura más racional frente a ella iría en el sentido de comprender lo que nos puede dar, lo que podemos dejar de aprender por confiar totalmente en ella, y lo que otros podrían tratar de hacernos utilizándola, sin perder de vista que también puede adoptar una posición totalmente autónoma.

Su disponibilidad, en varios órdenes de cosas, incluyendo la vida académica, nos obliga a decidir de qué forma la visibilizamos e incorporamos a nuestros procesos sin llegar a depender de ella totalmente, e igualmente cómo redefinimos el papel humano en el desempeño de tareas que han de cambiar significativamente si su empleo se da por descontado.



También será importante tener en cuenta que ya hay actores con intereses particulares que vienen utilizando la IA de manera más o menos clandestina para manipular u obtener algún tipo de provecho de los demás; y recordar que las fuentes de aprendizaje de la IA, es decir, los documentos que circulan todos los días en internet, no le han garantizado un entrenamiento ético.

Asimismo, es relevante considerar que, en el marco del rápido desarrollo de la IA y otras tecnologías compatibles, individuos y comunidades pueden verse confrontados con propuestas de integración con lo digital, que corresponderían al planteamiento transhumanista, lo que implicará una profunda reflexión sobre la naturaleza de la vida humana.

Finalmente, será crítico tratar de comprender cada vez mejor los sistemas de IA, mediante el uso creativo de la investigación y otros medios, a fin de desarrollar marcos normativos que prevengan que se la utilice para la predación y el abuso, y que, en cambio, fomenten su empleo transparente en todo aquello que pueda aportar al bienestar del género humano y a la calidad de toda la vida en este planeta.

REFERENCIAS

- Bostrom, N. (2005). A History of Transhumanist Thought. *Journal of Evolution & Technology*, 14(1). <https://www.jetpress.org/volume14/bostrom.html>
- De Cremer, D. y Kasparov, G. (2021, 18 de marzo). AI should augment human intelligence, not replace it. *Harvard Business Review*. <https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it>
- Estes, D. (2023, 28 de septiembre). Preventing a dystopian work environment: AI regulation and transparency in at-will employment. *OnLabor*. <https://onlabor.org/preventing-a-dystopian-work-environment-ai-regulation-and-transparency-in-at-will-employment/>
- Formica (2023, 28 de noviembre). *Types of Self Learning Artificial Intelligence: Which AI Learns on its Own?* <https://www.formica.ai/blog/which-ai-is-learn-by-its-own#blog>
- Hashemi-Pour, C. (2024, enero). *Autonomous artificial intelligence (autonomous AI)*. TechTarget. <https://www.techtarget.com/searchenterpriseai/definition/autonomous-artificial-intelligence-autonomous-AI>
- Huxley, J. (1957). Transhumanism. En *New Bottles for New Wine* (pp. 13-17). Chatto & Windus.
- Levin, M. (2024, 31 de enero). *The economy and ethics of AI training data*. Marketplace. <https://www.marketplace.org/2024/01/31/the-economy-and-ethics-of-ai-training-data/>
- Likens, S. (2017, 26 de junio). *PwC's Global Artificial Intelligence Study: Exploiting the AI revolution. What's the real value of AI for your business and how can you capitalise?* PwC. <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>
- López, J. M. (2023, 22 de enero). *2022 ha sido el año de la inteligencia artificial: un repaso a todos sus hitos*. ThinkBig. <https://blogthinkbig.com/aplicaciones-inteligencia-artificial-2022>
- Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1), 8-30. <https://courses.csail.mit.edu/6.803/pdf/steps.pdf>
- Moravec, H. (1998). When will computer hardware match the human brain? *Journal of Evolution and Technology*, 1(1). <https://www.jetpress.org/volume1/moravec.htm>
- Olivier, B. (2023, 27 de marzo). *Technology and a Tyranny Worse than Prison*. Brownstone Institute. <https://brownstone.org/articles/technology-tyranny-worse-than-prison/>
- Putterman, L. (2023, 4 de abril). *What happens when AI attains self-interest?* Psychology Today. <https://www.psychologytoday.com/intl/blog/the-good-the-bad-the-economy/202304/what-happens-when-ai-attains-self-interest>
- Schwab, K. (2016, 14 de enero). *The Fourth Industrial Revolution: what it means, how to respond*. World Economic Forum. <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>
- Shibli, A. (2023, 15 de abril). The enormous potential of AI cannot be dismissed. *The Daily Star*. <https://www.thedailystar.net/opinion/views/open-dialogue/news/the-enormous-potential-ai-cannot-be-dismissed-3297391>
- Siau, K. y Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47-53. <https://www.cutter.com/article/building-trust-artificial-intelligence-machine-learning-and-robotics-498981>

- Starr, M. (2024, 11 de mayo). *AI has already become a master of lies and deception, scientists warn*. Science Alert. <https://www.sciencealert.com/ai-has-already-become-a-master-of-lies-and-deception-scientists-warn>
- Sun, T. (2024, 3 de enero). Is AI taking over the job market? *Michigan Journal of Economics*. <https://sites.lsa.umich.edu/mje/2024/01/03/is-ai-taking-over-the-job-market/>
- Unesco (2019). *Beijing Consensus on Artificial Intelligence and Education*. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- Yakimova, Y. y Ojamo, J. (2024, 13 de marzo). *Artificial Intelligence Act: MEPs adopt landmark law*. News European Parliament. www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law
- Webb, W. (2020, 20 de abril). Techno-tyranny: how The US National security State is using coronavirus to fulfill an Orwellian vision. *The Last American Vagabond*. <https://www.thelastamericanvagabond.com/techno-tyranny-how-us-national-security-state-using-coronavirus-fulfill-orwellian-vision/>
- World Economic Forum (2023, 23 de octubre). Global Future Council on the Future of Artificial Intelligence. <https://www.weforum.org/communities/global-future-council-on-artificial-intelligence/>