

El lado oscuro de la inteligencia artificial generativa en educación médica: ¿Debemos preocuparnos?

The dark side of generative artificial intelligence in medical education: Should we be worried?

“Con la inteligencia artificial, estamos invocando al demonio”.

ELON MUSK

“Mitigar el riesgo de extinción debido a la IA debería ser una prioridad global junto con otros riesgos a escala social, como las pandemias y la guerra nuclear”.

CENTER FOR AI SAFETY¹

Probablemente nuestros amables lectores digan “¿otra editorial sobre inteligencia artificial?, ¡ya cambien de tema!”. En cierto sentido tienen razón, en virtud de que los medios de comunicación y las redes sociales durante todo el año 2023 nos bombardearon con noticias sobre la inteligencia artificial (IA), inteligencia artificial generativa (IAGen) y plataformas como ChatGPT, planteando escenarios desde los apocalípticos hasta aquellos que auguran que la IA logrará el bienestar total de la humanidad. Todo ello acompañado de cientos de análisis de sus implicaciones positivas y negativas en educación y atención de la salud, muchos de ellos contradicto-

rios, por lo que se ha generado cierto hastío^{1,2}. Sin embargo, por su importancia, vale la pena continuar reflexionando sobre el tema.

Escribo estas líneas unos días después del primer aniversario del lanzamiento de ChatGPT por la empresa OpenAI³. Al ver en retrospectiva el último año es difícil asimilar la magnitud y complejidad de los avances que han ocurrido en tan poco tiempo en el campo de la IAGen y sus aplicaciones en medicina y en educación. El desempeño de las versiones más recientes de estas herramientas es verdaderamente prodigioso, y es un hecho que la incorporación de la multimodalidad (imágenes, audio y video, además de texto) en sus funciones nos ha colocado en una situación de asombro continuo y progresivo, ante el incremento gigantesco de las aplicaciones descritas en los campos educativos y de la salud²⁻⁴. En editoriales anteriores hemos enfatizado los aspectos positivos de la IAGen, a veces con un optimismo excesivo, que nace de la emoción de atestiguar de cerca el potencial del uso de herramientas como ChatGPT y similares. En la Universidad Nacional Autónoma de México hemos realizado eventos académicos sobre el tema, en los que es inevitable ser positivos y promo-

ver principalmente los casos de éxito con el uso de las herramientas (por ejemplo, la I Jornada de IAGen en educación de la UNAM, cuyas videomemorias se encuentran disponibles en: <https://iagenedu.unam.mx>). Sin embargo, es importante ir haciendo altos en el camino y, aprovechando el inicio de un nuevo año, reflexionar sobre las diferentes aristas y matices de este complejo tema, para obtener un panorama más balanceado que incluya lo que podríamos llamar “el lado oscuro” de la IAGen⁵.

Los epígrafes con que inicia esta Editorial son representativos de la percepción de varias personas, académicos, políticos y algunos segmentos de la población en general, que emiten alertas para no caer en el “canto de las sirenas tecnológico” y adoptar a ciegas la IAGen en nuestra vida cotidiana. Incluso predicen escenarios catastróficos para nuestra especie si no usamos con racionalidad y prudencia estas herramientas^{1,5}. A continuación, hacemos una lista no exhaustiva de los problemas, limitaciones y potenciales efectos adversos del uso de la IAGen en medicina y educación²⁻⁶:

- Hipérbole y exageración sobre la IAGen a través de la mercadotecnia. Debemos tomar con un grano de sal todo lo que nos pretende vender la industria, para usar los recursos disponibles de forma adecuada.
- Proliferación de datos falsos, desinformación, errores y alucinaciones. Estos aspectos son particularmente relevantes para la atención de la salud y la educación.
- Aumento de conductas tramposas y faltas a la integridad académica, con estudiantes y académicos que hagan pasar trabajos elaborados por IA como propios. Este es uno de los principales factores que ha causado alarma en la comunidad universitaria, ya que se tienen pocas armas efectivas para enfrentarlo, incluso se habla de una época “post-plagio”⁷.
- Incremento de la brecha digital entre diferentes estratos socioeconómicos, con retraso en la adquisición de competencias digitales y en IA.
- Las consecuencias de la falta de regulación por gobiernos e instituciones sobre el uso de estas herramientas, lo que causa incertidumbre en

la población y propicia que se realicen acciones inapropiadas con la IA en escenarios de salud y educativos.

- La falta de conocimiento de nuestras profesiones sobre las limitaciones técnicas de la IAGen puede generar uso excesivo o infrautilización de estas herramientas. La IA no piensa, no siente, no es una persona, aunque a veces lo parezca. El fenómeno de la “caja negra” persiste, para muchos de nosotros la IA no es explicable, y la forma como toma decisiones y hace inferencias no es transparente.
- El exceso de confianza en la tecnología genera errores y problemas de todo tipo, que en el caso de la medicina pueden ser graves o fatales.
- Las fallas en la privacidad y seguridad de los datos, con revelación de información confidencial y delicada. Estos aspectos también son fundamentales en la atención de la salud y en las aulas universitarias.
- Suplantación de la personalidad, “*deepfakes*” de imágenes, videos y audios.
- Perpetuación e incluso amplificación de los sesgos de las plataformas informáticas, con efectos sociales dañinos en términos de discriminación, inequidad y daño a poblaciones vulnerables y aquellas no representadas en los datos con los que se entrenan las herramientas. Esto crea marginalización y daño a amplios sectores de la sociedad.
- Poca o nula participación en la planeación, diseño, aplicación y evaluación de estas herramientas por los usuarios finales de la misma: profesionales de la salud y la educación, pacientes y estudiantes. Hay un predominio de la toma de decisiones centralizada en personas y grupos que diseñan estos instrumentos con fines primordialmente comerciales.
- Retos epistemológicos y de valores, ya que la tecnología no es neutral, por necesidad está fundamentada y motivada por los valores de quienes la crean, la promueven y obtienen beneficios con su uso.
- Énfasis en los aspectos tecnológicos de la sociedad moderna, infravalorando los aspectos humanísticos, sociales y éticos de los seres humanos.
- Costos ambientales y ecológicos, de los que existe

cada vez más evidencia⁸. El desarrollo de uno de estos modelos de IA puede resultar en niveles de emisión de carbono equivalentes a los que produce la ciudad de Nueva York en un mes⁶. Es indispensable incorporar esta información en los argumentos a favor y en contra de la IAGen, para apostarle a la sustentabilidad del planeta.

- Conceptualizar a la inteligencia como algo que solo está en la mente y el cerebro, y que puede ser completamente reproducible por software y hardware, cuando en realidad la inteligencia forma parte de nuestros cuerpos, no está “desincorporada” de nuestro organismo⁹. Esto tiene muchas implicaciones biológicas, cognitivas y filosóficas, es necesario continuar en la investigación interdisciplinaria sobre la inteligencia humana.
- Huecos legales y de derechos de autor sobre todos los productos creados por el hombre, a nivel local y global. Continúa la incertidumbre en los creadores de obras de arte, música, tecnología e innovación, ante la falta de normatividad sobre la propiedad intelectual y las implicaciones de la IA.
- Toma de decisiones inadecuada por los diversos líderes institucionales, con reemplazo de personas por herramientas de IA, bajo el supuesto de eficiencia organizacional. La IAGen puede ayudar a liberar tiempo de las personas para que se emplee en otras actividades productivas y creativas, pero no puede, ni debe, reemplazar de golpe a trabajadores especializados sin un análisis sensato de los pros y contras personales y laborales.

Esta lista por necesidad no es completa, ya que el campo de la IAGen es muy dinámico, pero nos da una idea de la gran cantidad de variables y fenómenos a considerar cuando pretendemos utilizarla para la práctica de la medicina o la docencia. No se trata de retos menores, nunca como ahora debemos estar conscientes de los riesgos y peligros que entraña el uso irreflexivo de la IAGen. Uno de los elementos fundamentales para que la IAGen se utilice de manera prudente y racional, es incluir a humanos en todas las fases del proceso, e involucrar desde un principio a los principales participantes, en nuestro caso profesionales de la salud, docentes, pacientes y

estudiantes. Es tarea de todos avanzar firmemente y con prudencia en el complejo laberinto de la era moderna, en la que siguen existiendo cosas como el crimen organizado, las guerras, la violencia e inseguridad, la inequidad y discriminación, entre muchos otros fenómenos sociales de los que no podemos sentirnos orgullosos.

Por cierto, nuestra revista agregó recientemente a las Instrucciones para los autores la siguiente leyenda: *Los autores deben explicitar el uso de Inteligencia Artificial y tecnologías asistidas por IA en el proceso de escritura incluyendo una declaración en un apartado titulado: “Declaración de IA y tecnologías asistidas por IA en el proceso de escritura”*. La presencia de la IA en prácticamente todos los aspectos de la investigación es cada vez mayor, usemos estas herramientas con sabiduría y profesionalismo, aprovechando que comienzan a aparecer artículos con guías sobre su uso en investigación en educación médica¹⁰.

En cuanto al número actual de la revista, tenemos artículos originales sobre estos temas: malestar psicológico y rendimiento académico en estudiantes de medicina; acoso sexual en el internado médico; ansiedad en estudiantes de medicina ante el regreso a clases presenciales; situación de egresados de odontología; propuesta de enseñanza en fisioterapia; acreditación de programas de educación médica; estudiantes de enfermería y cápsulas educativas; factores asociados con elección de especialidad en medicina legal; y un taller en línea sobre razonamiento clínico y sesgos cognitivos para médicos residentes. Además, tenemos dos artículos de metodología, uno sobre la tecnología RSS y el otro sobre Investigación-Acción Participativa y Educación Basada en Evidencias en el campo de la salud. Por cierto, en las cartas al Editor tenemos dos que abordan el tema de la IAGen y los grandes modelos de lenguajes (LLM), esperamos en próximos números comenzar a recibir manuscritos de investigación sobre este fascinante tema. 🔍

Melchor Sánchez Mendiola

EDITOR EN JEFE

Facultad de Medicina, UNAM

REFERENCIAS

1. Center for AI Safety. Statement on AI risk. 2023. Disponible en: <https://www.safe.ai/statement-on-ai-risk>
2. Marc LW, Gunasekara A, Leigh J, Pallant J, Pechenkina E. Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*. 2023;21(2):100790. <https://doi.org/10.1016/j.ijme.2023.100790>
3. Jürgen R, Tan S, Tan S. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*. 2023;6(1):342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
4. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. *The New England Journal of Medicine*. 2023;388(13):1201-1208. <https://doi.org/10.1056/NEJMra2302038>
5. Ivanov S. The dark side of artificial intelligence in higher education. *The Service Industries Journal*. 2023;43(15-16):1055-1082. <https://doi.org/10.1080/02642069.2023.2258799>
6. Selwyn N. The future of AI and education: Some cautionary notes. *European Journal of Education*. 2022;57:620-631. <https://doi.org/10.1111/ejed.12532>
7. Eaton SE. Postplagiarism: transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *Int J Educ Integr*. 2023;19:23. <https://doi.org/10.1007/s40979-023-00144-1>
8. Dhar P. The carbon impact of artificial intelligence. *Nat Mach Intell*. 2020;2:423-425. <https://doi.org/10.1038/s42256-020-0219-9>
9. Mitchell M. Why AI is harder than we think. En: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '21)*. Association for Computing Machinery, New York, NY, USA; 2021, p. 3. <https://doi.org/10.1145/3449639.3465421>
10. Tolsgaard MG, Pusic MV, Sebok-Syer SS, et al. The fundamentals of Artificial Intelligence in medical education research: AMEE Guide No. 156. *Med Teach*. 2023;45(6):565-573. <https://doi.org/10.1080/0142159X.2023.2180340>