

Cómo afecta la inteligencia artificial generativa a los procesos de evaluación

How generative artificial intelligence affects assessment processes

Francisco José García-Peñalvo

Instituto Universitario de Ciencias de la Educación
Universidad de Salamanca
<https://orcid.org/0000-0001-9987-5584>
fgarcia@usal.es

Resumen

La irrupción de la inteligencia artificial generativa (IAG) en la educación exige redefinir los procesos de evaluación. Muchas tareas evaluativas ahora pueden ser realizadas por IAG, lo que subraya la necesidad de equilibrar tecnología y pedagogía, redefiniendo e innovando en los métodos de evaluación y un uso ético y educativo responsable de estas herramientas.

Palabras clave

Evaluación; Educación; Inteligencia Artificial Generativa; ChatGPT

Abstract

The emergence of generative artificial intelligence (GenAI) in education demands a redefinition of assessment processes. GenAI can now undertake many evaluative tasks, highlighting the need to balance technology and pedagogy, redefine and innovate assessment methods, and ensure ethical and responsible educational use of these tools.

Keywords

Assessment; Education; Generative Artificial Intelligence; ChatGPT

1. Introducción

La revolución digital que se ha experimentado en las últimas décadas ha dejado una marca indeleble en casi todos los ámbitos de nuestra vida cotidiana, con influencia en la forma en que nos comunicamos, trabajamos, nos divertimos, e incluso cómo aprendemos (Fundación Telefónica, 2022). Es precisamente en este último punto, la enseñanza y el aprendizaje, donde la inteligencia artificial (IA), y más concretamente la inteligencia artificial generativa (IAG), ha comenzado a desempeñar un papel protagonista, con especial énfasis desde finales de 2022 con la irrupción estruendosa de ChatGPT.

Pero, ¿qué es exactamente la IAG? Es un subcampo dentro de la IA que se centra en la creación de modelos capaces de generar nuevos contenidos basados en datos previamente aprendidos. Esto es, aplicaciones que pueden componer música, crear imágenes o escribir textos, de forma que, en principio, sería indistinguible de una creación humana. Modelos como GPT-4 (Liu et al., 2023) o LLaMA 2 (Touvron et al., 2023), son ejemplos representativos de este tipo de tecnología, capaces de generar textos coherentes y contextuales a partir de directrices específicas. De una manera más formal, se puede definir la IAG como la producción de contenidos sintéticos inéditos, en cualquier forma y para apoyar cualquier tarea, mediante modelización generativa (García-Peñalvo & Vázquez-Ingelmo, 2023).

En el ámbito educativo, la influencia de estas tecnologías es palpable y en constante evolución. Sin embargo, antes de adentrarnos en las innovaciones y los cambios impulsados por la IA, es esencial comprender cómo han funcionado usualmente los procesos de evaluación, con especial énfasis en los niveles educativos preuniversitarios.

Los procesos de evaluación se han basado en aproximaciones tradicionales que buscan medir el conocimiento y las habilidades del estudiantado. Estos métodos abarcan desde exámenes escritos y orales hasta tareas entregables (con especial mención a la entrega de trabajos y ensayos) y proyectos que se deben completar durante el curso escolar. El profesorado, como principal agente evaluador, tenía (y sigue teniendo) la responsabilidad de crear, aplicar y corregir estas evaluaciones, un proceso que consume tiempo y es susceptible a subjetividades.

Además de estos métodos de evaluación, el profesorado también se basa en observaciones directas y una realimentación continua para entender mejor las necesidades y progresos del alumnado. Aunque estos métodos son efectivos en muchos aspectos, no están exentos de limitaciones. Por ejemplo, la corrección manual de exámenes puede llevar a errores involuntarios y la creación de pruebas puede no reflejar adecuadamente el currículo o las habilidades que se pretenden evaluar. Además, estos métodos no siempre permiten una adaptación flexible a las necesidades individuales de cada estudiante. Por otro lado, la debilidad intrínseca de algunos de estos métodos ha salido a la luz con la aparición de las herramientas de IAG, fundamentalmente la capacidad potencial de que el estudiantado haga un uso inadecuado de estas herramientas para realizar sus tareas de una forma automatizada sin que revierta en un aprendizaje efectivo y por ende en la obtención de las competencias relacionadas.

Ante los nuevos escenarios que se presentan con la IAG planean las sombras de un uso indebido y un aprendizaje superficial, que se relacionan directamente con las actividades de evaluación. Sin embargo, también es cierto que la IAG ofrece la posibilidad de optimizar y automatizar algunos de estos procesos de evaluación, liberando tiempo y recursos. Pero, y más importante, plantea una innovación en los paradigmas de evaluación, para que estas pueden ser más personalizadas, adaptativas y, en última instancia, más justas.

Como telón de fondo a esta discusión, es vital comprender que la educación, en su esencia, es un reflejo de la sociedad. A medida que la sociedad evoluciona, impulsada por avances tecnológicos como la IA, la educación, como uno de sus pilares fundamentales, no puede quedarse atrás. La pregunta es: ¿estamos preparados para este cambio y cómo lo abordamos de manera efectiva y ética? (Flores-Vivar & García-Peñalvo, 2023).

2. Integración de la IAG en la evaluación: Beneficios y desafíos

El avance tecnológico que representa la IAG no es solamente una innovación de vanguardia; es, además, una herramienta con el potencial de transformar la educación en sus cimientos. Si bien sus aplicaciones son vastas, nos centraremos en cómo esta tecnología afecta específicamente los procesos de evaluación en el ámbito educativo preuniversitario.

2.1. Beneficios de la IAG en los procesos de evaluación

Automatización y eficiencia. Las herramientas de IAG permiten automatizar tareas repetitivas y laboriosas. En el contexto de la evaluación, esto significa que tareas como la corrección de exámenes escritos, particularmente de elección múltiple o respuestas cortas, o la generación de casos prácticos y ejercicios, etc. pueden realizarse con mayor rapidez, permitiendo al profesorado dedicar más tiempo a la instrucción y a la realimentación personalizada.

Evaluaciones personalizadas. La IAG tiene la capacidad de analizar el rendimiento y las necesidades individuales de los estudiantes para generar evaluaciones adaptadas a cada perfil. Esto no solo proporciona una experiencia de aprendizaje más individualizada, sino que también ayuda a identificar áreas específicas de mejora.

Repensar los procesos de evaluación. La insatisfacción con muchos de los métodos de evaluación tradicionalmente empleados ya se puso de manifiesto con el confinamiento durante la pandemia de la COVID-19 (García-Peñalvo et al., 2020). Cuando ya parecía que nos habíamos olvidado de estos momentos y, sobre todo de lo que supuso a los procesos de evaluación, la llegada de la IAG vuelve a poner encima de la mesa la oportunidad y necesidad de repensar e innovar en la evaluación, con una aproximación más cercana a la evaluación orientada al aprendizaje (Ibarra Sáiz & Rodríguez-Gómez, 2010) y la evaluación para el empoderamiento (Rodríguez-Gómez & Ibarra-Sáiz, 2015), se debe tener en consideración la naturaleza de las tareas de evaluación definidas, en las que no se puede obviar la existencia de la IAG.

2.2. Desafíos y preocupaciones

Riesgo de sobre dependencia tecnológica. A medida que las herramientas basadas en IAG se vuelven más prevalentes, existe el riesgo de que profesorado y estudiantes dependan demasiado de ellas, descuidando habilidades cruciales de razonamiento, análisis y síntesis.

Uso inapropiado. Una preocupación válida es que el estudiantado pueda hacer un uso inapropiado de estas herramientas para completar tareas y exámenes sin realmente aprender y adquirir las competencias esperadas.

Evaluación de habilidades abstractas. Si bien la IAG puede ser eficaz al evaluar respuestas objetivas, enfrenta dificultades cuando se trata de evaluar habilidades más abstractas, como la creatividad, la empatía o la habilidad de argumentación crítica.

Preocupaciones éticas. La recolección y el análisis de datos del estudiantado por herramientas de IAG pueden plantear problemas relacionados con la privacidad y el uso ético de la información. Además, ¿quién es realmente responsable cuando una herramienta de IAG comete un error o se utiliza inapropiadamente? (Crawford et al., 2023).

3. Recomendaciones para una implementación responsable y efectiva

La IAG ofrece oportunidades inéditas para mejorar y optimizar los procesos de evaluación. Sin embargo, su implementación en el entorno educativo debe hacerse con precaución y responsabilidad. Aquí se detallan algunas recomendaciones clave para garantizar una incorporación efectiva y ética de esta tecnología en la educación preuniversitaria.

Formación y capacitación para educadores y estudiantes. Es esencial que el profesorado esté bien informado sobre las capacidades y limitaciones de la IAG. Esto implica ofrecer cursos y talleres sobre cómo utilizar herramientas basadas en IAG y fomentar la comprensión de los algoritmos detrás de estas herramientas, para que los educadores puedan hacer un uso crítico y selectivo de ellas.

Pero tampoco podemos olvidarnos de alfabetizar al estudiantado para que sean conscientes también de los beneficios y peligros de la IAG (García-Peñalvo et al., 2024), es crucial promover la integridad académica entre el estudiantado para evitar el uso indebido de estas herramientas.

Uso ético de datos. La IAG, al depender en gran medida de los datos, plantea preocupaciones de privacidad y ética. Las instituciones educativas deben:

- Garantizar que los datos de las personas se manejen con privacidad y confidencialidad.
- Informar a estudiantes y familias sobre cómo y por qué se recopilan y utilizan sus datos.
- Implementar medidas de seguridad robustas para proteger la integridad y privacidad de los datos.

Flexibilidad y adaptabilidad en la evaluación. Para maximizar los beneficios de la IAG en la evaluación, es necesario repensar y adaptar los métodos tradicionales. Esto implica, por un lado, diseñar evaluaciones que vayan más allá de respuestas objetivas, centrándose en habilidades críticas y analíticas; y, por otro, fomentar la evaluación basada en proyectos y el trabajo colaborativo, que

son más difíciles de automatizar y ofrecen una visión más integral de las habilidades del estudiantado.

Fomentar el pensamiento crítico. La IAG, si bien es una herramienta poderosa, no debe ser vista como la solución definitiva para todos los desafíos educativos. Es crucial enseñar al alumnado a ser crítico con la información que generan estas herramientas y a no aceptar todo lo producido por la IAG como verdad absoluta, así como fomentar en ellos y ellas una mentalidad de aprendizaje continuo, donde la IAG es solo una herramienta más en su arsenal educativo.

Colaboración y discusión continua. La rápida evolución de la IAG significa que las implicaciones para la enseñanza y el aprendizaje también cambiarán con el tiempo. Las instituciones educativas deben establecer comités o grupos de trabajo centrados en la IAG y la educación para discutir regularmente las mejores prácticas, desafíos y oportunidades. Además, la colaboración entre el profesorado, expertos en tecnología, estudiantes y familias será necesaria para garantizar que la IAG se implemente de manera que beneficie a todos los involucrados. En este sentido, la creación de comunidades de prácticas centradas en la IA aplicada a la educación serán esenciales para avanzar en estos primeros momentos de expansión, como por ejemplo (Nerantzi et al., 2023).

4. Conclusiones

La revolución que supone la incorporación de la IAG en el ámbito educativo está redefiniendo el panorama de la evaluación. Ya no solo se enfoca en cuantificar el conocimiento, sino que su propósito es más profundo: comprender y apoyar el viaje de aprendizaje individual de cada estudiante.

Una de las propuestas más prometedoras de esta nueva era es la concepción de sistemas de evaluación adaptativa potenciados por IAG. Mediante algoritmos avanzados, estos sistemas tienen la capacidad de modificar en tiempo real tanto el grado de dificultad como el contenido de las evaluaciones, basándose en las respuestas inmediatas del estudiantado. Por ejemplo, cuando una persona demuestra un claro dominio en una temática, el sistema la desafía con cuestiones más complejas o la introduce en temas novedosos. En cambio, si detecta zonas de fragilidad en el aprendizaje, recalibra con cuestiones más elementales o proporciona recursos adicionales para reforzar el entendimiento.

Sin embargo, más allá de la fascinación tecnológica y las visiones futuristas, es vital abordar el presente. El estudiantado, de todos los niveles educativos, ya está integrando estas herramientas en su rutina académica, a menudo sin el entrenamiento adecuado. Estas herramientas, si bien avanzadas y útiles, pueden ser doble filo si no se utiliza un enfoque crítico. La confianza ciega en ellas puede generar desinformación o errores. La auténtica eficacia radica en lograr un equilibrio: la innovación tecnológica debe converger con el enfoque humano y pedagógico. La IAG, por impresionante que sea, debe ser interpretada como un complemento que potencia la enseñanza, no como un reemplazo del criterio y la experiencia docente. Los profesionales de la educación, por su parte, pueden y deben emplear la IAG como un aliado, enriqueciendo su práctica

docente y los métodos de evaluación, siempre con un enfoque centrado en el aprendizaje y el empoderamiento del alumnado.

Para cerrar, es fundamental subrayar que una adopción de la IAG en el sistema educativo preuniversitario requiere una perspectiva integral. Es esencial considerar tanto las ventajas como los desafíos inherentes a esta tecnología disruptiva. Solo con una formación pertinente, protocolos éticos sólidos y una filosofía de enseñanza centrada en las personas, la IAG podrá actuar como catalizador en la optimización del proceso educativo y, en consecuencia, de los métodos de evaluación.

Referencias

- Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Flores-Vivar, J. M., & García-Peñalvo, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4). *Comunicar*, 31(74), 37-47. <https://doi.org/10.3916/C74-2023-03>
- Fundación Telefónica. (2022). *Sociedad digital en España 2023*. Penguin Random House Grupo Editorial. <https://bit.ly/3MuZhoz>
- García-Peñalvo, F. J., Corell, A., Abella-García, V., & Grande-de-Prado, M. (2020). La evaluación online en la educación superior en tiempos de la COVID-19. *Education in the Knowledge Society*, 21, Article 12. <https://doi.org/10.14201/eks.23013>
- García-Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED: revista iberoamericana de educación a distancia*, 27(1). <https://doi.org/10.5944/ried.27.1>
- García-Peñalvo, F. J., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, In Press. <https://doi.org/10.9781/ijimai.2023.07.006>
- Ibarra Sáiz, M. S., & Rodríguez-Gómez, G. (2010). Los procedimientos de evaluación como elementos de desarrollo de la función orientadora en la universidad. *Revista Española de Orientación y Psicopedagogía*, 21(2), 443-461. <https://doi.org/10.5944/reop.vol.21.num.2.2010.11558>
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), Article 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Nerantzi, C., Abegglen, S., Karatsiori, M., & Arboleda, A. M. (Eds.). (2023). *101 creative ideas to use AI in education, A crowdsourced collection*. <https://doi.org/10.5281/zenodo.8355454>.
- Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2015). Assessment as Learning and Empowerment: Towards Sustainable Learning in Higher Education. In M.

Peris-Ortiz & M. L. J. (Eds.), *Sustainable Learning in Higher Education. Innovation, Technology, and Knowledge Management*. Springer.
https://doi.org/10.1007/978-3-319-10804-9_1

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Canton Ferrer, C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B., Gao, C., Goswami, V., Goyal, N., Hartshorn, A., Hosseini, S., Hou, R., Inan, H., Kardas, M., Kerkez, V., Khabsa, M., Kloumann, I., Korenev, A., Singh Koura, P., Lachaux, M.-A., Lavril, T., Lee, J., Liskovich, D., Lu, Y., Mao, Y., Martinet, X., Mihaylov, T., Mishra, P., Molybog, I., Nie, Y., Poulton, A., Reizenstein, J., Rungta, R., Saladi, K., Schelten, A., Silva, R., Smith, E. M., Subramanian, R., Tan, X. E., Tang, B., Taylor, R., Williams, A., Kuan, J. X., Xu, P., Yan, Z., Zarov, I., Zhang, Y., Fan, A., Kambadur, M., Narang, S., Rodriguez, A., Stojnic, R., Edunov, S., & Scialom, T. (2023). Llama 2: Open Foundation and Fine-Tuned Chat Models. *arXiv, Article* arXiv:2307.09288v2.
<https://doi.org/10.48550/arXiv.2307.09288>