

Inteligencia artificial generativa y autonomía educativa: metáforas históricas y principios éticos para la transformación pedagógica

Generative artificial intelligence and educational autonomy: historical metaphors and ethical principles for pedagogical transformation



-  Marc Alier-Forment - *Universitat Politècnica de Catalunya, UPC (España)*
-  María José Casañ-Guerrero - *Universitat Politècnica de Catalunya, UPC (España)*
-  Juan Antonio Pereira-Varela - *Universidad del País Vasco, UPV/EHU (España)*
-  Francisco José García-Peñalvo - *Universidad de Salamanca, USAL (España)*
-  Faraón Llorens-Largo - *Universidad de Alicante, UA (España)*

RESUMEN

Este artículo analiza la integración de la inteligencia artificial generativa en educación desde una perspectiva crítica, histórica y ética. Se identifica una creciente preocupación por la opacidad de las herramientas de inteligencia artificial actuales, especialmente en sistemas de aprendizaje. El trabajo utiliza un enfoque basado en metáforas para entender cómo las narrativas tecnológicas influyen en la adopción de innovaciones educativas. Se revisan metáforas históricas en las tecnologías aplicadas a la educación, desde Multivac y Matrix hasta el Bazar del software libre y la App Store, y se proponen nuevas imágenes conceptuales que podrían aplicarse al contexto actual en el que la inteligencia artificial irrumpe en la educación. A partir de este análisis metafórico, se plantean siete principios éticos para una adopción segura de la inteligencia artificial generativa en educación, centrados en la privacidad, la alineación pedagógica, la supervisión humana y la transparencia tecnológica. Estos principios se ejemplifican con el entorno LAMB (Learning Assistant Manager and Builder), un marco de código abierto que permite diseñar asistentes de aprendizaje basados en inteligencia artificial de forma ética y contextualizada. Se presentan casos reales de aplicación de LAMB en educación superior, incluyendo una experiencia controlada con estudiantes que muestran mejoras significativas en autonomía y coherencia pedagógica. Finalmente, se destaca cómo LAMB encarna los principios éticos propuestos y responde a las metáforas críticas identificadas, ofreciendo un modelo de integración tecnológica centrado en la autonomía de los docentes, la alineación con los principios y prácticas de la institución educativa y el aprendizaje significativo de los estudiantes.

Palabras clave: tecnología educativa; inteligencia artificial generativa; aprendizaje adaptativo; asistente de aprendizaje; ética; diseño instruccional.

ABSTRACT

This article examines the integration of generative artificial intelligence in education from a critical, historical, and ethical perspective. It highlights growing concerns about the opacity of current artificial intelligence tools, particularly in learning systems. The study adopts a metaphor-based approach to explore how technological narratives influence the adoption of educational innovations. It reviews historical metaphors used to describe educational technologies, from Multivac and Matrix to the free software Bazaar and the App Store, and proposes new conceptual frameworks that may better reflect the current context in which artificial intelligence is entering the educational sphere. Based on this metaphorical analysis, the article outlines seven fundamental ethical principles for the safe adoption of generative artificial intelligence in education, focusing on privacy, pedagogical alignment, human oversight, and technological transparency. These principles are illustrated through a practical application: the LAMB (Learning Assistant Manager and Builder) environment, an open-source software framework that enables the ethical and contextualized design of artificial intelligence-based learning assistants. The article presents real-world cases of LAMB implementation in higher education, including a controlled experience with students that demonstrates significant improvements in student autonomy and pedagogical coherence. Finally, it emphasizes how LAMB embodies the proposed ethical principles and responds to the identified critical metaphors, offering a model for technology integration centered on teacher autonomy, alignment with institutional values and practices, and meaningful student learning that prioritizes pedagogical control over technological determinism.

Keywords: educational technology; generative artificial intelligence; adaptive learning; learning assistant; ethics; instructional design.

INTRODUCCIÓN

En septiembre de 2024, se mantuvieron conversaciones con diez expertos en innovación educativa y tecnología sobre el “Manifiesto por una IA segura en la educación” (Alier, García-Peñalvo, Casañ et al., 2024). Estas discusiones giraron en torno a una propuesta de marco de evaluación para tecnologías educativas basadas en inteligencia artificial generativa (GenAI). Este marco dio como resultado un cuestionario para la evaluación de iniciativas educativas basadas en GenAI. El cuestionario asociado ponía el foco en aspectos clave: la privacidad de los datos, el control de acceso dentro de las instituciones y, sobre todo, la capacidad de las herramientas para alinearse con los objetivos pedagógicos. De forma sorprendente, más de la mitad de los expertos consideraban que ninguna solución basada en GenAI disponible actualmente cumpliría estos criterios o que fuera posible cumplirlos.

De estos intercambios surgió una observación relevante: muchos tecnólogos con amplia experiencia en el ámbito educativo no ven la GenAI, y en particular los sistemas de aprendizaje adaptativo basados en IA que prometen personalizar la experiencia del estudiante, como herramientas claras y moldeables, sino como artefactos potentes pero opacos, difíciles de entender o controlar. Este fenómeno nos llevó a proponer la metáfora del *Palantír* (Tolkien, 1954-1955).

En el universo narrativo de J.R.R. Tolkien, el *Palantír* es una esfera mágica que permite ver a través del tiempo y el espacio y comunicarse con usuarios de otros *Palantíres*. No obstante, el uso del *Palantír* también entraña riesgos significativos, como se ilustra con la corrupción del mago Saruman, que Sauron ejerce a través de este objeto. Esta metáfora, aplicada a la IA generativa en educación, captura dos sentimientos muy presentes: la fascinación ante su aparente poder de personalización y predicción, y el temor o escepticismo sobre quién la controla y con qué fines.

Llamaremos a esto la metáfora del *Palantír*: la GenAI entendida como una herramienta centralizada, poderosa, inevitable, pero en última instancia fuera del control de sus usuarios reales (profesores, estudiantes e instituciones). Esta imagen ayuda a explicar una actitud cada vez más común en el ámbito educativo: aceptar la implementación de sistemas de IA adaptativos tal como vienen, sin conocer en profundidad cómo se entrenan, qué sesgos contienen o cómo se toman las decisiones que afectan al aprendizaje individual.

Para efectos de este trabajo, utilizamos el término “asistentes de aprendizaje basados en GenAI” para referirnos específicamente a herramientas conversacionales que proporcionan apoyo personalizado al estudiante mediante procesamiento de lenguaje natural, distinguiéndolas de los sistemas adaptativos tradicionales que ajustan automáticamente rutas de aprendizaje basándose en análisis de rendimiento.

La idea de tecnología vista como “magia”, ya ha sido ampliamente discutida en el pasado. Por ejemplo, Arthur C. Clarke explicó en 1962 que una tecnología suficientemente avanzada podía convertirse en algo similar a la magia (Clarke, 1962).

Sin embargo, es necesario matizar. Aunque la metáfora del Palantír es útil para describir ciertas narrativas actuales, también tiene limitaciones. La inteligencia artificial no es magia, y describirla como tal puede reforzar una visión determinista de la tecnología, que inhibe la capacidad crítica y de comprensión de la misma.

Las metáforas, cuando se explicitan con claridad, son herramientas analíticas valiosas. Reflejan e influyen en cómo se imaginan, diseñan, implementan y, sobre todo, cómo se difunden las tecnologías. Como señalan Weller y Hofstadter, las metáforas y analogías no son solo recursos estilísticos, sino marcos de pensamiento que orientan nuestras decisiones (Weller, 2022; Hofstadter y Sander, 2018).

Por ejemplo, la metáfora de los “nativos digitales” moldeó políticas educativas durante años, muchas veces con resultados erróneos. De forma similar, la idea de una “Uber de la educación” inspiró modelos de negocio que ignoraban la complejidad del aprendizaje (Adell-Segura et al., 2018).

En la actualidad, se percibe una presión creciente para integrar la GenAI en los sistemas educativos. Ante esta situación, este trabajo presenta en el apartado dos, un recorrido crítico por la historia de la tecnología educativa, con el objetivo de analizar cómo distintas innovaciones han sido introducidas, qué desafíos han planteado y qué marcos conceptuales, a menudo en forma de metáforas, han guiado su comprensión y uso. A partir de este análisis histórico y conceptual, en el apartado tres se identifican principios éticos fundamentales que los autores consideran necesarios para una incorporación responsable de la GenAI en contextos educativos. A continuación, en la sección cuatro se propone una aplicación práctica que traduce estos principios en orientaciones concretas para el diseño y uso de sistemas de aprendizaje adaptativo basados en inteligencia artificial. Finalmente, se exponen las conclusiones del artículo.

MARCO ANALÍTICO: EVOLUCIÓN METAFÓRICA DE LAS TECNOLOGÍAS EDUCATIVAS

A lo largo de la historia, la introducción de nuevas tecnologías en el ámbito educativo no ha sido guiada únicamente por avances técnicos, sino también por narrativas, metáforas e imaginarios que han dado sentido a su adopción y modelado sus aplicaciones. En este apartado se explora cómo determinadas metáforas han moldeado la evolución de la tecnología educativa en las últimas décadas y qué lecciones se pueden extraer de esta historia para acabar explicando el enfoque actual respecto a la GenAI (Llorens-Largo, 2019). Como recuerda Weller (2022), “La tecnología suele venir envuelta en sus propias metáforas” (p. 59), y esas metáforas iniciales pueden tener profundas consecuencias sobre la forma en que dicha tecnología se desarrolla y se aplica. En esta línea, Mason (2018) demuestra empíricamente cómo las metáforas subyacentes estructuran el discurso de la tecnología educativa, sugiriendo que aunque es necesario usar metáforas para comprender nuevos fenómenos digitales, algunas de las metáforas específicas comúnmente empleadas pueden impedir

enfoques más reflexivos para conceptualizar e implementar nuevas tecnologías en el ámbito educativo.

Se ha dividido este apartado en cuatro momentos clave de la evolución de las tecnologías educativas, analizando las metáforas dominantes que acompañaron cada etapa y sus implicaciones en las formas de aprendizaje, especialmente en relación con la personalización y adaptación.

De Multivac al mainframe educativo: la metáfora de la educación como sistema centralizado

Isaac Asimov, en sus relatos sobre Multivac, imaginó una supercomputadora capaz de tomar decisiones cruciales para toda la humanidad. Esta visión refleja una idea poderosa: la tecnología como entidad centralizada que recopila datos, procesa información masiva y dirige las acciones de la sociedad (Asimov, 1955-1976).

Esta visión centralizada y autoritaria del ordenador también se reflejó en las primeras propuestas de la Enseñanza Asistida por Ordenador (EAO) o Instrucción Asistida por ordenadores. B. F. Skinner, figura clave del conductismo, planteó en su obra *The Technology of Teaching* (Skinner, 1968) que los ordenadores podían optimizar el aprendizaje mediante la aplicación sistemática de principios conductistas. Estos sistemas no concebían el aprendizaje como un proceso abierto o exploratorio, sino como una serie de trayectorias programadas y controladas desde un único centro.

Aplicado a los sistemas de aprendizaje basados en GenAI, Multivac representa una visión donde el sistema informático dicta el ritmo y el contenido que se presenta al alumno, relegando la adaptabilidad a un modelo de supervisión rígido. Se priorizaba la optimización de resultados predefinidos, no la personalización genuina basada en intereses, contextos o motivaciones individuales.

Esta metáfora sigue siendo relevante hoy cuando ciertos sistemas tienden a priorizar la eficiencia de la instrucción sobre la autonomía del estudiante, reproduciendo una lógica de control centralizado más que de aprendizaje autodirigido.

De Matrix a las píldoras educativas: la metáfora del aprendizaje como transferencia instantánea de conocimiento

En la película *The Matrix* (Wachowski y Wachowski, 1999), los personajes adquieren habilidades complejas, como pilotar helicópteros, mediante la descarga inmediata de programas en su cerebro. Esta metáfora de la “transferencia instantánea de conocimiento” se hizo muy influyente durante la era del CD-ROM y las primeras plataformas e-learning durante la segunda mitad de los 1990s.

En educación, se tradujo en el desarrollo de unidades de aprendizaje cerradas y estandarizadas, las llamadas “píldoras educativas”, que pretendían ofrecer formación

rápida, modular y autónoma. Bajo esta lógica, aprender se convirtió en “consumir” paquetes de información discretos, no en construir conocimiento activamente.

La relación con las herramientas de IA generativa contemporáneas es clara: muchos enfoques basados en IA siguen asumiendo que el aprendizaje es una cuestión de administrar la dosis adecuada de contenido, ajustada a la velocidad o nivel del alumno, pero sin necesariamente considerar la dimensión crítica, reflexiva o social del aprendizaje.

La metáfora de Matrix alimenta la promesa de la GenIA como solución mágica, pero también oculta los riesgos de simplificar el aprendizaje a un proceso meramente receptivo, pasivo y descontextualizado.

Del Bazar al Campus Virtual: la metáfora de la construcción colaborativa de saberes

La siguiente revolución en la tecnología educativa llegó de la mano de la web, y la web no puede entenderse sin la cultura del software libre. Este movimiento nació en los laboratorios del MIT, cuando Richard Stallman, en respuesta a las restricciones impuestas por las licencias de usuario de software (EULA), que anteriormente se compartían libremente entre programadores, publicó el *Manifiesto GNU* (Stallman, 1985) y fundó la *Free Software Foundation (FSF)* en 1985. La FSF proporcionó un marco legal fundamental: la licencia GPL, que permitía a los desarrolladores compartir su software garantizando cuatro libertades esenciales a los usuarios: usar, estudiar, compartir y mejorar el software. Esta filosofía, resumida en el lema “libre como en libertad, no como en cerveza gratis” (“*free as in freedom not as in free beer*”), facilitó la colaboración y el intercambio de código fuente a una escala sin precedentes (Ceruzzi, 2003).

En contraste con las visiones centralizadas anteriores, el movimiento del software libre, se puede sintetizar en la metáfora del Bazar (Raymond, 2001), que introdujo un nuevo paradigma en los sistemas de aprendizaje: sistemas abiertos, donde múltiples actores colaboran, innovan y adaptan soluciones de manera descentralizada y democrática.

En su influyente ensayo *The Cathedral and the Bazaar* (Raymond, 2001), Eric S. Raymond observa que las comunidades de software libre, donde el conocimiento y el código se comparten y los miembros ganan influencia en función de sus contribuciones más que de su estatus o posición, ofrecen una solución a la “crisis endémica del software” descrita por Pressman en *Software Engineering: A Practitioner’s Approach* (Pressman, 1991). Esta aplicación práctica de los principios del software libre genera una escisión en el movimiento con la aparición de la *Open Source Initiative (OSI)*. La OSI propuso el término “software de código abierto” (*Open Source Software*) como alternativa al “software libre”, introduciendo licencias más flexibles (como Apache o MIT) que resultaban más atractivas para las grandes empresas.

Esta filosofía inspiró la creación de entornos como Moodle y Sakai, plataformas de gestión de aprendizaje que, al ser de código abierto, permitieron a docentes, estudiantes, instituciones y desarrolladores adaptar las herramientas a sus necesidades específicas. No se trataba solo de consumir contenidos, sino de construir entornos de aprendizaje dinámicos, compartidos y configurables.

En este contexto, la metáfora del Bazar señala la posibilidad de diseñar sistemas que no sean cajas negras, sino espacios abiertos donde las adaptaciones se comprendan, negocien y reconfiguren colectivamente. Esta perspectiva refuerza la idea de que la adaptabilidad no debe ser solo algorítmica, sino también pedagógica y social.

De la App Store a la nube educativa: la metáfora de la dependencia invisible

Con la irrupción de los teléfonos inteligentes y de las tiendas de aplicaciones —la App Store de Apple y Google Play— se afianzó la metáfora de que “hay una app para todo”. La promesa era contar con soluciones inmediatas, personalizadas y accesibles, pero a cambio de una dependencia casi total de infraestructuras propietarias y opacas —plataformas— que venden seguridad y conveniencia al precio del control. Como recuerda Ben Thompson al explicar el verdadero motor de esas dinámicas: “la realidad es que las plataformas no son un problema huevogallina: lo primero que cuenta son los usuarios, que atraen a los desarrolladores a pesar de los obstáculos” (Thompson, 2025). En otras palabras, más que cada app individual, la plataforma misma es el centro de poder.

En educación, se dio una proliferación de apps educativas cuya información y gestión ya no se da en sistemas controlados por las instituciones, sino en “la nube” (una nueva metáfora), ofreciendo experiencias de aprendizaje a medida, pero a cambio de una pérdida de control sobre los datos, los procesos y los criterios de personalización. Los algoritmos deciden qué contenidos recomendar, qué rutas seguir, qué evaluaciones aplicar, casi siempre sin transparencia ni participación real de educadores o instituciones. Obsérvese como “la nube” y el “algoritmo” se convierten en los sujetos de análisis y crítica, cuando en realidad son instrumentos operados por las empresas que las diseñan y controlan.

Esta metáfora nos invita a preguntarnos: ¿quién adapta para quién? ¿Qué intereses comerciales, qué sesgos culturales o técnicos modelan los caminos que recorren los estudiantes?

La metáfora de la plataforma —el App Store— advierte que el aprendizaje personalizado puede convertirse en una personalización simulada, predefinida por intereses corporativos, si no se establecen mecanismos de soberanía tecnológica y pedagógica.

Metáforas alternativas para el aprendizaje adaptativo impulsado por inteligencia artificial

Si bien la metáfora del Palantír resulta útil para representar la fascinación y el temor que rodean el uso de la inteligencia artificial en contextos educativos, necesitamos marcos que nos permitan pensar en formas más críticas y pedagógicamente útiles de integrar estas tecnologías. Para ello los autores proponen dos metáforas alternativas que pueden ayudar a imaginar un sistema de apoyo al aprendizaje con GenAI más abierto, controlado por las comunidades educativas y centrado en el aprendizaje significativo: el conjunto de bloques de Lego y el dilema de Prometeo.

La metáfora de bloques de Lego se puede utilizar para explicar que, en lugar de concebir la IA generativa como un artefacto cerrado y mágico, la metáfora de Lego plantea una visión más accesible, modular y comprensible. Bajo este enfoque, la inteligencia artificial no es un “oráculo” que ofrece rutas de aprendizaje automáticas e incuestionables, sino un conjunto de piezas reutilizables que pueden combinarse de múltiples formas para diseñar experiencias de aprendizaje personalizadas, contextuales y con fundamentos pedagógicos.

Aplicado a las herramientas de aprendizaje basados en GenAI, esto implica que:

- Las herramientas basadas en IA generativa pueden ser configuradas por docentes para responder a las necesidades reales de sus estudiantes.
- El “algoritmo” no sustituye al educador, sino que le ofrece materiales con los que construir rutas alternativas.
- Las adaptaciones no se producen solo en función del rendimiento o la velocidad, sino también de intereses, lenguajes, motivaciones y trayectorias diversas.

Este enfoque conecta con una visión constructivista y participativa del aprendizaje, donde la tecnología no reemplaza las bases pedagógicas, sino que las amplifica. En este sentido, puede vincularse con la metáfora de la *construcción*, propuesta por Weller (2022), que entiende la tecnología educativa como un proceso en el que los usuarios (docentes y estudiantes) construyen activamente sus herramientas y entornos de aprendizaje, promoviendo la autonomía y la personalización. Además, permite imaginar futuros en los que los sistemas sean transparentes, auditables y creados por comunidades educativas, evitando así la dependencia de plataformas opacas o decisiones automatizadas que no pueden ser explicadas.

Una segunda metáfora relevante es la del dilema prometeico. En la mitología griega, Prometeo roba el fuego de los dioses y lo entrega a los humanos, un acto que representa tanto un avance para la civilización como un castigo potencial para la humanidad. “En el tallo de una caña escondí la chispa, el manantial del fuego robado, que para los mortales es maestra de todas las artes; y por ese delito pago ahora la pena, clavado con cadenas bajo el éter.” (Jonas, 1984).

Las herramientas impulsadas por GenIA pueden entenderse como sistemas con capacidad para ofrecer recomendaciones “personalizadas”, adaptar contenidos en tiempo real y predecir trayectorias educativas. Todo esto representa una forma de “fuego”, poderoso y transformador, pero también un riesgo si se usa sin tener en cuenta la ética, y la transparencia.

Esta metáfora alerta sobre:

- La posibilidad de delegar en exceso el juicio educativo a algoritmos automatizados, desplazando el papel del docente.
- El riesgo de generar trayectorias limitadas o cerradas para los estudiantes, guiados por patrones históricos de datos, en lugar de abrir caminos nuevos.
- Las implicaciones de reducir el aprendizaje a patrones de consumo o eficiencia, ignorando dimensiones más complejas como la creatividad, el conflicto o la exploración.

Así, mientras la metáfora del Lego apunta a la posibilidad de uso emancipador, la metáfora prometeica nos recuerda que toda tecnología poderosa conlleva tensiones éticas y políticas, especialmente cuando se despliega a gran escala en sistemas educativos.

PRINCIPIOS ÉTICOS PARA LA INTRODUCCIÓN SEGURA DE LA GENAI EN EDUCACIÓN

La creciente presión para integrar la GenAI en los sistemas educativos plantea no solo oportunidades, sino también desafíos éticos, legales y pedagógicos de gran calado. A nivel regulatorio, en la Unión Europea, los sistemas de GenAI deben garantizar una adopción segura, respetuosa con las normativas de privacidad de la Unión Europea (European Parliament & Council of the European Union, 2016). Además, su irrupción reciente ha impulsado el desarrollo de legislaciones específicas sobre inteligencia artificial en diversas regiones, como la Ley de IA de la Unión Europea (European Parliament, 2024) o las regulaciones emergentes en China (Cyberspace Administration of China et al., 2023).

Para garantizar una adopción segura, respetuosa con las normativas de privacidad y alineada con los valores, estrategias y prácticas institucionales, los autores proponen en esta sección una serie de principios que orienten la evaluación y despliegue de aplicaciones de GenAI en educación. Estos principios buscan asegurar que las tecnologías basadas en GenAI: 1) Se alineen con los objetivos educativos de las instituciones. 2) Mantengan niveles adecuados de seguridad, precisión e integridad ética. 3) Favorezcan un aprendizaje de calidad y equitativo, minimizando riesgos asociados al uso indebido, la privacidad o la desinformación.

Siguiendo estos principios, las instituciones educativas podrán aprovechar las oportunidades que ofrece la GenAI al tiempo que mitigan los riesgos asociados a su implementación en contextos de aprendizaje adaptativo.

Principios de GenAI segura en educación

Más allá del cumplimiento legal, la integración ética de la inteligencia artificial requiere principios prácticos y accionables que permitan evaluar tanto la adecuación pedagógica como la seguridad tecnológica de las soluciones adoptadas.

A partir de un análisis de los marcos legales y de la naturaleza técnica de la GenAI, Alier, García-Peñalvo y Camba (2024) han propuesto una serie de principios que, posteriormente, se han ampliado a siete directrices específicas para entornos educativos (García-Peñalvo et al., 2024). Estos principios permiten realizar una evaluación técnica, pedagógica y ética de las estrategias de integración de la GenAI en educación. A continuación, se exponen dichos principios.

- **(SAIE1) Garantía de confidencialidad:** los sistemas de GenAI deben proteger estrictamente los datos de los estudiantes, asegurando la privacidad de sus identidades, expedientes e interacciones educativas.
- **(SAIE2) Alineación con las estrategias educativas:** las herramientas de GenAI deben apoyar los objetivos institucionales y adherirse a las políticas de gobernanza tecnológica, facilitando el aprendizaje y la creación de conocimiento sin fomentar malas prácticas como el plagio o la elusión de controles académicos.
- **(SAIE3) Ajuste a las prácticas didácticas:** las aplicaciones de GenAI deben actuar bajo parámetros pedagógicos definidos. No deben resolver automáticamente problemas o tareas, sino guiar al estudiante en procesos de razonamiento alineados con el diseño instruccional.
- **(SAIE4) Precisión y minimización de errores:** los sistemas de IA deben priorizar la veracidad y relevancia de sus respuestas, especialmente para minimizar los riesgos de información errónea derivada de sesgos o alucinaciones de los modelos.
- **(SAIE5) Interfaz comprensible y comportamiento adecuado:** la GenAI debe presentar sus funcionalidades y limitaciones de manera clara para docentes y estudiantes, evitando falsas expectativas o interpretaciones erróneas de su nivel de fiabilidad.
- **(SAIE6) Supervisión humana y responsabilidad:** las decisiones basadas en GenAI deben ser supervisadas por personas. El profesorado debe mantener el control sobre los procesos educativos, asegurando la posibilidad de apelación y revisión de las decisiones automatizadas.
- **(SAIE7) Entrenamiento ético y transparencia:** los modelos de GenAI deben entrenarse de manera ética, garantizando transparencia sobre sus fuentes

de datos y metodologías, y minimizando los sesgos para asegurar resultados justos y explicables.

La aplicación de estos principios tiene importantes consecuencias prácticas para las instituciones educativas:

- **Confidencialidad (SAIE1):** las instituciones deben garantizar el control sobre las plataformas de IA utilizadas, evitando la obligatoriedad de registros en herramientas externas como ChatGPT o Gemini, lo cual podría vulnerar la privacidad estudiantil.
- **Alineación estratégica (SAIE2):** herramientas de propósito general, como los modelos de lenguaje de gran tamaño o Large Language Model en inglés (LLM), presentan dificultades de integración institucional debido a su complejidad de uso (Willison, 2023) y los riesgos asociados a la integridad académica (González-Geraldo y Ortega-López, 2024). Además, agregar complejidad al proceso de aprendizaje no es una buena práctica pedagógica, ya que aumenta la carga cognitiva del estudiantado (Chen et al., 2023).
- **Adaptación didáctica (SAIE3):** las aplicaciones deben integrarse de forma coherente con las metodologías educativas. Por ejemplo, en entornos como la formación médica (Hwang et al., 2024), los sistemas de GenAI deben actuar como asistentes en el razonamiento clínico, no como sustitutos automáticos de los procesos analíticos de los estudiantes.
- **Precisión y minimización de errores (SAIE4):** en entornos educativos, resulta fundamental diseñar sistemas de IA adaptativos capaces de citar fuentes y proporcionar mecanismos de validación de información (Tonmoy et al., 2024).
- **Interfaz y transparencia (SAIE5):** el diseño de las interfaces debe favorecer la comprensión del alcance y limitaciones de la GenAI, evitando la presentación de respuestas especulativas como verdades absolutas.
- **Supervisión y responsabilidad (SAIE6-7):** la supervisión humana no solo asegura la calidad educativa, sino que protege la equidad en el aprendizaje y garantiza que los sistemas adaptativos no reproduzcan sesgos o exclusiones sistemáticas.

Estos principios han dado lugar al desarrollo del Safe AI Manifesto (Alier, García-Peñalvo, Casañ et al., 2024), un documento vivo que recoge los compromisos necesarios para una integración ética y responsable de la GenAI en educación, disponible para consulta y suscripción en <https://manifesto.safeaieducation.org>. Además, los principios anteriores permiten realizar una evaluación técnica, pedagógica y ética de las estrategias de integración de la GenAI en educación, tal como se desarrolla en García-Peñalvo et al. (2024).

Relación de los principios de GenAI segura en educación y las metáforas educativas

Cada uno de los principios de GenAI segura que se proponen responden, de alguna manera, a los riesgos o aspiraciones representados por las metáforas históricas analizadas en el apartado anterior:

- **(SAIE1) Garantía de confidencialidad** combate el riesgo de la App Store y la nube educativa: La metáfora de la App Store muestra cómo la dependencia de plataformas opacas pone en riesgo la privacidad. La confidencialidad exige revertir esa lógica, asegurando que los datos de los estudiantes estén protegidos y bajo control institucional.
- **(SAIE2) Alineación con las estrategias educativas** responde al peligro de Matrix. Frente a la visión del aprendizaje como descarga pasiva de información predefinida, alinear la GenAI con las estrategias educativas permite construir sistemas que fomenten procesos activos de aprendizaje, en vez de automatizar la transmisión de contenidos.
- **(SAIE3) Ajuste a las prácticas didácticas** conecta con la metáfora del Bazar. Como en el Bazar del software libre, el principio busca que las herramientas de GenAI sean configurables y adaptables por los docentes a las prácticas pedagógicas específicas, no productos cerrados que impongan metodologías estándar.
- **(SAIE4) Precisión y minimización de errores** combate las ilusiones de la visión prometeica. El dilema de Prometeo nos recuerda que el “fuego” de la GenAI puede ser peligroso si no se gestiona adecuadamente. Minimizar errores y alucinaciones ayuda a controlar ese riesgo inherente a sistemas tan potentes como frágiles.

LAMB: UNA PROPUESTA PRÁCTICA PARA INTEGRAR LA GENIA SEGURA EN EDUCACIÓN

Tras analizar la evolución de las metáforas que han acompañado a las tecnologías educativas y definir los principios de GenAI segura (SAIE), resulta necesario mostrar ejemplos concretos que incluyan estos principios en su diseño y aplicación. En esta sección se presenta LAMB (Learning Assistant Manager and Builder), un marco de software diseñado específicamente para crear asistentes de aprendizaje basados en GenAI de manera segura, controlada y pedagógicamente alineada con los principios de la institución educativa. En primer lugar, se explica el concepto de asistente de aprendizaje basado en GenAI. A continuación, se define LAMB y como este marco permite crear asistentes de aprendizaje basados en GenAI. Después de presentar casos de uso de LAMB, como este marco está alineado con los principios presentados en la sección 3 y, finalmente, su relación con las metáforas educativas.

Asistentes de aprendizaje basados en GenIA

Un asistente de aprendizaje basado en GenIA es una herramienta conversacional que interactúa con los estudiantes para ofrecer apoyo educativo personalizado. A diferencia de un chatbot genérico, un asistente de aprendizaje está diseñado para cumplir funciones específicas dentro de un contexto educativo (Kochmar et al., 2020; Wollny et al., 2021). La aparición de herramientas como ChatGPT ha marcado un punto de inflexión en este campo, transformando las posibilidades de personalización e interacción educativa (Bettayeb et al., 2024; Chan y Lee, 2023):

- Responder a preguntas sobre el contenido de un curso.
- Guiar en la resolución de problemas aplicando métodos pedagógicos definidos.
- Sugerir recursos de apoyo pertinentes.
- Promover el razonamiento y la autoexplicación en los estudiantes.
- Acompañar procesos de evaluación formativa proporcionando retroalimentación ajustada.

Estos asistentes no buscan reemplazar al profesorado, sino amplificar su capacidad de acompañamiento y retroalimentación, especialmente en entornos de gran número de estudiantes o en modalidades híbridas y a distancia.

LAMB

LAMB es un entorno de software que permite a instituciones educativas y docentes crear asistentes de aprendizaje propios, adaptados a cada asignatura, curso o necesidad concreta. Este marco también permite integrar estos asistentes de manera sencilla en plataformas de gestión de aprendizaje (LMS) como Moodle, a través del estándar IMS LTI (Learning Tools Interoperability), de forma que los estudiantes acceden a los asistentes de aprendizaje a través del LMS de la institución (IMS Global Learning Consortium [IMS GLC], 2014). Además, LAMB permite controlar y configurar el conocimiento base que alimenta las respuestas del asistente, asegurando la alineación pedagógica con la materia. Adicionalmente, este marco permite gestionar la seguridad, privacidad y supervisión de los asistentes creados.

Alier, Pereira et al. (2024) presentan la descripción técnica, así como la arquitectura detallada de LAMB. De forma resumida y desde una perspectiva técnica, LAMB combina:

- Modelos de lenguaje de gran tamaño (LLMs) que generan respuestas en lenguaje natural.

- Técnicas de generación aumentada por recuperación o Retrieval-Augmented Generation en inglés (RAG), que limitan las respuestas del asistente a información verificada proveniente de fuentes controladas por la institución.
- Un sistema de plantillas de interacción y configuraciones pedagógicas que el docente define previamente.

De esta forma, LAMB permite aprovechar la potencia de la IA generativa, pero bajo condiciones de control, transparencia y responsabilidad humana.

Ejemplos de asistentes creados con LAMB

Para ilustrar la aplicación real de LAMB en contextos educativos, se presentan algunas de las experiencias realizadas.

En primer lugar, hay que destacar la experiencia desarrollada en la asignatura de Economía y Entorno económico, una asignatura obligatoria de segundo curso en el Grado de Ingeniería en Informática de la Facultad de Informática de Barcelona.

Antes de abordar experiencias piloto con estudiantes, es importante contextualizar el proceso de desarrollo de LAMB a través del asistente llamado “Macroeconomics Study Coach”, que funcionó como prototipo para el diseño y refinamiento del marco LAMB. Este asistente no fue sometido a evaluación formal con estudiantes, sino que sirvió como banco de pruebas técnico y pedagógico durante las iteraciones de desarrollo del sistema LAMB.

Este asistente fue creado a partir de 30 vídeos de conferencias del curso. Su funcionalidad principal consistió en responder preguntas de los estudiantes basándose en las transcripciones de los vídeos y documentos PDF complementarios, proporcionando citas específicas y enlaces a los vídeos con instantes precisos donde se abordaban las consultas realizadas.

Una vez consolidada la estabilidad del framework a través de este caso, se procedió a diseñar e implementar experiencias controladas con estudiantes, siendo el asistente PESTLE la primera de estas validaciones empíricas.

El asistente de aprendizaje del caso de PESTLE fue diseñado como recurso para la asignatura Fundamentos de Administración y Dirección de Empresas, impartida en el grado de Ingeniería en Informática en la EPSEVG (Escuela Politécnica Superior de Ingeniería de Vilanova y la Geltrú). El asistente creado ayudaba a los estudiantes a trabajar en un caso de evaluación de un proyecto empresarial impulsado por tecnología. El asistente había sido dotado de una base de conocimiento específica para el caso. Los estudiantes, organizados en equipos, debían analizar el caso utilizando la metodología PESTLE (Política, Economía, Sociedad, Tecnología, Legal y Medioambiente), lo cual implicaba que debían plantear preguntas desde varias dimensiones antes de redactar su informe final. Estas preguntas las plantearon al asistente de aprendizaje creado con LAMB. En cursos anteriores, los estudiantes

simulaban ser expertos con ayuda de motores de búsqueda, documentos específicos e incluso ChatGPT en el otoño de 2023 (Casañ et al., 2024).

La experiencia con el asistente para ayudar a los estudiantes en el caso de PESTLE se desarrolló durante el curso 2023-2024. La actividad involucró un total de 47 estudiantes de segundo curso del grado organizados en grupos de 6-8 miembros.

La experiencia se estructuró siguiendo una metodología de estudio de caso centrada en el análisis del robot humanoide Optimus de Tesla, utilizando el framework PESTLE (Political, Economic, Social, Technological, Legal, Environmental). La actividad se desarrolló durante dos sesiones de clase de 2 horas cada una, con una semana de intervalo entre ellas para permitir a los estudiantes continuar el análisis de forma autónoma.

En la primera sesión se realizó la presentación del caso y explicación de la metodología PESTLE, el trabajo colaborativo en grupos para analizar las seis dimensiones del PESTLE, así como la utilización del asistente para obtener información de expertos. Durante la semana los estudiantes tuvieron tiempo para seguir con el análisis utilizando el asistente desde casa. En la segunda sesión se introdujo el análisis SWOT integrado con PESTLE. A continuación, se pidió a los estudiantes que categorizaran los elementos o aspectos que habían encontrado en cada dimensión del PESTLE como “altamente relevante”, “relevante” o “poco relevante”. Para acabar los estudiantes elaboraron un informe final con conclusiones.

La metodología de recogida de datos para evaluar la efectividad del asistente, se administró mediante un cuestionario no obligatorio de 5 preguntas utilizando una escala Likert de 5 puntos (1=totalmente en desacuerdo, 5=totalmente de acuerdo). El cuestionario obtuvo **27 respuestas** de un total de 47 estudiantes (tasa de respuesta del 57,4 %). Las preguntas del cuestionario fueron las siguientes:

1. El asistente me ha ayudado a encontrar información relevante para el caso más rápidamente que si hubiera tenido que hacerlo yo mismo utilizando internet.
2. Las preguntas adicionales sugeridas por el sistema son buenas incitaciones para buscar más información sobre el caso.
3. Las respuestas a las preguntas sugeridas por el sistema proporcionan información útil.
4. Poder consultar las fuentes de datos ha sido útil.
5. La interfaz del asistente es sencilla y fácil de usar.

Los resultados muestran una recepción mayoritariamente positiva del asistente por parte de los estudiantes:

- **Pregunta 1** (Eficiencia en la búsqueda de información): 83,3 % de los estudiantes puntuaron con 4 o 5 (17 respuestas favorables de 27 total).
- **Pregunta 2** (Calidad de las preguntas sugeridas): 77,7 % puntuaron con 4 o 5.
- **Pregunta 3** (Utilidad de las respuestas): 85,2 % puntuaron con 4 o 5.

- **Pregunta 4** (Acceso a fuentes): 81,5 % puntuaron con 4 o 5.
- **Pregunta 5** (Usabilidad de la interfaz): 92,6 % puntuaron con 4 o 5.

Los resultados indican que ningún estudiante puntuó con 1 o 2 en las preguntas 1, 2 y 4, y solo un pequeño porcentaje lo hizo en las preguntas 3 y 5, evidenciando una ausencia significativa de valoraciones negativas.

Los resultados cualitativos y cuantitativos confirman varios beneficios pedagógicos:

En general, como retroalimentación positiva en la experiencia, las respuestas de los estudiantes destacaron beneficios como la consistencia y alineación. Se destacó la capacidad de los asistentes para ceñirse a la base de conocimiento proporcionada, sin generar información irrelevante o alucinaciones. En cuanto a la reducción de carga de trabajo, se valoró la utilidad de los asistentes para responder consultas repetitivas sobre procedimientos del curso. En el caso específico del asistente para ayudar a los estudiantes a resolver una evaluación de un proyecto empresarial, se observó una mejora en la calidad de los informes estudiantiles gracias al acceso a conocimiento experto.

Entre los retos y limitaciones cabe destacar la dependencia excesiva en este tipo de herramientas, ya que se advirtió sobre el riesgo de que los estudiantes usen el asistente como atajo, sin profundizar en los materiales del curso proporcionados por el profesorado. También se propusieron mejoras futuras de LAMB para una integración más fluida con el LMS, concretamente para permitir enviar consultas al docente cuando el asistente no tiene respuesta. Finalmente, se señaló la necesidad de apoyar lenguas minoritarias como el catalán o el euskera. En cursos anteriores, los estudiantes simulaban ser expertos utilizando motores de búsqueda, documentos específicos y, en el trimestre de otoño de 2023, ChatGPT de forma no controlada. La implementación del asistente LAMB proporcionó un entorno más controlado y pedagógicamente alineado, con información verificada y contextualizada específicamente para el caso de estudio.

Relación entre LAMB y los principios de GenIA segura en educación

El diseño de LAMB refleja de forma directa el cumplimiento de los siete principios de IA segura planteados en el apartado tres:

- **SAIE1 (Confidencialidad):** LAMB garantiza la privacidad de los datos estudiantiles al operar sobre servidores de la institución o mediante APIs controladas. La información sensible no se expone a plataformas externas no autorizadas.
- **SAIE2 (Alineación con estrategias educativas):** los asistentes creados con LAMB solo utilizan fuentes aprobadas por la institución, asegurando que

el contenido y las interacciones estén alineados con los objetivos y valores educativos institucionales.

- **SAIE3 (Ajuste a prácticas didácticas):** LAMB permite que los docentes definan plantillas de interacción y bases de conocimiento específicas para cada curso o actividad, integrándose de manera coherente con los diseños instruccionales, en lugar de imponer interacciones genéricas.
- **SAIE4 (Precisión y minimización de errores):** mediante técnicas de generación aumentada por recuperación (RAG), los asistentes basan sus respuestas en documentos de referencia fiables, reduciendo el riesgo de alucinaciones o errores típicos de los LLM generalistas.
- **SAIE5 (Interfaz comprensible y comportamiento adecuado):** LAMB proporciona asistentes con comportamientos predecibles y delimitados, aclarando su función y sus limitaciones, evitando la ilusión de infalibilidad que suele acompañar a muchos sistemas de IA.
- **SAIE6 (Supervisión humana y responsabilidad):** el profesorado diseña y supervisa los contenidos, estrategias y comportamientos del asistente, preservando su rol central como mediador pedagógico. Esta aproximación es consistente con experiencias prácticas donde los LLMs actúan como asistentes que automatizan tareas repetitivas (revisión de código, seguimiento de proyectos), pero manteniendo al docente en el rol de guía y supervisor del proceso educativo (Pereira et al., 2025).
- **SAIE7 (Entrenamiento ético y transparencia):** aunque los modelos subyacentes son de terceros, LAMB permite a las instituciones construir bases de interacción propias, que en el futuro podrán utilizarse para ajustar o entrenar modelos adaptados éticamente a sus contextos.

LAMB y las metáforas en educación

La implementación práctica de LAMB y los asistentes de aprendizaje en contextos reales ha revelado dificultades que, de manera significativa, validan las advertencias contenidas en el análisis metafórico desarrollado en la primera parte de este artículo. Esta convergencia entre teoría crítica y experiencia empírica refuerza la importancia de mantener una vigilancia activa sobre las narrativas que acompañan la integración tecnológica en educación.

Una de las preocupaciones más evidentes observadas en las experiencias con LAMB es el riesgo de dependencia excesiva por parte de los estudiantes, quienes pueden utilizar el asistente como un atajo que les evite profundizar en los materiales del curso proporcionados por el profesorado. Esta dificultad práctica materializa precisamente los peligros anticipados en la metáfora del Palántir.

Esta observación no invalida la utilidad de los asistentes, sino que refuerza la necesidad de marcos pedagógicos claros que orienten su uso. Como señaló uno de los expertos consultados: *“debe quedar claramente definido que su ‘conocimiento’*

no va más allá del contenido correspondiente y no trata contenidos más allá de las necesidades del profesor”. La metáfora del Palantír nos recuerda que el poder de la tecnología debe ir acompañado de estructuras de control y reflexión ética.

Las dificultades de integración con los sistemas de gestión del aprendizaje (LMS) observadas en las experiencias prácticas reflejan la tensión entre las metáforas del Bazar y la App Store analizadas anteriormente. Los estudiantes han solicitado mejoras para una integración más fluida con la plataforma Moodle (el LMS utilizado en la institución educativa), especialmente la posibilidad de realizar consultas al docente cuando el asistente no tenga respuesta adecuada.

Esta demanda ilustra la limitación inherente a los sistemas cerrados (metáfora de la App Store). Cuando las herramientas tecnológicas operan de manera aislada, sin capacidad de interconexión y colaboración, se genera frustración y se limita el potencial pedagógico. En contraste, la metáfora del Bazar sugiere la necesidad de sistemas abiertos, modulares y colaborativos donde diferentes herramientas y actores (estudiantes, profesores, sistemas) puedan interactuar de manera fluida.

LAMB, por diseño, intenta responder a esta tensión mediante su arquitectura modular y su integración con estándares abiertos como LTI. Sin embargo, las dificultades prácticas observadas subrayan que la apertura técnica no es suficiente: se requiere también una apertura pedagógica que permita la conexión entre la inteligencia artificial y la inteligencia humana del profesorado.

La necesidad identificada de soporte para lenguas minoritarias como el catalán y el euskera constituye una manifestación concreta de las limitaciones de la metáfora de Matrix discutida en el marco conceptual. La visión de Matrix sugiere que el conocimiento puede “descargarse” de manera instantánea y universal, ignorando las realidades contextuales, culturales y lingüísticas específicas.

La experiencia práctica con LAMB demuestra que esta universalidad aparente es ilusoria. Los modelos de lenguaje, por muy avanzados que sean, reflejan los sesgos y limitaciones de sus datos de entrenamiento, que tienden a favorecer las lenguas mayoritarias.

Desde una perspectiva conceptual, LAMB, rechaza la lógica de la App Store y del Palantír, donde los usuarios dependen de sistemas cerrados y opacos. En cambio, se alinea con la metáfora de Lego: los asistentes son herramientas modulares, personalizables y transparentes, adaptables.

En resumen, LAMB ofrece un ejemplo realista de cómo las instituciones pueden reapropiarse de la GenIA educativa como un recurso didáctico y no como una amenaza a su autonomía o a la agencia pedagógica.

CONCLUSIONES

Este artículo ha planteado una mirada crítica sobre la incorporación de la GenAI en educación, combinando análisis histórico, ético y conceptual. A través del uso de metáforas, se ha mostrado cómo las visiones sobre la tecnología educativa han

oscilado entre el control centralizado y la autonomía participativa. Metáforas como el Palantír, Matrix o la App Store ayudan a entender los riesgos actuales: opacidad, dependencia tecnológica y desalineación pedagógica. Frente a ello, se han propuesto marcos alternativos, como los bloques de Lego o el dilema prometeico, que invitan a una integración crítica, transparente y controlada por la comunidad educativa.

Sobre esta base, se han formulado siete principios para una GenIA segura, centrados en la confidencialidad, la alineación institucional, la supervisión humana y la transparencia. Estos principios no son teóricos, sino que se han aplicado en la práctica mediante LAMB, un entorno de software que permite a las instituciones diseñar y desplegar asistentes de aprendizaje configurables y auditables. Los casos analizados muestran beneficios claros en coherencia pedagógica, autonomía del alumnado y reducción de tareas repetitivas, pero también señalan retos como la dependencia o la integración con los LMS de las instituciones educativas.

En conjunto, este trabajo reivindica la necesidad de recuperar el control pedagógico sobre la tecnología, para asegurar que la inteligencia artificial en educación esté al servicio del aprendizaje significativo y no de intereses externos u opacos.

REFERENCIAS

- Adell-Segura, J., Castañeda, L. y Esteve-Mon, F. M. (2018). ¿Hacia la ubersidad? Conflictos y contradicciones de la universidad digital. *RIED-Revista Iberoamericana de Educación a Distancia*, 21(2), 51-68. <https://doi.org/10.5944/ried.21.2.20669>
- Alier, M., García-Peñalvo, F. J. y Camba, J. D. (2024). Generative artificial intelligence in education: From deceptive to disruptive. *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(5), 5-14. <https://doi.org/10.9781/ijimai.2024.02.011>
- Alier, M., García-Peñalvo, F. J., Casañ, M. J., Pereira, J. A. y Llorens-Largo, F. (2024). *Safe AI in Education Manifesto (v 0.4.0)*. <https://manifesto.safeaieducation.org>
- Alier, M., Pereira, J., García-Peñalvo, F. J., Casañ, M. J. y Cabré, J. (2024). LAMB: An open-source software framework to create artificial-intelligence assistants deployed and integrated into learning-management systems. *Computer Standards & Interfaces*, 92, 103940. <https://doi.org/10.1016/j.csi.2024.103940>
- Asimov, I. (1955–1976). *Historias de Multivac* [Serie de relatos]. Diversas publicaciones.
- Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z. y Dakalbab, F. (2024). Exploring the impact of ChatGPT: Conversational AI in education. *Frontiers in Education*, 9, 1379796. <https://doi.org/10.3389/feduc.2024.1379796>
- Casañ, M. J., Alier, M., Pereira, J. y García-Peñalvo, F. J. (2024). Asistentes de aprendizaje basados en inteligencia artificial: Principios de seguridad y experiencias de implementación en educación superior. En M. Navarro-Granados, J. J. Sánchez Amate, P. Berbel Oller y C. Rodríguez Jiménez (Eds.), *Investigación y conocimientos en la educación actual* (pp. 13-35). Dykinson.
- Ceruzzi, P. E. (2003). *A history of modern computing* (2ª ed.). MIT Press.

- Chan, C. K. Y. y Lee, K. K. W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Chen, O., Paas, F. y Sweller, J. (2023). A cognitive-load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review*, 35(2), 63. <https://doi.org/10.1007/s10648-023-09782-w>
- CyberspaceAdministrationofChina,National Development and Reform Commission, Ministry of Education, Ministry of Science and Technology, Ministry of Industry and Information Technology, Ministry of Public Security, & National Radio and Television Administration. (2023). *Interim measures for the management of generative artificial intelligence services*. Promulgadas el 13 de julio de 2023, en vigor desde el 15 de agosto de 2023. https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm
- Clarke, A. C. (1962). *Profiles of the future: An inquiry into the limits of the possible*. Harper & Row.
- European Parliament. (2024, 13 de junio). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (published in OJ on 12.7.2024)*. <https://data.europa.eu/eli/reg/2024/1689/oj>
- European Parliament, & Council of the European Union. (2016, 27 de abril). *Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (OJ L 119, 4.5.2016, pp. 1-88)*. <https://data.europa.eu/eli/reg/2016/679/oj>
- García-Peñalvo, F. J., Alier, M., Pereira, J. y Casañ, M. J. (2024). Safe, transparent, and ethical artificial intelligence: Keys to quality sustainable education (SDG 4). *International Journal of Educational Research and Innovation*, 22, 1-21. <https://doi.org/10.46661/ijeri.11036>
- González-Geraldo, J. L. y Ortega-López, L. (2024). Can AI fool us? University students' ability to detect ChatGPT. *Education in the Knowledge Society*, 25, e31760. <https://doi.org/10.14201/eks.31760>
- Hofstadter, D. R. y Sander, E. (2018). *La analogía: El motor del pensamiento* (R. Musa, Trad.). Tusquets.
- Hwang, G.-J., Tang, K.-Y. y Tu, Y.-F. (2024). How artificial intelligence supports nursing education: Profiling the roles, applications and trends of AI in nursing education research (1993–2020). *Interactive Learning Environments*, 32(1), 373–392. <https://doi.org/10.1080/10494820.2022.2086579>
- IMS Global Learning Consortium. (2014). *IMS Learning Tools Interoperability (LTI) implementation guide v2.0*.
- Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226850337.001.0001>
- Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V. y Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. En I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin y E. Millán (Eds.), *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco*,

- July 6-10, 2020, *Proceedings, Part II* (Lecture Notes in Computer Science, Vol. 12164, pp. 140-146). Springer. https://doi.org/10.1007/978-3-030-52240-7_26
- Llorens-Largo, F. (2019, 13 de febrero). Las tecnologías en la educación: Características deseables, efectos perversos. *Universidad*, sí. <https://www.universidadsi.es/las-tecnologias-en-la-educacion-caracteristicas-deseables-efectos-perversos/>
- Mason, J. (2018). A critical metaphor analysis of educational technology research in the social studies. *Contemporary Issues in Technology and Teacher Education*, 18(3). <https://citejournal.org/volume-18/issue-3-18/social-studies/a-critical-metaphor-analysis-of-educational-technology-research-in-the-social-studies/>
- Pereira, J., López-Gil, J. M. y Alier, M. (2025). The AI-powered classroom: LLMs as teacher assistants for enhanced software-engineering learning experiences. En R. Molina Carmona, C. J. Villagrà Arnedo, P. Compañ Rosique, F. García-Peñalvo y A. García-Holgado (Eds.), *Proceedings of TEEM 2024: The Twelfth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 115-124). Springer. (Lecture Notes in Educational Technology). https://doi.org/10.1007/978-981-96-5658-5_12
- Pressman, R. S. (1991). *Software engineering: A practitioner's approach* (3ª ed.). McGraw-Hill.
- Raymond, E. S. (2001). *The cathedral & the bazaar*. O'Reilly.
- Skinner, B. F. (1968). *The technology of teaching*. Appleton-Century-Crofts.
- Stallman, R. M. (1985). *The GNU manifesto*. Free Software Foundation. <https://www.gnu.org/gnu/manifesto.html>
- Thompson, B. (2025, 5 de mayo). *Platform power is underrated*. Stratechery. <https://stratechery.com/2025/platform-power-is-underrated/>
- Tolkien, J. R. R. (1954-1955). *The lord of the rings* (Vols. 1-3). George Allen & Unwin.
- Tonmoy, S. M. T. I., Mehedi Zaman, S. M., Jain, V., Rani, A., Rawte, V., Chadha, A. y Das, A. (2024). A comprehensive survey of hallucination-mitigation techniques in large language models. *arXiv preprint arXiv:2401.01313*. <https://doi.org/10.48550/arXiv.2401.01313>
- Wachowski, L. y Wachowski, L. (1999). *The Matrix* [Película]. Warner Bros.
- Weller, M. (2022). *Metaphors of ed tech*. Athabasca University Press. <https://doi.org/10.15215/aupress/9781771993500.01>
- Willison, S. (2023). *Prompt injection attacks: It's only going to get worse*. <https://simonwillison.net/2023/Apr/25/prompt-injection>
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M. y Drachsler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 654924. <https://doi.org/10.3389/frai.2021.654924>

Fecha de recepción del artículo: 1 de junio de 2025

Fecha de aceptación del artículo: 27 de agosto de 2025

Fecha de aprobación para maquetación: 26 de septiembre de 2025

Fecha de publicación en OnlineFirst: 16 de octubre de 2025

Fecha de publicación: 1 de enero de 2026