

Evaluación de modelos de Inteligencia Artificial generativa abierta y cerrada en educación superior

Paloma España-Naveira¹, Eva Luque-García², Luz del Pino Fernández-Valderrama Aparicio³, Daniel Cebrián Roble⁴

¹ Arte y Arquitectura, Universidad de Málaga, España

^{2,3} Proyecto Arquitectónicos, Universidad de Sevilla, España

⁴ Didáctica de la matemática, didáctica de las ciencias sociales y de las ciencias experimentales, Universidad de Málaga, España.

* Correspondencia: palomanaveira@uma.es

RESUMEN: Este estudio analiza el impacto de la Inteligencia Artificial (IA) generativa en la educación superior, contrastando el uso de modelos conversacionales abiertos con el de modelos cerrados. En el contexto de estudiantes de primer curso del Grado en Arquitectura (Universidad de Málaga y Universidad de Sevilla), bajo la metodología de Aprendizaje Basado en Proyectos, se evaluó la integración de NotebookLM, una herramienta basada en la Recuperación Aumentada por Generación (RAG). Los resultados mostraron que, aunque modelos abiertos como ChatGPT dominan el uso cotidiano del alumnado (90,9%), su uso irreflexivo fomenta la desinformación debido a las “alucinaciones” algorítmicas. En contraposición, la experimentación con NotebookLM mostró que el alumnado valora significativamente la fiabilidad, la precisión científica y el rigor documental de los sistemas cerrados. Sin embargo, esta arquitectura exige un rol metodológico activo: el estudiante-docente debe seleccionar críticamente sus fuentes antes de interactuar con la IA, lo que mitiga la brecha entre la fundamentación teórica y la praxis proyectual. Se concluye que las instituciones deben transitar de políticas restrictivas hacia una alfabetización en IA, promoviendo el uso de asistentes personalizados que combinen la fluidez generativa con un anclaje documental riguroso para fomentar el pensamiento crítico.

PALABRAS CLAVE: Inteligencia artificial; Pensamiento crítico; Educación superior; Desinformación; Asistentes personalizados de IA

ABSTRACT: This study analyzes the impact of generative artificial intelligence (AI) on higher education, comparing the use of open conversational models with that of closed models. In the context of first-year students in the Bachelor's Degree in Architecture (University of Málaga and University of Seville), using the Project-Based Learning methodology, the integration of NotebookLM, a tool based on Generation-Augmented Retrieval (RAG), was evaluated. The results showed that, although open models such as ChatGPT dominate students' daily use (90.9%), their unreflective use fosters misinformation due to algorithmic “hallucinations.” In contrast, experimentation with NotebookLM showed that students significantly value the reliability, scientific accuracy, and documentary rigor of closed systems. However, this architecture requires an active methodological role: the student-teacher must critically select their sources before interacting with the AI, which bridges the gap between theoretical grounding and project-based practice. It is concluded that institutions must transition from restrictive policies toward AI literacy, promoting the use of personalized assistants that combine generative fluency with rigorous documentary grounding to foster critical thinking.

KEYWORDS: Artificial intelligence; Critical thinking; Higher education; Misinformation; Custom AI assistants.

INTRODUCCIÓN

La Inteligencia Artificial (IA) generativa ha irrumpido en el panorama actual no solo como una disrupción tecnológica sin precedentes, sino también como una relevante Cuestión Sociocientífica (CSC). Dwivedi et al. (2023) advierten que la rápida adopción de estos sistemas trasciende la dimensión puramente instrumental para plantear debates éticos y cívicos profundos, relacionados con la opacidad de los algoritmos “caja negra”, la vulnerabilidad de la sociedad ante la facilidad de manipulación de la opinión pública y la capacidad de la IA para la generación masiva de desinformación. En el ámbito educativo, el dilema entre las enormes capacidades de estos modelos y su falta de transparencia exige abordar su integración desde el pensamiento crítico, cuestionando cómo las infraestructuras de IA modifican el pensamiento del alumnado (Bearman & Ajjaw, 2023).

Para comprender este impacto y planificar una actuación educativa, es crucial establecer una distinción tecnológica y cognitiva entre los sistemas de IA de interacción abierta o cerrada. Herramientas de uso masivo como ChatGPT operan bajo un modelo abierto, apoyados en Modelos de Lenguaje Grande (LLM) que analizan extensos conjuntos de datos para generar texto de naturaleza probabilística. El principal peligro educativo surge cuando el alumnado lo utiliza como búsqueda de la verdad sin llevar a cabo una verificación de la información, ya que estos sistemas sufren de “alucinaciones” sistemáticas (Bansal, 2025), produciendo información sintácticamente plausible, pero semánticamente falsa o carente de sustento epistemológico real (Ji et al., 2023). Este peligro sitúa al estudiante en una posición de vulnerabilidad frente al autoengaño y la desinformación, ya que probablemente carece del dominio conceptual necesario para aplicar la lente crítica que exige evaluar respuestas engañosas o erróneas (Lodge et al., 2023).

Frente a la dispersión y la incertidumbre epistémica del modelo abierto, existe el modelo de IA cerrado, que permite condicionar las respuestas a instrucciones y documentación, como en el caso de NotebookLM. Estas plataformas superan la mera generación probabilística al integrar una arquitectura de Recuperación Aumentada por Generación (RAG, por sus siglas en inglés). Desde un punto de vista técnico, la arquitectura RAG condiciona la generación de lenguaje, obligando al modelo a basar sus inferencias y respuestas en un contexto ampliado a partir de un corpus documental específico (Lewis et al., 2020). Pedagógicamente, este cambio es sustancial: al referirse a bases de conocimiento externas, se disminuye drásticamente el riesgo de generar contenido objetivamente incorrecto o de alucinaciones.

En este sentido, este trabajo presenta un estudio orientado a analizar la adopción y el impacto metodológico de la IA en un entorno cerrado con estudiantes de primer curso del Grado en Arquitectura. El objetivo principal de la investigación es contrastar la experiencia previa del alumnado con herramientas conversacionales de IA abiertas frente a cerradas. De este modo, se estudia cómo el modelo cerrado permite superar la brecha existente entre la fundamentación teórica y la praxis proyectual, fomentando el rigor documental, la fiabilidad académica y exigiendo un rol más activo y crítico por parte del estudiante en la validación de sus fuentes.

METODOLOGÍA

Para analizar el uso de herramientas de Inteligencia Artificial (IA), se realizó un estudio con una muestra de 39 estudiantes universitarios de primer curso de arquitectura. La recogida de datos se llevó a cabo mediante un cuestionario aplicado a tres grupos distintos correspondientes a asignaturas iniciales, de primer curso y primer cuatrimestre, del Grado en Arquitectura: el primer grupo pertenece a la asignatura Taller 1B de la Universidad de Málaga (UMA) con 10 estudiantes; el segundo y tercer grupo corresponden a la asignatura Proyectos 1 de la Universidad de Sevilla (US), divididos en el turno de mañana con 22 estudiantes y el turno de tarde con 7 estudiantes. Este perfil de estudiante se caracteriza por encontrarse en una fase de transición académica, donde convergen la necesidad de asimilar una nueva epistemología del diseño y la carencia de un bagaje previo en investigación bibliográfica especializada aplicada al proyecto arquitectónico. Habitualmente, el material de referencia (artículos teóricos, monografías, análisis de referentes y normativas) resulta ajeno o inabordable para estos estudiantes, generando una brecha entre la fundamentación teórica y la praxis proyectual.

Tanto en la UMA como en la US, el desarrollo docente se fundamenta metodológicamente en el Aprendizaje Basado en Proyectos (ABP) (Carcelén González, 2019). Si bien cada caso y asignatura persigue objetivos específicos, comparten una casuística común: el desafío de integrar la indagación teórica en la toma de decisiones del diseño. Tradicionalmente, la carga cognitiva que supone la lectura y asimilación de fuentes especializadas compite con la demanda de producción gráfica e ideación, relegando la investigación a un plano secundario. Para abordar esta brecha epistémica y adaptarse a las nuevas ecologías de la información de los estudiantes (altamente mediatizadas por el uso incipiente y a menudo acrítico de modelos generativos), se propone la integración de un asistente basado en IA. Se selecciona específicamente la herramienta NotebookLM por su arquitectura basada en la Recuperación Aumentada por Generación (RAG) sobre corpus cerrados, lo que permite un control riguroso de las fuentes. Al interactuar con el asistente, el estudiante establece un diálogo con los textos de la asignatura a través de citas trazables, redirigiendo su atención hacia las fuentes originales. Lejos de fomentar el consumo pasivo, la metodología exige la agencia crítica del alumnado al permitirles incorporar y justificar nueva bibliografía en sus proyectos.

El instrumento de recogida de datos incluyó dos dimensiones. En primer lugar, mediante la pregunta abierta “¿Qué IA habías usado?”, se permitió a los participantes indicar múltiples respuestas para conocer el tipo de IA previa que habían utilizado. Esta información se analizó mediante un proceso de categorización y consolidación semántica, agrupando las diversas variantes ortográficas de una misma plataforma (ej. “Chatgpt”, “Chat GPT”) para estandarizar los registros. Se realizó un análisis descriptivo de frecuencias y se calculó el porcentaje promedio de uso entre grupos, mitigando el sesgo estadístico derivado de la diferencia en el tamaño muestral de los tres grupos. En segundo lugar, para profundizar en la percepción de NotebookLM usado en clase y el impacto metodológico de su uso, se introdujo una pregunta cualitativa: “¿Qué diferencias has notado con el uso de NotebookLM respecto a la plataforma que estabas acostumbrada/o a usar?”. Es crucial destacar el contexto de novedad de esta herramienta para la muestra: al inicio del ejercicio, NotebookLM era prácticamente inédita, siendo conocida previamente por un único estudiante en cada uno de los grupos de Sevilla y por ninguno en el grupo de Málaga. Las respuestas abiertas fueron sometidas a un análisis de categorías para identificar patrones de percepción, ventajas y limitaciones frente a los asistentes conversacionales de uso habitual a la hora de procesar la información.

RESULTADOS

Resultados de la pregunta “¿Qué IA habías usado?”

Los resultados muestran un uso generalizado de la IA en las etapas formativas iniciales de arquitectura. Además, se observa una diferenciación entre el uso más popular de entre las distintas plataformas de IA (figura 1).

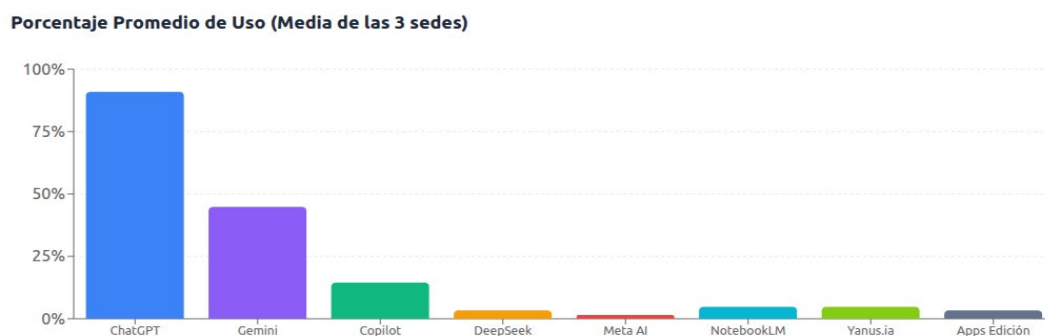


Figura 1. Porcentaje de uso de las distintas plataformas IA por el alumnado de arquitectura

ChatGPT se posiciona como la herramienta hegemónica, alcanzando un uso promedio del 90,9% al considerar los tres casos (con un uso del 100% en los grupos de la UMA y de la US Tarde). Le sigue *Gemini* con un 44,8% de uso promedio, lo que indica que una parte importante del alumnado no se limita a una sola opción, sino que ya usa múltiples IA. Asimismo, la aparición de IA secundarias como *Copilot* (14,5%), junto a menciones puntuales de IA más específicas o emergentes como *DeepSeek* o *NotebookLM*, sugiere una incipiente sofisticación en la selección tecnológica. Esto trae una profunda implicación pedagógica: la IA ha dejado de ser una innovación para convertirse en una herramienta básica y cotidiana. Por ello, las instituciones académicas deben superar el debate sobre la prohibición de la IA para centrarse en la transformación de sus metodologías. Esto implica rediseñar los talleres de proyectos, integrando herramientas de IA de forma estratégica y capacitando al alumnado en un uso ético, crítico y selectivo según cada necesidad educativa.

Percepción y contraste: El salto cualitativo hacia NotebookLM

Al analizar las diferencias percibidas, tras la interacción en clase con NotebookLM una herramienta nueva para el 90,9% del alumnado, emergen resultados claros que marcan una distinción respecto a las IA conversacionales como ChatGPT (tabla 1).

Tabla 1. Categorización, frecuencia y ejemplos representativos de diferencias percibidas en el uso de NotebookLM

Categorías	Subcategoría de análisis	Frecuencia	Ejemplo representativo del alumnado
Fiabilidad y Precisión	Anclaje documental	8	“Lo veo más fiable ya que se limita solo a lo que tú le aportas [...] no saca información de fuentes que yo no le haya aportado, no inventa nada.”
	Rigor y exactitud	9	“Una información más científica, redactada de manera más limpia [...] te da la página en la que ha encontrado esa información.”
Enfoque y Multimodalidad	Especificidad de las respuestas	13	“Ayuda en temas muy específicos y nichos [...] respuestas más precisas y profundas.”
	Recursos visuales integrados	6	“Una diferencia muy buena es la capacidad de generar vídeos, esquemas, tarjetas en base a la información [...] ChatGPT lo habría hecho peor.”
Limitaciones	Fluidez y profundidad	4	“Le falta fluidez a la hora de mantener la conversación [...] no profundiza tanto y repite muchas respuestas.”
Cambio de Paradigma	Selección crítica de fuentes	1	“Tienes que ser tú el que hace una vista crítica sobre qué documentos proporcionarle.”

Las distinciones más destacadas del alumnado son la fiabilidad y el anclaje documental. Las y los participantes valoraron de forma recurrente; por ejemplo, que la herramienta “se limita solo a lo que tú le aportas”, “no saca información de internet sin necesidad de adjuntar yo nada” y “no inventa nada fuera de los documentos”. Esto no se percibe como una limitación, sino como una mejora sustancial en la precisión y la fiabilidad académica, lo que, en palabras del alumnado, genera “información más científica” y “más fiable”.

En segundo lugar, destaca la especificidad de las respuestas y la multimodalidad. Frente a la generalidad que a veces se achaca a otras IA, NotebookLM es percibido como un entorno “más enfocado”, ideal para “temas muy específicos y nichos”. Además, el alumnado subraya el valor añadido de NotebookLM al aportar herramientas de estudio y visualización integradas (mencionando explícitamente la

generación de esquemas, mapas mentales, tarjetas y guías tipo podcast o “vídeos”), superando la interacción puramente textual y facilitando una mejor “organización visual” de la información. Sin embargo, también se identifican inconvenientes o limitaciones. Varios estudiantes perciben una menor “fluidez” para mantener conversaciones, una cierta tendencia a la “repetición de respuestas” y una menor capacidad autónoma para profundizar o redactar/sintetizar párrafos abiertos, en comparación con ChatGPT. Se destaca una respuesta que indica: “tienes que ser tú el que hace una vista crítica sobre qué documentos proporcionarle”, ya que captura la principal implicación pedagógica del uso de herramientas basadas en RAG en el aula: exige un cambio de paradigma en el estudiante. Se transita de la delegación de la búsqueda (propia de los prompts en ChatGPT) a la selección crítica de fuentes como paso previo e indispensable. Cuando el estudiante asume este rol activo en la selección de fuentes, la percepción generalizada es que la IA se convierte en un “ahorro de tiempo enorme” que “facilita mucho el trabajo”, alineándose de manera mucho más efectiva con el rigor exigido en el ámbito académico y de la investigación arquitectónica.

CONCLUSIONES

Este estudio evidenció la adopción de la IA por parte de la totalidad del alumnado de primer curso de arquitectura participante en este trabajo, la cual resultó ser una herramienta cotidiana. Sin embargo, al realizar una comparativa cualitativa, se identificó que el alumnado es capaz de discriminar entre la utilidad de un modelo conversacional abierto (rápido, pero generalista y propenso a inexactitudes) y el rigor de un entorno cerrado basado en fuentes primarias (NotebookLM), lo cual resulta importante para desarrollar el pensamiento crítico y la alfabetización en el uso de la IA. La valoración positiva de la fiabilidad, la precisión científica y las capacidades multimodales (generación de mapas y esquemas) señaladas por el alumnado como ventajas de la IA cerrada frente a la abierta es una muestra del valor que esta facilita para el estudio profundo y la organización visual de manera controlada, guiada y mediada por el profesorado de la asignatura.

A partir de estos hallazgos, se derivan implicaciones didácticas importantes para la enseñanza en talleres de proyectos y asignaturas teóricas afines:

1. La accesibilidad y el uso actual de la IA sugieren que las instituciones académicas abandonen políticas restrictivas y adopten una postura de alfabetización en torno a ella. Sería conveniente que la IA se integrara en las metodologías docentes, desde la guía docente.
2. El abordaje de problemas complejos mediante herramientas como NotebookLM puede forzar un cambio de rol en el alumnado. La competencia ya no reside únicamente en saber “cómo preguntar” a la IA, sino en la exigencia previa de seleccionar, evaluar y aportar documentos multidisciplinares sobre lo que se desea abordar desde la educación. La educación, así, puede contribuir a construir bases de conocimiento fiables (textos, normativas, referentes urbanos y sociológicos) que alimentarán a la IA y condicionarán sus respuestas.
3. Aprovechar las funciones de síntesis visual (esquemas, mapas mentales) de estas herramientas puede ser un buen punto de partida para las fases tempranas de análisis de problemáticas abiertas, facilitando el puente entre la lectura de textos complejos y la ideación gráfica de la propuesta.
4. Si las herramientas IA pueden ayudar a la síntesis y redacción inicial, los criterios de evaluación docente podrían orientarse hacia la capacidad crítica del estudiante para comprobar la información original, triangular fuentes y aplicar los resultados generados por la IA en resoluciones arquitectónicas tangibles que respondan al problema planteado.

A la luz de estos resultados preliminares y exploratorios se sugiere continuar investigando el uso de estas IA hacia una transición personalizada y cerrada de utilización de los mismos; por ejemplo, adoptando e investigando el uso de asistentes personalizados de IA (AIA) que permitan dar instrucciones a la vez que documentación cerrada a los asistentes. Además, se requiere estudiar un sistema que no sea ni abierto ni cerrado en el uso de la IA, en el que haya instrucciones y documentación que maneje la IA y que, mediante una personalización, se puedan construir AIA para tareas complejas en la educación.

Por ejemplo, a través de un AIA, el sistema orienta al estudiante, fomentando el desarrollo de sus habilidades cognitivas, metacognitivas y de aprendizaje autorregulado (Molenaar, 2022); siendo este, quizás, un buen punto para continuar a partir de las conclusiones de este trabajo.

AGRADECIMIENTOS

Proyecto I+D PID2024-156217NI-I00 Critical Education on Socioscientific Issues in Pre-service Science and Technology Teacher Training: Applications of Artificial Intelligence Assistants, financiado por Ministerio de Ciencia, Innovación y Universidades/Agencia Estatal de Investigación (MICIU/AEI/10.13039/501100011033) y Unión Europea (FEDER).

REFERENCIAS BIBLIOGRÁFICAS

- Bansal, V. (24 de noviembre de 2025). Estos trabajadores saben cómo funciona la IA y piden a sus amigos y familiares que no la usen: “Es un «no» rotundo”. *elDiario.es*. https://www.eldiario.es/tecnologia/trabajadores-funciona-ia-piden-amigos-familiares-no-usen-no-rotundo_1_12791733.html
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160–1173. <https://doi.org/10.1111/bjet.13337>
- Carcelén González, R. (2019). Metodologías de Aprendizaje Activo en Proyectos Arquitectónicos y su incidencia en la motivación del alumnado universitario. *Innovación Educativa*, (29), 95-108. <https://doi.org/10.15304/ie.29.5918>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1-38. <https://doi.org/10.1145/3571730>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Yih, W. T. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474.
- Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(2), 1-8. <https://doi.org/10.14742/ajet.8695>
- Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, 57(4), 632-645. <https://doi.org/10.1111/ejed.12527>