

Integración de Chatbots con IA Generativa en Educación Superior: Resultados de un Piloto en Programación de Sistemas

Carlos Alario-Hoyos, Rodrigo Prestes Machado, Carlos Delgado Kloos, Rebiha Kemcha, Iria Estévez-Ayres, Patricia Callejo

Introducción

En los últimos años, los chatbots basados en Inteligencia Artificial (IA) Generativa, especialmente aquellos soportados por *Large Language Models* (LLMs), tales como ChatGPT, Claude, o Gemini, se han convertido en herramientas cada vez más utilizadas por los estudiantes para resolver dudas, explorar conceptos y apoyar sus procesos de aprendizaje (Lo et al., 2024). Sin embargo, el uso de estos chatbots (en ocasiones también denominados asistentes de IA Generativa) no está exento de limitaciones. Por ejemplo, estos chatbots suelen carecer de conocimiento contextual específico para responder de forma más precisa a las preguntas planteadas por los estudiantes. Además, pueden generar respuestas incorrectas o engañosas (las llamadas “alucinaciones”). Finalmente, y lo que es más relevante desde la perspectiva educativa, su uso actual no deja trazabilidad ni ofrece datos accesibles para los docentes, lo que dificulta orientar y mejorar el aprovechamiento pedagógico de estos chatbots.

En este contexto, surge la necesidad de desarrollar soluciones más adaptadas al entorno educativo frente a la utilización de chatbots de propósito general (p. ej., ChatGPT). En particular, se propone la implementación de chatbots especializados para cada asignatura o materia, aprovechando la técnica de RAG (*Retrieval-Augmented Generation*) (Alario-Hoyos et al., 2024), con el objetivo de proporcionar respuestas más precisas y contextualizadas a los estudiantes. RAG combina la recuperación de información con la generación de texto para mejorar la calidad de las respuestas de un modelo de lenguaje, permitiéndole buscar y utilizar datos relevantes de una base de conocimiento externa antes de formular su respuesta final (Swacha & Gracel, 2025). De esta forma, el LLM no solo se basa en la información utilizada para su entrenamiento, sino que también incorpora información específica obtenida de dicha base de conocimiento para ofrecer contestaciones más precisas (Thüs et al., 2024). Además, RAG puede combinarse con la utilización de interfaces específicas para ofrecer el correspondiente chatbot a los estudiantes, de tal forma que el estudiante utilice dicha interfaz, la cual se conecta con la base de conocimiento y muestra la información devuelta por el LLM; de esta forma, se puedan recoger también datos de uso, incluyendo el *prompt* utilizado por el estudiante y la respuesta proporcionada por el LLM.

Es importante diseñar pilotos controlados que no solo evalúen la eficacia de los chatbots de IA Generativa, sino que también permitan recoger datos de los estudiantes. Estos datos pueden utilizarse para ayudar a que los estudiantes puedan desarrollar estrategias metacognitivas (Machado et al., 2025), y fomentar su adecuada alfabetización en el uso de la IA (*AI Literacy*).

Desarrollo del chatbot con RAG para la asignatura Programación de Sistemas

Con el objetivo de ofrecer a los estudiantes una herramienta más fiable y contextualizada, se ha desarrollado un chatbot basado en IA Generativa utilizando la técnica de RAG, de tal forma que se aprovecha el potencial de los LLMs junto con la capacidad de recuperar información específica y relevante desde una base de conocimiento previamente construida a partir de contenido de una asignatura concreta, permitiendo así generar respuestas más precisas.

Para construir la base de conocimiento, se ha recopilado y procesado todo el material didáctico disponible en una asignatura de programación (denominada *Programación de Sistemas*) de primer año de los grados en Ingeniería de Telecomunicación de la Universidad Carlos III de Madrid, España. Este material incluye los contenidos de un MOOC asociado, como los vídeos (concretamente, sus subtítulos), así como otros recursos como presentaciones y ejercicios (Alario-Hoyos et al., 2016). Todo este contenido textual ha sido vectorizado utilizando el modelo text-embedding-3-large de OpenAI, lo que permite representar semánticamente la información y facilitar su recuperación eficiente en función de las consultas realizadas. Una vez construida la base de conocimiento, se utiliza GPT-4o como motor generativo, integrando la información recuperada con la pregunta (*prompt*) del estudiante para elaborar respuestas más coherentes y alineadas con los contenidos del curso.

El sistema se ha desplegado a través de una plataforma web accesible de forma anónima para los estudiantes, donde pueden interactuar con el chatbot. Además, se ha desarrollado una interfaz específica para el profesorado, que permite consultar datos clave de las interacciones realizadas: el *prompt* original del estudiante, la respuesta generada por el modelo y los fragmentos de contenido recuperados de la base de conocimiento. Esta funcionalidad es esencial para entender cómo se está utilizando el chatbot, evaluar la calidad de las respuestas y orientar mejor el aprendizaje y la alfabetización en IA.

Piloto de utilización del chatbot en la asignatura Programación de Sistemas

Durante el segundo cuatrimestre del curso 2024-2025, se llevó a cabo un piloto con el objetivo de evaluar la integración del chatbot en un entorno real de enseñanza universitaria. El piloto se desarrolló desde febrero hasta mayo de 2025, y se ofreció a los estudiantes la posibilidad de usar libremente y de forma anónima el chatbot como herramienta de apoyo al aprendizaje. Desde el equipo docente se promovió activamente su utilización, integrándolo en las dinámicas del curso, particularmente en los laboratorios de la asignatura. Obviamente, nada impedía a los estudiantes utilizar otros asistentes de IA (p. ej., ChatGPT), aunque sí se reforzó el mensaje de que el chatbot ofrecido en la asignatura podía ofrecer respuestas más contextualizadas a la asignatura, recomendándolo fuertemente desde el equipo docente.

Los datos recogidos a lo largo del piloto reflejan un uso significativo de la herramienta por parte del alumnado (ver Figura 1):

- **Número total de conversaciones con el chatbot:** 342
- **Número total de mensajes intercambiados con el chatbot:** 1.783
- **Número medio de mensajes por conversación:** 5,21 (desviación estándar: 5,67)
- **Longitud media de los mensajes enviados por los estudiantes:** 406,8 caracteres (desviación estándar: 1.187,4)
- **Longitud media de las respuestas generadas por el chatbot:** 3.100,6 caracteres (desviación estándar: 1.570,2)

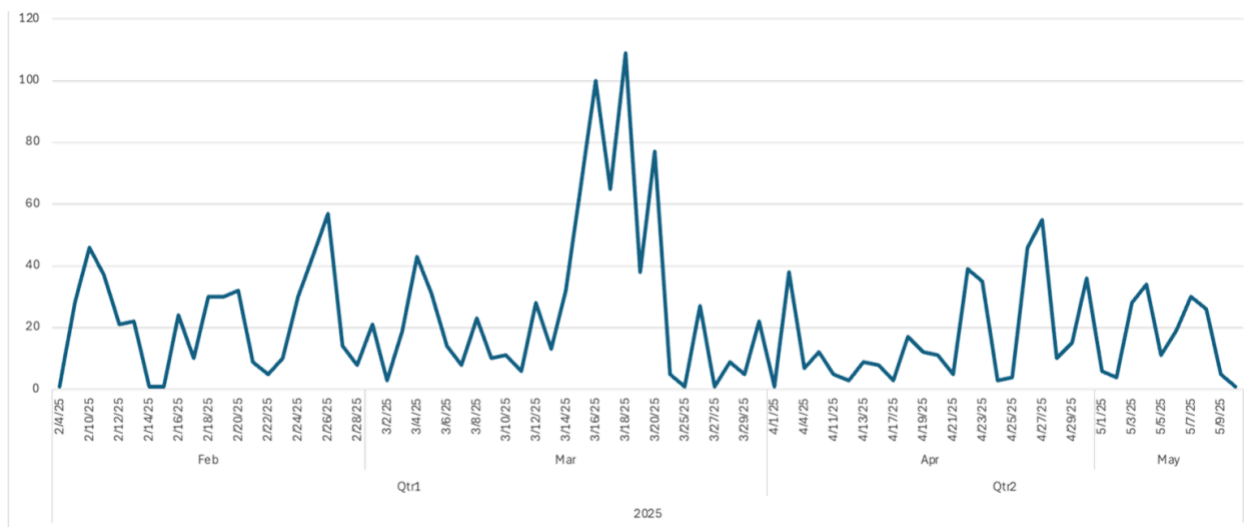


Figura 1. Número de mensajes intercambiados por día

Estos datos indican no solo una adopción razonable por parte de los estudiantes, sino también interacciones de cierta profundidad. La variabilidad observada sugiere una diversidad de usos: desde preguntas puntuales hasta consultas más complejas que requerían mayor contexto y desarrollo por parte del sistema. Sí se aprecia que los *prompts* de los estudiantes tienden a ser bastante cortos, con la excepción de aquellos casos en los que los estudiantes copian y pegan un ejercicio completo para ser resuelto por el chatbot o un trozo de código grande para ser explicado o depurado por el chatbot (ver Figura 2). También se aprecia un mayor uso del chatbot en las fechas cercanas a las pruebas de evaluación de la asignatura (ver Figura 1).

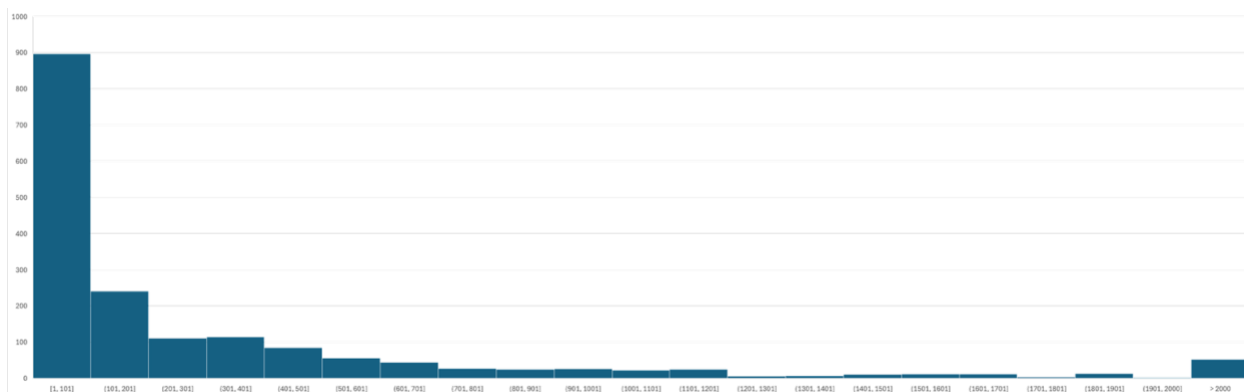


Figura 2. Longitud de los *prompts* de los estudiantes (número de caracteres)

Este piloto no solo ha servido para poner a prueba la viabilidad técnica de la solución implementada, sino también para explorar cómo los estudiantes interactúan con sistemas de IA Generativa dentro del marco de una asignatura de programación. Precisamente la industria del desarrollo de software y la programación de código están siendo impactadas fuertemente por los avances en los LLMs y por el desarrollo de herramientas y entornos que hacen uso de IA Generativa para facilitar las labores de desarrollo de software (Sauvola et al., 2024).

Los resultados de este piloto abren la puerta a futuras fases de análisis más detallado y a la posibilidad de integrar el uso de estos chatbots como herramientas permanentes de apoyo, así como recursos para fomentar competencias de alfabetización en IA.

Conclusiones

Este trabajo ha presentado los resultados de un piloto de utilización de un chatbot de IA Generativa que hace uso de la técnica de RAG en una asignatura de primer año de programación en ingeniería. El piloto no solo ha demostrado que un chatbot basado en IA Generativa con la técnica de RAG puede integrarse en una asignatura, sino que también ha puesto de manifiesto su valor como herramienta de apoyo al aprendizaje. El hecho de que el uso fuera libre y anónimo ha permitido observar comportamientos auténticos, lo que aporta una base para mejorar tanto la herramienta como su integración pedagógica. Además, contar con datos detallados sobre las conversaciones abre nuevas posibilidades para que el profesorado analice cómo los estudiantes piensan, preguntan y entienden, convirtiendo al chatbot no solo en un recurso de consulta, sino también en una fuente de información educativa valiosa y en un punto de partida para trabajar competencias en IA desde una perspectiva crítica y formativa. Precisamente, ya ha habido numerosos trabajos que han advertido de los efectos negativos que un uso excesivo y no controlado de la IA Generativa puede tener en el aprendizaje (Bastani et al., 2024; Gerlich, 2025).

Las líneas de trabajo futuro se centran en la comprensión del uso de chatbots en educación. En primer lugar, se realizará un análisis de la calidad de los resultados de las conversaciones, prestando especial atención a la precisión, relevancia y coherencia de las respuestas generadas por el chatbot; esto incluirá métricas para evaluar la satisfacción del usuario y la eficacia de la información proporcionada. En segundo lugar, se llevará a cabo un estudio de los aspectos metacognitivos del uso del chatbot con el objetivo de identificar y comprender las estrategias utilizadas por los estudiantes al interactuar con la herramienta, explorando, por ejemplo, cómo planifican sus consultas, monitorean el proceso de diálogo y evalúan la información recibida, con el fin de mejorar la adaptabilidad y el diseño de futuras interacciones.

Referencias

Alario-Hoyos, C., Kemcha, R., Kloos, C. D., Callejo, P., Estévez-Ayres, I., Santín-Cristóbal, D., ... & López-Sánchez, J. L. (2024, December). Tailoring Your Code Companion: Leveraging LLMs and RAG to Develop a Chatbot to Support Students in a Programming Course. In *2024 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)* (pp. 1-8). IEEE.

Alario-Hoyos, C., Delgado Kloos, C., Estévez-Ayres, I., Fernández-Panadero, C., Blasco, J., Pastrana, S., & Villena-Román, J. (2016). Interactive activities: the key to learning programming with MOOCs. *Proceedings of the European Stakeholder Summit on Experiences and Best Practices in and Around MOOCs, EMOOCS*, 319-328.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, O., & Mariman, R. (2024). Generative AI Can Harm Learning. *Available at SSRN, 4895486*.

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, 15(1), 6.

Lo, C. K., Hew, K. F., & Jong, M. S. Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 105100.

Machado, R. P., Alario-Hoyos, C., Callejo, P., Estévez-Ayres, I., & Delgado Kloos, C. (2025). Analyzing student use of spacing and interleaving strategies in interactions with GenAI-powered chatbots in programming courses. In *Proceedings of the 17th International Conference on Computer Supported Education (CSEDU)* (pp. 273–280).

Sauvola, J., Tarkoma, S., Klemettinen, M., Riekkı, J., & Doermann, D. (2024). Future of software development with generative AI. *Automated Software Engineering*, 31(1), 26.

Swacha, J., & Gracel, M. (2025). Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications. *Applied Sciences*, 15(8), 4234.

Thüs, D., Malone, S., & Brünken, R. (2024). Exploring generative AI in higher education: a RAG system to enhance student engagement with scientific literature. *Frontiers in Psychology*, 15, 1474892.