

Evaluating the response quality of LLM-driven university guidance applications

Evaluación de la calidad de modelos de IA generativa como aplicaciones de orientación universitaria

<https://doi.org/10.4438/1988-592X-RE-2025-411-729>

Estela Mayor-Alonso

<https://orcid.org/0000-0001-8226-3880>

University of León

Javier Vidal

<https://orcid.org/0000-0003-1060-6957>

University of León

Agustín Rodríguez-Esteban

<https://orcid.org/0000-0002-7409-5976>

University of León

Abstract

Recent technological advances are creating new educational challenges. The rise of Artificial Intelligence allows new useful tools for education to be implemented. For university guidance services, this includes applications such as Copilot and ChatGPT, based on multimodal language models. The aim of this study is to analyse the quality and reliability of the answers that Copilot and ChatGPT gave to questions posed by students in informal networks. The method was based on a qualitative, analytical validation approach to assess the accuracy of the answers. An observation tool was applied in Copilot and ChatGPT-4 for fifteen public universities. It consisted of 48 items, divided into four thematic blocks: access,

management, difficulty of courses, and employability. The responses in three of the thematic blocs, the exception being management, were deemed 100% suitable. Two items in the management block were found to be insufficiently suitable. Both were presented to the new multimodal language model, ChatGPT-4o, which gave improved answers. Subsequently, the answers provided by ChatGPT-4o and the information found on the websites were described. This highlighted confusion regarding information about the price of transferring credit on the websites and the difficulty finding academic credit limits when studying more than one course at the same time. The study concludes that Copilot and ChatGPT have potential as university guidance services. Their effectiveness will depend on the quality and accessibility of information on university websites. Universities must organise and update the information on their websites to improve the effectiveness of AI-based applications.

Keywords: artificial intelligence, guidance, recognition of studies, education technology, university

Resumen

Los recientes avances tecnológicos están creando nuevos desafíos educativos. El auge de la Inteligencia Artificial está posibilitando la implementación de nuevas herramientas útiles para la educación. En el caso de los servicios de orientación universitaria se destacan aplicaciones como Copilot o ChatGPT, basadas en modelos de lenguaje multimodal. El objetivo de este estudio es analizar la calidad y la fiabilidad de las respuestas que proporcionan Copilot y ChatGPT a preguntas planteadas por estudiantado en redes informales. El método se basó en un enfoque cualitativo y de validación analítica que evaluase la precisión de las respuestas. Se aplicó en Copilot y ChatGPT-4 una herramienta de observación conformada por 48 ítems, divididos en cuatro bloques temáticos: acceso, gestión, dificultad de estudios y empleabilidad, para quince universidades públicas. Se determinó un grado de ajuste suficiente del 100% para todos los bloques temáticos, excepto en gestión. En él, se encontraron dos ítems con un grado de ajuste insuficiente. Ambos fueron implementados en el nuevo modelo de lenguaje multimodal ChatGPT-4o y se detectó una mejora en el grado de ajuste. Posteriormente, se describieron las respuestas proporcionadas por ChatGPT-4o y la información encontrada en las páginas web, destacando la confusión respecto a la información del precio del crédito en las páginas web y la dificultad para encontrar el límite máximo de créditos a cursar simultaneando estudios. Se concluye que Copilot y ChatGPT tienen potencial como servicios de orientación universitaria. La eficacia de estos asistentes de IA dependerá de la calidad y accesibilidad de la información en las páginas web universitarias. Es fundamental que las universidades organicen y actualicen la información de sus páginas para mejorar la efectividad de aplicaciones basadas en IA.

Palabras clave: inteligencia artificial, orientación, reconocimiento de estudios, tecnología de la educación, universidad

Introduction

Technological changes in recent years have created new educational challenges. University education, an environment characterised by volatility, uncertainty, complexity and ambiguity (VUCA), is one of the most influential stages in students' professional development (Falcón-Linares & Arraiz-Pérez, 2017). The support and guidance services that universities offer—understood as the set of assistance and support strategies designed to guide students in their education processes and in their personal, professional and academic development (Sánchez Cabezas et al., 2018; Vieira et al., 2006)—must adapt in order to respond effectively to this changing environment.

Since 2010, guidance has been recognised as a fundamental right of university students and should be delivered through both individual and group approaches (González-Castellano et al., 2023; Royal Decree 1791/2010; Viñuela & Vidal, 2023). This right ensures that students receive the necessary support to address the academic and personal challenges they face throughout their time at university. Currently, this right continues to be acknowledged, and a more inclusive (Law 3/2022 on University Coexistence), more participatory environment is being promoted, where students actively collaborate in university activities and decisions (Organic Law 2/2023 on the University System).

Álvarez-Pérez et al. (2020) argued that the guidance process should begin in secondary and upper secondary education and continue at university. This guidance during the transition to university should be a collaboration between the various institutions. In order to promote effective guidance and counselling in the initial stages of students' time at university (Organic Law 2/2023 on the University System), issues such as volatility and uncertainty must be addressed. This is why both students and institutions need to be aware of the latest technological advances and be prepared to adapt to rapid and continuing changes in the educational environment (Chvanova et al., 2016; Cueva Gaibor, 2020).

In recent years, technological innovations have been one of the main drivers of change in society, so it would be interesting to integrate them into guidance, tutoring and support services (Flores-Vivar & García-Peñalvo,

2023). Digitising these services may make it easier for students to access support and improve how tailored it is, allowing for faster, more efficient intervention in certain situations.

In this regard, the emergence and rapid spread of artificial intelligence has had a significant impact on society, extending its influence to the field of teaching and learning (Bearman, 2022). However, despite the opportunities these innovations present, ethical issues must be addressed, especially in relation to data protection. This underlines the need to adopt strategies that promote the responsible use of artificial intelligence (Flores-Vivar & García-Peñalvo, 2023; UNESCO, 2021). In view of this situation, it is important to pay careful attention to the development of new regulations such as the European Union Artificial Intelligence Act (Regulation [EU] 2024/1689, 2024).

Artificial intelligence (AI) is the branch of computer science dedicated to the creation of intelligent systems capable of performing tasks automatically, simulating human behaviour (García-Peñalvo et al., 2024). It is characterised by tasks such as learning, reasoning, problem-solving, natural language comprehension, and decision-making support (Herrera-Ortíz et al., 2024). AI is divided into several disciplines, including machine learning, deep learning, and natural language processing (NLP) (Bearman & Ajjawi, 2023; Incio Flores et al., 2021; Vera-Rubio et al., 2023). Chatbots are particularly significant within NLP. They are services that simulate conversations with human users (Rodríguez Almazán et al., 2023). These conversations or interactions can be carried out with a pre-trained chatbot offering pre-programmed responses, or with AI based chatbots capable of understanding the message conveyed by a user and generating new responses (Guerrero-Bocanegra, 2022; Mayor-Alonso et al., 2024). Generative AI-powered chatbots can be adapted to a wide variety of contexts and needs, from customer service to education, offering personalised assistance and improving service efficiency.

In education, chatbots can be used in two main ways: as learning tools (Essel et al., 2022; Medrano et al., 2018) or as guidance tools (Artiles-Rodríguez et al., 2021). This paper aims to analyse the quality and reliability of two assistants based on large language models (LLMs), namely Copilot and ChatGPT-4, with the purpose of implementing these applications in

universities as a professional and academic support and guidance service. Although pages or services such as Watson Assistant and Decision Tree have been used previously for creating chatbots (Hiremath et al., 2018; Ogosi Auqui, 2021), this study will focus on Copilot and ChatGPT-4. Copilot is a generative AI virtual assistant developed by Microsoft (2023). It enables users to hold conversations or even generate images through DALL-3 model. In addition, Copilot is linked to OpenAI's generative language model, GPT-4, which has improved the chatbot's logical reasoning and conversational quality. Similarly, ChatGPT-4 is a generative AI virtual assistant and represents the penultimate multimodal language model released by OpenAI. The study will also consider the most recent model available at the time of the study, ChatGPT-4o, which introduced new features over GPT-4, such as enhanced text, audio and image comprehension, as well as faster response times.

In view of technological advances and the importance of knowing how to adapt guidance services, the following research question is posed: Could these language models specialise in university student guidance? In order to answer this question, the study aims to analyse the quality and reliability of the responses produced by Copilot and the ChatGPT-4 models. Quality and reliability, as detailed in the methodology, are understood in terms of how complete and accurate (or how suitable) the AI models' responses are. Within this context, the specific objectives are: (1) to analyse the suitability of the answers given by Copilot and ChatGPT to questions students posed in forums, and (2) to describe the content of the responses produced by ChatGPT-4o in relation to the information on university websites.

Method

The research method was qualitative and based on analytical validation (McMillan and Schumacher, 2005). The aim was to determine what responses generative AI provided and how well it agreed with the information on university websites. In other words, the aim was not to generalise results or make statistical inferences, but rather to assess content accuracy.

To that end, the analysis was based on an adaptation of the tool

developed by Mayor-Alonso et al. (2024) and its subsequent application in Copilot and ChatGPT-4. This study analysed the suitability of the answers produced by Copilot and ChatGPT-4 in 2024. Items with insufficiently suitable responses were selected and re-submitted to the more recent multimodal language model, ChatGPT-4o. Those newer responses were then compared with the information on the websites belonging to Spain's public universities.

Participants

The analysis used a sample of 15 public universities who offered a chatbot service on their websites that had previously been selected for a study by Mayor-Alonso et al. (2024) to analyse the suitability of their responses. Although that analysis was not directly relevant to the main objective of the present study, the sample was maintained in order to preserve its homogeneity and to be able to compare the effectiveness of the chatbots.

Instrument

The tool created by Mayor-Alonso et al. (2024) was adapted for data collection. The original tool consisted of 63 items divided into two tables: one related to the general procedural aspects of a chatbot (9 items) and another with the main questions to ask (54 items), extracted from an online forum (<https://yaq.es/>) where students ask questions about university. This dataset was obtained through web scraping as part of the 2021 Knowledge Generation Project on the use of AI in analysis of informal social networks for guidance in higher education, funded by the Ministry of Science and Innovation (PID2021-125405NB-I00).

The present study only used the items from the second table, as the focus was solely on the content of the answers. This information is presented in Annex I, Table IV, where for each thematic block the left-hand column lists the items, and the right-hand column describes the requirements for a complete answer.

In this adapted tool, the number of items was reduced to 48. Six items were discarded because they were too open-ended and could lead to multiple interpretations by generative AI models and researchers. This decision improved the validity of the instrument by eliminating items that could lead to subjective responses, rather than the informative and accurate information that chatbots are intended to provide.

The adapted tool maintains the original's thematic structure: access, management, difficulty of courses, and employability. Item selection and grouping into four thematic blocks was done inductively. This way of organising the tool also follows a chronological flow, reflecting the different phases of students' academic pathways: from access to university (before), through management, experience, and the difficulty of certain courses (during), to perceptions of employability (after).

The 48 items were posed to each of the 15 universities across two generative AI models, producing a total of 1,440 responses. Following the approach in Mayor-Alonso et al. (2024), response quality was graded on a dichotomous scale, in relation to the tool's descriptors: suitable (scored as 1 when a question is answered correctly), and unsuitable (scored as 0 when an answer is incorrect, either because it does not fit the descriptor, because quantitative information is not accurate, or when no answer is given).

Procedure

Two generative AI models were used: a) Microsoft Copilot, which enables conversations based on a predetermined context; and b) OpenAI's ChatGPT, which allows the creation of customised GPTs—generative AI models designed to generate text through deep learning.

To implement the tool in Copilot, a conversation was initiated for each university with the following message:

“Act as a guidance counsellor of University U to support students who are about to enter the university as well as those undertaking undergraduate, master's and PhD studies. To fulfil this role, search for all University U links you consider relevant in order to

respond to all questions.”

To implement the tool in ChatGPT, a database was generated with a minimum of 7 and a maximum of 10 public, accessible links from each university website, containing information related to the items extracted from the following sections: prospective students, access, pre-enrolment and registration, academic offer, undergraduate programmes, master’s degrees, PhD studies, scholarships and financial aid. Prompts were generated for the configuration and contextualisation of the GPTs. One GPT was created for each of the 15 universities, following these steps: (1) explore GPT, (2) create GPT, (3) configure the name, description and instructions, and (4) share privately. Once they had been created, they were given the following instruction:

“Act as a guidance counsellor of University U, addressing prospective students as well as those currently pursuing undergraduate, master’s and PhD studies. To fulfil this role, use the university web links specified below. These are related to the following topics: access, pre-enrolment and registration, scholarships and academic programmes.

You must answer all the questions posed by each student as if you were a guidance counsellor. All answers must be related to University U, ensuring that the information is relevant and specific to this institution. In cases where detailed answer cannot be provided, you should clearly inform the student and recommend seeking additional information through the university’s official channels.”

After obtaining the 1,440 responses, and in order to ensure the validity and reliability of the results, two researchers independently evaluated their suitability. The evaluation considered the descriptors for each of the items in Table IV in Annex I. Full agreement between the researchers (100%) was reached for ChatGPT-4, while in Copilot there was an initial discrepancy in 22 responses, yielding a Cohen’s Kappa coefficient of 0.842, considered to indicate almost perfect agreement (Cohen, 1968). The discrepancies were reviewed, which led to the conclusion that they stemmed from differing interpretations of the descriptors. After clarifying this issue, full agreement

(100%) was achieved in Copilot, allowing the analysis to proceed.

Considering that 100% of the responses in three of the four thematic blocks from both generative AI models were suitable, it was decided to examine why the same percentage was not achieved in the management thematic block.

After reviewing each of the responses in this block, two were identified as unsuitable:

- Item 25. What is the price for recognising credits from one degree programme to another?
- Item 26. If I am studying two degrees at the same time, how many credits can I take per year?

Because these are two complex questions requiring a specific quantitative response, and in light of the release of the new multimodal language model ChatGPT-4o on 13 May 2024, it was decided to examine how ChatGPT-4o would respond to them.

To obtain the 30 responses, and to streamline the research process, a Google Sheet integrated with ChatGPT was used, linking it to the new model. This tool, provided by the artificial intelligence company *Solver Intelligent Analytics*, enabled automatic generation of 30 answers to the two items in the ChatGPT-4o model directly from a spreadsheet. The process used *Google Apps Scripts* to make requests to the ChatGPT API and receive responses in real time.

The prompt used on this occasion was as follows:

"I want you to act as a guidance counsellor to support students who are about to enter university, as well as those undertaking undergraduate, master's and PhD studies. To fulfil this role, search for all the links and documents from University U that you consider relevant in order to answer the following questions: 1. What is the price for recognising credits from one degree programme to another? 2. If I am studying two degrees at the same time, how many credits can I take per year? I want an explicit answer that clearly states the price for recognising credits from one university degree to another and specifies how many credits I can take per year when pursuing two degrees simultaneously."

Once the responses were obtained, the suitability of each one was

determined, and a statistical analysis was carried out using the chi-squared test to establish whether the differences between Copilot, ChatGPT-4 and ChatGPT-4o were significant. In addition, the information provided by ChatGPT-4o was compared against the information on the university websites.

A manual search on each university website was conducted using the following key terms: “credit recognition” and “public prices” for item 25, and “simultaneous studies” and “regulations on enrolment continuity” for item 26. In almost all cases, no more than five steps were needed to access the information, and the information was found not only in PDF documents about public prices and regulations related to enrolment continuity, but also on the university websites themselves.

This procedure ensured a rigorous, systematic evaluation of the application’s responses, providing a comprehensive analysis of their accuracy and reliability in the context of public universities in Spain.

Results

The results are presented in three parts. First, indicating the suitability of the responses provided by Copilot and ChatGPT-4, highlighting the thematic blocks with suitable responses. Second, comparing the unsuitable responses to items 25 and 26 in the management block between Copilot, ChatGPT-4 and ChatGPT-4o. Finally, the analysis of the content in ChatGPT-4o’s answers to these two items is presented, alongside the information available on university websites. This structure allows for a detailed understanding of the results at each stage of the analysis.

Suitability of the responses produced by Copilot and ChatGPT-4

Table I shows the percentage of suitable answers in each of the thematic blocks. In three blocks, 100% of answers were suitable, indicating that Copilot and ChatGPT-4 answered all questions correctly.

TABLE I. Suitability of the responses produced by Copilot and ChatGPT-4 for each thematic block

Thematic blocks	Suitability of responses (%)	
	Copilot	ChatGPT-4
Access	100	100
Management	95	90
Course Difficulty	100	100
Employability	100	100
Total	98	97

Source: authors' own work

Examples of answers for each thematic block are given below. The responses presented for some items were chosen for their capacity to clearly show the type of information provided by the generative AI model. For reasons of space, some answers have been shortened, with ellipses (...) indicating omissions.

In the Access block, item 4 asks for information about the recommended baccalaureate for studying biochemistry or chemistry. Copilot and ChatGPT-4 took on a guiding role, offering advice on which type of baccalaureate would be most suitable. The responses not only indicated which baccalaureate would be most appropriate, but also explained the purpose of each option:

To study biochemistry or chemistry at U1, you would need a science baccalaureate. The university offers two specific pathways within the Science Baccalaureate: The Health Sciences Baccalaureate will allow students who take it to acquire basic knowledge and skills to pursue higher education focused on the health field. (...) The Scientific and Technical Baccalaureate is more related to computer science, engineering, and new technologies” (CopilotU1).

“To study biochemistry or chemistry at U12 or another university in Spain, it would be best to take a Science Baccalaureate, specifically the Health Sciences Baccalaureate or the Scientific-Technological Baccalaureate (...)” (ChatGPT-4U12)

In the Management block, item 23 asks whether a high grade point

average is needed in order to transfer from one university to another. The responses from Copilot and ChatGPT-4 provided guidance on the importance of achieving a good average grade, while clarifying that this is not the only determining factor that enables the transfer of an academic record.

“The transfer process may vary depending on the university and the degree program. In general, to request a transfer of records, it is necessary to have passed a minimum number of credits at your home university and to meet the admission requirements of the destination university. Regarding the grade point average, some universities take into account the average grade of the academic record for the transfer process. However, the weight of the average grade in the transfer process may vary depending on the university and the degree program” (CopilotU4).

“The requirement of having a high grade point average for transferring an academic record from one university to another varies depending on several factors, including the policy of the destination university and the demand for the programme of study you want to transfer to (...). Some universities may require a minimum grade point average, particularly for programmes with high demand or limited places (...)” (ChatGPT-4U2)

In the Study Difficulty block, item 33 asks for information about the difficulty of a degree in Computer Science. In the Copilot example, the responses were based on information presented in online surveys, and, like ChatGPT-4, provided guidance to consider:

“The difficulty of studying computer science can vary depending on the university and the country where you study. According to a survey of university students in Spain, computer engineering could be considered one of the most difficult degrees, given that it involves acquiring mathematical and computational knowledge. However, it is one of the most practical degrees available today, which makes learning easier. On the other hand, according to another source, a computer science or computer engineering degree is not difficult, it simply requires effort and consistency. If a student attends classes, pays attention (...)” (CopilotU5)

“The perception of the difficulty of studying computer science at U11 can vary widely depending on different individual and academic factors. Below are some aspects that may influence how the difficulty of this degree is perceived: 1. Interest and enthusiasm for technology and computer science (...) 2. Skills in mathematics and logic (...) 3. Commitment and dedication to study (...)” (ChatGPT-4U11)

In the Employability block, item 45 asks for information about career opportunities for a degree in Early Childhood Education. The examples from Copilot and ChatGPT-4 show that the responses indicated a wide range of professional pathways.

“Early childhood education graduates from U8 can choose from a wide variety of career opportunities. Some of the options include working as an early childhood educator in the first cycle of early childhood education, in nursery schools and public schools, in institutions with children (0-6 years old), in specific programs with children aged 0 to 6 who are at social risk, in childcare projects, in play centres, in family support projects, in environmental education in early childhood, in extracurricular and leisure activities, in school support, in children’s libraries, among others (CopilotU8).

“The Early Childhood Education degree prepares graduates for a variety of roles focused on the care and education of young children. Typical career opportunities include: 1. Teaching in early childhood education (...) 2. Special education (...), 3. Coordination and management of early childhood education centres (...)” (ChatGPT-4U14)

Comparison of suitability of responses from Copilot, ChatGPT-4 and ChatGPT-4o

As Table I shows, the only thematic block that did not achieve 100% suitability was management. The difficulty lay in items 25 and 26. Table II presents the percentages of suitable responses to these two items from Copilot, ChatGPT-4

and ChatGPT-4o. Although the differences were not statistically significant ($p>0.05$ in all cases), analysis of the absolute values does indicate a change that should be described.

TABLE II. Suitability of responses from Copilot, ChatGPT-4, and ChatGPT-4o

Suitability (%)	Copilot		ChatGPT-4		ChatGPT-4o	
	It25	It26	It25	It26	It25	It26
Insufficient	27	33	93	27	7	0
Sufficient	73	67	7	73	93	100

Source: authors' own work

Copilot gave a high percentage of suitable answers to items 25 (73%) and 26 (67%). For the latter, ChatGPT-4 also produced suitable answers 73% of the time. In contrast, it only produced suitable answers to item 25 in 7% of cases, because, as the example shows, the responses were incorrect, they did not explicitly mention a price, or they failed to match the descriptor.

“The price for the recognition of credits from one degree to another in Spanish universities, including U10, may vary depending on several factors, such as the Autonomous Community and the specific policy of the university (...)” (ChatGPT-4U10)

The responses from the subsequent model, ChatGPT-4o, were almost 100% suitable for item 25, and for item 26. In other words, all the responses provided quantitative, detailed data addressing the question.

A comparison of the information provided by the application and the information on university websites is presented below:

Item 25

Item 25 asks about the cost of credit recognition when changing from one degree programme to another, so the expected response should focus on providing an estimate of the associated costs. Comparing ChatGPT-4o’

answers with information found on university websites indicates that the information is presented in two ways: expressed in euros or in percentage terms.

None of the five university applications that reported the credit price in euros provided a correct answer (U1, U2, U3, U8, U15). This is because the price indicated refers to the credit cost of a course in first enrolment, rather than the credit recognition. For example, for U2, ChatGPT-4o stated:

“(...) For the 2022-2023 academic year, the price for credit recognition in undergraduate studies is approximately €12.62 per credit (...)” (U2)

On the university website, the public prices document specifies 30% of the amount established in Annex 1 (the annex lists the credit prices according to the number of enrolments), that is, 30% of €12.62 per credit. Therefore, if a subject has 6 credits and 30% must be applied to each of these credits, the cost for the recognition of the subject would be €22.74.

The ten universities (U4, U5, U7, U9, U10, U12, U13, U14) that responded with the percentage to be applied to each credit for recognition indicated that it should be 25%.

*“- **Undergraduate degree**”: The price per recognised credit is 25% of the price of the credit for first-time enrolment. For example, if the price of the credit for first-time enrolment is €25, the cost per recognised credit would be €6.25.*

*- **Master's degree**”: The price per recognised credit is 25% of the price of the credit for first enrolment. For example, if the price of the credit for first enrolment is €45, the cost per recognised credit would be €11.25.” (U10)*

Of these, only two (U6, U11) did not provide correct information. According to the public price established in the document, these two universities apply 30% rather than 25%.

Item 26

Item 26 asks how many credits can be taken per year when studying two

degrees simultaneously. The expected answer should focus on specifying the limits of credits permitted in concurrent studies. According to the university websites, the maximum number of credits for simultaneous studies is 90 credits.

Responses from the fifteen university applications indicated a maximum of 90 credits. However, three of them provided ranges. U1 provided a range of 90 to 120 credits, and U3 and U9 provided ranges of 78 to 90 credits, even distinguishing between simultaneous undergraduate and master's degrees.

"(...) in most Spanish universities, including U1, the limit is usually around 90-120 credits per academic year, although this may depend on the student's ability and the university's approval" (U1)

*"(...) Undergraduate students**: They can enrol in a maximum of 90 credits per academic year. Master's students**: They can enrol in a maximum of 75 credits per academic year"* (U9).

Three of the fifteen university applications (U2, U7, U9) mentioned a maximum of 90 credits, but this information could not be verified as it was not found on their respective websites.

This credit restriction, as mentioned in most responses and in the example (U15), aims to prevent students from facing an excessive academic workload.

"(...) It is important to note that this limitation is designed to ensure that students can manage their workload and maintain good academic performance" (U15).

Conclusions

This study analysed the suitability of responses provided by Copilot and ChatGPT-4 to student questions related to university guidance. Questions that produced unsuitable answers were then given to the new multimodal ChatGPT-4o and the responses were compared. The focus of the analysis was on the answers to items related to the price per credit recognition and to doing

simultaneous courses, which were in the management thematic block. The information provided by ChatGPT-4o was compared with the information on the websites of the fifteen universities, and an attempt was made to explore the possible reasons why the application failed to retrieve the information in some universities.

To address the first objective, the suitability of the responses provided by Copilot and ChatGPT-4 was analysed, indicating 100% suitable answers in three of the four thematic blocks. This suggested that the main focus of the study should be on the items with unsuitable answers, namely the two items that required quantitative information.

Subsequently, the study assessed whether ChatGPT-4o gave more suitable answers, which it did for both items, particularly in contrast to the less suitable answers from Copilot and ChatGPT-4 for both item 25 and 26. Despite that, a chi-square test showed that the difference was not statistically significant.

Regarding the second objective, content analysis was performed to describe the information in the responses in comparison to the information on the universities' websites. Consideration should be given to the difficulty that both AI and students face in locating such specific information.

Most of the information from the manual search was found in PDF documents about public prices and rules for remaining at university, or directly on the universities' websites. It was generally accessible in five steps or less. This suggests that the information is relatively available, but proper interpretation and presentation is crucial for guidance. The information could be organized to be more easily accessible on university websites, perhaps with a specific section referring to the number of credits that may be taken depending on whether a student is doing a single degree, a double degree, or simultaneous courses.

The initial research question was whether these language models specialise in university student guidance. ChatGPT-4o was shown to be a more accurate generative language model when it came to information on specific quantitative data in the two items analysed. However, given the continuous evolution of AI, both language models could specialise in university student guidance. Their potential as tools for the analysis and comprehension of

information opens new possibilities in education (González-Mayorga et al., 2024). Ongoing in-depth analysis of chatbot quality will be essential (Chiappe et al., 2025), as will continued verification of the information they provide against official sources to ensure accuracy. One implication of these models specialising in university guidance is the importance of maintaining up-to-date, well-structured information on university websites. University guidance must adapt to a new complex environment and take advantage of technological innovations, such as AI, to improve support and tutoring services (Bearman, 2022), thereby facilitating personalised attention (Viñuela & Vidal, 2023).

Implementing chatbots may be beneficial for educational guidance and learning (Rathore, 2022), as long as the ethics of using AI is taken into account (Flores-Vivar & García-Peñalvo, 2023). It is important to be aware that AI is still in development and, as this study shows, has not yet achieved maximum accuracy. It should be understood as one factor influencing educational quality (López Rodríguez del Rey et al, 2023; Sánchez Cabeza, 2017), a factor that the entire educational community must adapt to.

In academic year 2019-2020, 7% of students who began degree courses decided to change degree (SIIU, 2024), making it essential to optimise guidance services in order to facilitate administrative procedures such as credit recognition. Information must be structured so that AI can access it easily, which would also lead to greater student satisfaction when using these services (Segovia García, 2023). This would involve universities committing to updating their websites at least annually with the most relevant information and current regulations.

This study also confirmed that chatbots not only function as an information service, but also as a support, guidance and advisory service. This is particularly evident with questions related to experience, as in the thematic block on Course Difficulties.

In summary, although Copilot and ChatGPT show great potential as a university guidance service, their effectiveness is constrained by the quality of the information available on university websites. Therefore, universities must work to keep their websites up to date, in line with regulatory changes, deadlines and administrative procedures that may apply not only at the beginning of a university degree, but also throughout it. This would optimise

the use of AI-based applications.

One interesting future line of research would be to monitor the improvements being made in artificial intelligence as well as in university websites themselves. Moreover, given the large amount of information and regulations relating to universities, it may be worth developing chatbots devoted exclusively to managing administrative procedures and perhaps implementing them as a pilot guidance service at a university in order to gauge student satisfaction.

The need to continuously update information is also a limitation, as we live in a society that is constantly generating more knowledge, regulations and information. Another limitation lies in our perception of chatbots, which should be considered complementary rather than a replacement for guidance staff. Finally, the ethics of artificial intelligence must be considered (UNESCO, 2021). When implementing applications such as Copilot or ChatGPT on university websites, data confidentiality and data privacy must be ensured, bearing in mind that their development is still ongoing.

These conclusions and recommendations are essential in order to develop and adapt AI applications such as ChatGPT in the field of university guidance, ensuring that they are useful, accurate and reliable.

Bibliographical references

- Álvarez-Pérez, P.R., López-Aguilar, D., & Garcés-Delgado, Y. (2020). Vocational preferences, the transition and the adaption to university education: an analysis from the perspective of High School students. *Bordon. Revista de Pedagogía*, 72(4), 9-26. <https://doi.org/10.13042/Bordon.2020.80131>
- Artiles-Rodríguez, J., Guerra-Santana, M., Aguiar-Perera, M.V., & Rodríguez-Pulido, J. (2021). Agente conversacional virtual: la inteligencia artificial para el aprendizaje autónomo. *Pixel-bit*, 62, 107-144. <https://doi.org/10.12795/pixelbit.86171>
- 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](#)

- Bearman, M., Ryan, J., & Ajjawi, R. (2022). Discourses of artificial intelligence in higher education: a critical literature review. *Higher Education*, 1–17. <https://doi.org/10.1007/S10734-022-00937-2/TABLES/2>
- Chiappe, A., Sanmiguel, C., y Sáez Delgado, F.M. (2025). IA generativa versus profesores: reflexiones desde una revisión de la literatura. *Pixel-bit. Revista de Medios y Educación*, 72, 119-137. <https://doi.org/10.12795/pixelbit.107046>
- Chvanova, M. S., Hramov, A. E., Khramova, M. V., & Pitsik, E. N. (2016). Is it possible to improve the university education with social networks: The opinion of students and teachers. *2016 IEEE Conference on Quality Management, Transport and Information Security, Information Technologies, IT and MQ and IS 2016, October*, 33–38. <https://doi.org/10.1109/ITMQIS.2016.7751895>
- Cueva Gaibor, D. A. (2020). Transformación digital en la universidad actual. *Revista Conrado*, 16(77), 483–489.
- El uso de la IA en el análisis de redes informales para la orientación en Educación Superior. Ministerio de Ciencia e Innovación. Proyectos de Generación de Conocimiento 2021. Referencia PID2021-125405NB-I00.
- Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., y Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian highereducation. *International Journal of Educational Technology in Higher Education*, 19(1), 1–19. <https://doi.org/10.1186/s41239-022-00362-6>
- Falcón-Linares, C., & Arraiz-Pérez, A. (2017). Construcción eficiente y sostenible de la carrera: El portafolio profesional como recurso de orientación universitaria. *Revista Española de Orientación y Psicopedagogía*, 28(2), 8–29. <https://socioconstructivismo.unizar.es/>
- Flores-Vivar, J.-M., & García-Peñalvo, F.-J. (2023). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4). *Comunicar*, 31(74). <https://doi.org/10.3916/c74-2023-03>
- García-Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). The new reality of education in the face of advances in generative artificial intelligence. *Revista Iberoamericana de Educacion a Distancia*, 27(1), 9–39. <https://doi.org/10.5944/ried.27.1.37716>

- González-Castellano, N., Runte-Geidel, A., Berrios-Aguayo, B., & Muñoz-Galiano, I. M. (2023). Buenas prácticas docentes y tutoriales en el ámbito universitario: la visión del docente Good teaching practices and tutorials in the university context: the vision of the good teacher. *Estudios Sobre Educacion*, 45, 97–121. <https://doi.org/10.15581/004.45.005>
- González-Mayorga, H., Rodríguez-Esteban, A., y Vidal, J. (2024). El uso del modelo GPT de OpenAI para el análisis de textos abiertos en investigación educativa. *Pixel-bit. Revista de Medios y Educación*, 69, 227-253. <https://doi.org/10.12795/pixelbit.102032>
- Guerrero-Bocanegra, B. (2022). Tópicos frecuentes en los foros de acogida para el desarrollo de un chatbot de orientación inicial universitaria. *Revista de Psicología y Educación - Journal of Psychology and Education*, 17(2), 187. <https://doi.org/10.23923/rpye2022.02.225>
- Herrera-Ortiz, J. J., Peña-Avilés, J. M., Herrera-Valdivieso, M. V., & Moreno-Morán, D. X. (2024). La inteligencia artificial y su impacto en la comunicación: recorrido y perspectivas. *Telos: Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 26(1), 278–296. <https://doi.org/10.36390/telos261.18>
- Incio Flores, F. A., Capuñay Sanchez, D. L., Estela Urbina, R. O., Valles Corral, M. Á., Vergara Medrano, E. E., & Elera Gonzales, D. G. (2021). Inteligencia artificial en educación: una revisión de la literatura en revistas científicas internacionales. *Apuntes Universitarios*, 12(1). <https://doi.org/10.17162/au.v12i1.974>
- Ley 3/2022, de 24 de febrero, de convivencia universitaria. *Boletín Oficial del Estado*, núm. 47, de 24 de febrero de 2022, páginas 23201 a 23219
- Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario. *Boletín Oficial del Estado*, núm. 69, de 22 de marzo de 2023, páginas 34000 a 34050
- Mayor-Alonso, E., Vidal, J., y Rodríguez-Esteban, A. (2024). Los chatbots como herramienta de apoyo para la orientación universitaria. *Edutec. Revista Electrónica De Tecnología Educativa* (87), 188-203. <https://doi.org/10.21556/edutec.2024.87.2971>
- McMillan, J. H., & Schumacher, S. (2005). Investigación educativa. Prentice Hall / Pearson.
- Medrano, J.F., Castillo, C.A., Tejerina, M.A. (2018). *Empleo de un Chatbot en el aula como herramienta de apoyo al proceso de enseñan-*

- za-aprendizaje. V Jornadas de Innovación Educativa
- Ministerio de Ciencia, Innovación y Universidades (2024). *Datos y cifras del Sistema Universitario Español* <https://www.universidades.gob.es/publicaciones-e-informes/>
- Ogosi Auqui, J. A. (2021). Chatbot del proceso de aprendizaje universitario: Una revisión sistemática. *Alpha Centauri. Revista de Investigación Científica y Tecnológica*, 2(2), 29–43. <https://doi.org/10.47422/ac.v2i2.33>
- Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificial y se modifican diversos reglamentos y directivas (Ley de Inteligencia Artificial). Diario Oficial de la Unión Europea.
- Rathore, B. (2021). Exploring the Potential Impacts of Chatbot Software/Apps (ChatGPT) on Education: Benefits, Drawbacks, and Future Prospects. *International Journal of Multidisciplinary Innovation and Research Methodology*, 1(1), 2960-2968. <https://ijmirm.com/index.php/ijmirm/article/view/18>
- Real Decreto 1791/2010, de 30 de diciembre, por el que se aprueba el Estatuto del Estudiante Universitario. *Boletín Oficial del Estado*, núm. 318, de 31 de diciembre de 2010, páginas 109312 a 109352
- Rodríguez Almazán, Y., Parra-González, E.F., Zurita-Aguilar, K.A., Mejía Miranda, J. & Bonilla Carranza, D. (2023). ChatGPT: La inteligencia artificial como herramienta de apoyo al desarrollo de las competencias STEM en los procesos de aprendizaje de los estudiantes. *Revista electrónica de Computación, Informática, Biomédica y Electrónica*, 12 (1), 5-12
- Sánchez Cabezas, P. P., López Rodríguez del Rey, M. M., & Alfonso Moreira, Y. (2018). La orientación educativa en la actividad pedagógica profesional del docente universitario. *Revista Pedagógica de La Universidad de Cienfuegos*, 14(65), 50–57.
- UNESCO. (2021). Recomendación sobre la ética de la Inteligencia Artificial. En UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa
- Vera-Rubio, P. E., Patricia Bonilla-González, G. P., Quishpe-Salcán, A. C., & Campos-Yedra, H. M. (2023). La inteligencia artificial en la edu-

- cación superior: un enfoque transformador Artificial intelligence in higher education: a transformative approach Inteligência artificial no ensino superior: uma abordagem transformadora. *Polo Del Conocimiento*, 85(11), 67–80. <https://doi.org/10.23857/pc.v8i11.6193>
- Vieira, M. J., & Vidal, J. (2006). Tendencias de la Educación Superior Europea e implicaciones para la orientación universitaria. *Revista Española de Orientación y Psicopedagogía*, 17(1), 75-97. <https://bit.ly/3U6yfbg>
- Viñuela, Y., & Vidal, J. (2023). Guidance to students in Spanish public university degrees. *Revista Complutense de Educacion*, 34(3), 605–616. <https://doi.org/10.5209/rced.79759>
- Wagh, K. S., & Hiremath, G. (2018). Chatbot for Education System. *International Journal of Advance Research*. www.IJARIT.com

Annex I

TABLE III. Tool adapted from Mayor-Alonso et al. (2024)

Access	
Item 1. What are the entrance exams like?	Describe the university entrance exams, whether they are general or voluntary, and detail the admission process.
Item 2. How long is the break between one EBAU exam and another?	Provide an approximate break time
Item 3. How long is the EBAU mark valid for?	Describe the validity period, focusing on the general phase or the specific phase or both.
Ít4. To study biochemistry or chemistry, what would I need to do, a health sciences baccalaureate or a technology baccalaureate?	Describe which baccalaureate would be appropriate to take.
Ít6. How is the EBAU mark calculated?	Explain how to calculate the average mark and/or provide links to simulators
Ít10. Can I study a humanities degree after taking the voluntary biology and chemistry EBAU exams?	Explain how the choice of subjects in the EBAU affects career choices
Ít5. Can I apply for a place in more than one Autonomous Community?	Answer yes or explain how to apply in more than one Autonomous Community
Item 7. How do I pre-register for a degree?	Describe the process for pre-enrolling in a degree programme
Item 8. What are the deadlines for pre-enrolment?	Give an estimate of the deadlines (months or days) for pre-enrolment.
Ít9. If I do not reserve a place once the list of admitted students for the degree programme has been published, will I lose my place?	Answer yes or explain the policies and consequences of not reserving a place.
Ít11. How do I enrol?	Describe the process for enrolling in a degree programme.
Ít14. How do I pre-register for a master's degree?	Describe the process for pre-enrolling in a master's degree programme.
Management	
Ít15. If I enrol, do I have to pay anything? What if I am accepted onto the module in September, can I withdraw?	Explain the enrolment fees and the conditions for withdrawing
Ít16. We are halfway through the course and I want to drop out. Will I have to pay the enrolment fees?	Detail the financial obligations when dropping out of a course or possible exceptions
Q17. What scholarships can I apply for to study for a degree?	Describe the scholarships available for undergraduate studies.

Ít18. I have had a scholarship this year (I passed all 10 subjects) and I would like to know if I change degree programmes, could I receive the scholarship again next year?	Explain whether it is possible to receive the scholarship again and the conditions for renewal when changing degree programmes
Ít19. What scholarship can I apply for to study a master's degree?	Describe the scholarships available for master's degree studies
Ít20. I have a low grade point average and would like to start a PhD, but I know that I will not be eligible for the FPU, so I have thought about paying the tuition fees myself and applying for another scholarship later on. What other scholarships are available?	Explain alternative funding and scholarship options for doctoral students
Ít22. How do I transfer my academic record from one university to another?	Explain the procedure for transferring academic records and include the necessary documentation for the transfer
Ít23. Do I need to have a high grade point average to have my academic record transferred from one university to another?	Detail the requirements for transferring records, if any, and how the average grade affects the process.
Ít24. If you change universities, will you be awarded the same number of credits if you enter through the Selectividad exam as if you do so through a transfer of academic records?	Clarify the credit recognition policies in different scenarios of transfer between universities.
Ít25. What is the cost of recognising credits from one degree programme to another?	Provide an estimate of the costs associated with credit recognition and explain that this price may vary.
Ít26. If I am studying two degree programmes at the same time, how many credits can I take per year?	Explain the credit limits allowed when studying two degree programmes at the same time
Ít27. Even if I am in my first year of study, can I take subjects from other years?	Explain the regulations on choosing subjects from other years for first-year students.
Course Difficulty	
Ít29_40. (12 items) Is it difficult to study for a degree in [degree programme]?	Provide guidance on the difficulty or presents interesting statistics used to assess the difficulty
Employability	
Ít43_50. (12 items) What are the career opportunities for a degree in [degree]?	Present the career opportunities available

Source: authors' own work based on the original tool by Mayor-Alonso et al. (2024)

