International Journal of Innovative Technologies in Economy

e-ISSN: 2414-1305

Scholarly Publisher
RS Global Sp. z O.O.
ISNI: 0000 0004 8495 2390

Dolna 17, Warsaw,
Poland 00-773
+48 226 0 227 03
editorial_office@rsglobal.pl

ARTICLE TITLE	EVALUATING THE EFFECTIVENESS OF AI CHATBOTS IN HIGHER EDUCATION AND RESEARCH ACTIVITIES: PERSPECTIVES OF ALGERIAN UNIVERSITY PROFESSORS USING THE URPIC MODEL
ARTICLE INFO	Nadia Tahmi, Nadia Brahimi, Mahdjouba Benchohra. (2025) Evaluating The Effectiveness of AI Chatbots in Higher Education and Research Activities: Perspectives of Algerian University Professors Using The URPIC Model. *International Journal of Innovative Technologies in Economy*. 3(51). doi: 10.31435/ijite.3(51).2025.3865
DOI	https://doi.org/10.31435/ijite.3(51).2025.3865
RECEIVED	01 January 2025
ACCEPTED	15 April 2025
PUBLISHED	03 September 2025
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

EVALUATING THE EFFECTIVENESS OF AI CHATBOTS IN HIGHER EDUCATION AND RESEARCH ACTIVITIES: PERSPECTIVES OF ALGERIAN UNIVERSITY PROFESSORS USING THE URPIC MODEL

Nadia Tahmi
Laboratory of Economic strategies and policies, Management Department, Faculty of Economics, University of Mohamed Boudiaf, M'sila (Algeria)
ORCID ID: 0009-0004-5246-420X

Nadia Brahimi
Laboratory of Economic strategies and policies, Management Department, Faculty of Economics, University of Mohamed Boudiaf, M'sila (Algeria)
ORCID ID: 0009-0000-2453-5243

Mahdjouba Benchohra
Laboratory of Economic strategies and policies, Management Department, Faculty of Economics, University of Mohamed Boudiaf, M'sila (Algeria)
ORCID ID: 0009-0008-8605-2560

ABSTRACT

This study aims to evaluate the advantages of AI chatbots used by university teachers in Algeria for their education and research activities, considering increased human-machine interactions, especially with advances in natural language processing. The assessment is based on a model consisting of five dimensions: uses, reliability, privacy, integration, and constraints to use. The questionnaire items were developed based on this model, and the study sample included 246 university professors from various disciplines. The responses were analyzed using the SPSS 27 software. Among the most important findings of the research is the moderate evaluation of AI chatbots effectiveness by the study sample, based on an average level across the five aforementioned dimensions. The sample also placed significant emphasis on the criteria of reliability, privacy protection, and integration in order of importance.

KEYWORDS

AI chatbots, Higher Education, Scientific Research, Effectiveness, URPIC Model, Algerian Universities

CITATION

Nadia Tahmi, Nadia Brahimi, Mahdjouba Benchohra. (2025) Evaluating The Effectiveness of AI Chatbots in Higher Education and Research Activities: Perspectives of Algerian University Professors Using The URPIC Model. *International Journal of Innovative Technologies in Economy*. 3(51). doi: 10.31435/ijite.3(51).2025.3865

1. Introduction

In contemporary studies, the implications of artificial intelligence programs regarding higher education and scientific research have been garnering more interest. The integration of AI chatbots in higher education have changed the landscape of academic support, research, and education greatly. These areas of support, research, and education have radically changed, in ways unimaginable only a few years ago. AI Enabled tools can provide new opportunities based on accessibility, personalized learning experiences, and efficiencies for administrative tasks, but they can also have implications in the realms of ethics and data security and the balance of human interaction and artificial intelligence.

AI chatbots are one of the types of artificial intelligence tools, which have provided fertile ground for researchers, because they assist in improvements to administrative efficiency by helping to streamline processes like student admissions by providing 24/7 access to information and alleviating the burden on administrative staff (Kristophorus et al., 2024). They also contribute to an improved educational experience, as AI-powered chatbot programs offer personalized support, increase student engagement, and facilitate skill development due to their ease of use and effectiveness-thus enhancing the learning environment (Bolambao et al., 2025). Additionally, tools like ChatGPT are used to support smart teaching, personalized learning, and language assistance (van Rensburg & Reedy, 2024) .

While there are many benefits associated with the use of AI chatbot programs in higher education, there are also a number of challenges that must be carefully considered. These include both ethical and practical challenges, including bias, privacy, data security, and information verification (Bolambao et al., 2025; van Rensburg & Reedy, 2024) . Another concern relates to skills development, in that overuse of AI chatbots may inhibit both students' and professors' skills related to critical thinking and exploring, and potentially hurt academic integrity (Zafar et al., 2024) .

To help address these challenges, various guiding and instructional solutions have been suggested. This includes proposing that institutions develop transparent guidelines regarding the use of AI chatbots programs to address ethical challenges and promote ethical use (van Rensburg & Reedy, 2024). Also, AI chatbots programs should be designed for collaborative use and specifically trained on relevant datasets to maximize usefulness in an academic context (Khandakar et al., 2024).

Research Problem: Amid conflicting views between the benefits of artificial intelligence and concerns about its use, and in light of researchers' growing interest in studying the role of AI chatbots in higher education, where most previous studies have focused on evaluating their role based primarily on the benefits provided to students, this study seeks to address the problem of measuring the effectiveness of AI chatbots in the education and research activities of Algerian university professors. We propose a model that integrates both the advantages and limitations of their use, consisting of five dimensions: the AI chatbots uses, the reliability of the information they provide, their protection of user privacy, their integration with the professor's teaching and research functions, and finally, the constraints to their use.

Research Questions: Primary Research Question: To what extent are AI chatbots effective in educational and research activities based on the characteristics of the URPIC model, from the viewpoint of a sample of university professors in Algeria? Are there differences in the responses given by participants?

Sub-questions:

- To what extent do the study sample make use of AI chatbots?
- How trustworthy is the information provided by AI chatbots from the perspective of the study sample?
- To what extent do AI chatbots ensure data privacy according to the study sample?
- To what extent do AI chatbots integrate into the teaching and research activities of university professors according to the study sample?
- To what extent do constraints to the use of AI chatbots exist according to the study sample?

Research Hypotheses: To answer the research questions and based on previous studies, the following hypotheses are proposed:

Main Hypothesis 1: AI chatbots are characterized by a high and statistically significant level of effectiveness in education and research activities at the significance level ($sig \propto \leq 0.05$), from the perspective of

the study sample, based on the model's dimensions: usage, reliability, privacy, integration, and usage constraints.

- There is a high and statistically significant level at (sig∝≤0.05) for the **usage** dimension of AI chatbots from the perspective of the study sample.
- There is a high and statistically significant level at (sig∝≤0.05) for the **reliability** dimension of AI chatbots from the perspective of the study sample.
- There is a high and statistically significant level at (sig∝≤0.05) for the **privacy** dimension offered by AI chatbots from the perspective of the study sample.
- There is a high and statistically significant level at (sig∝≤0.05) for the **integration** dimension of AI chatbots from the perspective of the study sample.
- There is a high and statistically significant level at (sig∝≤0.05) for the **constraints** dimension related to AI chatbots usage from the perspective of the study sample.

Main Hypothesis 2: There are statistically significant differences in responses regarding the effectiveness of AI chatbots at the significance level (sig∝≤0.05), attributable to the demographic variables (gender, age, academic rank, qualification, experience, and field of specialization) of the Algerian university professors in the study sample.

Research Importance: The importance of our research topic stems from several factors, including the rapid transformation in research and teaching methods driven by the use of artificial intelligence in Algerian universities, particularly generative AI in its various forms and tools. Among the most well-known and widely used of these are AI chatbots, including general-purpose ones such as **ChatGPT**, **Microsoft Copilot**, **Google Gemini**, **Grok**, and **Deepseek**; those focused on accuracy like **Perplexity**; and others oriented toward business applications like **Claude AI**. This growing reliance highlights the need to establish clear criteria for evaluating the effectiveness of the AI chatbots used by university professors in their educational and research activities.

Research objectives:

- Evaluate the actual uses of AI chatbots by university professors in education and scientific research in Algeria;
- Evaluate the reliability of AI chatbots from the perspective of university professors in Algeria;
- Evaluate the ability of AI chatbots to protect data privacy from the perspective of university professors in Algeria;
- Identify the obstacles and limitations that prevent university professors in Algeria from effectively using AI chatbots.

Research Contribution: This research study provides a significant contribution by introducing a model that could be utilized to measure the effectiveness of AI chatbots in the research and teaching activities of university professors. Although most of the previous studies were concentrated on the use of AI chatbots by students, the use of these intelligent chatbots is as important (if not more) to a university faculty member in the research role through potentially supporting research, enhancing quality, and improving the standards by which academics and intellectuals exist within their disciplines.The research sample consisted of 246 university professors from different areas of study, and academic ranks. At the empirical level, this research developed a valid measurement tool. Each dimension of the proposed model has been studied and ranked empirically in different research, which enhances the model as a tool for future experimental studies into the effectiveness of AI catbots, especially the importance these add to faculty services and academic outputs.

In practice, the model may also have an important part to play in assisting the development of AI chatbots that are contextually relevant to the working practices of university professors within Algerian universities.

2. Background and Literature Review

Artificial intelligence aims to make computers sophisticated enough to mimic the cognitive functions of humans (Hassan & Ablahad, 2023, p. 434) and AI chatbot is a software application designed to conduct online chat conversations via text or text-to-speech, rather than providing direct contact with a human representative in real time. Chatbots systems are developed to convincingly simulate human behavior as conversational

partners, often requiring continuous tuning and testing. Many of them, however, are still unable to carry out proper conversations or meet industry standards, such as passing the Turing test .

Estimates of AI chatbots efficiency involve analyzing various performance metrics and methodologies that assess their effectiveness in real-world applications. Studies show AI chatbots can help improve operational efficiencies, decrease processing times, and enhance user satisfaction. A useful key performance indicator is response time reduction, AI chatbots can reduce response times by at least 70% for straightforward inquiries, allowing human personnel time to work on more complex problems (Uzoka et al., 2024, p. 2485). Responsive bots, by using advanced forms of natural language processing (NLP), have helped achieve better accuracy and better user satisfaction (Ortiz-Garces et al., 2024, p. 01).

Previous studies indicate that well-designed AI-Chatbots can enhance accessibility by offering tailored support services for underrepresented student groups, thereby promoting equity within academic environments (Bobro, 2024). Among these studies is the research conducted by Jomar Elizabeth Guzman Seraquive and colleagues, which evaluated user experience with chatbot programs using a proposed four-factor model. This model was later streamlined into a simpler two-factor model focusing on "Chatbot Response Quality" and "User Experience and Satisfaction." The study found that the two-factor model explained a greater proportion of variance in user experience, highlighting its effectiveness in assessing the impact of chatbots on education and informing future improvements in learning experiences (Elizabeth et al., 2024).

In another study concerning an AI-based chatbot located on the website of the University of Namibia, researchers uncovered various insights. One finding was the successful chatbot role in answering frequently asked questions, or previously accessed information, that needed time-consuming solutions (phone calls or emails). The capabilities of switching to a chatbot to get a quicker resolution and lessen the wait time for users, students or staff, is an added benefit. The course of creation for the chatbot, like many UX processes, was through an iterative process involving users - as end users who were using the chatbot to provide feedback regarding designs, features and function, the chatbot was tailored and meaningful to people regarding their needs for accurate - and timely- access to information. This is crucial to any support system's success (Hashiyana & Kamati) .

As for the survey study looking at the potential use of ChatGPT in higher education involving 133 students, it notes there was a high level of cognizance of ChatGPT amongst the participants. The majority of people in the study were acquainted with the AI-powered chatbot. Otherwise, the survey found a more complicated picture about the use of ChatGPT in higher education; that there is potential for AI to augment learning, there continues to be significant concerns, and traditional flows suggest these old ways of doing things were still significant. This suggests there is a trade-off between adopting strategies that are effective (Tick, 2024)

Among the studies related to the research topic is the study by Rosy Jan, which addressed one of the dimensions we are studying reliability in the context of academic research, applied to the ChatGPT engine, with a particular focus on its ability to handle the retrieved literature related to COVID-19. The significant findings were that there are potential uses of ChatGPT in academic research, but its current limitations in identifying the literature that has already been retrieved and employ predictive logic are currently barriers to the technology, which will require improvements in AI so that ChatGPT can be used with reliability and accuracy in a scientific context (Jan, 2024).

A different study evaluated the intelligent chatbot ChatGPT for 9 different questions related to breast cancer surgery as presented in the FAQ section of the American Society of Breast Surgeons. In the study, four breast cancer surgeon experts rated the accuracy and trustworthiness of the ChatGPT responses using a Likert scale and the Patient Education Materials Assessment Tool (PEMAT). ChatGPT was found to be reliable answering breast cancer surgery questions, with only minor harmless errors. The answers provided by ChatGPT (Chat Generative Pre-Trained Transformer) were accurate, clear, and understandable (Roldan-Vasquez et al., 2024)

Upon assessment of the previous studies, we find that they deal primarily with the reliability of intelligent chatbots, speed of answers, how well they comprehend the underlying meaning of users' questions, and how we must properly balance the traditional modes of education and with tools and resources that AI provides to us. They also stress user experience and user satisfaction. The value added to our research that we obtain from this evaluation is building a model to evaluate the performance of intelligent chatbots for use in future studies, and to derive benefit from the results and continue moving research forward.

Table 1. Key Insights from Literature Review

Aspect	Findings	Citations
Participation and Learning Outcomes (Usage)	- Chatbot programs improve academic performance and knowledge retention; - Increase student engagement and motivation.	(Bettayeb et al., 2024; Ilić et al., 2024)
Privacy and Security	Ensuring data protection and compliance with privacy regulations is essential.	(Segovia-García, 2024)
Integration of Chatbots into Learning Systems (Integration)	- Smoother and more effective learning experiences for students; - Providing instant support and personalized feedback; - Enhancing collaborative learning environments.	(Bobro, 2024; Saifullah et al., 2024)
Reliability	- Good reliability in responding to inquiries; - Minor harmless errors present; - Answers are accurate, clear, and easy to understand.	(Jan, 2025; Roldan-Vasquez et al., 2024)
Ethical Considerations and Technical Limitations	- Issues such as bias and accuracy need to be addressed; - Chatbot programs may struggle with complex inquiries; - May lack human-like empathy.	(Bettayeb et al., 2024; Ilić et al., 2024; Sandu et al., 2024)

Source: Prepared by the researchers based on previous studies

3. Methodology and research tool:

We adopted a model identified by five dimensions. We built the model and research instrument in the form of a questionnaire, using the following processes:

3.1. Systematic Literature Review: A search was completed in multiple academic databases to obtain recent relevant literature. It focused on qualitative studies that highlighted advantages and fields of application of intelligent chatbots in higher education to provide a dataset that is broad enough, but comprehensive enough, to meet the research purpose.

3.2. Model Construction: The study model includes five factors for examining how intelligent chatbots work in higher education in Algerian universities. The model is grounded in previous research relating to this field, where previous studies tend to only target one or two dimensions if anything, as summarised in the table below:

Table 2. The Five Dimensions of the URPIC Model

Dimension	Definition
Uses	Intelligent chatbots assist instructors in creating educational materials, designing tests and exams, developing course syllabi and lesson plans, constructing learning assessments, brainstorming ideas and outlining research papers, designing creative teaching techniques, encouraging autonomy and engagement, and providing a virtual learning environment — showcasing their versatility in enhancing educational experiences (Mbwambo & Kaaya, 2024).
Reliability	The presentation of information enables the recipient to understand and interpret it, which includes accuracy, timeliness, completeness, and formatting. Users must perceive that the chatbot is capable of accurately understanding their concerns and providing appropriate responses (Elizabeth et al., 2024, p. 3). Chatbots can offer reliable, accurate, and consistent information, though performance may vary significantly depending on the context and the specific AI model used.
Privacy	Self-disclosure is a primary privacy concern, as users often share personal and sensitive information during interactions with chatbots .This raises challenges in data protection and risks of privacy breaches. Chatbots may also perpetuate existing biases in their (Gumusel, 2025) responses, raising ethical concerns regarding fairness and equity in educational contexts (Williams, 2024). Users may be unaware of the extent to which their data is shared with third parties. This lack of transparency can lead to distrust and anxiety regarding the safety of their personal information. Users fear the potential misuse of their data by chatbots for malicious purposes without their consent (Gumusel, 2025)
Integration	The need for cautious integration: Many researchers advocate for the use and integration of intelligent chatbots into the educational system, but emphasize that this must be done cautiously to ensure it does not harm educational values or student learning (Mbwambo & Kaaya, 2024)
Constraints	Integrating intelligent chatbots into scientific research and higher education imposes several constraints that may affect research quality and integrity. These include contextual understanding, information verification, and ethical considerations. Chatbots lack deep contextual understanding, which may lead to misinterpretations of complex scientific concepts (Giray et al., 2024, p. 40), affecting content accuracy. They may also generate incorrect or misleading information, potentially undermining research credibility (Sedaghat, 2024, p. 1). Furthermore, the absence of reliable citations in some chatbots can pose challenges in verifying the accuracy of the information provided and maintaining academic integrity (Waduge et al., 2024).

Source: Prepared by the researchers based on previous studies.

Accordingly, the study model can be summarized in the following figure:

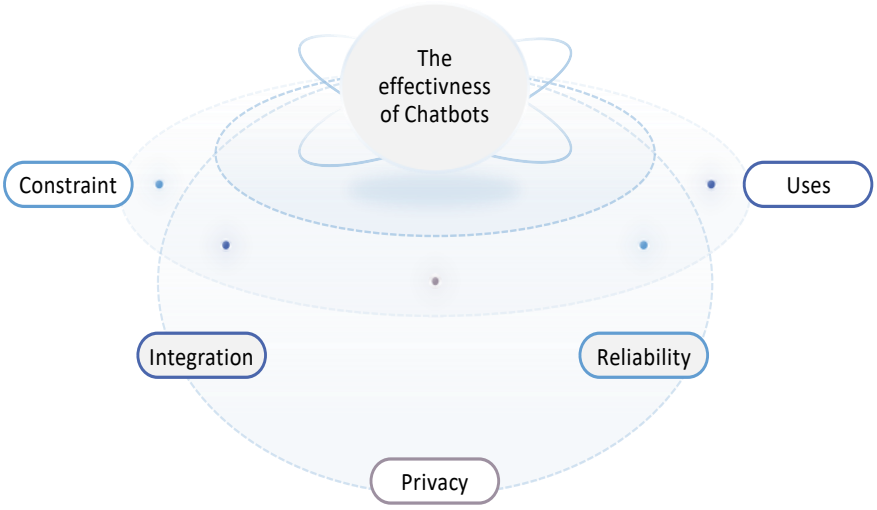

Fig. 1. *The central role of AI chatbots effectiveness in education and research processes based on the main dimensions of the URPIC Model*
Source: Prepared by the researchers based on the findings of previous studies.

4. The experimental framework for the study

4.1. Procedures and Methods

The descriptive approach was adopted to describe the study variables on the theoretical side by referring to sources and references. Field data were then collected using a questionnaire as the primary measurement tool, and the data were analyzed and hypotheses tested using SPSS v27.

4.2. Study Population & Sample

The study population was higher education professors in the Algerian universities. A random sample was used to increase the ease of collecting data and allow for results that can be generalized. An online questionnaire was sent to 300 professors in various Algerian universities, and 250 questionnaires were distributed in paper format. 246 questionnaires were usable and analyzed, a 54.66% response rate based on the number distributed.

4.3. Research Instrument

The questionnaire is divided into two sections: the first section covers the demographic variables of the sample, while the second section includes 25 items distributed across the five dimensions of the URPIC model. A five-point Likert scale was applied, with weights: "Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree" for each statement.

4.4. Statistical Procedures

The data were processed using the Statistical Package for the Social Sciences **(SPSS), version 27**. After coding and entering the data, the following statistical tests were applied:

- Cronbach's Alpha Coefficient was used to evaluate reliability of the research instrument;
- Kolmogorov-Smirnov Test was used to evaluate normality of distribution;
- Descriptive Statistics: means and standard deviations;
- One-Sample Wilcoxon Test for assessing single-sample responses;
- Mann-Whitney U Test for comparing two independent sample;
- Kruskal-Wallis H Test for comparing more than two independent samples, for evaluating hypotheses of differences;
- Levene's Test for homogeneity of variances;
- Post-hoc tests (Scheffé and LSD) to compute the source of differences.

4.5. Reliability and Validity of the Research Instrument

The research instrument was evaluated by a group of experts, and their comments were considered. The internal consistency of the scale was confirmed by Cronbach's Alpha Coefficient, and Self-validity Coefficient as described below:

Table 3. Reliability Coefficient – Cronbach's Alpha

Dimension	Number of Items	Sample Size (N)		Cronbach's Alpha	√Cronbach's Alpha
Uses	5	246		0.879	0.937
Reliability	5	246		0.805	0.897
Privacy	5	246		0.773	0.879
Integration	5	246		0.886	0.941
Constraints	5	246		0.843	0.918
Total	25	246		0.950	0.974

Source: Prepared by the researchers based on SPSS 27 output.

The variable "Effectiveness of Chatbots" (in its five dimensions: reliability, privacy, integration, use, and constraints), discussed in Table (03), had Cronbach's Alpha coefficients between [0.77–0.95] with all five coefficients greater than 0.70, which illustrates the instrument is reliable and suitable for use in the study; the coefficient of internal validity (self-validity) of the study was also high, with coefficients of [0.879–0.974]. Thus, the instrument may also be considered valid. Based on the results, the study may safely use the questionnaire for analyzing results and testing hypotheses.

4.6. Presentation and analysis of study data:

4.6.1 Demographic Characteristics of the Sample: These characteristics are presented using **percentages and frequencies**, as shown in the following tables:

Table 4. Description of Demographic Characteristics (Gender and age)

Variables	**Gender**		**Age**			
Categories	Male	Female	<30 years	from 30 to <40 years	From 40 to 50 years	>50 years
Total N (246)	110	136	23	93	73	57
(%)	44.7	55.3	9.3	37.8	29.7	23.2

Source: Prepared by the researchers based on SPSS 27 output.

From Table (04), we observe that the proportion of females is higher than that of males, accounting for 55.3%, which indicates that Algerian universities are keen on employing women in higher education, given their positive impact on the educational process. As for the age variable, the distribution of age groups is as follows: 30 to 40 years, 40 to 50 years, over 50 years, and finally under 30 years, with respective percentages of 37.8%, 29.7%, 23.2%, and 9.3%. This reflects that the young age group constitutes the majority of the teaching staff in Algerian universities.

Table 5. Description of Demographic Characteristics (Academic Rank, Academic Qualification, Specialization, Experience)

Variables	Academic Rank			Academic Qualification		Specialization		Experience			
Categories	Assistant professor	Lecturer professor	Full Professor	PhD	Master	Technical Sciences	Human sciences+ Economic Sciences	< 5 years	5 to <10 years	10 to 15 years	> 15 years
Total N (246)	76	123	47	189	57	117	129	86	58	49	53
(%)	30.9	50	19.1	76.8	23.2	47.6	52.4	35	23.6	19.9	21.5

Source: Prepared by the researchers based on SPSS 27 output.

Regarding the academic rank variable, the category of Associate Professors ranked first, representing 50% of the study sample—half of the total respondents. This indicates that the majority of the professors hold a PhD degree. This group is followed by Assistant Professors at 30.9%, and finally Full Professors at 19.1%.

As for the academic qualification, the PhD holders outnumbered the Master's degree holders, accounting for 76.8%. This is due to the transition of Algerian universities from the classical system to the LMD system (Licence-Master-Doctorate) in 2008, which has gradually reduced the number of Master's degree holders.

With regard to the experience variable, the category of less than 5 years ranked first with 35%, followed by the 5 to 10 years category at 23.6%, and then the more than 15 years category at 19.9%. This aligns with the high proportion of Associate Professors (A and B) in the sample.

As for the specialization variable, the field of Humanities, Social, and Economic Sciences surpassed Technical Sciences, accounting for 52.4% of the sample.

4.6.2 Normality Test for the Study Variables

Since the study sample exceeds 50 respondents, we rely on the results of the Kolmogorov-Smirnov test for the various variables. By comparing the calculated significance level (sig) in the table with the adopted significance level set at 0.05 (sig = 0.05), we obtain the results shown in the following table:

Table 6. Results of the Normality Test for Study Variables (Tests of Normality)

Variables	Kolmogorov-Smirnova			Data Distribution
	Statistic	df	Sig.	
Uses	0.100	246	0.000	Does not follow normal distribution
Reliability	0.118	246	0.000	Does not follow normal distribution
Privacy	0.148	246	0.000	Does not follow normal distribution
Integration	0.148	246	0.000	Does not follow normal distribution
Constraints	0.117	246	0.000	Does not follow normal distribution
Effectiveness of Chatbots	0.125	246	0.000	Does not follow normal distribution

Source: Prepared by the researchers based on SPSS v27 output.

From the results of Table 6, it is observed that the significance level is less than 0.05 for the variable "effectiveness of chatbots" and for each of its five dimensions (Uses, Reliability, Privacy, Integration, and Constraints), with a Sig(α) = 0.00 for every dimension. This indicates that the data do not follow a normal distribution; therefore, we rely on non-parametric tests for further analysis.

5. Results: Hypothesis Test Summary

5.1. Testing the Main Hypothesis and Its Sub-Hypotheses:

The five model dimensions: **Uses, Reliability, Privacy, Integration, and Constraints**.

To address the research problem and test the various hypotheses, the attitudes of the study participants were statistically measured, with a focus on **mean scores** and **standard deviations**. In the analysis, we relied on the relative level of respondents' opinions by classifying the arithmetic means according to the following scale:

- [1.00 – 1.80]: Strongly Disagree
- [2.61 – 3.40]: Moderately Agree (MA)

- [4.21 – 5.00]: Strongly Agree

Additionally, we used a range-based classification into three levels to interpret overall mean scores:

To verify the significance and validity of the hypotheses, we applied the One-Sample T-Test to assess the main hypothesis, and used the non-parametric Wilcoxon test to evaluate the five sub-hypotheses. The results are presented in Table 6.

Table 7. Results of the Test of Sample Members' Perception Levels Regarding the Study Variables

Statements	Mean	Std. Deviation	Rating	Level	Rank	Wilcoxon Sig
I am well aware of the positive contribution of chatbots in higher education and research.	3.48	0.95	Agree	Moderate		0.00
I used chatbots in my research and educational activities as soon as they became available.	3.43	0.97	Agree	Moderate		
I use chatbots regularly in my research and educational activities.	2.94	1.10	Moderately Agree	Moderate		
I find chatbots useful in supporting my educational and research activities.	2.85	1.16	Moderately Agree	Moderate		
My use of chatbots for research and education has recently increased.	3.28	1.21	Moderately Agree	Moderate		
Uses (average)	3.19	0.89	Moderately Agree	Moderate	3	
Chatbots are capable of understanding complex questions in my academic/research field.	3.07	1.18	Moderately Agree	Moderate		0.00
The information provided by chatbots is accurate and reliable.	3.25	0.99	Moderately Agree	Moderate		
Chatbots provide up-to-date information in my field of education and research.	3.01	0.93	Moderately Agree	Moderate		
I can interact very easily with chatbots.	3.30	1.00	Moderately Agree	Moderate		
Chatbots respond quickly to all my academic inquiries.	3.46	1.10	Agree	Moderate		
Reliability (average)	3.21	0.78	Moderately Agree	Moderate	1	
Chatbots provide safe advice and information.	3.34	1.03	Moderately Agree	Moderate		0.00
Human interaction should be part of the higher education experience alongside chatbot use.	2.82	0.82	Moderately Agree	Moderate		
Chatbots provide objective inquiries about educational and research contexts.	3.00	1.04	Moderately Agree	Moderate		
Using chatbots raises concerns about the misuse of my research data.	3.42	1.08	Agree	Moderate		
I use the information obtained from chatbots ethically.	3.36	1.01	Moderately Agree	Moderate		
Privacy (average)	3.20	0.97	Moderately Agree	Moderate	2	
Chatbots effectively complement traditional teaching and research methods.	3.48	1.00	Agree	Moderate		0.00
Chatbots enhance the role of the professor rather than replacing them.	3.14	1.18	Moderately Agree	Moderate		
Chatbots provide support and assistance in my educational and research activities.	2.97	1.35	Moderately Agree	Moderate		
Using chatbots increases my participation in educational and research activities.	3.18	1.13	Moderately Agree	Moderate		

Chatbot usage improves the quality of higher education and scientific research.	3.22	1.16	Moderately Agree	Moderate		
Integration (average)	3.20	0.97	Moderately Agree	Moderate	2	
I face technical difficulties when using chatbots.	3.22	1.12	Moderately Agree	Moderate		0.00
Language limitations hinder my effective use of chatbots.	3.22	1.19	Moderately Agree	Moderate		
I find chatbots limited in dealing with specific topics in my field.	3.25	1.19	Moderately Agree	Moderate		
There is monopolization of chatbot usage information in my professional environment.	3.11	1.11	Moderately Agree	Moderate		
The university impedes the dissemination of chatbot usage culture among students and professors.	3.11	1.07	Moderately Agree	Moderate		
Constraints (average)	3.18	0.89	Moderately Agree	Moderate	4	
Overall Chatbot Effectiveness	3.19	0.73	Moderately Agree	Moderate		0.00

Source: Prepared by the researchers based on the output of SPSS 27

Table 7 shows that the average level of responses from the sample individuals regarding the variable "Effectiveness of Chatbots" took a moderate degree of agreement, meaning there is a moderate effectiveness of intelligent chatbots in the educational and research process at Algerian universities from the perspective of the research sample. The arithmetic mean reached (3.197), a value that falls within the range [2.34 - 3.67], with a standard deviation of (0.737). This indicates a low variation among the responses of the sample individuals, due to the moderate level of awareness by Algerian university professors regarding the five dimensions: "Reliability, Privacy, Integration, Usage, and Constraints." The mean values fell within the range [2.34 - 3.67] respectively at (3.217, 3.20, 3.20, 3.19, 3.11) with standard deviations of (0.78, 0.97, 0.97, 0.89, 1.07). All of this indicates that Algerian university professors perceive the use of chatbots in educational and research processes with a moderate level of effectiveness. In addition, the significance level value for all dimensions of chatbot effectiveness from the sample's point of view appears at (Sig = 0.00), which is less than (0.05). Therefore, we accept the null hypothesis (H0) for the first main hypothesis, which states: "Chatbots are characterized by a high level of effectiveness in educational and research activities with statistical significance at (sig ≤ 0.05) from the perspective of the research sample, based on the model's dimensions represented by: usage, reliability, privacy, integration, and usage constraints."

We also accept the null hypothesis (H0) for the five sub-hypotheses related to the first main hypothesis, which state:

- There is a statistically significant moderate level at (sig ≤ 0.05) for the dimension of chatbots usage from the perspective of the research sample;
- There is a statistically significant moderate level at (sig ≤ 0.05) for the reliability dimension that characterizes chatbots from the perspective of the research sample;
- There is a statistically significant moderate level at (sig ≤ 0.05) for the privacy dimension provided by chatbots from the perspective of the research sample;
- There is a statistically significant moderate level at (sig ≤ 0.05) for the integration dimension with chatbots from the perspective of the research sample;
- There is a statistically significant moderate level at (sig ≤ 0.05) for the constraints dimension when using chatbots from the perspective of the research sample.

5.2. Testing the second main hypothesis:

There are statistically significant differences in responses regarding the effectiveness of chatbots at a significance level of (sig ≤ 0.05), attributed to demographic variables (gender, age, rank, academic qualification, experience, specialization) that characterize the research sample of Algerian university professors. To test the difference hypotheses, it is necessary to:

- Confirm the existence of differences in responses toward the variable of chatbot effectiveness. Here, non-parametric tests for two independent samples (Mann-Whitney test) are used for (gender, academic qualification, and specialization), and for more than two independent samples (Kruskal-Wallis H test) for (age, rank, experience).
- Identify the source of the differences:

- Use Levene's test to check the homogeneity of groups for demographic variables that show differences;
- Choose the most appropriate post-hoc test to determine the source of the differences.

5.3. Testing the hypothesis of differences and their sources between the means of two independent samples for the variables gender, academic qualification, and specialization

The following table expresses the outputs of the Mann-Whitney test, Levene's test, and the test identifying the source of differences:

Table 8. Hypothesis Testing for Differences and Their Source Between the Means of Two Independent Samples

Variables		Difference Test	Homogeneity Test	Source of Difference Test				
		Mann-Whitney Sig	Levene Sig.	Test Statistics	Categories	Mean	Sig	Decision
Effectiveness of Chatbots	Gender	0.012	0,013	ANOVA	male	3,054	0.006	Significant differences in favor of females
					female	3,312		
	Educational Qualification	0.251	-	-	-	-	-	No significant differences
	Specialization	0.003	0,008	ANOVA	Humanities & Economics	3.058	0.000	Significant differences in favor of Technical Sciences
					Technical Sciences	3.350		

Source: Prepared by the researchers based on SPSS v27 output.

It is evident from Table 8 that the significance level (Sig.) value for the variable "effectiveness of chatbots" using the Mann-Whitney test is less than 0.05, specifically (Sig = 0.012). This indicates the presence of statistically significant differences at the 0.05 significance level in the respondents' perceptions regarding the effectiveness of chatbots based on gender.

As for the homogeneity of variance, the significance level for Levene's test is also less than 0.05 (Sig = 0.013), which leads to the rejection of the null hypothesis that assumes equal variances. Thus, the variance between the two groups (male and female) is unequal. Accordingly, the source of the differences was tested using ANOVA, which revealed significant differences in the participants' responses toward the effectiveness of chatbots between the male and female groups, in favor of females.

Based on the above, the null hypothesis is rejected, and the alternative hypothesis of the first sub-hypothesis under the second main hypothesis is accepted, namely:

There are statistically significant differences in the responses regarding the effectiveness of chatbots at the significance level (sig $\propto \leq 0.05$) attributable to the gender variable, in favor of the female group.

As shown in Table 8, the significance level (Sig.) value of the "effectiveness of chatbots" variable using the Mann-Whitney test is (Sig = 0.012), which is less than 0.05. These results suggest there are statistically significant differences at the 0.05 level, which indicates that the perceptions of the respondents of the effectiveness of chatbots differ based on gender.

In terms of the assumption of homogeneity of variance, the significance level of Levene's test is also below the .05 level (Sig = 0.013), allowing for the rejection of the null hypothesis which stated equal variances. Thus, the variance is indeed unequal between both of the groups (male and female). Accordingly, the source of the differences was tested using ANOVA, which revealed significant differences in the participants' responses toward the effectiveness of chatbots between the male and female groups, in favor of females.

Based on the above, the null hypothesis is rejected, and the alternative hypothesis of the first sub-hypothesis under the second main hypothesis is accepted, namely:

There are statistically significant differences in the responses regarding the effectiveness of chatbots at the significance level (sig $\propto \leq 0.05$) attributable to the gender variable, in favor of the female group.

For the third variable, which is the field of specialization, Table 7 shows that the significance level (Sig.) value for the effectiveness of chatbots using the Mann-Whitney test is less than 0.05, specifically (Sig = 0.003).

This indicates that there are statistically significant differences at the 0.05 level in the participants' responses regarding the effectiveness of chatbots based on their field of specialization.

Furthermore, the significance level for Levene's test was also less than 0.05 (Sig = 0.008), which indicates that the null hypothesis of equal variances has been rejected. As such, there are unequal variances between the two groups and differences should be further interpreted using ANOVA.

The ANOVA results demonstrate that there are significant differences in the participants' perceptions of the effectiveness of chatbots between the two specialization groups: Humanities, Social and Economic Sciences, and Technical Sciences — with Technical Sciences groups having more favorable perceptions of effectiveness.

As a result, it can be concluded that the field of specialization is a highly significant predictor of how participants perceive chatbot effectiveness, with those in technical specializations having more favourable perception of effectiveness. Thus, we reject the null hypothesis and accept the alternative hypothesis for the second sub-hypothesis of the second main hypothesis, which states that:

There are statistically significant differences in the sample's responses regarding the effectiveness of chatbots at a significance level ($sig\propto \leq 0.05$) attributed to the specialization variable, in favor of the technical sciences group.

5.4. Testing the hypothesis of differences and their source between the means of more than two independent samples for the age variable:

The following table shows the outputs of the Kruskal-Wallis test, Levene's test for homogeneity, and the post hoc Scheffe test:

Table 9. Test of Differences and Their Source Between the Means of More Than Two Independent Samples for the Age Variable

Variables		Test of Differences	Test of Homogeneity	Test of Source of Differences Multiple Comparisons					
		Kruskal-Wallis H	Levene	Test Statistics	(I)	(J)	Mean Difference (I-J)	Sig.	Decision
Effectiveness of Chatbots	Age	0.000	0.460	Scheffe	Less than 30 years	30 to <40 years	0.110	0.919	Significant difference in favor of (under 30 years) and from 30 to <40 years and 40 to 50 years categories
						40 to 50 years	0.272	0.410	
						Over 50 years	0.887*	0.000	
					30 to <40 years	Less than 30 years	-0.110	0.919	
						40 to 50 years	0.161	0.495	
						Over 50 years	0.777*	0.000	
					40 to 50 years	Less than 30 years	-0.272	0.410	
						30 to <40 years	-0.161	0.495	
						Over 50 years	0.615*	0.000	
					Over 50 years	Less than 30	-0.887*	0.000	
						30 to <40 years	-0.777*	0.000	
						40 to 50 years	-0.615*	0.000	

Source: Prepared by the researchers based on SPSS 27 output.

Table 9 shows that the significance level (Sig) for the variable effectiveness of chatbots, based on the Kruskal-Wallis test, was less than 0.05, with a value of (Sig = 0.00). This indicates the presence of statistically significant differences at the 0.05 significance level in the participants' responses regarding the effectiveness of chatbots, depending on the age variable. As for the homogeneity of variances, the significance level for Levene's test was found to be greater than 0.05, with a value of (0.460), which leads to accepting the null hypothesis that assumes homogeneity of variances. This means that the variances across all groups are equal, allowing for the use of post hoc tests that assume homogeneity. Accordingly, the Scheffe test was selected.

The results revealed significant differences in the participants' responses regarding the effectiveness of chatbots between the following age groups:

- Less than 30 years and Over 50 years,
- 30 to less than 40 years and Over 50 years,
- 40 to 50 years and Over 50 years,

in favor of the groups:

- Less than 30 years,
- 30 to less than 40 years,
- and 40 to 50 years, respectively.

Based on the above, the null hypothesis is rejected, and the alternative hypothesis for the second sub-hypothesis of the second main hypothesis is accepted. It states that:There are statistically significant differences in the research sample's responses regarding the effectiveness of chatbots at a significance level (sig $\propto \leq 0.05$), attributed to the age variable, in favor of the age groups: less than 30 years, 30 to less than 40 years, and 40 to 50 years.

5.5. Testing the Hypothesis of Differences and Their Source Between the Means of More Than Two Independent Samples for the Rank Variable

The following table presents the outputs of the **Kruskal-Wallis** test, **Levene's** test for homogeneity, and **LSD** post hoc tests.

Table 10. Test of Differences and Their Source Between the Means of More Than Two Independent Samples for the Rank Variable

Variables		Test of Differences	Test of Homogeneity	Test of Source of Differences Multiple Comparisons					
		Kruskal-Wallis H	Levene Sig	Test Statistics	(I)	(J)	Mean Difference (I-J)	Sig.	Decision
Effectiveness of Chatbots	Rank	0.000	0.073	LSD	Assistant lecturer	Lecturer	0.38231*	0.000	Significant difference in favor of Assistant Lecturer
						Professor	0.64738*	0.000	
					Lecturer	Assistant Lecturer	-0.38231*	0.000	
						Professor	0.26508	0.029	
					Full Professor	Assistant Lecturer	-0.64738*	0.000	
						Lecturer	-0.26508	0.029	

Source: Prepared by the researchers based on SPSS 27 output.

Table 10 shows that the significance level (Sig) for the variable effectiveness of chatbots, based on the Kruskal-Wallis H test, was less than 0.05, with a value of (Sig = 0.00). This indicates the presence of statistically significant differences at the 0.05 significance level in the participants' responses regarding the effectiveness of chatbots depending on the academic rank variable. As for the homogeneity of variances, the significance level for Levene's test was greater than 0.05, with a value of (0.073), which leads to accepting the null hypothesis that assumes variance homogeneity. This means that variances across all groups are equal, allowing the use of post hoc tests based on homogeneity. Therefore, the LSD test was selected.

The results revealed statistically significant differences in participants' responses regarding the effectiveness of chatbots between the Assistant Lecturer group and both the Lecturer and Full Professor groups, as well as between the Lecturer and Full Professor groups—in favor of the Assistant Lecturer and Lecturer groups, respectively. Based on the above, the null hypothesis is rejected and the alternative hypothesis for the third sub-hypothesis of the second main hypothesis is accepted, which states: There are statistically significant

differences in the sample's responses regarding the effectiveness of chatbots at a significance level (sig $\propto \leq$ 0.05), attributed to the academic rank variable, in favor of the Assistant Lecturer and Lecturer groups.

5.6. Testing the Hypothesis of Differences and Their Source Between the Means of More Than Two Independent Samples for the Experience Variable: The following table presents the outputs of the **Kruskal-Wallis** test, **Levene's** test for homogeneity of variances, and **Scheffe** post hoc tests.

Table 11. Test of Differences and Their Source Between the Means of More Than Two Independent Samples for the Experience Variable

Variables		Test of Differences	Test of Homogeneity	Test of Source of Differences Multiple Comparisons					
		Kruskal-Wallis H	Levene	Test Statistics	(I)	(J)	Mean Difference (I-J)	Sig.	Decision
Effectiveness of Chatbots	Experience	0.000	0.103	Scheffe	Less than 5 years	5 to less than 10 years	0.075	0.938	Differences exist in favor of (Less than 5 years) and (5 to less than 10 years)
						10 to 15 years	0.419*	0.011	
						More than 15 years	0.645*	0.000	
					5 to less than 10 years	Less than 5 years	-0.075	0.938	
						10 to 15 years	0.344	0.091	
						More than 15 years	0.569*	0.000	
					10 to 15 years	Less than 5 years	-0.419*	0.011	
						5 to less than 10 years	-0.344	0.091	
						More than 15 years	0.225	0.443	
					More than 15 years	Less than 5 years	-0.645*	0.000	
						5 to less than 10 years	-0.569*	0.000	
						10 to 15 years	-0.225	0.443	

Source: Prepared by the researchers based on SPSS 27 output.

It is evident from Table (11) that the significance level for the variable **effectiveness of chatbots**, based on the **Kruskal-Wallis H** test, was less than 0.05, with a value of (Sig = 0.00). This indicates that there are statistically significant differences at the 0.05 significance level in the responses of the sample individuals regarding the level of effectiveness of chatbots according to the variable of experience.

As for homogeneity, the significance level (Sig) for **Levene's test** was greater than 0.05, with a value of (0.103), which leads to accepting the null hypothesis that the variances are homogeneous and thus equal across all groups. Consequently, we chose the post hoc tests based on homogeneity, using the **Scheffe** test. We found differences in the responses of the sample regarding chatbot effectiveness between the categories of (less than 5 years and both 10 to 15 years and more than 15 years) and between the categories of (5 to 10 years and more than 15 years), in favor of the groups (less than 5 years) and (5 to 10 years), respectively.

Based on the above, we reject the null hypothesis and accept the alternative hypothesis for the fifth sub-hypothesis of the second main hypothesis, which states that:

There are statistically significant differences in the sample's responses about the level of chatbot effectiveness at a significance level ($sig \propto \leq 0.05$) attributable to the experience variable, in favor of the groups (less than 5 years) and (5 to 10 years).

6. Discussion of Results

Contrary to most previous studies that focused on the extent of (ChatGPT) use in higher education (Ilić et al., 2024; Roldan-Vasquez et al., 2024; Tick, 2024; Waduge et al., 2024) , and assessed its accuracy and reliability, our study is based on applying a model to evaluate the effectiveness of chatbots in higher education and scientific research in general, without specializing in either domain. This evaluation is conducted from the perspective of university professors in Algeria, considering that university professors represent an elite group in any country.

All results showed a moderate level in evaluating the five dimensions of the (URPIC) model and in the overall evaluation of the model. Differences in the sample responses were attributed to variations in their demographic data, except for the variable of academic qualification, which showed no difference in responses. This indicates that the studied sample is characterized by a close level of competencies and scientific awareness based on the obtained degrees.

In our first hypothesis, we expected a high level of effectiveness for chatbots in research and educational activities from the perspective of the study sample. Our assumption was based on the widespread use of chatbots in various aspects of human life. Additionally, it was grounded on a considerable number of recent studies in this field, which confirm the increasing use of intelligent chatbots in education at various levels, including higher education.

For example, the study by Wangsa et al. (2024) conducted a systematic review of 20 scientific articles examining the potentials and drawbacks of artificial intelligence and the impact of chatbots in higher education. All findings were favorable, thus indicating some advantages of generative AI in higher education, and some ethical difficulties and problems with its irresponsible use of chatbots.

Likewise, the review undertaken by Khandakar et al. (2024) provided a literature review examining current research on AI in education, intelligent chatbots, and their integration with the context of higher education. This review also collected many articles on intelligent chatbots and examined the ethical implications and ensuing challenges of using chatbots, and likewise integrated them into a discussion of the literature reviewed which was also drawn from various sources and articles related to academic fields.

The ranking of the importance of the five dimensions of the model according to the study sample was as follows: the most important by far was the criterion of reliability of chatbots, with a mean score of (3.21). This was followed by the privacy criterion related to data protection, and the integration criterion concerning the compatibility of chatbots with the educational and research functions of university professors, both with equal mean scores of (3.20). Based on the importance of these three criteria, the level of chatbot usage ranked fourth with a mean of (3.19), and usage constraints ranked fifth with a mean of (3.18).

This indicates that the reliability of chatbots, their ability to protect data privacy, and their capacity to complement the role of university professors in higher education and scientific research are the most important factors in evaluating their effectiveness from the perspective of the study sample.

Regarding the level of effectiveness of intelligent chatbots from the perspective of the research sample, it was average, with differences in evaluation being greater among the age groups under 30 years and between 30 to 40 years. This is logical because the younger group of university professors is currently the most engaged with

artificial intelligence technology. Additionally, there were differences in the evaluation of effectiveness favoring the female category, which is natural since 55.3% of the sample members are female university professors.

As for the differences found in favor of the technical sciences category in judging the effectiveness of chatbots, this is because this group of professors is more knowledgeable about the precise scientific and technical details of chatbots due to their scientific specialization, even though the majority of the sample members were from the humanities professors (52.4%). Therefore, the professor specializing in technical sciences is more experienced in assessing the reliability of these chatbots, their ability to protect data privacy or not, and the extent of the need to integrate them into higher education.

In conclusion, we believe that chatbots powered by advanced artificial intelligence models have the potential to elevate the quality of higher education. However, their effectiveness will not be without challenges and limitations, including ethical concerns and technical constraints. Addressing these challenges through responsible use, ethical safeguards, and continuous improvement of chatbot capabilities will be crucial to maximizing their benefits. As technology continues to evolve, chatbots are likely to play an increasingly important role in shaping the future of higher education in Algeria and around the world.

Conclusions

This research provides a measurement tool for future experimental studies on the effectiveness of AI chatbots, focusing on the importance of the services they offer to university professors, the reliability of their use in the outcomes of their teaching and research work, as well as the extent to which their roles integrate with the professors' traditional roles, and their ability to protect the privacy of their data within the constraints that may hinder their use. However, the current dimensions of the (URPIC) model face a challenge in that it was used for the first time in an Algerian environment characterized by limited resources and somewhat weak technological infrastructure, unlike most of the previous studies reviewed, which were conducted in high-income, technologically advanced countries compared to Algeria. This reveals the scarcity of such research in developing countries. Therefore, more research should be conducted on this topic in developing countries to broaden the understanding of how artificial intelligence can enhance learning in resource-limited societies.

Additionally, the use of the model was for assessing the effectiveness of AI-chatbots in general, which highlights the uniqueness of using the model to evaluate the effectiveness of each individual intelligent chatbot separately. This opens the door for conducting such research in the future using the (URPIC) model.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

1. Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: conversational AI in education. Frontiers in Education ,
2. Bobro, N. (2024). Use of chatbots based on artificial intelligence to optimize the educational process in a higher education institution. Grail of Science(43). https://doi.org/https://doi.org/10.36074/grail-of-science.06.09.20 24.040
3. Bolambao, F. J., Pogoy, A. M., & Plaisent, M. (2025). AI Chatbots in Higher Education: A Meta-Synthesis. In *Emerging Technologies Transforming Higher Education: Instructional Design and Student Success* (pp. 53-74). IGI Global. https://doi.org/10.4 6-3904-3693-8-979/018
4. Elizabeth, J. G. S., Patricio, Á.-M., Kerly, P.-Z., Dennis, G., & Marco, F.-H. (2024). Unveiling the core constructs: A statistical approach to evaluating user experience with Chatbots in higher education (A case study from a university in Ecuador. Journal of infrastructure, policy and development, 01-24, 24(01). https://doi.org/10.24294/jipd.v8i10.6381
5. Giray, L., Jacob, J., & Gumalin, D. L. (2024). Strengths, Weaknesses, Opportunities, and Threats of Using ChatGPT in Scientific Research. International Journal of Technology in Education, 7(1), 40-58. https://doi.org/10.46328/ijte.618
6. Gumusel, E. (2025). A literature review of user privacy concerns in conversational chatbots: A social informatics approach: An Annual Review of Information Science and Technology (ARIST) paper. Journal of the Association for Information Science and Technology, 76(1), 121-154. https://doi.org/10.1002/asi.24898
7. Hashiyana, & Kamati. *An Artificial Intelligence-Powered Chatbot-Integrated Web-Based Application*. Advances in business strategy and competitive advantage book series. https://doi.org/10.402024. 18/979-8-3693-9586-8.ch008, 11 10
8. Ilić, J., Ivanović, M., & Milićević, A. K. (2024). *The Impact of ChatGPT on Student Learning Experience in Higher STEM Education: A Systematic Literature Review.* 21st International Conference on Information Technology Based Higher Education and Training (ITHET). Paris, France : IEEE, November

9. Jan, R. (2025). Examining the Reliability of ChatGPT: Identifying Retracted Scientific Literature and Ensuring Accurate Citations and References. In *Impacts of Generative AI on the Future of Research and Education* (pp. 367-392). IGI Global. https://doi.org/10.4018/979-8-3693-0884-4.ch014
10. Khandakar, H., Fazal, S. A., Afnan, K. F & ,.Hasan, K. K. (2024). Implications of artificial intelligence chatbot models in higher education. Int J Artif Intell ISSN, 2252(8938), 3809. https://doi.org/10.11591/ijai.v13.i4.pp3808-3813
11. Kristophorus, H., Felix, A., Aji, S., & Sandy, I. (2024). Chatbots Implementation for Students Admision. Journal of software engineering and simulation, 11(1), 44-52. https://doi.org/ 10.35629/3795-10114452
12. Mbwambo, N. M., & Kaaya, P. B. (2024). ChatGPT in education: Applications, concerns and recommendations. Journal of ICT Systems, 2(1), 107-124. https://doi.org/10.56279/jicts.v2i1.87
13. Ortiz-Garces, I., Govea, J., Andrade, R. O., & Villegas-Ch, W. (2024). Optimizing chatbot effectiveness through advanced syntactic analysis: A comprehensive study in natural language processing. Applied Sciences, 14(5), 1737. https://doi.org/10.3390/app14051737
14. Roldan-Vasquez, E., Mitri, S., Bhasin, S., Bharani, T., Capasso, K., Haslinger, M., Sharma, R., & James, T. A. (2024). Reliability of artificial intelligence chatbot responses to frequently asked questions in breast surgical oncology. Journal of Surgical Oncology, 130(2), 188-203. https://doi.org/10.1002/jso.27715
15. Saifullah, S., Yawan, H., Syafitri, N., & Nurhaliza, S. (2024). INTEGRATING AI CHATBOT INTO LEARNING MANAGEMENT SYSTEM: ENHANCING STUDENT ENGAGEMENT AND LEARNING OUTCOMES. Dharmas Education Journal (DE_Journal), 5(2), 1346-1359. https://doi.org/10.56667/dejournal.v5i2.1605
16. Sandu, R., Gide, E., & Elkhodr, M. (2024). The role and impact of ChatGPT in educational practices: insights from an Australian higher education case study. Discover Education, 3(1), 71 .
17. Sedaghat, S. (2024). Plagiarism and wrong content as potential challenges of using chatbots like ChatGPT in medical research. Journal of Academic Ethics, 1-4 .https://doi.org/10.1007/s10805-024-09533-8
18. Segovia-García, N. (2024). Optimizing Student Support. A Review of the Use of AI Chatbots in Higher Education. European Public & Social Innovation Review, 9, 1-20 .
19. Tick, A. (2024). Exploring ChatGPT's Potential and Concerns in Higher Education. 2024 IEEE 22nd Jubilee international symposium on intelligent systems and informatics (SISY), 10.1109/SISY62279.2024.10737595
20. Uzoka, A., Cadet, E., & Ojukwu, P. (2024). Leveraging AI-Powered chatbots to enhance customer service efficiency and future opportunities in automated support. Computer Science & IT Research Journal. P-ISSN, 2709-0043 . https://doi.org/10.51594/csitrj.v5i10.1676
21. van Rensburg, H., & Reedy, N. (2024). ChatGPT in Higher Education: A Comprehensive Study of the. Proceedings of The 23rd European Conference on e-Learning, https://doi.org/10.34190/ecel.23.1.3126
22. Waduge, A. O., Kulasooriya, W. K. V. J. B., Ranasinghe, R. S. S., Ekanayake, I., Rathnayake, U., & Meddage, D. P. P. (2024). Navigating the Ethical Landscape of ChatGPT Integration in Scientific Research: Review of Challenges and Recommendations. Journal of Computational and Cognitive Engineering, 3(4), 360-372. https://doi.org/10.47852/bonviewJCCE42023238
23. Wangsa, K., Sandu, R., Karim, S., Elkhodr, M., & Gide, E. (2024). A Systematic Review and Analysis on the Potentials and Challenges of GenAI Chatbots in Higher Education. 2024 21st International Conference on Information Technology Based Higher Education and Training (ITHET), https://doi.org/10.3390/fi16070219
24. Williams, R. T. (2024). The ethical implications of using generative chatbots in higher education. Frontiers in Education, https://doi.org/10.3389/feduc.2023.1331607
25. Zafar, S., Shaheen, F., & Rehan, J. (2024). Use of ChatGPT and generative AI in higher education: opportunities, obstacles and impact on student performance. IRASD Journal of Educational Research, 5(1), 01-12 . https://doi.org/10.52131/jer.2024.v5i1.2463