IMPLEMENTATION AND EVALUATION OF A CHATBOT IN A BUSINESS COURSE IN HIGHER EDUCATION

Pedro Isaias[1], Tania Hoque[1] and Paula Miranda[2]
[1]*Universidade Aberta – Portuguese Open University*
Lisbon, Portugal
[2]*Instituto Politécnico de Setúbal, ESTSetúbal*
Setúbal, Portugal

ABSTRACT

While the higher education sector is continuously searching for innovative technologies, the use of chatbots requires extensive research and careful consideration of their pedagogical value. The lessons learned from lecturers who experiment with chatbots can constitute important evidence to support their use. This paper presents a chatbot specifically designed for a Higher Education course, named RESOURCEbot, its mechanics and workflows. This chatbot was used in a university course to assist the students with recommendations of relevant research papers. It is also presented the chatbot´s evaluation. The evaluation of the RESOURCEbot derived from a questionnaire that was distributed among the students to assess their opinions about the experience of using a chatbot, its performance and their intention to use chatbots in educational settings. The results highlighted some of RESOURCEbot's limitations, such as some difficulty in understanding the students' prompts, but overall they reflected the students' positive opinions concerning its ease of use, and the value and pertinence of its recommendations.

KEYWORDS

Chatbots, Artificial Intelligence, Higher Education, Technology Evaluation

1. INTRODUCTION

Chatbots present unprecedented opportunities for education in terms of communication and information. They're accessible and offer scalability, while providing important possibilities for personalised learning (Wollny et al., 2021). Chatbots can be defined as "intelligent computer systems designed to mimic human conversation to enable automated guidance and support [using] natural language processing and machine learning" (Caldarini et al., 2022, p. 1), with applicability to various areas (Smutny & Schreiberova, 2020). Chatbots, can improve the engagement of students, assist the personalisation of the learning process, support lecturers and provide valuable information about students' behaviour (Kuhail et al., 2023). They have the potential to offer a more equative and interactive learning experience (Koyuturk et al., 2023) as well as to provide valuable feedback to the learners (Sáiz-Manzanares et al., 2023). At the same time, more studies are highlighting the need to examine the ethical and security implications of their use (Fulgencio, 2024).

A fundamental aspect of using technology for educational purposes is to assess if its impact was, in fact, pedagogically beneficial. This paper presents the evaluation results of RESOUCEbot, a chatbot, implemented, in a curricular unit, within an undergraduate course, to provide recommendations on which research papers, the students should use. It intends to examine the students' perception about its use in a twofold research question: what were the strengths and limitations of RESOURCEbot and what improvement can be introduced in the future to enhance its educational value? The paper begins by examining the use of chatbots in higher education and exploring students' perceptions of their use. It then describes the implementation of RESOURCEbot and outlines the research methods. The ensuing section presents the results and precedes the discussion and conclusion.

2. CHATBOTS FOR EDUCATIONAL PURPOSES

Since chatbots can provide support to multiple users at the same time, they are a more productive and more affordable alternative to human operators. The evolution of technology has facilitated their implementation improving their flexibility and simplifying their maintenance. In addition, these advancements in artificial intelligence have enhanced their capacity to emulate human dialogue (Caldarini et al., 2022).Chatbots can be categorised into informative, chat-based or conversational, and task-based (Adamopoulou & Moussiades, 2020). Chatbots are being used in education to improve student services, to improve engagement, and to reduce the financial cost of administrative tasks (Abbas et al., 2022). Wollny et al. (2021) study concluded that chatbots are being used in education mainly for the improvement of skills, to enhance the efficiency of education and to increase learners' motivation, and they have three pedagogical roles: learning, assistance and mentoring. Some examples include Coding Tutor, which assists students with writing software code (Hobert, 2019); Mondly, Andy, John Bot, and Buddy.ai, designed for language teaching (Belda-Medina & Kokošková, 2023); and Jill Watson, a chatbot that acts like a teaching assistant to address students questions regarding the course content (Kakar et al., 2024). In learning settings, chatbots can be teacher-oriented or service-oriented (Abbas et al., 2022).

There are several aspects at the origin of some scepticism concerning chatbots, including mistrusting its capacities, and insufficient socio-emotional intelligence (Bilquise et al., 2024). The challenge of the deployment of chatbots, in educational environments, relates to those associated with the use of AI in general, and that is to achieve a balance between harnessing its potential and fostering an implementation that is both responsible and ethical (Adıgüzel et al., 2023). When implemented in practice, for educational purposes, chatbots present some challenges, they lack sufficient resources to answer appropriately to the students' questions, there can be a mismatch between the type of chatbot and the task requirements, poor conversational design, low-quality content (Tlili et al., 2023).

Several studies propose methods for the evaluation of chatbots (Belda-Medina & Kokošková, 2023)(Koyuturk et al., 2023)(Abbas et al., 2022). There is a panoply of methodologies that can be used for evaluating the implementation of the chatbots, some studies portray an evaluation that is more focused on the students, while others concentrate on the actual performance of the chatbot, using a variety of strategies. Radziwill and Benton (2017) study presented a summary of previous research on chatbot evaluation, listing the quality attributes, which can be classified into efficiency (performance), effectiveness (functionality and humanity) and satisfaction (affect, ethics and behaviour, and accessibility). The evaluation criteria can also be mainly focused on the efficiency of the chatbots outputs and assess the completeness and accuracy of its responses (Govender, 2024). The assessment of the interactions of the students with the chatbot can use posts from discussion boards, students' grades, course access and conversation logs (Song et al., 2019).

Students encouraging perceptions and their willingness to use chatbots will have a positive impact in its deployment (Bilquise et al., 2024). The research on this subject has voiced different perspectives, taken from several studies that describe implementation experiments. Some studies have also appraised students views and concluded that their satisfaction levels were only moderate with students highlighting the need for improvement in areas such as adjusting to users' needs, including interactive multimedia, and enhancing speech functionalities (Belda-Medina & Kokošková, 2023). Some students remain reticent regarding the use of chatbots and find it difficult to trust them for essential support, to obtain information and for guidance (Bilquise et al., 2024).

In contrast, there is previous research that portrays mainly positive views. Sáiz-Manzanares et al. (2023) examined students' satisfaction with the use of chatbots for self-regulated learning and concluded that the students considered the chatbot to be a valuable tool for concentrating their questions on the specific concepts of the subject and for reflecting on their own learning process. In Hobert (2019) paper the students rated a chatbot that assist them with computer coding, as being a valuable addition to the course. The chatbot was characterised as being helpful and a useful support in understanding the content. On the other hand, they highlighted some limitations at the level of the excessive guidance, the capacity to understand their questions and the type of feedback that was offered. Similarly, Huang et al. (2019) conclusion of the evaluation of students perception demonstrates that they were satisfied with the chatbot in terms of minimising their feelings of isolation, increasing collaborative learning and improving conceptual learning. At the same time, some design limitations became evident, such as insufficient support when the students answered incorrectly and the quality of the content of the chatlogs.

3. RESOURCEBOT

RESOURCEbot (Resource Optimization and Selection for Education bot) is a pedagogical innovation that has been developed under the project HYBOT - Enhancing hybrid teaching in higher education through chatbots, a EU funded project of which Universidade Aberta (Portuguese Open University) is a consortium member, having also the following partners: Fachhochschule des Mittelstands (FHM) - University of applied Sciences, Germany; Université Côte d'Azur (UCA), France; Tallinn University, Estonia; Kaunas University of Technology (KTU), Lithuania; Trainings-Online Gesellschaft für E-Portale mbH, Germany (associated partner). RESOURCEbot has been specifically developed for the course Innovation and Knowledge Management, a semester course of the Master's Degree in Management offered by Universidade Aberta. RESOURCEbot´s goals are to act as a recommendation system to help students to select resources for their assignment under a specific course activity. Students interacted with the bot in English in order to enquire it regarding the available selected resources for that activity. Both natural language as well as choice menus (prompts) were available for students to interact with the system.

3.1 RESOURCEbot Mechanics

An immediate and accurate response can be generated including offering suggestions of resources. This virtual assistant can support lecturers and learners in numerous ways. Many researchers identified the tremendous benefits of integrating AI chatbots in educational settings (Labadze et al., 2023). The RESOURCEbot was developed to assist students in finding relevant resources, as is described in the system flow chart in figure 1.

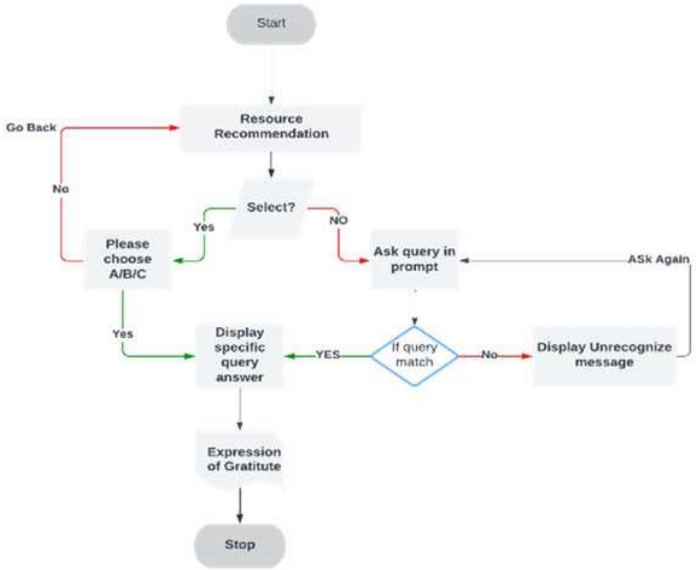

Figure 1. Flow chart of RESOURCEbot

The chatbot window was available on the Moodle webpage. The welcome message and recommendations for resources appeared on the display and then the recommendation feature interpreted the menus to inspire and guide students to interact with the chatbot. The menus can boost engagement for learners as they can minimise the potential confusion (Hew et al., 2023). The menus have several selection criteria, filters, and results. At this point, it is not necessary to ask questions in a prompt. However, if the users want to return to the start menu, they can use the back option. The user can ignore the offered menus and ask queries in the prompt directly. If the input query matches the intent, then the result will given to the user. Otherwise, the system will display an "unrecognised message" text, asking the user to try again. Once the desired result is obtained the chatbot displays a message of gratitude.

ISBN: 978-989-8704-61-0 © 2024

3.2 Chatbot Internal System Workflow

The RESOURCEbot´s internal system workflow had six phases (figure 2).

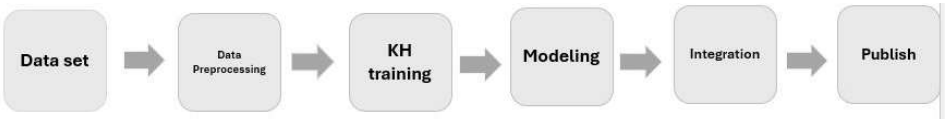

Figure 2. Chatbot´s internal system workflow diagram

Firstly, a data set was prepared to implement this chatbot. Dataset preparation is one of the crucial stages of building any system (Abdelmoiz et al., 2024). The resources were to be recommended according to the search query, therefore an Excel sheet was prepared with the dataset: 13 resources were analysed and categorised according to their abstract, keywords, title, publication date, resource type, author name, etc. The second step, data preprocessing, in this case, entailed the normalization of a dataset. Cleaning and removing corrupt or inaccurate records from the dataset is necessary to improve a system's quality and performance. Modifying and replacing any irrelevant information is part of data preprocessing. Finally, the data in the Excel sheet was converted to text data. With concern to Knowledge Hub (KH) training, the Melibo framework developed an AI-driven tool named Knowledge Hub to facilitate a wide range of information. The user interface of this platform is easy to use which can be beneficial in increasing engagement for non-technical people. It doesn't require high knowledge of technology. Three options are available to train this hub: text, web and file analyser. In the training phase, text analyser was used to increase knowledge and the preprocessed text data in the text analyser section was inputted.

The modelling step involved designing a chat flow using the Melibo framework and involved several phases. Firstly, a chat flow was selected in the create chat option, to design the model. Figure 3 describes the connecting nodes with the knowledge hub. Node one contains the possible questions asked by students. In this step, a large number of potential questions should be input to enrich the chatbot knowledge (Hew et al., 2023).

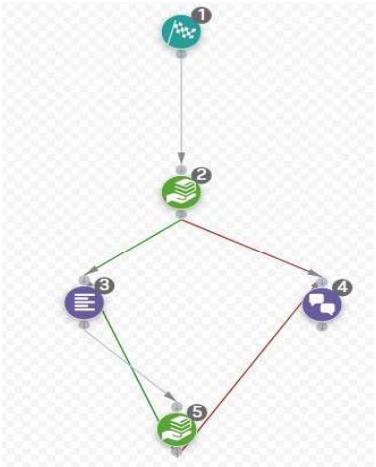

Figure 3. Chat flow Model

The training phase should continue until it gives the green signal. The red signal means the training is poor, orange indicates not sufficient. Node number two contains the knowledge hub symbol. In node number two the knowledge pieces and variables are selected. The configuration of this node is the search setting and general setting. The search settings function provides the custom search results of the input query. General settings have the function of GPT version selection, in this case the GPT-4 version was used.

The fifth phase, integration, entails merging with the Moodle webpage once the modelling is completed. It is possible to customise the chatbot according to the desired design, by choosing the name, colour, language, and other options. In the HTML tag, the API was merged in Moodle webpage. Finally, the last phase is where the users can see the chatbot logo on the webpage. After clicking the logo the chat window will open and users can see the console window. Initially, potential menus were given to the students to inspire them to conduct the chatbot (figure 4).

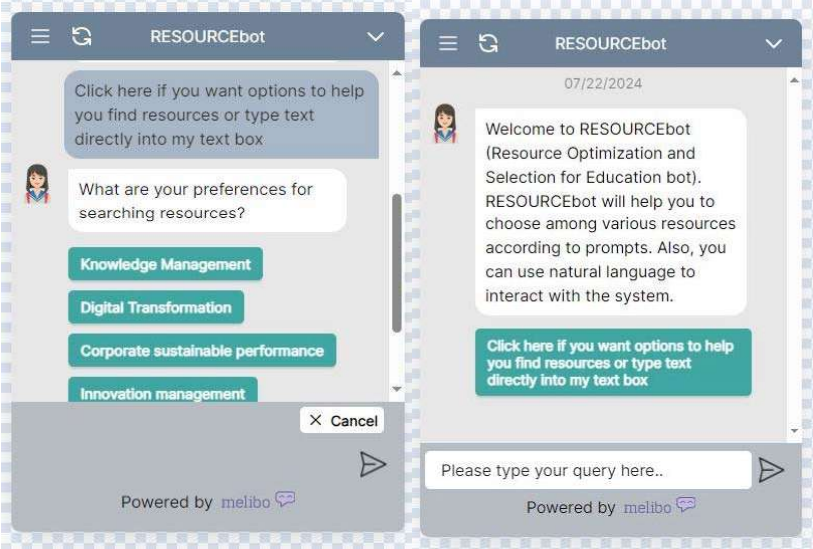

Figure 4. Outlook of RESOURCEbot

After publishing, every chat was monitored by using a Melibo tool named Bot Gym. It is possible to teach the bot if the output result gives wrong information (Hew et al., 2023). The improvement continued after the publication of the chatbot. It was important to understand how students were interacting with the system, and what kind of keywords they were using. The chatbot was trained in possible keywords, such as ID, resource number, document number, but not all keywords could be anticipated. Many students were searching using the word “article”, for example, “find me the article number 3 details”. In this case, since the bot was unfamiliar with the word, it displayed an “unrecognised message” in return. The RESOURCEbot was then trained with the word “article”. This process was continuous during the test case of the chatbot.

4. RESOURCEBOT EVALUATION: STUDENTS’ PERSPECTIVE

Following the application of the RESOURCEbot, a questionnaire was distributed among the students to assess their experience with the chatbot, and its performance. A total of 16 students completed the questionnaire. Most students (12) had already used chatbots, 6 of them only rarely and the other 6 on a regular basis. Only 2 students had no knowledge of what chatbots were, and 2 students knew what they were, but had never used them before. Hence, most were familiar with chatbots, prior to RESOURCEbot’s implementation. In addition, most students (10) used them in the personal arena, 6 for educational purposes, 4 students used them professionally and 1 student stated that he/she used them on e-commerce websites.

The remaining questions were focused on the experience with the RESOURCEbot specifically. Firstly, the students responded to a Likert scale question, which assessed several dimensions of the RESOURCEbot performance (figure 5).

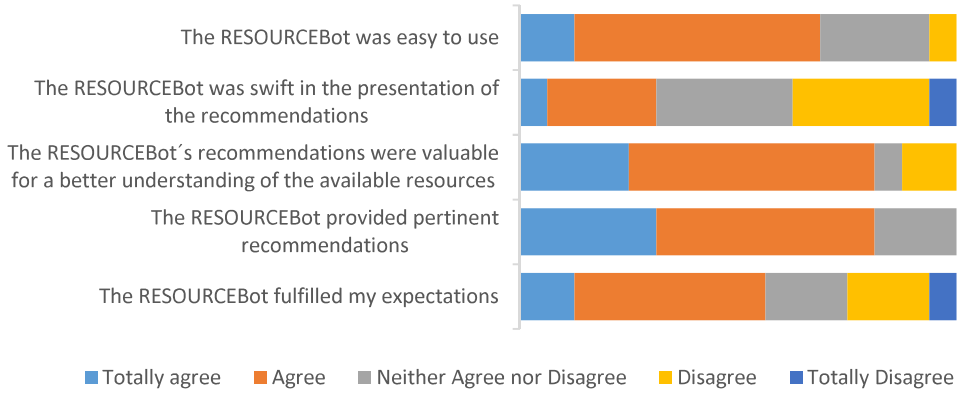

Figure 5. Students´ views about the RESOURCEbot performance

As can be seen in the chart above, most students found the RESOURCEbot easy to use (11), though 4 were neutral about its ease of use and 1 disagreed. With regard to swiftness, only 5 students considered the chatbot to be fast, while 5 were neutral in this question, 5 disagreed and 1 totally disagreed. The value of the recommendations for a better understanding of the available resources, on the other hand, were a subject that the majority agreed (13). The neutral position was selected by 1 respondent and 2 disagreed. Most students also found the recommendations to be pertinent (13) and only 3 students were neutral. Furthermore, 9 respondents stated that the RESOURCEbot fulfilled their expectations, but 3 students were neutral, 3 disagreed and 1 totally disagreed. The performance of RESOURCEbot was equally assessed with 2 open-ended questions regarding its strengths and limitations.

In the open-ended question about its positive aspects, the students reinforced its ease of use and the fact that it provided access to valuable information. They also mentioned its role in saving time, since it presented a swifter strategy to access the relevant papers, and its efficiency in providing pertinent recommendations: “direct recommendation of the most adequate papers” (student 9). In addition, some students highlighted its innovative nature, its flexibility and its role in supporting the task “innovation and a supporting tool” (student 12). The main trends in these open-ended responses were coherent with the opinions voiced in the Likert scale items. When asked about the limitations, the students focused on the difficulties that the chatbot had with understanding the requests, due to language restrictions. Some of the terms that were used by the students weren’t understood by the chatbot, as described by some students: “Sometimes it doesn’t understand what we write” (student 11); “ I had some language problems with the chatbot. For example, I asked "Do you have any resource on the topic X?" and the chatbot was not able to process the request…and I found out that it was enough to change the term “resource” (…) I had this problem three times, where I’ve used basic terms and the chatbot did not understand, I had to guess the term the chatbot knew, through synonyms” (student 4). Furthermore, some respondents highlighted the incompleteness of the responses: “it should give 8 resources for the determined criteria and it only gives 6 or 7” (student 10). Others cited its technological immaturity and mentioned that there was not enough time to explore the chatbot, due to the deadline of the activity.

After appraising the strengths and limitations of the RESOURCEbot, the students were asked to suggest additional resources. Some responded by numerating general improvements based on limitations, such as better interface or more accurate responses. Nonetheless, the majority of the students did suggest specific resources: communication in natural language, higher number of papers, available in other languages, such as Portuguese, access to previous questions, and proactive recommendations. As student 15 stated “the capacity to interact with quality in natural language” and student 13 added “I would like that it had more papers. It only has 13”. A few students also suggested the addition of other functionalities related to the papers it recommends, more specifically, a summary of the papers, analysis of tables and figures, ability to take notes and the possibility of translation.

The Likert scale question also included two items to assess the students future intention to use chatbots in the curricular unit where the RESOURCEbot was deployed and in the course in general (figure 6).

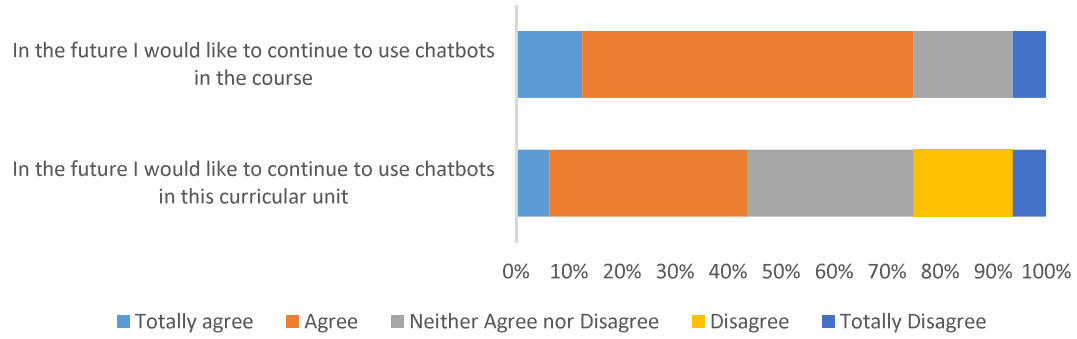

Figure 6. Students' intention to continue to use chatbots

The students showed a stronger intention to use chatbots in the course rather than in the curricular unit. The prospective use of chatbots in the course was supported by 12 of the students, 3 respondents were neutral and 1 totally disagreed. On the other hand, only 7 students showed their intention to continue to use chatbots in the curricular unit, while 5 were neutral, 3 disagreed and 1 totally disagreed. The students were asked if they were likely to recommend the RESOURCEbot to a peer. Only 3 students stated that they were extremely likely to recommend and 4 stated that they were very likely to recommend. Three students selected a more moderate position, stating that they were somewhat likely to recommend. On the opposite side of the spectrum, 5 students said that they were not very likely and 1 not likely at all.

The final question of the survey asked the students to leave any additional comments to the use of RESOURCEbot. Six students did not provide comments. The remaining students said that they liked the experience, that it was a new for them and that it was a good initiative. They thanked the lecturer and stated that it has room for improvement, but that in the future it can be a valuable tool for the students. As student 16 stated: "I understand RESOURCEbot in this context is still growing and being changed, but I strongly believe it can truly be a force for objectivity and assistance to students (…) I have explored, and I believe, I have taken from this bot a lot of information that was very helpful and insightful".

5. DISCUSSION AND CONCLUSION

Chatbots can be used in higher education for numerous purposes, such as to enhance communication and to access more relevant information. This paper focused on the use of RESOURCEbot, a chatbot that was developed to act as a recommendation system.

The analysis of the questionnaire shows that the use of RESOURCEbot was perceived by most students as being mainly positive. The students highlighted its ease of use, the value and relevance of its recommendations and they felt that their expectations were met. On the other hand, they identified some limitations, as in previous studies (Tlili et al., 2023), such as its difficulty in understanding some of the prompts. Overall, the majority of the students stated that they would like to continue to use chatbots in the course.

Despite this study's limitations, more specifically, its limited scope and small sample size, it constitutes an example of AI use in higher education and it can contribute to assist the design of an educational chatbot. The results revealed student's openness to the use of AI in educational settings and they illustrated the potential of this technology. This study can be used as a stepping stone for prospective research on learning and teaching with AI and it can encourage other lecturers to venture into the deployment of this technology in their courses. On the other hand, the difficulties reported by the students seem to support the scepticism that exists around chatbots' actual capacities, as argued in the literature (Bilquise et al., 2024). These difficulties can equally be examined as valuable insight that can pave the way for future improvements in chatbots development and implementation.

Given that the outcomes of the RESOURCEbot's deployment were mainly positive, future research can focus on improving the characteristics of the chatbot, to enhance its performance and improve its interaction with the students. Moreover, future studies can examine the use of chatbots for other purposes in higher education, with more in-depth research methods, and with other stakeholders, such as lecturers, developers and higher education institutions.

REFERENCES

Abbas, N., Whitfield, J., Atwell, E., Bowman, H., Pickard, T., & Walker, A. (2022). Online chat and chatbots to enhance mature student engagement in higher education. *International Journal of Lifelong Education*, *41*(3), 308-326.

Abdelmoiz, M. M., Mostafa, M. M., & Soliman, T. H. (2024). Developing an Educational Chatbot for Scientific Data Management Course Using DialogFlow. *Appl. Math*, *18*(3), 629-640.

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with applications*, *2*, 100006.

Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*.

Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education*, *20*(1), 62.

Bilquise, G., Ibrahim, S., & Salhieh, S. E. M. (2024). Investigating student acceptance of an academic advising chatbot in higher education institutions. *Education and Information Technologies*, *29*(5), 6357-6382.

Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information*, *13*(1), 41.

Fulgencio, S.-V. (2024). Developing Effective Educational Chatbots with GPT: Insights from a Pilot Study in a University Subject. *Trends in Higher Education*, *3*(1), 155-168.

Govender, R. G. (2024). My AI students: Evaluating the proficiency of three AI chatbots in completeness and accuracy. *Contemporary Educational Technology*, *16*(2), ep509.

Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). Using chatbots to support student goal setting and social presence in fully online activities: learner engagement and perceptions. *Journal of Computing in Higher Education*, *35*(1), 40-68.

Hobert, S. (2019). Say hello to 'coding tutor'! design and evaluation of a chatbot-based learning system supporting students to learn to program. In *Fortieth International Conference on Information Systems, Munich*.

Huang, W., Hew, K. F., & Gonda, D. E. (2019). Designing and evaluating three chatbot-enhanced activities for a flipped graduate course. *International Journal of Mechanical Engineering and Robotics Research*.

Kakar, S., Maiti, P., Taneja, K., Nandula, A., Nguyen, G., Zhao, A., Nandan, V., & Goel, A. (2024). Jill Watson: Scaling and Deploying an AI Conversational Agent in Online Classrooms. International Conference on Intelligent Tutoring Systems,

Koyuturk, C., Yavari, M., Theophilou, E., Bursic, S., Donabauer, G., Telari, A., Testa, A., Boiano, R., Gabbiadini, A., & Hernandez-Leo, D. (2023). Developing effective educational chatbots with chatgpt prompts: Insights from preliminary tests in a case study on social media literacy. *arXiv preprint arXiv:2306.10645*.

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1), 973-1018.

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, *20*(1), 56.

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*.

Sáiz-Manzanares, M. C., Marticorena-Sánchez, R., Martín-Antón, L. J., Díez, I. G., & Almeida, L. (2023). Perceived satisfaction of university students with the use of chatbots as a tool for self-regulated learning. *Heliyon*, *9*(1).

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151*, 103862.

Song, D., Rice, M., & Oh, E. Y. (2019). Participation in online courses and interaction with a virtual agent. *International Review of Research in Open and Distributed Learning*, *20*(1).

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart learning environments*, *10*(1), 15.

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in artificial intelligence*, *4*, 654924.