Education and Mental Health Through Artificial Intelligence Virtual Assistant

David-Theodor Ioniță, Ioana Corina Bogdan(✉), Horia Alexandru Modran, and Alexandru Dinu

Department of Electronics and Computers, Transilvania University of Brasov, Brasov, Romania
david.ionita@student.unitbv.ro, {corina.bogdan,horia.modran, alexandru.dinu}@unitbv.ro

Abstract. With the evolution of technology in daily life, there is a growing opportunity to leverage Artificial Intelligence (AI) to support students. Existing solutions demonstrate that AI can personalize learning and assist with health management, but there is still a gap in integrated, companion-based AI systems that can seamlessly operate in both domains. This paper introduces EduPal, an innovative virtual assistant designed to enhance educational experiences and provide psychological support. Leveraging advancements in AI, EduPal offers personalized learning experiences and mental health assistance, tailoring its interactions based on user-specific profiles and psychological assessments. The system employs the GPT-4 language model, facilitating natural language interactions through voice and text, and integrates advanced algorithms for real-time response generation. EduPal aims to assist in academic subjects, generate tailored exercises, and provide timely psychological interventions, thereby improving educational outcomes and emotional well-being. Future work will focus on enhancing system capabilities, ensuring data privacy, and broadening the scope of its applications. The potential of EduPal to transform personal education and health management underscores the significant impact of AI-driven virtual assistants in modern society.

Keywords: AI Companion · Education · Personalized Learning · Emotional Monitoring

1 Introduction

Nowadays, in the context of an engaging environment, with many domestic duties, job related tasks, and a desire for relaxing activities, people seek ways to fast access information when they need it. In this context, digital assistants gain more and more ground. The most employed digital assistants are represented by already mature platforms based on AI models like ChatGPT [1]. These models are trained on vast amounts of data to understand and generate human-like responses. The messages and questions sent by users to the digital assistants are sent to the servers where the model is stored. These servers have the necessary computational resources in order to parse the questions, extract and interpret the main ideas, keywords, and details, and provide an appropriate answer. The

M. E. Auer and T. Rüütmann (Eds.): ICL 2024, LNNS 1281, pp. 600–609, 2025.
https://doi.org/10.1007/978-3-031-83520-9_55

evolution of these models has progressed from text processing to image processing, and it is anticipated that, in future iterations, it will possess an enhanced capacity to understand and interpret video clips [2].

Recently, there has been a significant increase in the integration of virtual assistants into both robotic control systems and social robots. An example of this is RoboGPT, a robot that is based on ChatGPT. This is a robotic arm with seven degrees of freedom that provides assistance and guidance to human operators in tasks such as pick-and-place operations. At the same time, these operators can interact with and control the robotic arm using natural language [3].

Social robots have found applications in various fields, including the medical field, especially in the context of the COVID-19 pandemic [4, 5], and in education as personal tutors. Robots such as the Phillip K. Dick android [6], Sophia, Jiajia, and Ameca could potentially emulate a ChatGPT-based virtual assistant, allowing for more natural interaction—engaging in conversation and interaction with a humanoid robot rather than merely exchanging dialogue with a display.

In this context, EduPal is introduced, which is a unique contribution of this work. EduPal is a virtual assistant designed primarily to provide students with an effective tool for both educational work and psychological support. EduPal aims to make the learning process of a skill or subject more interactive by providing analogies and exercise examples tailored to the specific subject. Before utilizing EduPal, a virtual user profile can be created to include various information (e.g., pre-university or university level, etc.). EduPal is designed to develop a psychological profile of the user to adapt more effectively to the user's abilities, ensuring personalized and effective interaction.

The purpose of the project is to develop an artificially intelligent, powered companion assistant that can provide personalized support in education and healthcare. Moreover, the aim of this project is to design and implement a system that not only adapts to individual demands, learning styles, or educational needs, but which also is able to monitor psychological status and, finally, offer timely ethic, psychological support, and appropriate solutions to improve student or user capabilities. The smart system is giving the users the ability to have an interactive chat using their voice with the AI, making the interaction easier to give commands and receive a suitable answer. Furthermore, to minimize the likelihood of misinterpretation, the system will provide responses in text format as an alternative method and will also employ visual aids such as images and diagrams to articulate the response appropriately.

This work is structured into five sections: Initially, a succinct introduction is provided, presenting the relationship between digital assistants and social robots, along with their participation in diverse activities. The subsequent section encompasses a discussion about AI virtual assistants and the incorporation of EduPal within the realm of humanoid robots. The third and fourth sections delve into the methodologies, strategies, and algorithms employed to effectively implement AI virtual assistants. The final section presents the corresponding results, the conclusion of this study, together with further developments.

2 AI Virtual Assistants

2.1 Artificial Intelligence in Virtual Assistants

In the rapidly evolving landscape of technology, a variety of general-purpose assistive systems have emerged, each with its own unique capabilities and applications. These systems primarily aim to expedite the process of information retrieval for the user. A selection of these systems is illustrated in Fig. 1 [8], and their functionalities are detailed below:

- Amazon Alexa: This system is particularly beneficial for managing various devices within a smart home environment. It allows users to control different appliances, enhancing the convenience and efficiency of home automation.
- Siri: As a digital assistant, Siri is designed to facilitate seamless interaction with Apple devices. It recognizes and responds to vocal commands, thereby providing a hands-free and intuitive user experience.
- Bixby: This digital assistant is tailored for Samsung device owners and incorporates voice control services. It offers a personalized user experience and enables efficient device management through voice commands.
- Woebot Health: This system provides mental health support through a chat-based interface. It is designed for individual use as well as for implementation within corporate environments.
- Replika: This system offers psychological support for individuals. It employs advanced algorithms to mimic user behavior, fostering a more personalized and friendly interaction.

Fig. 1. Virtual Assistants [8].

These assistive systems, while diverse in their functionalities, share a common goal of improving user experience and efficiency through rapid information delivery and personalized interaction. Their continued development and integration into everyday life underscores the significant role of technology in enhancing human-computer interaction.

2.2 EduPal and Humanoid Robots

Numerous humanoid robots have been designed to emulate human behavior and engage with individuals in a congenial manner [9]. The primary objective of these interactions is to enhance students' academic achievements and foster their professional development. These humanoid robots are capable of engaging through both vocal and digital mediums, such as text, images, and pre-recorded sounds, with their human interlocutors [10, 11]. Platforms like EduPal can be integrated into these robots, eliminating the need for users to manually record every message, thought, or question they wish to convey or share. Responses can be received via an auditory signal and/or a text prompt. Humanoid robots such as Phillip K. Dick, Eve, Sophia, Ameca, Alter 3, Jiajia, and Nadine serve as tangible examples where EduPal can be implemented (Fig. 2). In various global experiments, humanoid robots [12, 13] are trained to replicate human behavior during interactions. Competitions are even held to determine if individuals can discern whether they have conversed with a robot or a real person.

Fig. 2. (a) Humanoid Robot [6, 13] (b) EduPal, interaction with Humans [8]

3 Methods and Approach

3.1 Available Methods

The preliminary iteration of the proposed software system effectively facilitated auditory interaction between a human operator and a computer serving as a digital assistant. This system demonstrated proficiency in accurately addressing a diverse range of general knowledge inquiries, thereby laying the groundwork for the supplementary enhancements delineated in this research.

The software application, as presented in this manuscript, was developed utilizing the Python programming language and incorporated several libraries, each serving a distinct purpose:

- The *tkinter* library was harnessed to construct the graphical user interface (GUI) of the application, as depicted in Fig. 3. This library provided the necessary tools to create an intuitive and user-friendly interface, enhancing the overall user experience.
- The *openai* library was utilized to gain access to the GPT-4 based platform, which served as the primary resource for retrieving answers to the posed questions. This advanced AI model played a crucial role in the system's ability to provide accurate and relevant responses.
- The pyaudio library was integrated to incorporate the speech recognition feature. This library enabled the system to interpret and process vocal inputs from the user, transforming spoken language into a format that the system could understand and respond to.
- The *threading* library was employed to parallelize the operation of the program. This allowed for the continuous capture of voice input at any given moment, ensuring that the system was always ready to receive and process user commands.
- The *setuptools* library was also leveraged for voice capture and for establishing communication with the input and output sound devices. This ensured seamless audio input and output operations, contributing to the system's overall performance and reliability.

Each of these libraries played an important role in the successful implementation of the software system, and their combined use resulted in a robust and efficient digital assistant capable of vocal communication.

The computational infrastructure employed for project execution is anchored by an AMD Ryzen 7 5800H processor, operating on the Windows 10 platform. It is equipped with 16 GB of DDR4 RAM and powered by a GeForce RTX 3070 8 GB graphics processing unit.

The application has specific timing requirements. The startup time is set at 2 s, followed by an audio capture duration of 3.36 s. The speech recognition process takes 0.97 s, and the time to return a response is 1.24 s. In total, the application requires 5.57 s to generate a response.

Additionally, there are data constraints to consider. The application can generate responses only based on pre-existing information. For instance, when asked about the capitals of European countries, it successfully provided the correct answers. However, when queried about real-time topics such as weather conditions, the stock market in America, or other current events, it cannot produce accurate responses. Instead, it directs the user to credible information sources to find answers to such questions.

3.2 Technical and Social Solutions

The platform conceived during this project aims to provide problem-solving support for undergraduate/graduate subjects such as mathematics, physics, and electrical engineering. Moreover, it is designed to help in programming by providing explanations or code examples [9–11]. It also must be able to generate tasks or tests and provide feedback concerning the student's results. The student's results could be adapted considering the level of the user, in order to provide a more eloquent picture of the situation (for a student in 12^{th} grade, several errors at simple additions could reveal some abnormal problems,

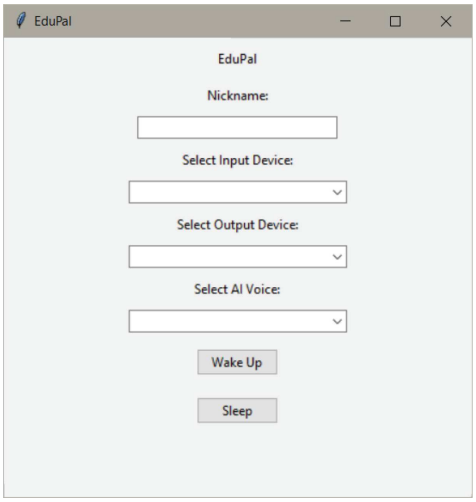

Fig. 3. Current GUI of the application

while for students in 1st or 2nd grade, this situation is more acceptable). Also, by using this platform, strategies to improve the assimilation and retention of information could be implemented, by providing suggestions to the user and explaining the best practices or solutions concerning the subject in focus.

As mentioned above, the expected advantage of EduPal is that the user can talk to it to feel a clearer and more human interaction. It aims to help teenagers or young people suffering from social anxiety or who are shyer to have a be helped to better communicate in day-to-day interaction. If psychological problems are detected, the platform might guide the person to seek specialized help.

3.3 Approach

The approach described employs GPT-4, an advanced AI language processing model, augmented with bespoke algorithms that incorporate natural language processing, machine learning, and knowledge representation. The system undergoes training on a variety of datasets, sourced from both user interactions and its own database. These datasets encompass educational content and healthcare acquisitions, with the objective of cultivating and assessing a comprehensive understanding of user queries, delivering pertinent and precise assistance, and disseminating scientific information.

4 AI System: Algorithm and Functioning

4.1 EduPal Functioning

The application workflow is described in Fig. 4. First, the user needs to open the application, providing requested information such as a nickname, selected available voices, output devices, and input devices, respectively. Upon the successful completion of this

level of queries, the application progresses towards the initialization phase, specifically to the Listening function. At this point, it can be distinguished into two cases: (a) The Listening function maintains an 'idle' state until the completion of the initialization process; (b) The execution of the Listening function concludes, followed by the delivery of voice commands to the AI system. Based on the voice message, the information could be retrieved or not, thus several types of messages could be sent:

(i) Once the information is analyzed, a response is provided. Then the program is returning to the initial step of initializing the Listening function,
(ii) If the message is misunderstood, an error message is sent, and the program returns to the initial step of initializing the Listening function,
(iii) In the event that the system cannot generate an answer, the program is returning to the initial step of initializing the Listening function.

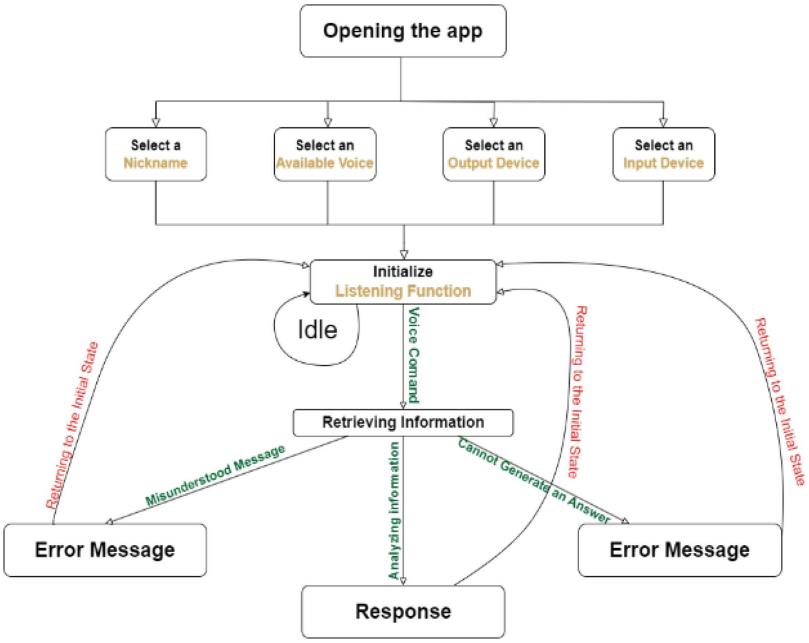

Fig. 4. Application workflow.

4.2 EduPal Algorithm

The operational process of the EduPal system is presented as is given below:

EduPal Algorithm

Input: Initialization Parametrs**,** Import libraries (*Tkinter*, *OpenAI, PyAudio, Setuptools, Threading, Speech_Recognition, Pyttsx3*)
Output: Responses Generated by the Assistant

1. ***EduPal algoritm*** initialization:

    Initialize the assistant's devices:
    Set nickname
    Configure output devices (e.g., speakers)
    Configure input device (e.g., microphone)
    (Set up device configurations, Establish communication protocols)

2. Main Loop:

    Operates continuously, awaiting trigger (awaiting activation)
    Responds upon recognizing its name
    Interprets the message and generates a response

3. Listening for ***Voice Commands***:

    Utilizes the microphone to capture voice commands (Records input)

4. Processing ***Voice Commands***:

    Verifies command understanding:
        If not understood:
            Generates an error message
            Returns to the initial listening state

        If understood:
            Proceeds with command processing

    (Checks command validity, Handles errors or proceeds with valid commands)

5. Analysing ***Information:***

    Uses the GPT-4 engine to generate responses based on the database

6. Providing ***Response***;

    Outputs the audio response through the speaker

    (Plays the generated response, Resets to listen for the next command)

Returns to the initial listening state.

5 Conclusion and Future Work

This study presents EduPal, an innovative AI-driven virtual assistant designed to enhance academic performance and mental health management. Although still in the early stages of development, it is anticipated that EduPal will help in providing tailored educational support and psychological assistance. By utilizing advanced natural language processing and machine learning algorithms, EduPal can interpret user inputs, recognize vocal tones, and respond with appropriate solutions. These interventions include offering immediate feedback, defusing potential issues, and providing encouragement and alternative explanations to improve user comprehension and outcomes.

EduPal has the potential to revolutionize individual learning and health management, especially in home environments or situations where traditional support systems are unavailable. By delivering customized educational content and emotional guidance, EduPal addresses the unique needs of each user, fostering a more effective and engaging learning experience.

Despite its promising capabilities, further research and development are essential to optimize EduPal's functionalities. Future work will focus on enhancing the system's adaptive learning algorithms, expanding its knowledge base, and improving its ability to handle a wider range of educational and psychological scenarios. Additionally, ensuring the privacy and security of user data remains a top priority. Rigorous testing and validation are necessary to establish EduPal's reliability and efficacy across diverse user populations.

Future work on the application includes a few key enhancements. First, an upgrade to the engine, specifically to GPT-4o, is planned to achieve more accurate and faster responses. Additionally, for educational and emotional support purposes, access to a comprehensive database filled with academic information will be developed. By utilizing machine learning and collaborating with educators and mentors, the goal is to transform AI into an effective study partner. This AI will be thoroughly trained to provide accurate and context-appropriate responses, making it a valuable tool for educational and emotional support. Furthermore, more refined settings for the nature of the responses will be introduced. The application must be capable of generating concise and clear audio cues that are easily understandable and accessible to students, avoiding overwhelming them with lengthy responses. Finally, a more modern and user-friendly graphical user interface (GUI) will be designed to enhance the overall user experience, making the application more accessible and attractive to a wide range of users.

In conclusion, EduPal represents a significant advancement in the integration of AI technology within education and mental health support. Continued development and refinement of this system could lead to substantial improvements in educational outcomes and emotional well-being, highlighting the transformative potential of AI-powered virtual assistants in modern society.

References

1. Pan, C., Banerjee, J. S., De Sarigiannidis, D., Chakraborty, A., Bhattacharyya, S.: ChatGPT: A OpenAI platform for society 5.0, In: Bhattacharyya, S., Banerjee, J.S., De, D., Mahmud, M. (eds.) Intelligent Human Centered Computing. Human 2023. Springer Tracts in Human-Centered Computing. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-3478-2_33
2. Wang, J., et al.: Chatvideo: a tracklet-centric multimodal and versatile video understanding system, arXiv preprint arXiv:2304.14407, https://doi.org/10.48550/arXiv.2304.14407 (2023)
3. Ye, Y., You, H., Du, J.: Improved trust in human-robot collaboration with ChatGPT (arxiv.org), https://doi.org/10.48550/arXiv.2304.12529 (2023)
4. He, H.: RobotGPT: from ChatGPT to robot intelligence, TechRxiv. April 10, 2023, https://doi.org/10.36227/techrxiv.22569247.v1 (2023)
5. Janamla, V., Daram, S.B., Rajesh, P., et al.: Response of ChatGPT for humanoid robots role in improving healthcare and patient outcomes. Ann. Biomed. Eng. **51**, 2359–2361 (2023). https://doi.org/10.1007/s10439-023-03267-1
6. Bogdan, I.C., Simion, E.: Games theory involvement in the social robots' facial expressions, UPB. Sci. Bull., Series C, vol. 85, Iss. 1, ISSN 2286–3540 (2023)
7. York, A.: 10 Best AI virtual assistants to improve productivity in 2024, https://clickup.com/blog/ai-tools-for-virtual-assistants/, Accessed 13 May 2024
8. Syntetic data generated with AI, https://chatgpt.com/, Accessed 28 May 2024
9. Budiharto, W., Cahyani, A.D., Rumondor, P.C., Suhartono, D.: EduRobot: intelligent humanoid robot with natural interaction for education and entertainment. Procedia Comput. Sci. **116**, 564–570 (2017). https://doi.org/10.1016/j.procs.2017.10.064
10. Mina, P.N.R., et al.: Leveraging education through artificial intelligence virtual assistance: a case study of visually impaired learners. Int. J. Educ. Innov. Res. **2**(1) (2023). https://doi.org/10.31949/ijeir.v2i1.3001
11. Mariani, M., Hashemi, N., Wirtz, J.: Artificial intelligence empowered conversational agents: a systematic literature review and research agenda, J. Bus. Res. **161**, 113838 (2023). ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2023.113838
12. Loideain, N., Adams, R.: From Alexa to Siri and the GDPR: the gendering of virtual personal assistants and the role of data protection impact assessments. Comput. Law Secur. Rev. **36**, 105366 (2020). ISSN 0267-3649, https://doi.org/10.1016/j.clsr.2019.105366
13. Habib, A., et al.: Lerning human-like facial expressions for the android. In: Dick, P.K. (ed.) IEEE International CASE 2014, Taipei, Taiwan (2014). https://doi.org/10.1109/CoASE.2014.6899473