

A Large Language Model Implementing an AI Assistant in a Higher Education Setting

Anna Mavroudi^{1,†}, Gerald Torgersen^{2,*}

¹ The University of Oslo (UiO), Boks 1072 Blindern, NO-0316 Oslo, Norway

Abstract

We are reporting on a piloting phase of creating, using, and evaluating an AI assistant for university students in the context of dental education. The assistant was meant to be used as a conversational partner with the students and it was instructed to implement the Socratic method; that is, avoiding giving answers to students, and help them obtain own understanding. We conducted a focus group discussion with a small number of students that had used the assistant. The results of our pilot pinpoint that the students used the AI assistant both as an information source as a conversation partner. Also, to new skills for both the teacher and the students to effectively use this type of technology.

Keywords

Large Language Model, chatGPT, artificial intelligence assistant, higher education

1. Introduction

Lately, Large Language Models (LLM) have attracted the attention of several researchers worldwide working not only in Artificial Intelligence (AI), but also in other domains, such as: business, security, and education. The main reason for the awakening of the interest of the research community concerns the latest advancements on chatbots (for example chatGPT from OpenAI) implementing LLM freely available online. In 2023, many prominent researches and experts in AI signed an open letter calling for a six-month pause on development of AI systems more capable than OpenAI's latest GPT-4 arguing that AI is advancing quickly and unpredictably. The letter garnered over 50,000 signatures [10]. ChatGPT is built upon the LLM technology called "Generative Pretrained Transformer" (GPT). The use of tools that resemble chatbots, such as chatGPT, along with the LLM that they incorporate has accelerated in most domains. Yet, LLM are in their infancy and we are in a very early stage of understanding and using them [1]. The focus of this paper is in higher education.

Proceedings of the 8th International Workshop on Cultures of Participation in the Digital Age (CoPDA 2024): Differentiating and Deepening the Concept of "End User" in the Digital Age, June 2024, Arenzano, Italy

* Corresponding author.

† These authors contributed equally.

✉ anna.mavroudi@iped.uio.no (A. Mavroudi); gerald.torgersen@odont.uio.no (G. Torgersen)

ORCID [0000-0002-3930-6336](https://orcid.org/0000-0002-3930-6336) (A. Mavroudi); [0000-0001-8335-7053](https://orcid.org/0000-0001-8335-7053) (G. Torgersen)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This paper discusses the case of an AI assistant for university students that incorporated chatGPT with the aim of helping students with their assignments in a dental education course. On the one side, we know from existing literature that chatGPT can generate human-like responses that are more natural compared with most educational technologies. This makes the tool appropriate for applications, such as AI assistants. On the other side, its limitations include the need for large amounts of data and computational resources to train and run the LLM [2].

In this paper, we are describing the whole lifecycle of the AI assistant starting from its conception, design, and development, to its use and empirical evaluation. With respect on the first three phases, we highlight particularly the different roles that interplay and their main tasks focusing on the division between professional developers and end users and their collaborative interactions. Finally, we conclude with a discussion of preliminary results on its use and evaluation. This involves the results of the thematic analysis of a focus group discussion with final users i.e. a small group of university students. The results indicate some interesting insights on additional learning demands that empower students as end users to assess LLM possibilities and limits. Our hypothesis is that the development of critical reflection around the use of this technology would be essential on behalf of the students. The main reason for our hypothesis stems from the problem of overreliance of end users in AI systems on the one hand [3] and on the other hand on the fact that LLM frequently hallucinate [4]. The use of chatGPT shows that its output in many cases cannot be regarded as secure knowledge, it is imprecise, or too general, which has consequences for the learning activities. Academically sound knowledge must then be secured in other ways [5]. Critical assessments must then be based on users' previous knowledge and their work with multiple information and knowledge resources [4].

2. Background

LLMs facilitate user interaction with generative transformers through conversational interfaces. These transformers, such as the GPT, are constructed using reinforcement learning and leverage large-scale online data, including human dialogues, to simulate natural language responses. The GPT's network architecture allows it to pretrain extensive language models using web-based textual resources like Wikipedia [7]. Pretrained models undergo initial training on general-domain datasets, followed by fine-tuning for specific tasks. While the GPT can generate contextually relevant text, it operates by discerning statistical patterns rather than truly comprehending word meanings. Consequently, it can produce coherent and contextually suitable language responses but lacks true semantic comprehension.

Prompt Engineering (PE) is the process of designing (or fine-tuning) effective questions or instructions ("prompts") for AI language models with the aim of producing the desired results [8]; herein, to help students develop their own understanding and ideas. In the context of higher education, PE is important because it pinpoints to the fact that the success of AI language models, like chatGPT, is not merely determined by their algorithms or data, but also on the skills of the who creates the prompts [8]. Therefore, this technology

introduces a new role for the university teacher (prompt engineer) along with the associated skills, a topic which is still under-researched.

The use of pedagogical conversational agents in digital learning environments is not new since it dates back to the 70s. There exists a body of literature that discusses their functions and their advantages. Regarding the former, their main teaching functions have been categorized in [12] as: 1) motivation, 2) information, 3) information processing, 4) storing and retrieving (information), 5) transfer of information and 6) monitoring and guiding students. Regarding the latter, main advantages include [11]: their 24/7 availability and their ability to respond naturally through dialogue-based systems.

3. Lifecycle of the AI assistant in the piloting phase

The lifecycle of the AI assistant comprises the following phases: 1) conceptualization and requirement analysis, 2) design and development, and 3) use and evaluation. We use two actors named "Alex" and "Dora" who are actual professionals that cooperated to set up the AI assistant.

1. **Conceptualization and requirements analysis:** Our intention was to support students in their learning process, using an approach in which the role of the tool is primarily to respond by asking questions to promote student learning mimicking as much as possible the Socratic method and avoiding as much as possible to provide answers to students' questions. We were also aiming in our approach to help students critically reflect on the use of chatGPT for learning purposes. To do that, we needed an interdisciplinary development team with a set of competences within AI, digital pedagogies, and the disciplinary territory of the subject matter which was dental hygiene.

2. **Design and development:** The development was based on an idea by "Alex," the project leader, who had two roles: a university teacher and a supporter of the dental hygiene course. Alex aimed to create an AI assistant for the dental hygiene course he was teaching. Recognizing the importance of pedagogical integrity, Alex enlisted "Dora," a specialist in digital pedagogy and AI in education, to support the project's educational effectiveness and assess its outcomes. As mentioned already, our intention was not to create an agent that would answer students' questions, but a conversation partner. This was possible by using the function "instruction" of the privacy-friendly chatGPT version 4 offered by the university that both Alex and Dora work along with the following system prompt:

"You are an intelligent teaching assistant, expert on medical physics and radiation protection in oral radiology. You are supposed to help students learning and are not supposed to give final answers. Give hints on the next step and inspire students to find the solution. If the student is stuck, provide a little more help to avoid frustration. You should not give complete answers, but help the students one step at a time to solve the problem. All answers and communication must take place in Norwegian, and you must use questions and answers from the given file as a source of knowledge when relevant. You must also be careful not to convey information that is not supported by the documents in the file. The aim is to help students think

independently and critically, by giving them the tools they need to understand and solve problems on their own."

3. **Use and evaluation:** Students got help from the AI assistant to write an essay for their course. We conducted the evaluation with a focus group interview using an interview protocol inspired by [6]. It comprises of the following stages:

- Introduction: purpose of the study, objectives, ethical considerations
- Main part: prior experience, quality of responses, functionality, pedagogical value, difficulties, affective domain (engagement, motivation), intentions of use in the future
- Closing: final questions, acknowledgment, considerations

The project team received ethical approval from the Norwegian Centre for Research Data and the participant students gave written consent. A focus group discussion took place using the interview protocol with five first year students that have used the AI assistant. The focus group discussion was recorded and transcribed using a combination of a dedicated AI tool offered by the university for the creation of the transcript and human intervention for making corrections whenever needed in the transcript. The transcript was qualitatively analyzed using thematic analysis comprising six phases [9]: 1) critically read the transcript, 2) generating initial codes, 3) search for themes, 4) review themes, 5) defining themes, 6) write up the results.

4. Preliminary results of the focus group discussion analysis

The main themes generated are:

- The difference in terms of student experience between the printed book, google search, and the AI assistant. This involves the role of other existing knowledge sources and what is unique for the assistant in comparison with the other sources.

*"Sometimes it can take you a while to find it on google. And the information you need isn't always on google. So chatGPT can help you with that. You just get a **general overview** of what you're looking for. So it's straight to the point."* (Student 3)

*"You want easy answers that you understand. Very simple and easy to explain. And you don't always get that in textbooks or the internet. So you look up chatGPT, and you get short and simple answers that you understand. And **then you understand what's written in the textbook**."* (Student 2)

*"No. And if you don't understand, it often says "it's not what you mean. Did you mean that?" And then it's easy to say, "no, I meant this". It's very rare that it doesn't understand. It's a **conversation** you have with the chatbot."* (Student 1)

*"It's like you have **a teacher who asks...**" (Student 4)*

*"It's like a different person. **An assistant who helps.**" (Student 5)*

- In terms of learning demands posed, the students say that it requires critical thinking from them. Issues of trust were central to the discussion and in particular a) the degree to which the tools that implements the LLM can be considered a reliable knowledge source and b) what constitutes plagiarism in such a learning environment

About point (a):

*"You have to be careful with learning from it, that you can't... How do I say this? **Source critical?** Yes, to be source critical. Because something has been wrong. Not everything is correct. So you have to be very critical of what you get there." (Student 5)*

*"But you have other sources that you refer to. I didn't **trust** chatGPT blindly, but I trusted it enough to use it, to help me to understand better. You read the original text, or watch it on YouTube, or on TV." (Student 4)*

About point (b):

*"I've heard that someone has **cheated** with chatGPT. Because they have copied and pasted. Yes, I agree. And that's not okay." (Student 1)*

*"You can use it to **steal information**. Because it's easily accessible... You can tell it to write a text about this and that. And then it can write it. So, you can just take this text. If you have an assignment." (Student 3)*

*"It has more advantages in the learning experience (than disadvantages). And in the student experience. Even though it's so strict with cheating. So not many people dare to **abuse** the platform." (Student 2)*

5. Discussion

We report on a pilot phase of an AI assistant in the context of higher education that was implemented using chatGPT4 along with an instruction to support and guide the students in their assignment while avoid providing answers. The paper concludes on changes regarding the role of end users in the era of LLMs. The main conclusions can be summarized as follows: firstly, this new technology introduces a new role for the university teacher, that of prompt engineer. On behalf of the students, the use of our AI assistant required critical reflection. The students used it as a source of information, but also as a conversation partner and this combination constitutes a different learning experience compared to other popular means, namely the printed book and information fetched via google search.

Future plans include fine-tuning a standard LLM or using Retrieval-Augmented Generation (RAG) with such a model. RAG is a method that improves the quality of generated text by incorporating relevant information from external sources. To tailor the AI model, a copyright-free dataset has been compiled, drawing from various reliable sources, including government documents, publicly available internet materials, and Alex's teaching content. Originally, we aimed in fine-tuning the AI assistant to the specific context and subject matter of the course, but due to time restrictions in the piloting phase, this was not possible.

Limitations include the fact that we cannot claim generalization of findings since this was a small piloting phase. The results are indicative of what other researchers might be interested in focusing on in the future, since the field is still in its infancy.

References

- [1] F. Khaddage and K. Flintoff, Say Goodbye to Structured Learning chatGPT in Education, is it a Threat or an Opportunity?, in Society for Information Technology & Teacher Education International Conference, March 2023, pp. 2108-2114. Association for the Advancement of Computing in Education (AACE).
- [2] Thurzo, M. Strunga, R. Urban, J. Surovková, and K. I. Afrashtehfar, Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, vol. 13, no. 2, p. 150, 2023.
- [3] F. Leiser, S. Eckhardt, M. Knaeble, A. Maedche, G. Schwabe, and A. Sunyaev, From chatgpt to factgpt: A participatory design study to mitigate the effects of large language model hallucinations on users. In *Proceedings of Mensch und Computer 2023*, 2023, pp. 81-90. <https://dl.acm.org/doi/abs/10.1145/3603555.3603565> G.
- [4] Bansal, V. Chamola, A. Hussain, M. Guizani, and D. Niyato, Transforming Conversations with AI—A Comprehensive Study of ChatGPT. *Cognitive Computation*, pp. 1-24, 2024. <https://link.springer.com/article/10.1007/s10639-023-11834-1>
- [5] S. Ludvigsen, A.I. Mørch, and R.B. Wagstaffe, Lett å bruke – Vanskelig å forstå. Studenters samtaler og bruk av generativ KI: chatGPT. Institutt for pedagogikk. UV, UiO. Rapport, 2023.
- [6] M. Rodriguez-Arrastia, A. Martinez-Ortigosa, C. Ruiz-Gonzalez, C. Roperio-Padilla, P. Roman, and N. Sanchez-Labraca, Experiences and perceptions of final-year nursing students of using a chatbot in a simulated emergency situation: A qualitative study. *Journal of Nursing Management*, vol. 30, no. 8, pp. 3874-3884, 2022.
- [7] D. Luitse and W. Denkena, The great transformer: Examining the role of large language models in the political economy of AI. *Big Data & Society*, vol. 8, no. 2, 2021. <https://doi.org/10.1177/205395172111047734>

- [8] Jacobsen, Lucas J., and Kira E. Weber. "The Promises and Pitfalls of chatgpt as a Feedback Provider in Higher Education: An Exploratory Study of Prompt Engineering and the Quality of Ai-driven Feedback." OSF Preprints, 29 Sept. 2023. Web. <https://doi.org/10.31219/osf.io/cr257>
- [9] V. Braun, and V. Clarke. Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological*, pp. 57–71, American Psychological Association, 2012. <https://doi.org/10.1037/13620-00>
- [10] I., Struckman and S. Kupiec. Why They're Worried: Examining Experts' Motivations for Signing the 'Pause Letter'. arXiv preprint arXiv:2306.00891, 2023. URL: <https://arxiv.org/abs/2306.00891>
- [11] F., Weber, T., Wambsganss, D. Rüttimann, and M., Söllner, 2021, December. Pedagogical Agents for Interactive Learning: A Taxonomy of Conversational Agents in Education. In *International Conference on Information Systems (ICIS)*.
- [12] S., Heidig, S. and G., Clarebout. Do pedagogical agents make a difference to student motivation and learning?. *Educational Research Review*, vol. 6, no. 1, pp.27-54, 2011.