"We Need To Talk About ChatGPT": The Future of AI and Higher Education

Michael Neumann
University of Applied Sciences and Arts Hannover
Hannover, Germany
michael.neumann@hs-hannover.de

Maria Rauschenberger, Eva-Maria Schön
University of Applied Sciences Emden/Leer
Emden, Germany
{maria.rauschenberger,eva-maria.schoen}@hs-emden-leer.de

Abstract—On November 30th, 2022, OpenAI released the large language model ChatGPT, an extension of GPT-3. The AI chatbot provides real-time communication in response to users' requests. The quality of ChatGPT's *natural speaking* answers marks a major shift in how we will use AI-generated information in our day-to-day lives. For a software engineering student, the use cases for ChatGPT are manifold: assessment preparation, translation, and creation of specified source code, to name a few. It can even handle more complex aspects of scientific writing, such as summarizing literature and paraphrasing text. Hence, this position paper addresses the need for discussion of potential approaches for integrating ChatGPT into higher education. Therefore, we focus on articles that address the effects of ChatGPT on higher education in the areas of software engineering and scientific writing. As ChatGPT was only recently released, there have been no peer-reviewed articles on the subject. Thus, we performed a structured grey literature review using Google Scholar to identify preprints of primary studies. In total, five out of 55 preprints are used for our analysis. Furthermore, we held informal discussions and talks with other lecturers and researchers and took into account the authors' test results from using ChatGPT. We present five challenges and three opportunities for the higher education context that emerge from the release of ChatGPT. The main contribution of this paper is a proposal for how to integrate ChatGPT into higher education in four main areas.

Index Terms—ChatGPT, GPT-3, large language model, higher education, AI influences, position paper

I. Introduction

ChatGPT[1] has attracted a lot of attention in recent months due to the quality of its language model. Worldwide media is reporting on what the tool can do, what possibilities it offers, and what disruptive changes may occur in a wide variety of areas. Since its launch, the tool has also been the subject of intense discussion on social media sites such as Twitter, Discord, and LinkedIn [2]. These discussions cover potential application in a wide range of areas, from software development to daily life. By now, ChatGPT has fundamentally changed the perception of already existing possibilities and potentials of AI tools in many communities, including software engineering (SE).

Due to the strong interest in ChatGPT, it is only a matter of time before students will widely use this tool. One may assume the tool will soon be considered state-of-the-art, which will lead to several consequences for higher education. Hence,

we execute tests with ChatGPT[2]. Although we did not specify detailed test cases, we created a structured test design that consists of seven categories: definitions; explanation of methods, tools, and models; solve specific problems (*e.g.,* maths and programming); solution options and their evaluation of specific tasks; creation of (scientific) texts based on specific topics; creation of (scientific) texts with references of scientific sources based on specific topics assuming you (ChatGPT) are a student; derivation of research questions for a given context and/or phenomenon. Due to space limitations, we use the term *paper* to summarize artifacts of scientific writing in the context of higher education, such as theses, term papers, take-home exams, protocols of experiments, and synopses (exposés).

Our tests show how high-quality, natural, and context-dependent the answers are, despite there being limitations (*e.g.,* nonexistent references or code security issues). Nevertheless, the results also raised some concerns regarding appropriate usage. These concerns have been reinforced by the fact that the plagiarism detectors licensed by our universities and/or departments could not identify the texts generated by ChatGPT. Furthermore, the results of the plagiarism checks have not given us any reason to perform a manual validation check of the text generated by ChatGPT. Although we knew it was possible to validate the ChatGPT-generated texts with AI output detectors, we did not include this in our testing.

The above leads us to the objective of this paper: Based on the findings extracted from grey literature and informal talks with other lecturers and researchers, we present five challenges and three opportunities for higher education that have emerged due to ChatGPT. We then offer recommendations for integrating ChatGPT into higher education, which we cluster in four main areas: *Area 1 - Teaching*, *Area 2 - Papers*, *Area 3 - Curricula* and *Area 4 - Regulations*.

II. Results from the Grey Literature Review

The intense interest in ChatGPT since its release and our own experience using the tool led us to the following question: Are there lessons to be learned from the research community? Due to the short time period between the release of ChatGPT (at the end of November 2022) and the creation of this paper, we could not find any relevant peer-reviewed literature;

---

[1]The tool created and released by OpenAI is a large language model and a fine-tuned extension of GPT-3 [1].

[2]Examples of our test prompts can be found in [3].

however, we did find an increase in preprints on Google Scholar. Although a systematic literature review (SLR) is usually conducted on peer-reviewed literature, the well-known SLR guidelines [4], [5] can be easily adapted to suit grey literature. The analysis of ChatGPT in education extracted from grey literature is valuable, as a discussion of how ChatGPT should be integrated into higher education is urgently needed. Research has already shown possible ways to misuse ChatGPT (*e.g.,* scientific abstracts, plagiarism detector [6]), and humans can no longer tell the difference. [7]. Since this is an urgent issue, the analysis of small data sets is valuable until larger data sets are available [8].

We performed at least two search runs per week starting in mid-December and ending the first week of January 2023 (2023/01/03, 79 total results) using the following search string [3]: *<<chatgpt AND higher education AND software engineering>>*. Out of the 79, 55 are preprints, and the rest are editorials, reports, or white papers. We identified five preprints related to ChatGPT and higher education in software engineering (SE). Over the weeks, we observed an increasing number of results in Google Scholar (with the latest check on 2023/01/17 revealing 150 results). Since we know preprints have limitations, we checked the quality of the papers, *e.g.,* using a methodology approach. However, we could not identify any serious limitations in the included preprints; thus, we did not exclude any other preprints.

A brief overview of the results from the included preprints is given below:

Haque et al. [2] conducted a mixed-methods study using Twitter data to analyze the thoughts, opinions, and feelings of early adopters of ChatGPT. Their results show that the two largest groups of ChatGPT users in the first weeks were students and academics/researchers. Besides intensively positive sentiments related to disruptive changes in software development and creativity, they also present findings related to educational contexts. For instance, users discussed the impact on literature reviews and grading papers as well as arising concerns related to plagiarism in writing assessments.

In their study, Gao et al. [6] evaluated the quality of 50 research abstracts generated with ChatGPT based on a comparison with the original ones. For their comparison, the authors used both tools (plagiarism checker and AI output detector) as well as human reviewers. Their results show that AI output detection tools (66%) and human reviewers (68%) can often identify the texts generated by ChatGPT. Nevertheless, the findings also show that neither type of verification (tool- or human-based) can reliably and consistently identify generated texts.

Susnjak et al. [9] deal with the impact of ChatGPT on online exams. The authors found that the high quality of texts generated by ChatGPT creates a higher risk that students will cheat on online exams. They anticipate a shift toward more oral exams and point out the need to integrate AI tools like ChatGPT into higher education to teach the skills needed to use it successfully.

In his experience report, Zahi [10] piloted ChatGPT by writing an academic paper on an example topic. He notes that the quality of ChatGPT is excellent and has an impact on the writing process. This raises the obvious question of how the tool can be integrated into university teaching. The author makes several suggestions for doing so. For example, he proposes that lecturers should question and adapt their forms of examination, or, in case of doubt, the publication of ChatGPT must teach further competencies in scientific work.

Similarly, Qadir [11] conducted tests and experiments with ChatGPT to discover potential benefits and challenges for higher education. He identified and formulated proposals regarding plagiarism detection, the importance of differentiating between acceptable and unacceptable use of the tool, and possibilities for integrating ChatGPT into teaching.

In summary, although we have only analyzed preprints, there are challenges and opportunities to be addressed in a discussion.

III. EMERGING CHALLENGES AND OPPORTUNITIES

We identified emerging challenges and opportunities based on our findings from the grey literature (see II). Table I presents an overview of the challenges and opportunities. In the first column, we added an identifier C- for a challenge, O- for an opportunity, and an index for short reference.

In summary, we identified five emerging challenges and three opportunities. It is worth mentioning that the challenges *unknown potential* and the opportunities *innovation potential* are closely related to each other. The current unknown potential of the tool leads to uncertainty about what aspects to consider when integrating ChatGPT into higher education. Nevertheless, we believe AI tools like ChatGPT have great innovation potential for our primary tasks. For example, they can provide new ideas for preparing a lecture or written assignment. However, we need further research and a deeper understanding of such tools. Therefore, in the following section, we discuss four areas in higher education to which the challenges and opportunities apply.

IV. FAST FORWARD INTO THE FUTURE

For decades, we have known that the integration of tools in educational contexts is important [12], as students´ knowledge and skills change over time. In addition, the prior knowledge of students varies immensely.

In the near future, people who use AI tools will work much more efficiently than those who do not use these tools because ChatGPT is assistive technology for them. In order to create equal opportunities in education, the use of AI tools must be anchored in the curricula of institutes of higher education. Ignoring AI tools and their advances would not lead to a better understanding among students and would hinder improvements to higher education. Although some propose blocking AI tools in schools and universities [13] and prohibiting their use, these measures will not turn back time. We have to face the current reality that the quality of AI tools is rapidly increasing and that this is a game changer for our day-to-day lives. Therefore, we must start the discussion now

TABLE I
OVERVIEW OF THE IDENTIFIED CHALLENGES AND OPPORTUNITIES

Challenge/Opportunity	Description
Unknown handling by students - C1	If the rules for using ChatGPT are unknown or do not exist, students may not know how and when they are allowed to use them, which may lead to asynchronous responses to assignments (*i.e.,* identify, discuss, or solve source code errors).
Heterogeneous evaluation - C2	Variations in how lecturers of different courses handle this issue may confuse students about "what is allowed" and "what is prohibited" regarding the use of ChatGPT (*e.g.,* in papers).
Acceptable/Unacceptable use of ChatGPT - C3	This challenge has several aspects. First, what should be acceptable depends on various factors and should be thoroughly discussed in the research communities to be able to adapt existing guidelines or create new ones. Furthermore, we see an increased risk of cheating, especially on papers or online exams (especially since plagiarism detectors do not cover ChatGPT-generated text). Thus, AI detection tools must be evaluated and prepared for increased use.
More time-consuming assessments - C4	Integrating ChatGPT into teaching and assessment will immediately increase the workload for the lecturer since they will need to adapt the exams, exercises, and other material (*e.g.,* by adding oral exams to courses evaluated by papers).
Unknown potential - C5	The unknown potential of AI tools such as ChatGPT is a challenge, as we can only discuss aspects, challenges, risks, as well as advantages and new ideas related to the current version of the tool. In recent weeks, we have observed impressive numbers of new ideas and upcoming challenges. Another aspect to consider is the highly dynamic market; thus, we assume that other more powerful AI tools will be released in the near future.
Increased virtual tutoring system - O1	The quality of virtual tutoring will increase when ChatGPT is used in addition to the lecturer's approach. The students can use ChatGPT as their personal tutor and receive support based on their prior knowledge, skills, and experiences (*e.g.,* they can have it explain a definition, translate explanations into simple text, or verify an artifact such as source code). This may lead to increased fundamental knowledge of a specific topic, since individuals can learn in personalized ways.
Engaged creativity - O2	This opportunity relates to the creativity for both lecturers and students. For students, the tool can increase the creativity to find new ideas or approaches to solving problems. The creativity of lecturers can be improved to adapt the didactic approaches in use, identify new didactic methods, or create new exercises or lecture materials.
Innovation potential - O3	The potential of upcoming ideas like new or adapted didactic approaches is high. We see a manifold adaption potential for existing didactic approaches such as flipped classroom or problem-based learning. We also expect new didactic methods and exam types alongside increasing digitization.

before the use of these tools gets out of control as well as teach our students the necessary skills for integrating ChatGPT into their studies, work, and personal lives.

Based on intensive discussions about emerging challenges and opportunities, we have categorized selected recommendations along four main areas: *teaching*, *papers*, *curricula*, and *regulations* of higher education. An overview of the potential impacts of ChatGPT on higher education is shown in Figure IV. Below we discuss the recommendations for each area.

Area 1 - Teaching: First, we consider the important fundamentals that students learn in the first few terms (*e.g.,* programming fundamentals or requirements engineering). At the start of a course, teachers must specify where (*in which areas*) and how (*the didactic approach*) ChatGPT may be used their course (addressing Table I, C1). This should achieve transparency about the awareness of ChatGPT (*e.g.,* wrt. functions) as well as its features and limitations. There are various approaches for how to do this. For example, lecturers already provide handouts (*e.g.,* rules for the use of tools) in courses with exam papers. These can easily be adapted to include ChatGPT-related aspects. Also, coordination among the lecturers can prevent redundant or even contradictory guidelines (addressing Table I, C2). Another possibility is to integrate ChatGPT into teaching (addressing Table I, C5). We recommend that students practice using the tool to explore the possibilities and limitations for specific use cases. From a student's perspective, there is a wide variety of possible use cases, *e.g.,* creating unit tests, analyzing source code, or explaining something for different skill levels (see Table I, O1). The opportunities to integrate ChatGPT into didactic state-of-the-art approaches such as problem-based learning or

flipped classroom are manifold (addressing Table I, O1-3). Furthermore, we highly recommend inviting practitioners to the lecture when integrating ChatGPT, especially for courses related to programming or software processes. Practitioners are also currently challenged by such tools and can provide interesting insights from a practical perspective on how to integrate these tools into their day-to-day work. The opportunities we include emphasize the relevance for practice and increased transparency surrounding usage.

Area 2 - Papers: A major challenge to integrating ChatGPT into higher education may be the influence on scientific writing (see Table I, C1-4). This is particularly relevant for the parts of a paper that replicate existing knowledge or theories. In our view, this includes the introduction, theoretical background, related work and literature reviews, and the description of the research design. We identified several options for detecting plagiarized text created by ChatGPT. First, our results show that AI detectors can reliably identify ChatGPT-generated text in many cases [6]. According to the idea of *acceptance by design* [14], tools need to address users' or stakeholders' concerns to be successful. However, we will be challenged with suspected plagiarism cases beyond ChatGPT in the future. Therefore, we recommend the combined use of plagiarism checkers and AI detection tools. If the use of such tools is not possible (*e.g.,* due to copyright and data protection laws in Germany), the manual examination of the texts remains, which is necessary for the evaluation anyway. Here, we recommend a thorough check of the references and a validation of the referenced literature. Our test cases and the results from the grey literature show that ChatGPT using GPT-3 has an easily-identifiable fingerprint (*e.g., it references nonexistent literature*). Likewise, the quality of the text, especially wrt.

Fig. 1. ChatGPT's impact on higher education
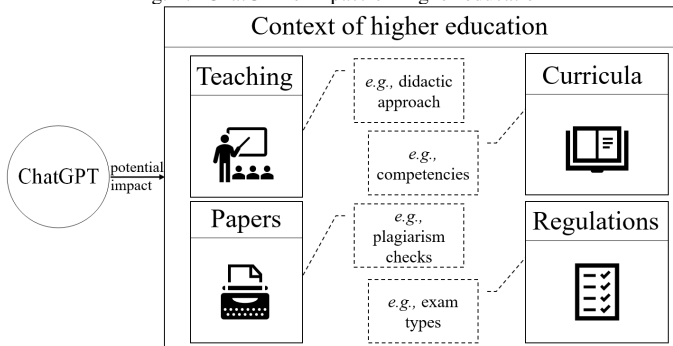
the lines of argumentation, is notably flawed in certain areas [11]. In these cases, we recommend an additional oral examination. Another, less time-consuming option is to require documentation of the examination process. Protocols can be used for this purpose, which in turn can also be evaluated and thus be considered in the exam grade. Finally, we also see a potential added value in the expected increased focus on the research design and results sections of papers. Provided that this aspect is well prepared, *e.g.,* by focusing on it in courses during the study or supervision during the preparation of paper, we assume a higher quality of scientific education.

Area 3 - Curricula: Adapting a curriculum is usually a time-consuming process because it requires a thorough analysis of the side-effects for other courses in the program and compliance with underlying laws and (institutional) rules. We expect that lecturers will thoroughly discuss this topic, as some may not want to integrate ChatGPT into their lectures or have other opinions about the tool in general (see Table I, C2). However, these discussions are of high importance, as they offer an opportunity for lecturers to learn from each other and find solutions to the emerging challenges.

Area 4 - Regulations: Here, we point to the need to evaluate official regulation documents (*e.g.,* examination regulations). Several legal aspects (*e.g.,* copyright or data protection) based on the underlying understanding of the evaluation of the different exam types must be considered in order to integrate ChatGPT into teaching. Thus, we recommend reevaluating the existing examination regulations to provide consistent rules for the students (see Table I, C2). The second aspect relates to the specific regulations of a study program. We see the need for thorough discussion among the lecturers of a study program to identify adoption potentials (*e.g.,* objectives of the course, underlying theoretical background, or even exam types) for the specific courses (addressing Table I, C4).

Consideration of the four areas may lead to the successful integration of ChatGPT into university teaching, thus reducing uncertainties and focusing on future-oriented teaching.

V. FUTURE WORK

The dynamic of emerging trends and disruptive technologies has increased in recent decades. Thus, SE educators have been challenged by various aspects.

Our results show the disruptive potential of AI-based chatbots such as ChatGPT for higher education, especially regarding scientific writing. Several open questions remain. Is text generated by ChatGPT a suspected plagiarism case? How should one reference text generated by ChatGPT? What proportion of text generated with ChatGPT in relation to the total scope is acceptable? Various further questions will probably arise and should be discussed in the SE community as they do.

Furthermore, we are already seeing that ChatGPT is very popular in practice, *e.g.,* in software development teams. The application possibilities of ChatGPT are so manifold that the extent of the impact is currently unpredictable. Taking all these aspects into account, we assume that this technology will not disappear in the future. Instead, AI tools will support software engineers in practice, education, and research. As lecturers educate the experts of tomorrow, there is a need to determine how we can teach students the necessary skills for sensible use of ChatGPT. It is therefore obvious that we need to address the question of how we will integrate such AI tools into higher education. Our contribution is a starting point for raising further questions, stimulating discussions, and finding solutions.

REFERENCES

[1] OpenAI, "Chatgpt: Optimizing language models for dialogue." 2022. [Online]. Available: https://openai.com/blog/chatgpt/

[2] M. U. Haque, I. Dharmadasa *et al.*, ""i think this is the most disruptive technology": Exploring sentiments of chatgpt early adopters using twitter data," 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2212.05856

[3] M. Neumann, M. Rauschenberger, and E.-M. Schön, "Research protocol," 2023. [Online]. Available: https://figshare.com/s/07fefa4448f6373ab710

[4] B. Kitchenham and Charters S., "Guidelines for performing systematic literature reviews in software engineering," 2007.

[5] A. Yasin, R. Fatima *et al.*, "On using grey literature and google scholar in systematic literature reviews in software engineering," *IEEE Access*, vol. 8, pp. 36 226–36 243, 2020.

[6] C. A. Gao, F. M. Howard *et al.*, "Comparing scientific abstracts generated by chatgpt to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers," 2022. [Online]. Available: https://doi.org/10.1101/2022.12.23.521610

[7] H. Else, "Abstracts written by chatgpt fool scientists," 1 2023. [Online]. Available: https://www.nature.com/articles/d41586-023-00056-7

[8] M. Rauschenberger and R. Baeza-Yates, "How to Handle Health-Related Small Imbalanced Data in Machine Learning?" *i-com*, vol. 19, no. 3, pp. 215–226, 2020.

[9] T. Susnjak, "Chatgpt: The end of online exam integrity?" 2022. [Online]. Available: https://doi.org/10.48550/ARXIV.2212.09292

[10] X. Zhai, "Chatgpt user experience: Implications for education," 2022. [Online]. Available: https://doi.org/10.2139/ssrn.4312418

[11] J. Qadir, "Engineering education in the era of chatgpt: Promise and pitfalls of generative ai for education," 2022. [Online]. Available: https://doi.org/10.36227/techrxiv.21789434.v1

[12] G. Salomon, "On the nature of pedagogic computer tools: The case of the writing partner," *Computers as Cognitive Tools*, pp. 179–196, 1993.

[13] M. Elsen-Rooney, "Nyc education department blocks chatgpt on school devices, networks," 2023. [Online]. Available: https://cacm.acm.org/news/268457-nyc-education-dept-blocks-chatgpt-on-school-devices-networks/fulltext

[14] M. Rauschenberger, "Acceptance by Design : Voice Assistants," in *1st AI-DEbate Workshop : workshop establishing An InterDisciplinary pErspective on speech-BAsed TEchnology*, 2021. [Online]. Available: https://opendata.uni-halle.de//handle/1981185920/38717