CHALLENGES AND OPPORTUNITIES PRESENTED BY GENERATIVE AI-CHATBOTS: EFFECTS ON AN UNDERGRADUATE COURSE ON DATABASES

Patricia Brockmann
Nuremberg Institute of Technology
GERMANY

Abstract

The recent revolution in chatbot technology based on artificial intelligence poses new problems in education. These machine learning systems have the ability to convincingly mimic human intelligence to an extent which makes it difficult to differentiate between answers generated by the chatbot versus answers written by a human expert. Not only news articles and essays, but also answers to homework assignments and even computer code can be generated with relatively little effort by using a generative AI chatbot. For instructors at institutions of higher education, ignoring these possibilities is not a realistic option. This paper presents an overview of the available relevant literature. Critical questions for instructors of computer science are raised. Experiences in an undergraduate course on databases are described. Initial lessons learned and resulting possibilities on how to integrate chatbot technology into teaching methods are presented.

Introduction

The term "chatbots" was first defined as a short-term moniker for a "chatter-bot" to denote an artificial intelligence application which could answer simple user questions (Mauldin, 1994). These initial chatbots were not very advanced. They were pre-programmed to answer a specific set of question with simple, predetermined answers. They were initially used to automate business processes, such as customer service.

Recently, a new generation of chatbots, such as Chat-GPT by OpenAI (OpenAI, 2023) and Bard (now Gemini) by Google (Google, 2023), have taken the media by storm. This revolution in chatbot technology is based on an artificial intelligence method called generative Large Language Models. These machine learning systems have the ability to convincingly mimic human intelligence to an extent which makes it difficult to differentiate between answers generated by the chatbot versus answers written by a human expert. Not only news articles and essays, but also answers to homework assignments and even computer code, can be generated with relatively little effort by using a generative AI chatbot.

For instructors of computer science, this technological revolution raises a number of critical questions:

1. Can instructors determine if students have actually completed homework assignments on their own, as opposed to just copying answers from a chatbot?
2. What effect does the availability of generative AI chatbots have on digital examinations? Are digital exams still feasible?
3. What should instructors be teaching their students in the future? Do students even have to learn how to program computers anymore?

First, an overview of the available literature will be presented. Next, experiences in an undergraduate course on databases will be described, examining solutions to lab assignments generated by ChatGPT. Initial lessons learned and possibilities on how to integrate comparing Chatbot technology into teaching methods will be presented. Finally, questions for further inquiry will be posed.

Relevant Literature

Generative Artificial Intelligence Chatbots

Large language models have been developed with advanced natural language capabilities. These artificial intelligence systems use deep learning methods to train language models on unstructured, unlabelled datasets, such as Wikipedia (Floridi & Chiriatti, 2020). Large language models are trained using statistical patterns of language on huge amounts of text available on the internet. These systems are called generative artificial intelligence because not only can they mimic human-like behavior as chatbots, but they can also perform intelligent searches, summarize text, and generate essays and computer code (Tamkin et al., 2021). The most widely known large language model chatbot is ChatGPT, which was released for free in November of 2022 (OpenAI, 2023). One reason for its widespread use was the fact that it was made available for free. Microsoft has also announced its investment in OpenAI, with the intention of incorporating it into its word processing software (Microsoft, 2023).

The rapid growth in the capabilities of generative artificial intelligence presents a number of opportunities for science and society. Some authors argue that generative AI chatbots can greatly speed up the process of writing and revising scientific articles. Pividori and Green (2023) demonstrated how an AI-based large language model could be used to suggest revisions to scholarly text. They predict that this development will revolutionize the performance of knowledge work.

On the other hand, Stokel-Walker and Van Noorden (2023) point out challenges caused by the use of generative AI. The unreliability of the results output by

ChatGPT and similar large language models is inherent to their method of learning. Because they are trained on content readily available on the internet, outdated facts, untruths, and inherent biases are part of their training data. Especially on topics with a low amount of training data, chatbots often return errors or misleading information. An editorial on ChatGPT in the journal *Nature Machine Intelligence* (2023) states that ChatGPT "cannot be trusted to get facts right or produce reliable references" (p. 1). As a result, a number of scientific journals, such as *Nature* and *Science*, have forbidden the use of any text generated by ChatGPT or any other AI tool (Thorpe, 2023).

A further concern is that because they are trained on the existing content of the internet, AI chatbots could further reinforce historical biases. To prevent this from happening, OpenAI installed filters to try to prevent ChatGPT from producing inappropriate content to hateful user prompts. Additional human moderators were required to flag certain text as toxic. Open AI has been criticized for allegedly paying Kenyan laborers less than $2 per hour to train ChatGPT (Perrigo, 2023).

Chatbots in Higher Education

Even before the advent of generative AI, chatbots have been successfully used in educational contexts. Molnár and Szüts (2018) discuss the use of chatbots in education. Simple retrieval-based systems were initially used to help students find information about study programs and curricula. With the advent of more complex artificial intelligence, such as IBM's Watson, it became possible to integrate a chatbot into a project-based learning class to teach about human-computer co-creativity (Goel et al., 2015). Watson was used to teach how to use biological systems as an inspiration to design new technological systems.

Shortly after the release of ChatGPT in November of 2022, Zhai (2022) reflected on the potential impact of generative AI chatbots on education. He demonstrated that with the aid of ChatGPT, it was possible to write a coherent, informative, and systematic research paper within 2-3 hours, without prior professional knowledge. The accuracy of the information used was not completely correct. He concludes that it will be necessary to adjust student learning goals. The use of AI tools can significantly increase the speed of conducting subject-domain tasks. He recommends that instructors should instead focus on improving students' creativity and critical thinking skills.

Qadir (2022) explored the challenges and pitfalls presented by ChatGPT in engineering education. He points out the potential advantages of using chatbots as intelligent virtual tutors to offer individual students personalized, adaptive learning. One of its most dangerous limitations, however, is its lack of reliability and tendency to output so-called "hallucinative" misinformation (p. 8). The major

challenge lies in differentiating between acceptable and non-acceptable use of chatbots, especially with regard to assessments.

The problem which chatbots pose with respect to plagiarism in higher education was illustrated by an editorial written by ChatGPT in response to prompts from King (King & ChapGPT, 2023). The danger of college students abusing ChatGPT to cheat on assignments was deemed unethical by ChatGPT itself. The chatbot recommended that college professors should incorporate a wide variety of assessment methods which go beyond traditional essays, such as oral presentations or group projects. Hands-on activities which require students to demonstrate their knowledge and skills would be an interactive and engaging assessment method that might help avoid cheating.

Rudolph et al. (2023) identify a number of challenges presented by generative AI chatbots. In addition to problems identifying plagiarism, they are also concerned about ChatGPT's inability to evaluate the relevance or accuracy of information shared. They classify generative AI as a new, disruptive educational technology and stress the role of educators in managing this new situation. They caution that the failure to address these challenges can lead to the exposure of inadequate pedagogical methods.

Instead of focusing solely on the dangers posed by this new technology, one university instructor (Guerra-Pujol, 2023) decided to positively integrate AI chatbots into his course. As a first step, students are explicitly assigned to look up answers to discussion questions using AI chatbots. They post screenshots of the answers delivered by the chat bot to the course learning platform. Their task is to cross-check the chatbot's response for accuracy and to suggest factual, substantive, or stylistic revisions which would be necessary to improve this answer. In this way, students learn to work with this new technology responsibly, while at the same time developing their critical thinking skills. McMurtrie (2022) argues that with the integration of generative AI tools in word processing software, chatbots will become a normal part of writing, just as calculators and computers have become an essential part of math and science. ChatGPT can also open positive opportunities for student-centered pedagogies, such as experiential learning or game-based learning (Rudolph et al., 2023). They advise against a policing approach which completely forbids the use of AI tools. Instead, they recommend embracing this technological disruption to improve teaching methods. Instead of simple reproduction-centered assignments, instructors should strive to inspire students' creative and critical thinking abilities. By incorporating AI tools into discussions and assignments, they can learn the limits and faults of generative AI and learn to use them judiciously (Fyfe, 2022).

Observations in an Undergraduate Course in Databases

During the COVID-19 pandemic, contact restrictions necessitated online teaching and examinations. Many educators count themselves as lucky that ChatGPT did not yet exist at that time. The release of ChatGPT demonstrated that old-fashioned assessment methods could no longer be used. This section describes initial investigations into the effect of generative AI chatbots on an undergraduate course in databases. The free version of ChatGPT available in May of 2023 was used. Lab assignments were typed into ChatGPT and then the answers provided by ChatGPT were subjected to same grading scale used for students.

The first lab assignment for this course is to perform requirements engineering for a case study. Students are assigned to write user stories (Cohn, 2004) from the perspective of database users for a web shop. If a student feels unsure of how to write a user story, ChatGPT provides an initial introduction, as shown in Figure 1.
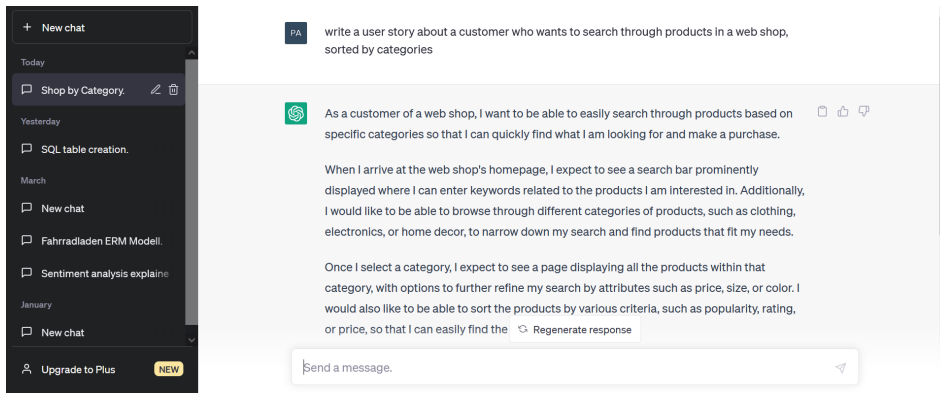
Figure 1

Introductory Text about User Stories Output by ChatGPT (OpenAI, 2023)
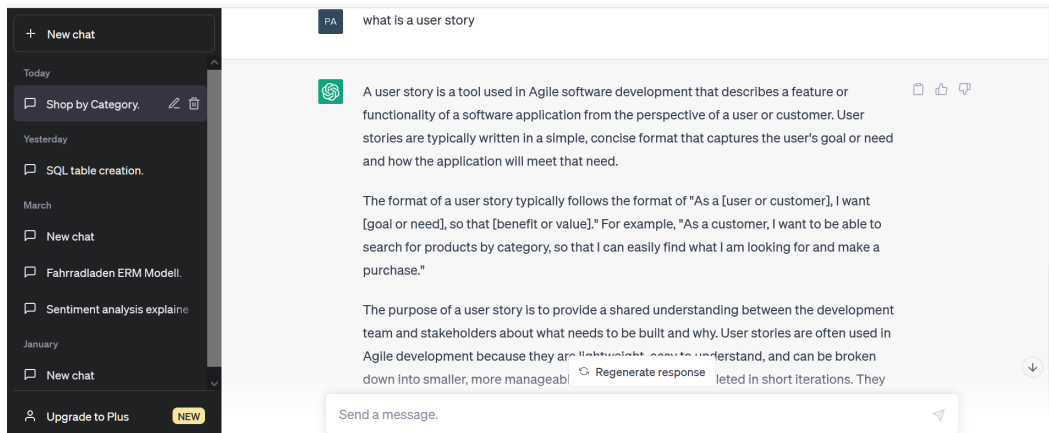
As shown in the first sentence of Figure 2, ChatGPT then correctly generated a user story utilizing the formal format as specified in agile software requirements engineering. The second and third paragraphs output by ChatGPT provide additional descriptive text and would not be considered part of the user story. The inclusion of scenario description which resembles UML notation rather than a user story would have led to a lower grade of "B", good, but not perfect.

Figure 2

Solution to Text-based Lab Assignment by ChatGPT (OpenAI, 2023)

The next step in the development of a database would be to develop a conceptual model, called an Entity-Relationship Diagram (Bagui & Earp, 2011). As shown in Figure 3, ChatGPT was able to generate a primitive Entity-Relationship Diagram for a webshop, but not a specific one for the case study presented. For this lab assignment, the solution would receive a barely passing grade of "D". While this solution was created by the free version of ChatGPT in May of 2023, the licensed, Plus version may produce better results for graphic diagrams.

Figure 3

Solution of Entity-Relationship Diagram by ChatGPT (OpenAI, 2023)
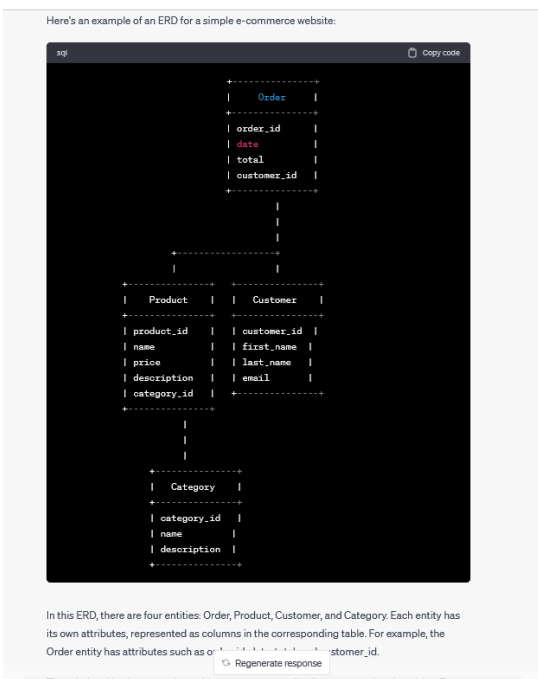
For the next lab assignment, a relational data model (Codd, 1980) was derived from this ER-diagram. The solution generated by ChatGPT was so general that it did not reflect the case study and thus also earned a grade of "D".

Relational Algebra (Prade & Testemale, 1984) is an extremely formal topic in database theory, which many students have difficulty understanding. For this lab assignment, ChatGPT was able to produce a partially correct solution (Figure 4). The selection operator is missing, correlation operands are missing, and the sequence of the operators presented in the solution is inefficient. For this lab assignment, ChatGPT would get a grade of "C", sufficient.
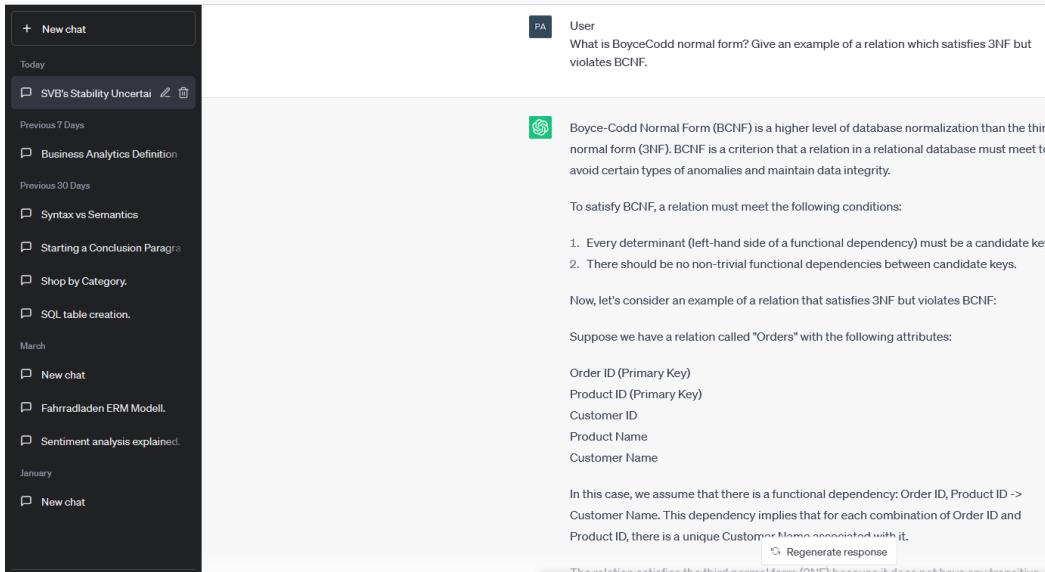
Figure 4

Solution to Relational Algebra Output by ChatGPT (OpenAI, 2023)
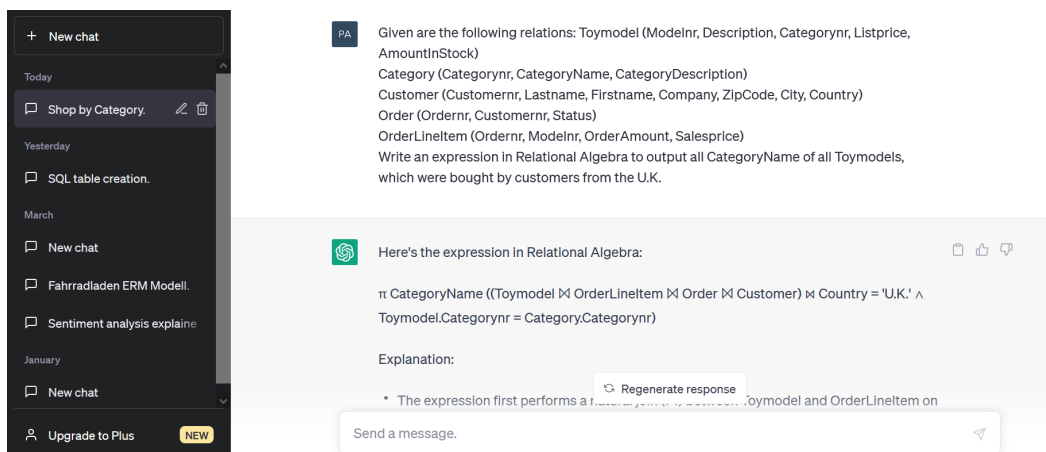
Normalization of a relational data model is a theoretical topic based on set theory (Codd, 1980). As shown in Figure 5, ChatGPT returns correct definitions.
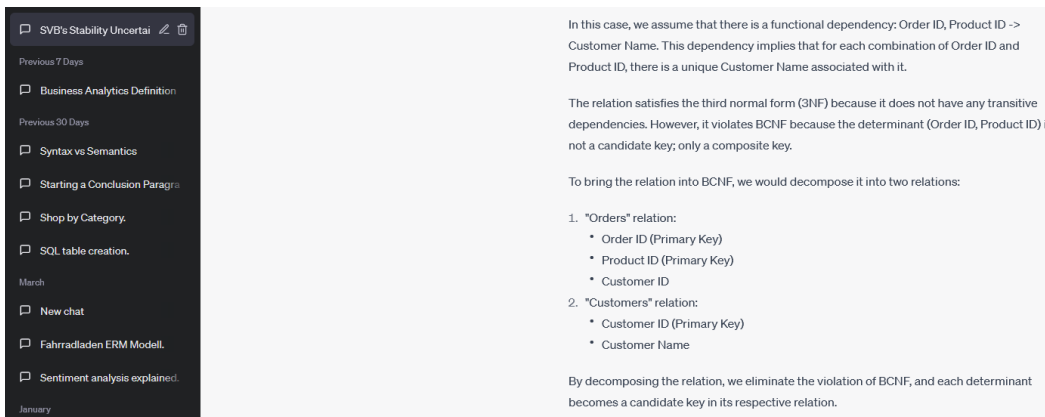
Figure 5

First Steps to Solve Normalization Output by ChatGPT (OpenAI, 2023)

However, as shown in Figure 6, ChatGPT gives incorrect answers when asked to actually solve the lab problem on normalization. Both functional dependencies and normal forms are incorrectly identified in the lab problem. This solution would receive the grade of "D", poor.

Figure 6

Incorrect Solution to Normalization Output by ChatGPT (OpenAI, 2023)


Programming assignments were easily solvable by ChatGPT. As shown in Figure 7, ChatGPT was asked to generate a query in the programming language SQL to calculate the number of customers in each country of the database and then to select only those countries which have at least five customers and sort the results.

Figure 7

Solution to Lab Assignment Provided by ChatGPT (OpenAI, 2023)
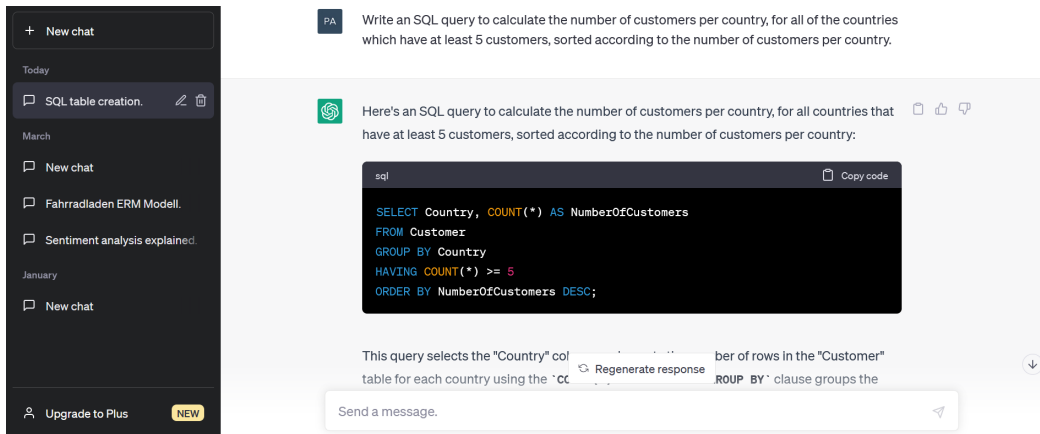
Undergraduate students who are first learning to program databases in the language SQL for the very first time typically take 15 to 30 minutes to produce the correct answer. By simply typing their assignment into ChatGPT, they could obtain the correct answer within seconds. In addition, ChatGPT provided an explanation of the results, with a description of each operator used, as shown in Figure 8. For this assignment, ChatGPT was awarded the grade of "A", excellent.

Figure 8

Explanation Provided by ChatGPT (OpenAI, 2023)
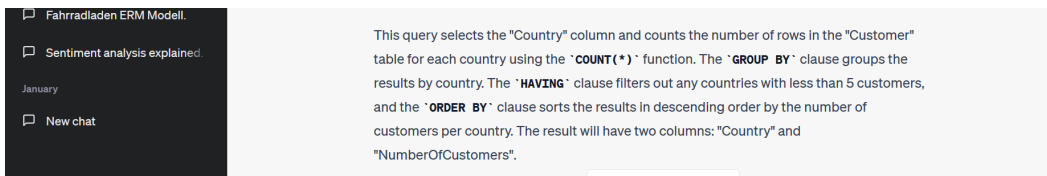
The final set of lab assignments deal with the concept of transaction processing in multi-user databases (citation). ChatGPT was also able to generate mostly correct solutions for this assignment. A grade of "B", good was awarded.

Discussion of Results

The results of the grades which ChatGPT received on the lab assignments are summarized below in Table 1. For a total of 10 lab assignments which students have to complete during the semester, ChatGPT would have received the grade of 2.7, equivalent to a grade of B-. This ChatGPT grade is considerably better than the average grade (1.7 C-) achieved by students over the last 10 years.

Table 1

Grades Received by ChatGPT on Lab Assignments for Course on Databases

Nr.	Assignment	Grade	Nr.	Assignment	Grade
1	Requirements Engineering	B	6	SQL Prog. Basic	A
2	Entity-Relationship Model	D	7	SQL Prog. Middle	A
3	Relational Model	D	9	SQL Prog. Difficult	A
4	Relational Algebra	C	9	SQL Prog. Expert	A
5	Normalization	D	10	Multi-user Transactions	B

What does this mean specifically for this undergraduate course in databases? A return to the initial research questions is warranted.

1. Can instructors determine if students have actually completed homework assignments on their own, as opposed to just copying answers from a chatbot?

   For assignments which require conceptual modelling, the highly general nature of the answers suggests they were written by a chatbot. For technical assignments such as programming, it is not easily possible to detect whether the solutions examined here were generated by a chatbot.

2. What effect does the availability of generative AI chatbots have on digital examinations? Are digital exams still feasible?

   For online exams conducted without any monitoring software, it would be close to impossible to detect cheating. The use of monitoring cameras or software may be considered an invasion of privacy in some countries.

3. What should instructors be teaching their students in the future? Do students even have to learn how to program computers anymore?

   This question cannot be yet be fully answered by this initial investigation into ChatGPT. Automation of simple, commonly used programming tasks may be feasible. More complex problems, which need an analysis of requirements and the development of models, cannot yet be adequately automated by ChatGPT.

Conclusions and Future Work

ChatGPT poses a significant technical disruption to educational practices. As educators, it will not be possible to ignore the effects of generative AI chatbots. Recognition of plagiarism and cheating in student assessments is more difficult. Simple programming skills may no longer be adequate for their future careers.

Instead of completely forbidding the use of chatbots, finding methods to integrate this new tool into pedagogic practices would be advisable. Future work will include experiments which explicitly assign students to investigate how well ChatGPT can perform on lab assignments. The quality of the answers which the chatbot generates will then be a topic for classroom discussion. Is this answer correct? Could this answer be improved? This mindful use of generative AI could help students not only learn how to use this new technology, but also further develop their critical and analytical thinking skills.

References

Bagui, S., and Earp, R. (2011). *Database design using entity-relationship diagrams*, (3rd ed.). Crc Press

Codd, E. F. (1980, June). Data models in database management. In *Proceedings of the 1980 workshop on Data abstraction, databases and conceptual modeling* (pp. 112-114). New York: Association for Computing Machinery. https://dl.acm.org/doi/10.1145/800227.806891

Cohn, M. (2004). *User stories applied: For agile software development*. Addison-Wesley Professional.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, *30*, 681-694.

Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI & Society,* 1-11.

Goel, A., Creeden, B., Kumble, M., Salunke, S., Shetty, A., & Wiltgen, B. (2015, September). Using Watson for enhancing human-computer co-creativity. In *2015 AAAI fall symposium series (pp. 22-29).* Association for the Advancement of Artificial Intelligence. https://cdn.aaai.org/ocs/11713/11713-51257-1-PB.pdf

Google. (2023). *Gemini*. https://gemini.google.com

Guerra-Pujol, F. E. (2023). How I Learned to Love GPT-3. *UCF Faculty Focus*, *22*(1), 25.

King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, *16*(1), 1-2.

Mauldin, M. L. (1994, August). Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition. In *AAAI* (*94*), 16-21).

McMurtrie, B. (2022, December 13). AI and the future of undergraduate writing. *The Chronicle of Higher Education*. https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing

Microsoft (2023). https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership/

Molnár, G., & Szüts, Z. (2018, September). The role of chatbots in formal education. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 000197-000202). IEEE.

Nature Machine Intelligence. (2023, January). The AI Writing on the Wall. *Nature Machine Intelligence*, (2023). *5*, p. 1. https://doi.org/10.1038/s42256-023-00613-9

OpenAI. (2023). *Introducing ChatGPT*. https://openai.com/blog/chatgpt

Pividori, M., & Greene, C. S. (2023). A publishing infrastructure for AI-assisted academic authoring. *bioRxiv*, 2023-01.

Prade, H., & Testemale, C. (1984). Generalizing database relational algebra for the treatment of incomplete or uncertain information and vague queries. *Information sciences*, *34*(2), 115-143.

Qadir, J. (2022). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. In: *IEEE Global Engineering Education Conference (EDUCON)*, Kuwait, Kuwait, 2023, pp. 1-9. https://doi.org/10.36227/techrxiv.21789434.v1

Perrigo, B. (2023, January 18). Open AI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic. *Time Magazine.* https://time.com/6247678/openai-chatgpt-kenya-workers/

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 342-263. https://doi.org/10.37074/jalt.2023.6.1.9

Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. *Nature*, *614*(7947), 214-216.

Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the capabilities, limitations, and societal impact of large language models. *arXiv preprint arXiv:2102.02503*.

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science 379*(6630), 313. DOI: 10.1126/science.adg7879

Zhai, X. (2022). ChatGPT user experience: Implications for education. http://dx.doi.org/10.2139/ssrn.4312418

Author Details

Patricia Brockman
Nuremberg Institute of Technology
Germany
patricia.brockmann@th-nuernberg.de