



# Challenges for Computer Science Education Arising from new AI Systems like ChatGPT

Robert Mischak<sup>1</sup>, Robert Pucher<sup>2</sup>

University of Applied Sciences - FH JOANNEUM, Institute of eHealth Austria, Styria<sup>1</sup>  
University of Applied Sciences – Technikum Wien, Department of Computer Science, Austria, Vienna<sup>2</sup>

## Abstract

*Intelligent AI technologies like ChatGPT can significantly impact traditional teaching methods, diminishing the value of familiar elements like knowledge testing and coding exercises. With ChatGPT's rapid production of high-quality solutions, it becomes challenging for teachers to distinguish student work from machine-generated output. This issue especially affects computer science students, as ChatGPT can produce code for most computer languages. ChatGPT, powered by transformer networks, accepts natural language input and delivers grammatically correct answers in multiple languages. Trained on vast internet data until September 2021, its responses often exhibit extensive knowledge. However, these answers often follow a discernible pattern, which trained teachers may identify. The difficulty lies in proving that the answers originate from the student as they are generated on the spot, eluding traditional plagiarism detection tools. Consequently, computer science educators must develop new instructional approaches to ensure fair and accurate assessment of student performance or find meaningful ways to incorporate ChatGPT's potential into lessons. This necessitates considering general principles and providing practical recommendations for teaching in this context.*

**Keywords:** *artificial intelligence in education, ChatGPT, Large Language Model (LLM), computer science, educational recommendations*

## 1. Introduction

New AI-based technologies have the potential to limit severely the effectiveness and significance of familiar didactic elements, such as knowledge testing through seminar papers or the submission of source code as an exercise. This is the case when students are supported in their work by AI applications such as ChatGPT. ChatGPT is referred to as a "Large Language Model" (LLM) and produces astonishing and high-quality solutions in a matter of seconds - without any significant input from the students themselves. It is then often very difficult for the teachers to recognize whether the work was done by the student himself or by a machine. This phenomenon is not limited to computer science students, but it can be assumed that they are even more affected because of their relevant technical competencies. ChatGPT is based on an artificial neural network technology, called "transformer networks". This technology allows not only to take natural language as an input but also gives the answers in several languages in grammatically correct sentences which are structured like similar sentences written by humans. The network was trained by using most data available on the entire internet with a cutoff date of September 2021. Therefore, many answers seem to reflect profound knowledge. Our first assessment of ChatGPT-generated answers showed that in many cases these answers follow a specific pattern. This pattern might allow trained teachers to recognize the source. The challenge is that it is very difficult to prove the answers do not come from the student himself, as the answer is generated and not stored on the internet. All standard tools used to check for plagiarism fail utterly to detect the source of the text.

Thus, it is necessary for teachers of computer science courses to develop new didactic concepts that allow a fair and appropriate performance assessment of the students or to integrate the potential of ChatGPT into the lessons in a meaningful way. Therefore, some general considerations as well as practical recommendations for teaching are given

## 2. Example – "What are the benefits and risks of LLMs in the context of teaching?"

LLMs (Large Language Models) have several benefits and risks when it comes to their application in the context of teaching. Here are some of the key points to consider:



### Benefits:

- **Access to vast knowledge:** LLMs can provide access to a vast amount of information and knowledge. They have been trained on large datasets, making them capable of answering a wide range of questions and providing detailed explanations on various subjects.
- **Personalized learning:** LLMs can adapt to individual student's needs and provide personalized learning experiences. They can assess the proficiency level of students and tailor their responses accordingly, offering targeted explanations and guidance.
- **24/7 availability:** Since LLMs are digital entities, they can be available at any time, enabling students to access information and receive assistance whenever they need it. This accessibility can be particularly beneficial for students who require additional support outside of regular class hours.
- **Multilingual support:** LLMs can assist students who are learning a second language by providing translations, explanations, and examples in their native language. This feature can help overcome language barriers and enhance comprehension for students with diverse linguistic backgrounds.

### Risks:

- **Lack of human interaction:** One of the primary risks of relying solely on LLMs in teaching is the absence of human interaction. Human teachers bring empathy, emotional support, and a deeper understanding of individual student's needs that may be lacking in an automated system.
- **Limited contextual understanding:** Although LLMs can generate accurate and relevant responses, they may lack contextual understanding. They may not grasp the subtleties of a student's question or the specific educational goals of a particular lesson, potentially leading to irrelevant or misleading information.
- **Potential biases:** LLMs are trained on vast datasets that can contain biased or unverified information. As a result, they may inadvertently perpetuate biases or present inaccurate information. This risk highlights the importance of critical thinking and human oversight when using LLMs in educational settings.
- **Reduced creativity and critical thinking:** While LLMs excel at providing factual information, they may struggle to foster creativity and critical thinking skills in students. These qualities are often developed through interactive discussions, brainstorming sessions, and hands-on activities, which are better facilitated by human teachers.
- **Overreliance on technology:** There is a risk of overreliance on LLMs and technology in general, leading to a diminished role for human teachers. This can potentially hinder the development of social and emotional skills that are best nurtured through personal interactions and mentorship.

It is important to strike a balance between utilizing LLMs as a valuable tool for information retrieval and leveraging the expertise and guidance of human teachers to create a holistic and effective learning environment.

What's wrong with the above answer? It seems to be a good summary of many key risks and benefits. There is nothing wrong with the summary, except, it was entirely written using ChatGPT May 24 Version. As input the question "What are the benefits and risks of LLMs in the context of teaching?" was used, and the output had to be adjusted minimally.

### 3. The technology behind ChatGPT



ChatGPT is built on a technology called "transformer networks," which form the backbone of its AI capabilities. Transformer networks are a type of artificial neural network architecture specifically designed for processing sequential data, such as language. They revolutionized natural language processing tasks by introducing a mechanism called self-attention [1].

Self-attention allows the model to weigh the importance of different words or tokens within a given context, enabling it to capture long-range dependencies effectively. This mechanism helps ChatGPT understand the relationships between words and generate coherent and contextually appropriate responses.

The training process of ChatGPT involves feeding it a vast amount of text data from the internet, up until a cutoff date (September 2021). Through a process known as unsupervised learning, the model learns to predict the likelihood of a word given its surrounding words, thereby acquiring language understanding and generating responses that mimic human language. A detailed description of how ChatGPT works can be found here [2]. The astonishing performance of LLMs like ChatGPT is partly based on their complexity. The number of parameters in ChatGPT version 4 is in the same order of magnitude as the number of synapses found in the human brain [3]. However, a direct comparison to the human brain is difficult; only to simulate a single biological neuron, one needs at least 1000 artificial neurons [4].

ChatGPT underwent a supervised fine-tuning process. Humans provide conversations where they play both the user and an AI assistant. They have access to model-written suggestions to aid their responses. This dialogue dataset is mixed with the InstructGPT dataset, which is transformed into a dialogue format. The model is then fine-tuned on this combined dataset using supervised learning techniques. This was necessary to make answers of ChatGPT less toxic. The huge work was mainly done by underpaid workers in Kenia [5].

## 4. Implications for education

### 4.1. Didactic framework

Students are supposed to acquire competencies during their studies, which are well known in education according to Bloom's Taxonomy [6], in which six levels of the cognitive domain are defined as follows: knowledge, comprehension, application, analysis, synthesis, and evaluation, starting with the lowest level, which is knowledge. In Table 1 these six levels are described more in detail [7]

**Table 1. Bloom Cognitive Competency [7]**

Level	Description
Knowledge	<ul style="list-style-type: none"> <li>• There is no real understanding</li> <li>• The concept is known</li> </ul>
Comprehension	<ul style="list-style-type: none"> <li>• Understand how the concepts are working</li> </ul>
Application	<ul style="list-style-type: none"> <li>• Able to apply the knowledge in the new scenario by example guided by the lecturer</li> </ul>
Analysis	<ul style="list-style-type: none"> <li>• Able to analyze the problem</li> <li>• Able to solve the problem on its own</li> </ul>
Synthesis	<ul style="list-style-type: none"> <li>• Able to generalize what he learned for a new problem</li> <li>• Able to make a conclusion</li> </ul>
Evaluation	<ul style="list-style-type: none"> <li>• Able to compare various solution</li> <li>• Able to make a conclusion based on various solutions</li> </ul>

In a more modern version, the six levels are: remembering, understanding, applying, analyzing, evaluating, and creating [6]. Nevertheless, the progress in the development of the students has to be checked continuously. For this purpose, different examination methods are in use: written and oral



examinations, multiple- and single-choice examinations, seminar papers, presentations, project work, and scientific publications. It should be noted that these examination methods may differ depending on the organizational form of the study program. For example, in the case of part-time courses, a significantly higher proportion of e-learning and blended learning must be taken into account. It is also important to distinguish whether you are in a bachelor's or master's program because it makes a difference which level of competencies can fully be achieved. It is to be considered that in bachelor's programs, the top level will not be fully achieved. Here's in table 2 is what this means for teachers when applying Bloom's competencies to types of exams

**Table 2. Bloom Cognitive Competency applied to types of exams**

Level	Exams
remembering	single and multiple-choice tests
understanding	writing seminar papers, homework, oral and written exams
applying	project work
analyzing	presentations with discussions
evaluating	colloquium as a discussion between lecturers and students
creating	bachelor's, master's thesis, writing scientific posters and papers

#### 4.2. Control ChatGPT with software

Conspicuous features of AI-generated texts are, for example, the language level in terms of grammar and spelling, which is probably above the level of average students. Also, references are often hallucinated by the AI, this can at least be verified. In the case of computer source code, a certain style may be recognized by the teacher.

It does not seem purposeful that software for checking AI-generated texts, codes, etc. can be used in a meaningful way in teaching similar to plagiarism software. Therefore, the focus should be on avoiding such practices of the students already in advance, which leads to the following recommendations.

#### 4.3. Recommendations

For texts or codes that are partially or even completely created by AI-generated results, it is generally true that humans have responsibility and liability for their actions. Nevertheless, in assessing the final product, an instructor cannot see the extent to which students have used KI. This consequently means a shift in the focus of examinations from product quality to process quality. Scientific theses (bachelor's or master's theses) are to be reconsidered in their overall structure, especially introductory chapters, etc.

It is therefore incumbent upon instructors to examine the quality and appropriateness of a student's process to a final product. With good process quality, this should also have good product quality anyway.

The lessons themselves should be as open to AI as possible so that online research on the Internet, translations into other languages, suggestions for code parts or scientific sources, standards, and laws, or other interesting content can be carried out at any time. The use of ChatGPT should be integrated, examples of this are described [8]. Human educators in education are very important and the supporting role of technology should not be exaggerated [9]. There is a lot of research going on addressing the teacher's perspective as well as the learner's perspective [10].

For teachers, this means a significant additional effort, since simple forms of examination that only evaluate a product at the end are no longer sufficient. To sensibly integrate AI into the lessons, the teachers themselves need didactic and professional further training.





## 5. Conclusions

Applications of generative AI cannot be stopped and should not be stopped. Tools like ChatGPT should be accepted as a technological advance, with all its opportunities and dangers. Therefore, students should be empowered to use these technologies responsibly and competently. This of course has consequences for the design of the lessons and the performance review, which should be adequate and fair.

Interactive performance assessments are to be preferred to simply hand in seminar papers or homework assignments.

To maintain the competitiveness of our society, computer science education must proactively integrate this issue. Therefore, teachers must also be involved in the sense of awareness raising and professional and didactic training. We are at the beginning of a process, a new stage of digital transformation, which has now also massively reached teaching. The motto of *lifelong learning* has never been more relevant than today!

## References

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). "Attention Is All You Need" <https://doi.org/10.48550/arXiv.1706.03762>
- [2] Mark Riedl. (2023). "A Very Gentle Introduction to Large Language Models without the Hype", <https://mark-riedl.medium.com/a-very-gentle-introduction-to-large-language-models-without-the-hype-5f67941fa59e>, Accessed May 2023.
- [3] Alberto Romero (2023). "GPT-4 Will Have 100 Trillion Parameters — 500x the Size of GPT-3" <https://towardsdatascience.com/gpt-4-will-have-100-trillion-parameters-500x-the-size-of-gpt-3-582b98d82253>, Accessed Apr. 2023.
- [4] David Beniaguev, Idan Segev, Michael London (2021). "Single cortical neurons as deep artificial neural networks", *Neuron* Volume 109, Issue 17, P2727-2739.E3, September 01, 2021, <https://doi.org/10.1016/j.neuron.2021.07.002> .
- [5] Eva Wolfangel (2023). "Ausgebeutet, um die KI zu zähmen" <https://www.zeit.de/digital/2023-01/chatgpt-ki-training-arbeitsbedingungen-kenia>, Accessed Apr. 2023.
- [6] M. Forehand, "Bloom's Taxonomy: Original and Revised", Accessed: May 26, 2023. [Online]. Available: <http://www.coe.uga.edu/epltt/bloom.htm> .
- [7] A. Lajis, H. Md Nasir, and N. A. Aziz, "Proposed assessment framework based on bloom taxonomy cognitive competency: Introduction to programming," *ACM International Conference Proceeding Series*, pp. 97–101, Feb. 2018, <https://doi.org/10.1145/3185089.3185149> .
- [8] B. Qureshi, "Exploring the Use of ChatGPT as a Tool for Learning and Assessment in Undergraduate Computer Science Curriculum: Opportunities and Challenges," Apr. 2023, Accessed: May 22, 2023. [Online]. Available: <https://arxiv.org/abs/2304.11214v1>
- [9] A. Bozkurt *et al.*, "Speculative Futures on ChatGPT and Generative Artificial Intelligence (AI): A Collective Reflection from the Educational Landscape," *Asian Journal of Distance Education*, vol. 18, no. 1, Jan. 2023, Accessed: May 22, 2023. [Online]. Available: [https://digitalcommons.odu.edu/teachinglearning\\_fac\\_pubs/199](https://digitalcommons.odu.edu/teachinglearning_fac_pubs/199)
- [10] E. Kasneci *et al.*, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learn Individ Differ*, vol. 103, p. 102274, Apr. 2023, <https://doi.org/10.1016/j.lindif.2023.102274> .