ChatGPT in Education – Let's Think Step by Step

Michael Schwarz
schwarz.michael@posteo.net

Abstract

Since ChatGPT was released by OpenAI, there has been a lot of discussion about its educational applications. That is why, I investigate the prerequisites and constraints for using ChatGPT in education [RQ1], practical integration methods for learners and lecturers [RQ2], and the effectiveness of detecting AI-generated content in academic writing [RQ3], on the basis of the latest data until November 2023.

The results show that ChatGPT became accessible to educational institutions with ChatGPT Enterprise. The use of ChatGPT opens up new methodological possibilities, which will change teaching and learning in the future, which offers significant benefits for both learners and lectureres. However, lecturers need to be supported in their further training, as they currently lack AI skills. In addition, students need to be taught the necessary skills to use AI correctly and they need rules for its use. Furthermore, there are concerns about copyright and the risk of plagiarism. Current recognition tools are still at an early stage or unreliable, projects such as CheckGPT have shown promising recognition rates (98-99%) in benchmarks, which indicates that it is manageable, but also highlights the need for further research in this area.

1 Introduction

Large Language Models (=LLM) like Generative Pre-trained Transformer 4 (=GPT-4) have the potential to significantly impact education as we know today. With the introduction of ChatGPT, which was offered by OpenAI [1] for free, many people who had not interacted with conversational agents before could get first-hand experience with an advanced conversational agent. This development led to extensive discussions about the possible applications of ChatGPT in educational institutions, from primary schools to universities.

I investigate whether *the prerequisites for the use of ChatGPT in the field of education are fulfilled and which restrictions must be followed* [RQ1]. In particular, *how ChatGPT can be practically integrated in educational settings for learners and lecturers* [RQ2]. Furthermore, given the challenges in distinguishing between human-written and AI-generated content, the question arises *how effective current detection mechanisms are in academic writing* [RQ3]. This research does not collect any new data and therefore relies on data from studies conducted by others, focusing on information that is as up-to-date as possible (until November 2023). Furthermore, In this work, I use DeepL and ChatGPT, two popular AI-tools. For transparent use of ChatGPT, the full conversation log with ChatGPT can be found at the following Zenodo repository: https://doi.org/10.5281/zenodo.10252663. Results show that ChatGPT is age restricted and in most cases not GDPR-compliant. With the release of ChatGPT Enterprise OpenAI meets all safety standards and opened up opportunities for schools and universities to make ChatGPT accessible for each student and lecturer. Practical application examples of ChatGPT outline that ChatGPT has many added values for learners and teachers in and out of the classroom. Furthermore, concrete guidelines are presented that teachers and students can follow in order to successfully use ChatGPT.

Importantly, teachers should encourage students to use ChatGPT. By teaching students how to effectively use these tools, educators can equip them with important skills while also emphasizing the risks and restrictions of ChatGPT. Unfortunately, the study (Polak et al., 2022) shows that teachers currently know little about AI, which can hinder students digital competence development. On the other hand, lecturers are very motivated and have the necessary understanding of the importance of the topic to be willing to learn and teach AI skills. It is also clear that AI will change the way of teaching and assessing.

---

[1] https://openai.com

In academia, concerns arise regarding copyright and the risk of plagiarism. Text generated by ChatGPT are difficult for humans to recognize as well as current detector tools are still at an early stage or simply not reliable enough. However, projects like CheckGPT demonstrate in benchmark tests (Liu et al., 2023) with an average recognition rate of 98%-99% accuracy that the problem is manageable but also that more research is needed in that domain.

2 Prerequisites for the Use of ChatGPT

2.1 OpenAI Conditions

Before ChatGPT can be used in education, the platform must meet some legal requirements, which are regulated by the General Data Protection Regulation (=GDPR) and data protection regulations of the federal states in Germany. Especially schools must contract with a OpenAI Data Processing Addendum (=DPA) according to Art. 28 GPDR, that the data will only be processed for the purpose and on the instruction of the school.

The Europe Terms of Use (14 Nov 2023) govern the use of OpenAI services like ChatGPT for personal and non-commercial purposes in the European Education Area[2] (=EEA). [3] Therefore, OpenAI provides users with a legal guarantee for their services under EEA consumer law. To be able to use ChatGPT, users have to create a free or paid account connected to a mobile number. Especially for users in Germany this a problem, because every mobile number has to be verified by proof of identification, resulting that the user data will be uniquely identifiable. Additionally, users must be at least 13 and those under 18 need parental or guardian permission to use the services. By using the online platform, users retain ownership of their input and receive all rights to the output, as well as OpenAI is allowed to use the user content globally for purposes such as service improvement or model training. Additionally, OpenAI collects personal information like account information or log data, what can be viewed in the privacy policy.[4]

The Business Terms of Use (14 Nov 2023) govern the use of OpenAI services like ChatGPT Enterprise and OpenAI APIs for businesses and developers.[5] OpenAI grant enterprise accounts extensive

security, privacy and performance improvements. Thereby, OpenAI does not review any business data and has been audited for SOC 2 Type 1 compliance which includes GDPR, DPA and more.[6] Additionally, enterprise accounts do have features like admin console to manage members, single-sign-on, domain verification and large-scale deployments, which is important for educational facilities.

2.2 School and University Conditions

According to current information from the Bavarian State Ministry of Education and Cultural Affairs (=StMUK), relevant regulations must be observed in the context of teaching, including school law, e.g. Art. 56 para. 5 BayEUG and § 46 BaySchO. School teachers may use, on a voluntary basis, ChatGPT within the scope of their pedagogical scope of action, subject to compliance with the terms of use and data protection. Since there are different rules in all federal states in Germany, there seems to be no uniform regulation, which is why teachers and schools have to clarify the use individually. Likewise, there is no common understanding how ChatGPT should be used for a pedagogically meaningful teaching, which hardly depends on the training course.(StMUK, 2023)

The universities specify different rules for the use of ChatGPT and is ultimately dependent on the lecturer. For example, Augsburg University of Applied Sciences focus on best-practices and general knowledge, whereas University of Augsburg also recommends rules for dealing with AI systems.(DZM, 2023c) (DigiLLAb, 2023) In contras, the published document "Rules for Tools"(Spannagel, 2023c) by Prof. Dr. Spannagel from University of Heidelberg provides clear rules on how students should deal with AI systems such as ChatGPT in his courses, which got referenced by many universities already.

The example of data privacy shows how problematic the inadequate rules at many universities are for students, teachers and researchers: Already there are a growing number of anecdotes flying around of researchers who have naively uploaded human subjects data or drafts of papers into ChatGPT, giving away protected data and their intellectual property to OpenAI.(Maynard, 2023)

---

[2]https://education.ec.europa.eu/about-eea/the-eea-explained
[3]https://openai.com/policies/eu-terms-of-use
[4]https://openai.com/policies/privacy-policy
[5]https://openai.com/policies/business-terms

[6]https://openai.com/enterprise-privacy

2.3 Opportunities to Use ChatGPT

Before the release of ChatGPT Enterprise a regular offering by institutions for individual student accounts was not possible. But now, OpenAI has created the conditions for a sustainable use in education. However, schools and universities can not yet claim this at an institutional level. One positive aspect is that the use of ChatGPT is possible in principle, but currently teachers and students still have to deal with very individual conditions:

1. Students create handwritten prompts that the teacher enters into ChatGPT.(König et al., 2023) This may be a sensible approach for children under the age of 13.

2. Students use the teacher access on a school device for independent interaction.(König et al., 2023) Caution: OpenAI prohibits the disclosure of login data to students or the creation of multiple logins.

3. Use of personal accounts from students.(König et al., 2023) This would be the worst option because not all students can afford a paid account or have limited access to technology, leading to increasing inequality.

4. Institutions or service provider deploy their own GPDR compliant application connected to the OpenAI API, which can be used by each student.

5. Institutions use an enterprise account and thus offer access for every student in school or university to ChatGPT for personal learning, professional use, class assignments and research. This marks a significant step forward institutions can offer students and is a real game changer in most discussions regarding the use of ChatGPT in the field of education.

2.4 A Survey on the Teachers' Perspective on AI in Education

Teachers play a crucial role in classroom innovation. Using the "Will, Skill, Tool" (=WST) model, the research by (Polak et al., 2022) aimed to guide the development of educational content and platforms for AI instruction.
The WST model is a theoretical framework by (Petko, 2012) that outlines the conditions affecting the acceptance of digital tools in classrooms:

- *Will*: Represents a teachers attitude towards technology. A positive "Will" believes technology enhances learning, while a negative "Will" may fear potential risks.

- *Skill*: Represents the digital competencies of a teacher. Regarding the study by (Petko, 2012) the skill element has been found to have the largest explanatory power among the elements of the WST model.

- *Tool*: Refers to the actual availability of digital tools, such as computers, laptops, tablets, and software.

Based on the WST model an online questionnaire was distributed across professional educational network in Bulgaria, Greece, Italy and Romania. A total of 135 people from 89 schools responded to the online questionnaire. Due to the non-normality of the data, the Mann-Whitney U test was used to investigate the difference in agreement with the statements regarding teachers specialities between science (S) and non-science teachers (NS). Furthermore, the group was split into those who perceive their level of digital competence to be between low (low DC) and high (high DC).
First respondents were asked about their perceived level of digital competence according to the levels described in the DigCompuEdu framework[7]. Teachers were also asked about their level of satisfaction and which specific competencies they would like to improve. Secondly, the respondents were asked about their perceived level of students digital competence according to the levels described in the DigComp framework[8], both in general and for specific competencies. The last section focused on respondents viewpoint towards AI.
The results, see figure 1, show a general agreement with all the statements. Between the answers from the teachers competencies levels is no statistically difference. However, there is a statistical difference in agreement between science and non-science teachers. Science teachers are more of the opinion that the basics of AI and its concrete use should be learned by all teachers (ID 1 and 2) as well, that that students should be taught how to use AI (ID 7). Nevertheless, the difference between

---

[7]https://joint-research-centre.ec.europa.eu/digcompedu_en
[8]https://joint-research-centre.ec.europa.eu/digcomp/digcomp-framework_en

the teachers specializations does not emerge as an important factor.

- *Will*: There was significant interest among teachers to learn more about AI and integrate the topic into their classes. Teachers agreed on the importance of understanding the impacts of AI in daily life and developing critical views about AI. They slightly disagreed on the necessity of teaching basic AI concepts in schools and mostly disagreed that AI should be a separate topic.

- *Skill*: Most teachers had heard of terms such as machine learning or strong AI, but have to be supported to develop digital competencies, as their lack can hinder students digital competence development (Nascimbeni and Vosloo, 2019). Correspondingly, the StMUK, for example, offers a wide range of training courses on AI, chatbots, deep learning and learning analytics via ALP Dillingen[9] as well as universities offer further training courses like ProfiLehre[10] by University of Augsburg.

- *Tool*: In the survey, little attention was paid to the tool factor, which is why further research is needed to make a statement. Respondents did not report significant issues regarding the availability of resources.

3 ChatGPT for Learners

ChatGPT is already frequently used by students. The study by (Sidoti and Gottfried, 2023) shows that 1 in 5 U.S. teens who have heard of ChatGPT have used it for schoolwork and also for many different study tasks according to (DZM, 2023a).

3.1 Students in Schools

In an exemplary English lesson, students should practise the use of the Past Tense Simple with the help of ChatGPT via the exercise "Amazing Persons".(König et al., 2023) Accordingly, each pupil chooses a famous person, finds out what makes them special and summarises this in form of an essay, where the checklist in figure 2 provides guidance. During the lesson, the students will learn that to instruct ChatGPT, reviewing as well as revising the answers are their responsibility:(König et al., 2023)

1. *Prompting*: Students should recognize what kind of information ChatGPT needs to generate relevant answers, to formulate this information clearly and precisely, and to make consecutive adjustments.

2. *Reviewing*: To revise the text in a familiar environment alongside the conversation with ChatGPT, a simple procedure is to copy the answer into a text editor.

3. *Reflecting*: Students have to combine their written text with the generated results from ChatGPT to a complete coherent essay.

4. *Referencing*: Students have to account transparently for the use of ChatGPT, for example by attaching the transcription of the conversation and marking generated text passages.

In addition, students must also be aware of the ethical context and new challenges:(Nuxoll, 2023)

1. *Neutrality*: Students can learn about bias, for example, in a German class. Therefore, they write about an everyday situation from a male and female perspective (König et al., 2023). Afterwards, ChatGPT solves the task, too. The students will learn that ChatGPT reproduces gender stereotypes.

2. *Misinformation*: Students need to be able to assess information for correctness and completeness and the quality of sources. Via "Group Puzzle" (König et al., 2023) the students get split up into three research teams: "Team Books", "Team ChatGPT" and "Team Search Engine". In the first phase, the groups each work in their assigned research strategy. Afterwards, the results are getting compared with each other.

3. *Disinformation*: Students need to be aware that the advancement of AI can amplify fake news and disinformation in order to become responsible and informed digital citizens.

4. *Inequalities*: Students in regions or households that do not have the necessary infrastructure or technical know-how likely be at a disadvantage compared to their better-equipped peers.

---

[9]https://alp.dillingen.de/themenseiten/unterr

icht-ki

[10]https://www.uni-augsburg.de/de/organisation/einrichtungen/qa/profilehre/

ID	Statement	M (SD)	S vs NS	Low vs High DC
1	Artificial Intelligence is a topic that should be taught in my school	3.75 (1.06)	U= 1533, p = 0.08	U= 2329, p = 0.76
2	Teachers/educators in my school should learn the basic concepts behind AI technology	3.91 (1.05)	**U= 1410, p = 0.02**	U= 2597, p = 0.12
3	Teachers/educators in my school should learn how to actually use AI technology	3.87 (1.10)	**U= 1480, p = 0.04**	U= 2504, p = 0.26
4	Teachers/educators in my school should learn about the impacts of AI in everyday life	4.00 (1.04)	U= 1492, p = 0.05	U= 2562, p = 0.16
5	Teachers/educators in my school should develop a personal critical view on AI	3.95 (1.09)	U= 1493, p = 0.05	U= 2542, p = 0.19
6	Students in my school should learn the basic concepts behind AI technology	3.83 (1.08)	U= 1485, p = 0.05	U= 2476, p = 0.32
7	Students in my school should learn how to actually use AI technology	3.83 (1.07)	**U= 1407, p = 0.02**	U= 2506, p = 0.26
8	Students in my school should learn about the impacts of AI in everyday life	3.93 (1.05)	U= 1520, p = 0.07	U= 2547, p = 0.19
9	Students in my school should develop a personal critical view on AI	3.92 (1.07)	U= 1500, p = 0.06	U= 2641, p = 0.08

Figure 1: Ratings on Likert statements about teachers' attitude towards AI Education and comparisons between Science (S, N=39) and non-Science teachers (NS, N=96) and between teachers with reported low (low DC, N= 73) and high (high DC, N= 62) Digital Competences (© Polak et al., 2022)

Figure 2: Checklist for students (© Friedrich Verlag, 2023)

5. *Dependence*: It is important that students do their homework as the main author and not simply hand them over to ChatGPT. It should be understood as a tool that can be used to support homework and learning.

3.2   Students in Universities

Build on top of the knowledge from school, conversational agents are valuable tools for university students, aiding with academic work, offering accessibility, improving critical thinking and language skills (Atlas, 2023). Given the back-

ground that DMZ speaks about "Prompt Crafting" as an important future skill, students should experiment with structured approaches like CREATE by (Berrett, 2023) and frameworks such as CONTEXT, CRISPE, PROSE, CLEAR, STAR or APPROACH (Helaluzzaman, 2023) to exploit the full potential of ChatGPT.(DZM, 2023a) The THA highlights two frameworks for their students:(DZM, 2023b)

- ChatGPT Mega-Prompt by Rob Lennon is a six-step prompt structure, which places Chat-GPT in a concrete scenario, as shown in figure 3, instead of a simple instruction.
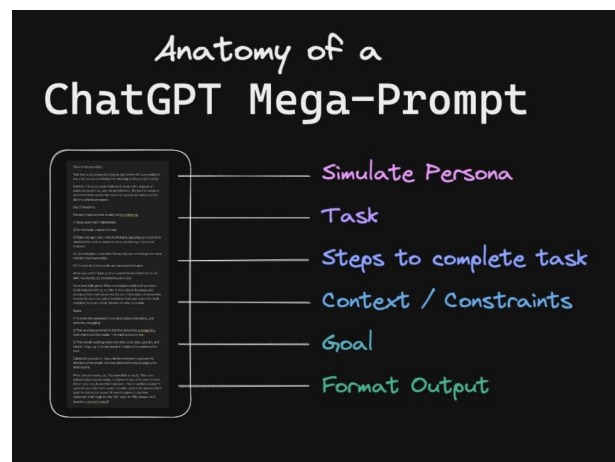

Figure 3: Setup of a ChatGPT Mega-Prompt (© Rob Lennon 2023, @thatroblennon/Twitter).

- By (?) a distinction is made between four objectives: learning, chatting, data transformation and text generation. Each objective needs a different approach for an effective prompt. For example, regarding objective learning, students should use the approach "starting broad then zooming in"; see figure 4.

In view of the fact that the context and structured input contribute significantly to the output, Ope-
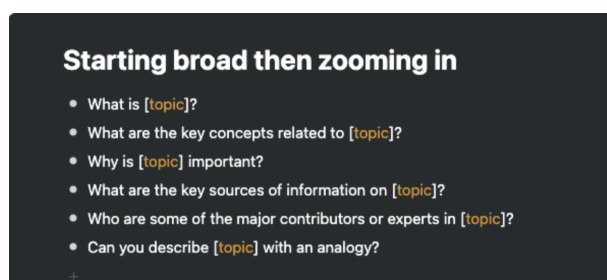
Figure 4: Chain prompts via least-to-most tactic for the goal learning (© Workflowy, 2023)

nAI released *Custom Instructions*[11] and *GPTs*[12] to support users with a better user experience. Additionally, University of Hohenheim (Gimpel et al., 2023) recommends students to consider et al. the following while working with ChatGPT:

1. *Review the university's rules and regulations regarding generative AI like ChatGPT*: Students must follow the rules regarding quotations and good scientific practice. It is currently advisable to ask the lecturer directly whether and how ChatGPT should be used. It is also recommended to check whether there are relevant examination regulations of the university.

2. *Understand the capabilities and limitations of ChatGPT*: Students should be aware, for example, that a summary by ChatGPT can always miss details and that these details might be important. ChatGPT can and should not replace critical thinking and creativity, which are essential components of writing assignments. Therefore, students should only use ChatGPT as a supplementary tool such as translation programs like DeepL (Spannagel, 2023b).

3. *Check if the use of ChatGPT is clever or if the task requires the learning of basic knowledge*: By reflecting on the desired learning outcomes, students can determine how to integrate ChatGPT into their education best. The answers that ChatGPT provides are not verifiable, so the students are still required to add further knowledge. If the result is unsatisfactory, students must be able to provide detailed feedback to improve the response (Atlas, 2023).

---

[11] https://openai.com/blog/custom-instructions-for-chatgpt
[12] https://openai.com/blog/introducing-gpts

4. *Verify that the results given by ChatGPT are trustworthy and accurate (Atlas, 2023)*: Students should reflect on each generated outcome and are responsible for their use of tools.(Rademacher, 2023) Currently, there are issues with accurate citation of generated sources. ChatGPT creates non-existent sources and false information, as well as increases substantial risks relating to accidental plagiarism.

5. *Consider which topics could be cleverly linked to produce novel insights*: ChatGPT is accessible 24/7 which provides students with new opportunities to help them acquire new knowledge or test existing knowledge for exam preparation like discussed in the presentation by (DZM, 2023a). Figure 5 lists some exemplary prompts.

Activity to be supported	Exemplary prompts
Generation of pattern solutions from old exam tasks	"Please generate a pattern solution for the following task: **[past your task]**"
Create a mind map to gain a quick overview of a new topic	**[paste your topic]** "Create a mind map on the topic above, list the central idea, main branches, and sub-branches"
Explanation of concepts (e.g., mathematical equations)	"I want you to act as a math teacher. I will provide some mathematical equations or concepts, and it will be your job to explain them in easy-to understand terms. My first question is: I need help to understand how **[paste your concept]** works"
Vocabulary acquisition	"Could you please provide me with terms related to **[paste your text]**"
Create flashcards	"Topic: **[paste your topic]** Please help me create a two-column spreadsheet with questions and corresponding answers on the topic above"
Self-testing of specific knowledge	"Topic: **[paste your topic]** Please ask me five questions on the topic above. I will then respond to it. After my response, you will tell me if my answer was right or wrong and provide an explanation"

Figure 5: Exemplary prompts on learning with ChatGPT (© University of Hohenheim, 2023)

lea

4 ChatGPT for Lecturers

The emergence of new educational technologies often arouses strong emotions, ranging from doomsday predictions to endless euphoria (Rudolph et al., 2023). In the case of ChatGPT, experts also speak of an "educator's dilemma" between banning these technologies or promoting their use.(Teubner et al., 2023) A school or universitAy without lecturers will continue to be a bad school or university - but so will a school or university without AI in the future. To do so, there must be room for experimentation and innovation to find out how best to integrate AI.

4.1 Teaching

It is still important to teach basic knowledge, as people operate ChatGPT and are responsible for further processing. However, ChatGPT helps students offloading basic tasks to be able to focus on more specific tasks, like a calculator in mathematics. Therefore, lecturers are able to ask more complex questions, which will raise the general level of learning. To motivate the students to learn basic knowledge, lectures can argue with the self-determination theory by (Ryan and Deci, 2000). Competence, autonomy and relatedness are imported factors for human beings to feel in control of their own lives and experience a sense of belonging with other people. In the context of ChatGPT, this would mean that students must be able to solve a task together in a team without ChatGPT. In concrete terms for teaching, this means that teaching must change as we know it. One idea is the concept of "flipped classroom"6. Using the flipped
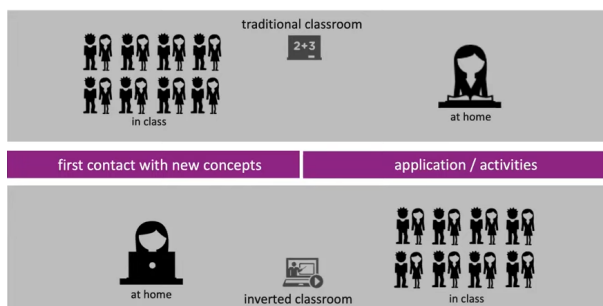

Figure 6: Flipped classroom concept (© Spannagel, 2023)

classroom concept as an example, at level remembering and understanding based on the revision of Bloom's Taxonomy of Educational Objectives by (Hanke, 2023a), students should use ChatGPT beside other sources to learn and understand new concepts outside the classroom. During the class, students should apply their know-how and evaluate the results without ChatGPT. This shows that traditional teaching situations will no longer optimally support students in their learning phases in the future, as the tasks for the use of AI can be used in the early phases instead of frontal teaching.
Beside basic knowledge, the specialist skills will change, too. Skills like media literacy, computer science, critical thinking and prompting are getting more important in the future.(Spannagel, 2023a) Therefore, teachers must reflect on which learning objective they are pursuing in their teaching. Lectures should also encourage students to use Chat-

GPT and teaching students how to effectively use these tools, to equip them with important skills for their future careers, while also emphasizing the importance of academic integrity.
Beside the pedagogical part, ChatGPT can create custom exercises and quizzes, offer feedback and generate tailored educational materials. Moreover, ChatGPT can assist in developing lecture ideas and lesson plans. OpenAI released, for example, a free guide for teachers on how to use ChatGPT in the classroom including suggested prompts.[13]

4.2 Assessment

Some lecturers are concerned that written assignments are being outsourced to ChatGPT without being detected by a plagiarism detector. Universities should therefore focus on ensuring responsible use by addressing potential challenges related to ChatGPT rather then implementing policies that restrict use.(Gimpel et al., 2023). On the other hand it needs rules for the use of AI tools like ChatGPT, otherwise students would have unclear conditions, which leads to more mistakes and possible serious consequences. Copying generated text into a self-written document carries the risk of plagiarism and copyright infringement. Universities should therefore implement, with the right measure, an AI policy like the one by (Spannagel, 2023b) "Rules for Tools" to inform their students about the risks and their responsibility, as well as examples how to cite ChatGPT correctly. Current detector tools are far from perfect, but will improve in the future. More on that topic in the section *Detecting ChatGPT-generated Academic Writing*.
Closed book exams without access to the internet might be the easiest way to test the students knowledge in the future. According to (Hanke, 2023b) classic paper-pencil test are still very effective to test the understanding of the topic. However, during Covid-19 pandemic open exams and take at home exams become necessary and popular. When creating open exams where technical aids are also allowed, care should be taken to prevent student misbehavior (Koenders and Prins, 2023):

- *Declaration of how ChatGPT was used*: Monitoring which tools are used in a take home exam is impossible and not all students might be truthful. For this purpose the Cambridge University Press has already implemented

---

[13]https://openai.com/blog/teaching-with-ai?ref=viden.ai

guidelines that require authors to disclose their use of AI-based tools when writing articles. For example, lectures might require students to provide a full transcript of the conversation. (Pardey, 2023).

- *Innovate assessment formats*: (**?**) suggest steering clear of standardized assessments, that could quickly be completed by a computer. Instead lectures should design assessments that promote students ability to think creatively and critically, as well as require students to reflect on their learning. Additionally, encourage oral presentations to assess students public speaking skills and their understanding of the material (McCormack, 2023).

- *Supervision Process*: (Hanke, 2023b) suggest requiring additional information in form of a report, on work steps during the process from students. For example, results of the students analyses and quality in form of process report and solutions in the form of a reflection report.

- *Evaluation Criteria*: Students can assess the content of the text while placing less emphasis on the structure and writing style, which are areas where ChatGPT tends to excel. (Gimpel et al., 2023) considers to focus more on evaluation criteria such as quality and individuality of the research questions, correct references, contributions that go beyond a summary of the literature, and a personal statement are some of the criteria that lecturers should pay more attention to.

5 Detecting ChatGPT-generated Academic Writing

The usage of ChatGPT for academic writing started a significant amount of discussions and concerns in the community regarding copyright and plagiarism. Detecting LLM-generated content (=LLM content) is still in an early stage of exploration. The study from (Liu et al., 2023) present investigations of detecting LLM content specifically to academic writing, which we will have a closer look at below.

5.1 GPABenchmark Dataset

The corpora for ChatGPT text classification mainly focus on question and answering (=Q&A) dialogues. ChatGPT is designed as a chat interface, but the usage scenarios of ChatGPT have significantly expanded beyond that. Furthermore, human generated conversations often contain subjective, personal and emotional opinions. In contrast, academic writing is more formal and objective, as well meet higher standards for fluency, clarity and grammatical correctness as well as a formal organized style. Accordingly, academic writing mostly refers to highly specialized topics which lead to different term distribution from conversational dialogues. The unique characteristics of academic writing lead to a new ChatGPT-generated corpus, which is necessary to benchmark GPT detectors, namely the GPABenchmark.

There are three main scenarios where ChatGPT is used/misused in academic writing. In summary GPABenchmark has 150000 samples in each category:

1. *GPT-written full abstract (GPT-WRI/WRI)*: The author will paste a title to ChatGPT and asks it to write an abstract from scratch. A sample prompt is: *Please generate an abstract for a research paper titled "XYZ"*.

2. *GPT-completed abstract (GPT-CPL/CPL)*: The author writes a few sentences to ChatGPT asks to complete the rest of the paragraph. A sample prompt is: *Please complete the abstract with X words*.

3. *GPT-polished abstract (GPT-POL/POL)*: The author pastes the entire abstract to ChatGPT of polishing. A sample prompt is: *This is an abstract of a research paper. Please rewrite for clarity*.

4. *Human written (HUM)*: Abstracts that are completely written by a human author.

5.2 Detecting GPT-Generated content by Humans and ChatGPT Detectors

The study benchmarked open source and commercial detectors as well as the newly developed CheckGPT detector. Additionally, a survey was made to see how good humans are in detecting LLM-content in academic writing. Figure 7 shows the evaluation of open source and commercial ChatGPT detectors. The accuracy for GPT-generated abstract appears to be very low via GPTZero. However, GPTZero demonstrates very high accuracy for human written abstract. ZeroGPT detects fully

GPT-written abstract fairly accurate. However, accuracy for GPT-completed or GPT-polished abstracts appears to be very low. The classifier from OpenAI detects GPT-generated content slightly better than GPTZero and ZeroGPT. Overall existing open source and commercial GPT detectors provide unsatisfactory performance on GPABenchmark, especially for GPT-polished text. CheckGPT on the other side shows very promising results. According to figure 8 CheckGPT achieve an average classification accuracy of 98% to 99%.

	T1. GPT-WRI			T2. GPT-CPL			T3. GPT-POL		
	CS	PHX	HSS	CS	PHX	HSS	CS	PHX	HSS
(a) Classification accuracy (in %) of GPTZero.									
GPT	30.3	25.3	72.0	17.0	6.0	43.7	1.7	2.3	20.3
Human	99.3	99.7	100	99.7	99.7	94.3	99.7	95.7	95.7
(b1) Detection accuracy (in %) of ZeroGPT									
GPT	67.4	68.4	92.3	25.3	10	62.4	3.3	2.7	24.7
Human	100	98.4	95	99.7	99.7	94.7	98.3	98.6	92.7
(b2) Average score reported by ZeroGPT. 0:human, 8:GPT									
GPT	5.43	5.39	7.41	2.26	0.97	4.97	0.35	0.29	2.15
Human	0.09	0.13	0.52	0.08	0.04	0.47	0.20	0.14	0.64
(c.1) Detection accuracy (in %) of OpenAI's detector									
GPT	80.7	70	63	63.7	23.7	27.3	6.3	4.3	6
Human	51.0	69.7	84.0	35.3	59.7	79.6	50.7	69.0	88.0
(c.2) Average score reported by OpenAI. 0:human, 4:GPT									
GPT	3.11	2.89	2.72	2.70	2.12	2.04	1.75	1.59	1.52
Human	1.42	1.17	0.59	1.71	1.35	0.68	1.38	1.14	0.52

Figure 7: Performance of open-source and commercial GPT detectors. Values in red: detection accuracy <50%, or average score on the wrong side of the decision threshold. (© Zeyan et al., 2023)

	T1. GPT-WRI			T2. GPT-CPL			T3. GPT-POL		
	CS	PHX	HSS	CS	PHX	HSS	CS	PHX	HSS
TPR	99.96	100.0	99.94	99.21	97.71	98.58	98.43	98.93	98.45
TNR	99.98	99.99	99.92	99.43	98.14	99.00	98.37	99.22	98.49
Acc	99.97	99.99	99.93	99.32	97.93	98.79	98.39	99.08	98.47

Figure 8: CheckGPT's classification accuracy (in %) for each task and each discipline: TP rate, TN rate, overall accuracy. (© Zeyan et al., 2023)

In addition to GPT-detectors a user study showed, how good humans are in detecting GPT-generated content in academical writing. The results show, that humans are slightly better than random guesses and that it is extremely challenging for human users to distinguish between human written and GPT-generated paper abstract. Accordingly, the result confirms the public opinion that ChatGPT achieves human like language style and quality.

6 Conclusion

[RQ1] The use of ChatGPT in educational institutions, both at school and university level, is possible. Legal problems arise from the lack of compliance with the GDPR when using the web application and age restrictions limit its use. However, OpenAI's release of a GDPR and DPA compliant usage via API and Enterprise offers solutions for institutions. In schools and universities, there are still a number of hurdles that need to be overcome before teachers and students have equal access and receive information on consistent rules for AI tools. Until educational institutions introduce technical offerings, teachers have to use their personal access and may share it with the students and/or students need their own access. The digital divide remains an issue as access to the technology is unequal and the capabilities of the free and paid versions of ChatGPT are different. The development of digital competencies among teachers is very important as their current knowledge gap can hinder students digital competence development. Fortunately, there are already great training opportunities.

[RQ2] Practical integration of ChatGPT varies between school and university settings. School students require foundational skills and ethical awareness regarding ChatGPT use. These competencies can be learned in a hands-on experience during class, which structured guidelines and hands-on activities like "amazing persons" and "group puzzle" have proven. University students should reflect on their learning goals to determine how to integrate ChatGPT into their education as a supplement. Most importantly, students must inform themselves how to integrate ChatGPT in their academic work properly, to comply with guidelines and rules for AI tools. Generally for all students, prompt competences is an important future skill, which students can practice by using ChatGPT as a writing or learning partner. There are many sources with prompting best practices like frameworks and examples that provide support.

Recommendations for educators focus on leveraging ChatGPT for curriculum development, assessment modifications, and fostering academic integrity. As ChatGPT excels in writing and offloads work for student students, lecturers are able to ask more complex questions, which raises the general level of learning. New possibilities to learn new learning materials at home with AI, will change teaching as we know it today. One idea is the con-

cept of "flipped classroom" which focuses on an interactive application in the classroom instead of traditional frontal teaching. Additionally, ChatGPT can help lectures by creating exercises and developing lecture ideas and lesson plans. Assessments should be expanded by requesting students to reflect on their learning in writing or orally. Most important, in order to create the required conditions for the use of ChatGPT, teachers must introduce guidelines and rules how ChatGPT can be used by students.

[RQ3] A lack of knowledge can lead to misuse and learners may choose to use ChatGPT's quick answers instead of their own investigations and solutions, which leads to risks like plagiarism. Unfortunately, detecting GPT-generated content is quite challenging for many tools like GPTZero, ZeroGPT, OpenAI classifier and mostly for humans. Therefore, ideas exist to implement something like watermarks to help to detect GPT-generated content faster in the future. Another important finding is, that ChatGPT is build for a conversation, which differs from academic writing in several characteristics. For example, human generated conversations often contain subjective personal opinions while academic writing is more formal and objective. New tools like CheckGPT are aware of that knowledge and manage to detect GPT-generated content with a high accuracy (figure 8).

7 Discussion

Despite the many challenges, it is clear that there is not only a lot of movement in the development of ChatGPT, but also that many ideas and structures have emerged in the field of education. Likewise, the challenges mentioned are all already known and can be solved. In my opinion, Germany is still falling behind in the technical modernisation of the education system. The question I personally ask myself is how we as a society will manage to keep up with the pace of digitalisation in the future.

One major challenge is the digital divide, where not all students have access to the necessary infrastructure, devices and applications due to the legacy of digitalisation. Also, ensuring the privacy and security of students data in a digital learning environment is crucial. Especially a lot of software and infrastructure from commercial resources is currently being used like Microsoft, Zoom and of course ChatGPT. These challenges hinder the adoption of digital tools in education massively and make the most important infrastructure dependent on large companies.

To keep up, Germany needs to invest in expanding and improving its digital infrastructure that connects students, teachers, parents, institutions and digital learning resources. To make this infrastructure safely usable and tailored to its use cases and legal framework conditions, more open source software for self-hosting should be supported and further developed. There are many projects, for example, like the LLM Bloom [14], collaboration and messaging app like Element [15], video conferencing app like Jitsi [16], learning platform like Moodle [17] and data infrastructure initiative Gaia-X [18], that we can and should put into practice.

At the end, all factors have to come together to be able to reach an independent modern education system, which supports lifelong learning and accessible tools for all, what I believe is part of the answer to keeping up with the pace of digitalisation.

References

Stephen Atlas. 2023. Chatgpt for higher education and professional development: A guide to conversational ai. https://digitalcommons.uri.edu/cba_facpubs/548 [Retrieved on October 20, 2023].

Tom Berrett. 2023. Uplevel your prompt craft in chatgpt with the create framework - tom barrett. https://edte.ch/blog/2023/01/22/create-framework/?v=3a1ed7090bfa [Retrieved on October 20, 2023].

DigiLLAb. 2023. Ki-systeme wie chatgpt in der hochschullehre (faq). https://digillab.uni-augsburg.de/ki-systeme-wie-chatgpt-in-der-hochschullehre-faq/ [Retrieved on December 01, 2023].

DZM. 2023a. Chatgpt in der hochschullandschaft. https://www.youtube.com/watch?v=jTQXafSBWAo [Retrieved on December 01, 2023].

DZM. 2023b. Chatgpt-promptkompetenz. https://www.tha.de/ChatGPT-Promptkompetenz.html [Accessed: on December 01, 2023].

DZM. 2023c. Ki in der lehre. https://www.tha.de/KI-in-der-Lehre.html [Retrieved on December 01, 2023].

---

[14] https://en.wikipedia.org/wiki/BLOOM_(language_model)
[15] https://en.wikipedia.org/wiki/Element_(software)
[16] https://en.wikipedia.org/wiki/Jitsi
[17] https://en.wikipedia.org/wiki/Moodle
[18] https://en.wikipedia.org/wiki/Gaia-X

Henner Gimpel, Kristina Hall, Stefan Decker, and Steffen Vandirk. 2023. Unlocking the power of generative ai models and systems such as gpt-4 and chatgpt for higher education a... *ResearchGate*.

U. Hanke. 2023a. Lernen und prüfen in einer welt mit chatgpt mit hilfe der lernzieltaxonomie. `https://hochschuldidaktik-online.de/wp-content/uploads/sites/3/2023/02/Lernen-Pruefen-mitChatGPT-Lernzieltaxonomie_neu.pdf` [Retrieved on December 01, 2023].

U. Hanke. 2023b. Lernen und prüfen in einer welt mit chatgpt mit hilfe der lernzieltaxonomie. `https://www.youtube.com/watch?v=re5j1I6jHTE` [Retrieved on December 01, 2023].

A. Helaluzzaman. 2023. 7 chatgpt prompt frameworks for better output, creativity and productivity. `https://datayon.substack.com/p/7-chatgpt-prompt-frameworks` [Retrieved on December 01, 2023].

L. Koenders and F. Prins. 2023. Chatgpt in education: can you still use take-home exams and essays? `https://www.uu.nl/en/education/educational-development-training/knowledge-dossier/the-influence-of-chatgpt-on-assessment-can-you-still-use-take-home-exams-and-essays` [Retrieved on October 29, 2023].

A.(Hrsg.) König, D. Thiede, J. Mosbach, A. Bankhofer, C. Potthoff, J. Brüggemann, E. Gutschlag, F. Hanff, E. Müller, S. Wilsmann, and T. Kantereit. 2023. Künstliche intelligenz: Wie chatbots & co. den unterricht verändern. *Friedrich Verlag GmbH*.

Zeyan Liu, Zijun Yao, Fengjun Li, and Bin Luo. 2023. Check me if you can: Detecting chatgpt-generated academic writing using checkgpt. *arXiv (Cornell University)*.

A. Maynard. 2023. Chatgpt enterprise could be a game changer for universities. `https://futureofbeinghuman.com/p/chatgpt-enterprise-game-changer` [Retrieved on December 01, 2023].

G. McCormack. 2023. Chat gpt is here! – 5 alternative ways to assess your class! `https://gavinmccormack.com.au/chat-gpt-is-here-5-alternative-ways-to-assess-your-class/` [Retrieved on October 29, 2023].

Fabio Nascimbeni and Steven Vosloo. 2019. Digital literacy for children: Exploring definitions and frameworks. *ResearchGate*.

F. Nuxoll. 2023. Ki in der schule. `https://www.bpb.de/shop/zeitschriften/apuz/kuenstliche-intelligenz-2023/541500/ki-in-der-schule/` [Retrieved on December 01, 2023].

C. Pardey. 2023. Renommierter britischer verlag regelt umgang mit chatgpt. `https://www.forschung-und-lehre.de/zeitfragen/renommierter-britischer-verlag-regelt-umgang-mit-chatgpt-5473?` [Retrieved on October 29, 2023].

Dominik Petko. 2012. Teachers' pedagogical beliefs and their use of digital media in classrooms: sharpening the focus of the 'will, skill, tool' model and integrating teachers' constructivist orientations. *Computers & Education*, 58(4):1351–1359.

Sara Polak, Gianluca Schiavo, and Massimo Zancanaro. 2022. Teachers' perspective on artificial intelligence education: An initial investigation. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*.

M. Rademacher. 2023. Warum chatgpt nicht das ende des akademischen schreibens bedeutet. `https://digiethics.org/2023/01/03/warum-chatgpt-nicht-das-ende-des-akademischen-schreibens-bedeutet/` [Retrieved on October 29, 2023].

J. Rudolph, S. Tan, and Tan S. 2023. Chatgpt: Bullshit spewer or the end of traditional assessments in higher education? *Journal of applied learning and teaching*, 6(1).

Richard M. Ryan and Edward L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78.

O. Sidoti and J. Gottfried. 2023. About 1 in 5 u.s. teens who've heard of chatgpt have used it for schoolwork. `https://www.pewresearch.org/short-reads/2023/11/16/about-1-in-5-us-teens-whove-heard-of-chatgpt-have-used-it-for-schoolwork/` [Retrieved on December 01, 2023].

C. Spannagel. 2023a. Chatgpt & co. in der hochschullehre. `https://www.youtube.com/watch?v=aM6fZuH1cGw` [Retrieved on December 01, 2023].

C. Spannagel. 2023b. Chatgpt und die zukunft des lernens: Evolution statt revolution. `https://hochschulforumdigitalisierung.de/blog/chatgpt-und-die-zukunft-des-lernens-evolution-statt-revolution/` [Retrieved on October 29, 2023].

C. Spannagel. 2023c. Rules for tools. `https://csp.uber.space/phhd/rulesfortools.pdf` [Retrieved on December 01, 2023].

StMUK. 2023. Arbeiten mit künstlicher intelligenz. `https://www.km.bayern.de/schule-digital/orientierungsrahmen-ki-und-schule/arbeiten-mit-kuenstlicher-intelligenz.html` [Retrieved on December 01, 2023].

Timm Teubner, Christoph M. Flath, Christof Weinhardt, Wil Van Der Aalst, and Oliver Hinz. 2023. Welcome to the era of chatgpt et al. *Business & Information Systems Engineering*, 65(2):95–101.

List of Figures