

Using ChatGPT in Higher Education – an Assessment of the AIDAHO Project Team¹

Jens Vogelgesang, Johannes Bleher, Christian Krupitzer, Anthony Stein, Robert Jung

ChatGPT has caused a lot of (public) attention in a short time since its release. This text aims to provide a first assessment and better understanding of the opportunities and risks of using ChatGPT in the higher education context. The remarkable dynamics of the current discussion around ChatGPT has resulted in the need for multiple updates to this text during its writing. The current version of the text is dated February 7, 2023.²

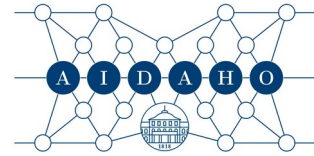
Functioning

The software prototype named [ChatGPT](#), released by the US company [OpenAI](#) on November 30, 2022, is an AI-based dialogue system. This dialogue system is based on a Large Language Model (LLM) called GPT-3 version 3.5. GPT (Generative Pre-trained Transformer) are deep neural network architectures of the “Transformer” type, which were proposed by [Vaswani et al.](#) in 2017. They belong to the class of generative models because they are capable of automatically generating new text passages, i.e., attaching the next most likely word to a word sequence. This way, text summaries, scientific definitions, poems in the style of Rainer-Maria Rilke, and program code in various computer languages can be generated. Since the texts generated by probabilistic language models such as GPT-3 involve chance, each text is unique. Nevertheless, these generated texts are based on previously learned patterns extracted from a huge number of human-generated texts, known as training data. This makes the arrangement of words plausible and fitting in terms of theme, so that the texts generated by ChatGPT do not seem to differ from those written by humans.

The ability of GPT-3 to generate different types of text that seem meaningful is based on one hand on a very large corpus of training texts, on which the AI is trained, and on the other hand – according to OpenAI – on a deep neural network model with 175 billion parameters, which are adjusted through machine learning. Regarding the number of parameters, GPT-3 is currently far from being the largest language model, as shown by the chart from the 2022 AI Association in the appendix.

¹ The Artificial Intelligence and Data Science Certificate Hohenheim (AIDAHO) is a continuing education program funded by the BMBF and supported by the federal state of Baden-Württemberg. The certificate program is offered by the authors of this text to Bachelor students at the University of Hohenheim since the winter term 22/23.

² This text has been translated from German using ChatGPT and manually checked afterwards.



The functionality of ChatGPT appears limitless. If you ask the chatbot, for example, a simple arithmetic problem, you will usually get a correct answer. Given this, you might assume that ChatGPT can perform all arithmetic operations, but this is currently not the case. As long as ChatGPT is “only” based on a language model, it can only correctly solve mathematical problems that were part of the training material, such as those provided by universities on public pages to students. A simple problem like $2 + 2 = 4$ can be solved by ChatGPT because this information was included in the training data. However, ChatGPT can sometimes fail to multiply decimal numbers with several decimal places (which would not be noticeable without checking). This weakness of ChatGPT can already be compensated for today by using a complementary semantic search engine called [Wolfram|Alpha](#), as indicated in the blog by search engine founder Stephen Wolfram with the post "[Wolfram|Alpha as the Way to Bring Computational Knowledge Superpowers to ChatGPT.](#)"

The special thing about ChatGPT is that, unlike using the underlying GPT-3 language model, no special knowledge of the Application Programming Interface (API) is now required. ChatGPT is now available as a dialog system software. Through the dialogic interaction with the chatbot, you can concretize your own inputs and “together” iteratively improve the solution proposed by the AI system.

So far, the use of ChatGPT was freely available after registration. However, the strong demand has now led to the ChatGPT servers no longer being able to keep up with all input requests in real time. A premium version, which can apparently be acquired for \$20 soon according to recent media reports, is supposed to limit access to the program and continue to guarantee high response speeds.

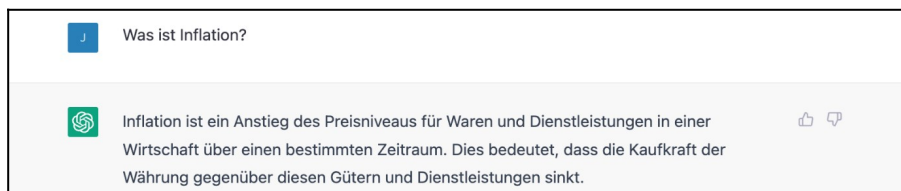
The input request from ChatGPT is very easy to use, as the following screenshot shows:



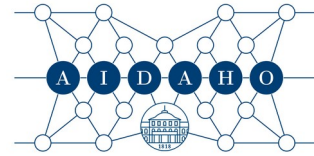
ChatGPT Jan 30 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

Source: <https://chat.openai.com/chat>

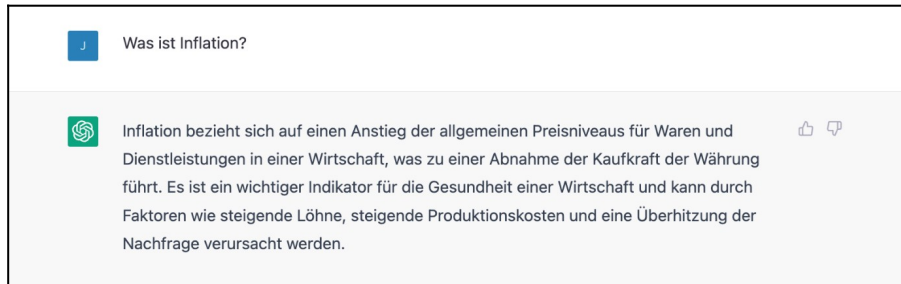
After entering the question ‘What is inflation?’, one receives the following answer:



ChatGPT screenshot from February 2nd, 2023



The repeated input of the question “What is inflation?” illustrates the stochastic functioning of the ChatGPT underlying language model: thus, for most inputs (prompts), the same answer is never given twice.



ChatGPT screenshot from February 2nd, 2023

This example also shows the enormous dynamics of the development of ChatGPT. Only a few weeks ago, the system was only able to respond adequately to English inputs.

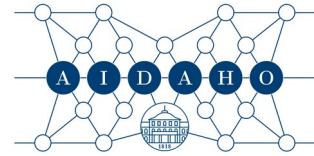
However, ChatGPT does not continue to learn and is not “connected” to the Internet, as OpenAI writes in its own FAQ. This means that after training the underlying GPT-3.5 model, ChatGPT can generate very little plausible text about events after 2021 - such as the war in Ukraine. However, it is assumed that the inputs of the users, along with the evaluations for the generated texts in the form of “thumb up and down” votes, will be used for later versions of the chatbot.

Due to the combination of performance, ease of access, and user-friendly dialog format, KI expert Doris Weßels from the Kiel University of Applied Sciences referred to the ChatGPT software as a “[milestone in KI development](#)” in the January issue of *Forschung & Lehre*. Whether a milestone or not, we can no longer avoid starting to discuss the implications of using ChatGPT in teaching and research at the universities.

ChatGPT & Technology Panics

Before discussing the opportunities and risks of using ChatGPT in teaching and research at the university, here are some preliminary remarks:

The history of media shows that fears and concerns are often expressed when new technologies or computer software are introduced. In this context, communication researcher Amy Orben from the University of Oxford writes in “[Sisyphean Cycle of Technology Panics](#)”. This cycle always begins with great fears about negative consequences when a new technology gains societal significance.



For example, in 2000, the Guardian warned in the article “[Cheats are scaring us stiff](#)” of a “plagiarism epidemic”: “Is the work I see in front of me the work of a student, or just a pre-made, chargeable download from the internet? This is a question that is beginning to concern professors and will not disappear.” In 2014, The Atlantic asked with concern: “[Is Google Making Students Stupid?](#)”

The “plagiarism epidemic” did not materialize. Students are not made stupid by Google either. There were also appropriate warnings at the time of the spread of Wikipedia, which also sparked debates about its use in teaching and research. However, Amy Orben’s article also makes it clear that the almost ritualistically expressed drastic fears when introducing new media technologies never actually occur in the form feared.

Our assessment of the opportunities and risks of using ChatGPT is based on the awareness that we are in the first phase of a cycle of fears and concerns, given the widespread penetration of AI technology in society. This phase is strongly characterized by anecdotal evidence of what ChatGPT is capable of:

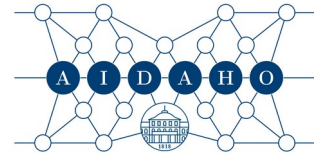
- In December, the Wall Street Journal columnist Joanna Stern proclaimed: “[ChatGPT Wrote My AP English Essay—and I Passed](#)”.
- At the same time, the New York Times predicted: “[A New Chat Bot Is a ‘Code Red’ for Google’s Search Business](#)”.
- In late January, CNN Business headlined “[ChatGPT Passes Exams from Law and Business Schools](#)”, after ChatGPT participated as an unknown candidate in four module exams at the University of Minnesota Law School and received a “C+” for its answers.

We will now discuss the opportunities and risks of using ChatGPT in higher education.

Teaching

As outlined in the previous chapter, the availability of ChatGPT for a wide user base presents new challenges for academic and university education.

The following two identical questions we asked ChatGPT illustrate how problematic the use of ChatGPT can be in scientific work:



J Nenne die drei wichtigsten deutschsprachigen wissenschaftlichen Werke, die den Begriff der Inflation erklären.



1. "Makroökonomie" von Walter Eucken (1940) - Dieses Buch bietet eine grundlegende Einführung in die makroökonomischen Konzepte, darunter Inflation, und beschreibt, wie sie mit den Konjunkturschwankungen in der Wirtschaft zusammenhängen.
2. "Geld, Kredit, Banken" von Alfred Müller-Armack (1947) - In diesem Buch wird erklärt, wie das Geld- und Bankensystem die Inflation beeinflussen kann und wie eine angemessene Kontrolle des Geldangebots eine wichtige Rolle bei der Bekämpfung der Inflation spielen kann.
3. "Ökonomie" von Hans-Werner Sinn (2014) - Dieses Buch bietet eine umfassende Einführung in die Ökonomie und erklärt, wie die Inflation entsteht und wie sie durch politische Entscheidungen beeinflusst werden kann. Es bietet auch eine Analyse der Auswirkungen von Inflation auf die Wirtschaft und die Gesellschaft.

ChatGPT-Screenshot vom 03.02.2023

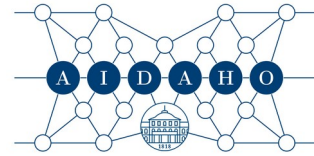
J Nenne die drei wichtigsten deutschsprachigen wissenschaftlichen Werke, die den Begriff der Inflation erklären.



1. "Geldtheorie und Inflationsproblematik" von Hans Ott (1962) - Dieses Buch bietet eine umfassende Analyse der Inflationsproblematik und erklärt, wie Inflation durch Veränderungen im Geldangebot und den Ausgabenentscheidungen der Regierung verursacht werden kann.
2. "Geld, Kredit, Banken" von Alfred Müller-Armack (1947) - In diesem Buch wird erklärt, wie das Geld- und Bankensystem die Inflation beeinflussen kann und wie eine angemessene Kontrolle des Geldangebots eine wichtige Rolle bei der Bekämpfung der Inflation spielen kann.
3. "Makroökonomik" von Ulrich van Suntum (2008) - Dieses Buch bietet eine umfassende Einführung in die Makroökonomik und erklärt die Konzepte und Theorien zur Inflation, einschließlich der Rolle der Geldpolitik bei der Inflationskontrolle. Es bietet auch eine Analyse der Auswirkungen von Inflation auf die Wirtschaft und die Gesellschaft.

ChatGPT-Screenshot vom 03.02.2023

It is easy to see that Müller-Armack (1947) is always mentioned second – otherwise the information differs. As easy as ChatGPT can be used in scientific work, its use requires, first and foremost, training in scientific work to understand the source-technical principle of the answers. Secondly, and just as important, a response to the question of what the three most important German-speaking scientific works are that explain the concept of inflation can only be evaluated based on a sufficient scientific education. In other words, if one cannot competently ask ChatGPT, one will not receive the best possible answer within the technological framework of the language model. Simple adoption of a ChatGPT answer is likely not sufficient in the university context, as the complexity of scientific matters always requires critical reflection.



Just as one worried about a plagiarism epidemic in the year 2000, the question now arises to what extent students will use ChatGPT in the future when writing texts and how the acquired competence can be determined in exams.

Due to the stochastic function of ChatGPT in generating texts, the use of a classic plagiarism software is completely doomed to fail in proving scientific dishonesty. Ironically, OpenAI itself offers a system called [AI Text Classifier](#), which is intended to serve as a check when examining the question of whether a text has typical characteristics of AI-based text generation.

If the five scientific sources described above on the concept of inflation are submitted as text blocks to OpenAI's classification system, the assessment is that this text block was most likely generated by an AI.

Text

"Geldtheorie und Inflationsproblematik" von Hans Ott (1962) - Dieses Buch bietet eine umfassende Analyse der Inflationsproblematik und erklärt, wie Inflation durch Veränderungen im Geldangebot und den Ausgabenentscheidungen der Regierung verursacht werden kann.

"Geld, Kredit, Banken" von Alfred Müller-Armack (1947) - In diesem Buch wird erklärt, wie das Geld- und Bankensystem die Inflation beeinflussen kann und wie eine angemessene Kontrolle des Geldangebots eine wichtige Rolle bei der Bekämpfung der Inflation spielen kann.

"Makroökonomik" von Ulrich van Suntum (2008) - Dieses Buch bietet eine umfassende Einführung in die Makroökonomik und erklärt die Konzepte und Theorien zur Inflation, einschließlich der Rolle der Geldpolitik bei der Inflationskontrolle. Es bietet auch eine Analyse der Auswirkungen von Inflation auf die Wirtschaft und die Gesellschaft.

By submitting content, you agree to our [Terms of Use](#) and [Privacy Policy](#). Be sure you have appropriate rights to the content before using the AI Text Classifier.

Submit

Clear

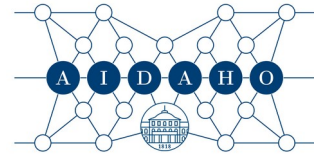
The classifier considers the text to be **likely** AI-generated.

AI Text Classifier screenshot from February 3, 2023

In contrast to plagiarism detection, whose evidence of identical text passages is clear, a probability-based examination is not sufficient to prove the fact of dishonest scientific work.

What does this mean for teaching in academia?

The University of Tübingen's rectorate clearly positioned itself in an internal university email on January 31, 2023: "Texts created by ChatGPT may not be used by students as part of written academic and exam performances (exams; final papers; etc.), unless justified by the exam task and legally allowed in individual cases."



In our opinion, there is no need for an explicit ban of this kind. A look at the [Statute for Ensuring Good Scientific Practice of the University of Hohenheim](#) is enough to ensure that the use of ChatGPT without making this transparent is dishonest. In the obligations on the general principles (§1) it says:

“The University of Hohenheim establishes with this statute, approved by the Senate, the rules for good scientific practice. It makes them known to its members and obliges them to comply with them, considering the peculiarities of the relevant discipline. Each researcher bears the responsibility for ensuring that their behavior complies with the standards of good scientific practice. The principles include, in particular,

- to work according to professional standards and norms,
- to maintain strict honesty with respect to their own contributions and those of others,
- to consistently question their own results, and
- to allow and promote critical discourse in the scientific community.“

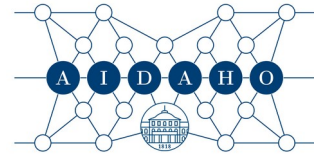
The statute obliges all members of the university, including students when they are undergoing exams, to adhere to the standards of scientific work.

Currently, on behalf of the Ministry of Culture and Science of North Rhine-Westphalia, a [legal opinion](#) is being prepared at Ruhr University Bochum to clarify questions of copyright and examination law when using AI tools in universities.

So, what to do with the use of ChatGPT by students?

The response from Ethan Mollick of the Wharton School at the University of Pennsylvania to this question is quite provocative: “[All my classes suddenly became AI classes: We can't beat AI, but it doesn't need to beat us \(or our students\).](#)” Mollick advocates for the conscious use of technologies such as ChatGPT in education. In his AI policy, it states: “I expect you to use AI (ChatGPT and image generation tools, at a minimum), in this class.” In a pre-print entitled “[New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments,](#)” Mollick explains his pedagogical proposals.

While a corresponding deep education in the field of AI and more systematic knowledge of the workings of the GPT-3.5 language model would be desirable for the conscious use of ChatGPT in higher education, a solid education in subject matter and scientific work is sufficient to be able to evaluate whether the generated texts meet scientific standards or accurately reflect the subject matter. Nevertheless, the need for specialists who master both, AI system competence and solid subject-specific education, will become increasingly important in our opinion. The AIDAHO certificate has exactly this goal: to equip students in the fields of their studies with AI knowledge and skills.



However, we see great opportunities beyond the generation of texts. For example, it is very common to use online forums such as [stackoverflow](#) when programming and to adapt documented solutions there. However, the exact solution for one's programming problem is not always found in these forums. Here, ChatGPT not only offers the possibility to generate initial program code that students can then adjust individually, but the chatbot technology also allows the programming problem to be solved interactively with the AI.

We assume that the attempt to de jure prohibit the use of AI technologies within written examination forms such as term papers and theses is doomed to failure de facto. Instead, in education, the transparency principle as a scientific honesty principle should be repeatedly emphasized. Students must learn the scientifically sound use of these new tools and be able to critically question the AI-generated texts.

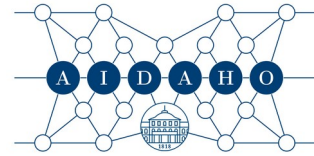
Even though this may lead to great disappointment: written homework and thesis may have served their purpose as valid indicators of what students “have learned” in the ChatGPT era. In the future, those who want to know what knowledge and skills students have acquired during a semester will probably have to resort to alternative forms of examination – be it the tried-and-true exams, oral exams, or novel formats that may even explicitly incorporate the use of chatbots like ChatGPT.

The debate about which forms of examination are still up-to-date and appropriate at universities considering the availability of ChatGPT is just beginning, so it is difficult to make a final assessment at this point. It will be necessary to discuss what standards of originality and transparency will guide the formulation of examination regulations in the future.

Research

There are already first scientific publications such as “[Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse?](#)” where ChatGPT is listed as a co-author. The editor-in-chief of *Nature*, Magdalena Skipper, has published specific guidelines for the use of AI in this context in a recently published editorial. According to these guidelines, an AI cannot be a co-author of articles. Skipper also demands complete transparency in the use of AI in all stages of the scientific research process.

A study recently published in *Finance Research Letters* by Dowling and Lucy entitled “[ChatGPT for \(Finance\) research: The Bananarama Conjecture](#)” shows how the dialogue system can be used in research. The authors argue that ChatGPT “can generate, even in its basic state, plausible-seeming research studies for well-ranked journals.” Dowling and Lucy also identify five general areas in which chatbot systems like ChatGPT can support the research process:



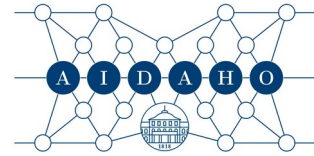
1. Idea generation
2. Literature synthesis
3. Determining appropriate data sources and preparing them for empirical analysis
4. Implementation of empirical analysis (design of test strategy, implementation, etc.)
5. Analysis of results

Undeniably, the results that ChatGPT currently produces for these five areas require human control, evaluation, and possible revision. ChatGPT is therefore not to be considered as an autonomous research system in its current form.

The change in the scientific research process is already foreseeable today. Provided a language model is specifically trained on scientific texts, the creation of text frameworks using software such as ChatGPT should be much faster in the future. Literature reviews could also be researched much faster if the training material is of adequate quality. ChatGPT already demonstrates a high level of performance in optimizing programming code. The AI assistant does what it's supposed to do: it assists.

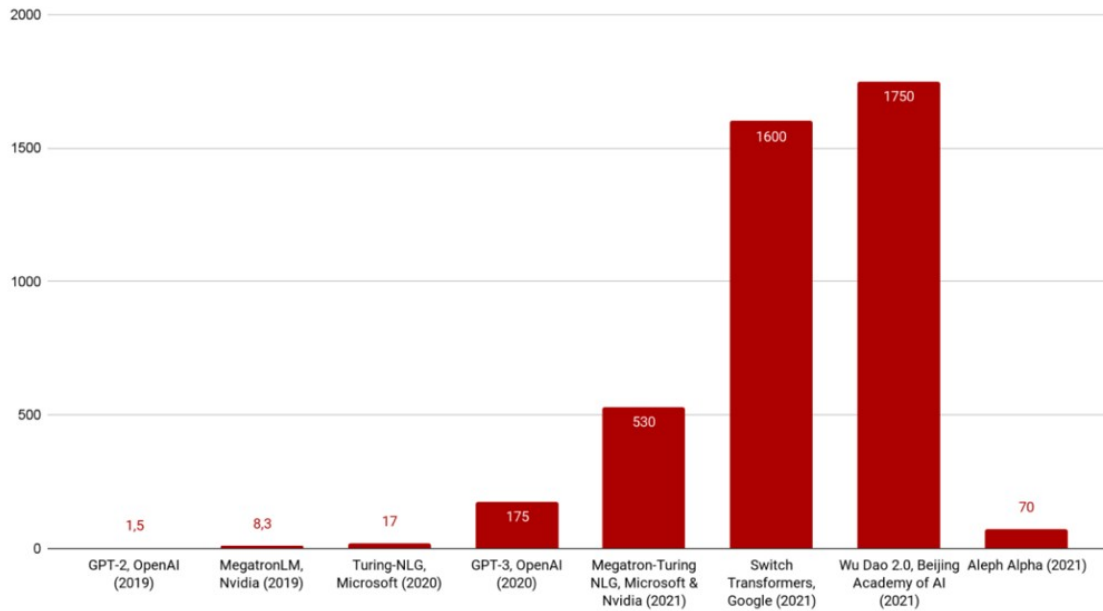
The consequences associated with the use of AI systems in research are likely to be far-reaching. ChatGPT and further developed AI models will positively accelerate the progress of knowledge in all scientific fields as assistance and support systems. At the same time, the attainment of reputation based on outstanding publication performance in the era of AI-assisted support systems will become more difficult: Is the knowledge gain of a scientific text generated with the help of ChatGPT still to be evaluated as an independent achievement? We deliberately leave this rhetorical question unanswered and undiscussed at this point.

We share the view of the *Nature* editorial board that AI systems cannot be co-authors of scientific texts. We also join Nature's call for maximum transparency in documenting the use of AI systems in the scientific research process. Just as the use of design in a study, the instruments used for data collection, or the statistics software used are transparently addressed today in the method section of a scientific work, the use of AI systems will also have to be addressed in the future.



Appendix

Anzahl Parameter großer KI-Modelle in Milliarden



Source: [Bundesverband Künstliche Intelligenz](#)