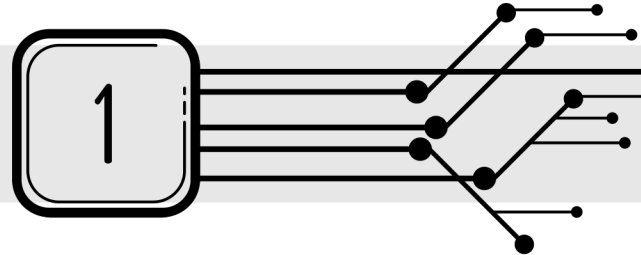


APPENDIX

1



Handling Text-Generating AI Systems: A Guide for Action

**Department for Schools and Education of the State of
North Rhine-Westphalia, Germany**

Ministerium für Schule und Bildung
des Landes Nordrhein-Westfalen
Völklinger Straße 49
40221 Düsseldorf
Telefon 0211 5867 - 40
poststelle@msb.nrw.de
www.schulministerium.nrw
© MSB 02/2023

Translated by Heinrich Niemann, from the original German text:
https://www.schulministerium.nrw/system/files/media/document/file/handlungsleitfaden_ki_msb_nrw_230223.pdf

Introduction

Text generators based on artificial intelligence (AI) are capable of producing texts of such quality that it is often difficult to tell whether they were produced by a human or not. Since the end of November 2022, when the text generation AI ChatGPT was released and made freely accessible, many questions have arisen for everyday school life. What exactly are text-generating AI applications? How can the new possibilities offered by AI applications be handled? Can AI applications be used in the classroom? What legal framework applies? Do I as a teacher have a direct benefit for my own activities? Can AI be used in the classroom in the future and how will it change teaching? Although the questions cannot be comprehensively answered at this time, we would like to provide you as educators, school management, teacher training institutions, and members of school supervisory bodies with initial information and guidance on AI applications with this action guide. Due to the dynamic developments in the field of AI, it will be necessary to update the action guide for dealing with text-generating AI systems regularly.

The commercial products mentioned in the document (e.g. ChatGPT, DeepL, you.com, BARD) are mentioned in this text as examples of the underlying technology, but this is not associated with an implicit or explicit advertisement or evaluation of these products.

What are text-generating AI (such as ChatGPT) and what can they achieve?

Even though ChatGPT currently dominates the discussion and is therefore mentioned as an example in this text, there are numerous other AI applications in the fields of art and video production, research, translation, or as planning and structuring aids that are changing our work and learning processes and are therefore also important for teaching and learning in schools. It is expected that further text-generating AI applications will quickly come onto the market; in the past few weeks, for example, the AI-powered search engine you.com or BARD have already attracted attention.

ChatGPT is a chatbot that can answer questions, summarize or evaluate texts, write poems or computer programs, translate texts, create multiple-choice tests, and much more in various languages. It is noteworthy that ChatGPT is capable of taking into account the context between consecutive text inputs, creating the impression of a conversation.

The chatbot generates extensive responses to the user's text inputs based on the language model GPT (generative pre-trained transformer). The responses vary with each input, as the text is generated anew each time. Simply put, the algorithm has learned which character strings frequently follow each other in a text.

Text creation: With clear instructions and questions, texts can be created in a predetermined text form or style.

Text editing: The AI text generator can also be used to edit existing texts. The chatbot can offer a possible structure for a text, help with formulations, or suggest corrections.

Text evaluation: Likewise, the AI text generator can evaluate created texts according to predetermined criteria (AES - Automated Essay Scoring). This makes it possible to receive quick and direct feedback and to improve one's own skills as well as to steer writing processes through direct feedback.

The term "text" is to be understood broadly here: ChatGPT can also offer mathematical equations or generate and analyze computer programming.

Why can text-generating AI applications lead to misinformation?

Texts can be created much faster, easier, and on a larger scale than before through the use of text-generating AI. Even if an AI-generated text contains mostly reliable facts, it may also contain false statements that appear to be based on facts. However, as these tools become better, they will be more difficult to detect.

Therefore, it is particularly important to focus on the topic of "misinformation" in education. Students should be made aware that the generated responses may be flawed or incomplete. For example, ChatGPT's training was completed in the summer of 2021, so the world after that time is still unknown. Text-generating AI such as ChatGPT also fills "knowledge gaps" with new combinations of text fragments or embellishments. The underlying language model is trained to create texts that read as if they were written by a human. Since the AI application learns by machine, it is also difficult for developers to predict which response will be generated in response to which input. If training were optimized to only output reliable facts, it could often lead to no response being given, which is not the intended purpose of the language model. The system thus focuses more on eloquence than on truth. In addition, there is also a risk of making incorrect or incomplete statements through text generation itself, as the creation of the text is based on probabilities. The produced content must therefore not be accepted uncritically but must be controlled by learners. This is only possible if learners have acquired sufficient knowledge about the subject beforehand.

This makes media literacy even more important. The ability to distinguish fake news from facts based on one's own secure knowledge will become increasingly important.

Does AI provide neutral and unbiased answers?

The training data for ChatGPT is based to a large extent on English-language texts. It can therefore be assumed that this alone creates a predominantly Western view of the world as the basis for the texts generated by the AI. However, precise statements on this cannot be made, as the model is not transparent to the scientific community. Nevertheless, there are already approaches to training AI with data sets in different languages and from many different countries to create a more diverse language model. However, the problem remains that very large data sets are needed for training, which can only be found on the internet. As a result, it

often happens that widespread prejudices and ethically morally questionable texts are included in the data used to train the AI. Filtering out unwanted training data is not possible, as many suitable texts would also be filtered out. Alternatively, the outputs of the chatbot could be checked and filtered so that offensive content is no longer displayed or replaced by predetermined answers. However, targeted user inputs can still generate insulting text snippets.

The "human" filter remains crucial here. A certain amount of problematic content can be prevented through subsequent training. However, it is more important to be aware when using the chatbot that articulate texts are not necessarily true and, above all, are not an indication of the intelligence or even consciousness of the AI.

How should this new capability be approached in education?

The accessibility of AI text generators is currently difficult to assess in terms of their impact. Experts from science and research unanimously emphasize that the effects of such AI on the education system and the world of work will be immense. In light of the educational and pedagogical mandate of schools, the education system must therefore address these new developments and their effects: Depending on the perspective, AI can offer opportunities and open up new perspectives or pose new challenges or even dangers. Both perspectives must be taken into account. We, therefore, ask you to engage openly and constructively with the new possibilities and to address them in the classroom.

Can a text-generating AI be used in the classroom?

The use of AI applications will increasingly gain importance in many areas of life and work. In the interest of the educational and pedagogical mandate of schools, which also includes a responsible and safe handling of media in the digital world (§ 2, 6 no. 9 Education Law of North-Rhine-Westphalia SchulGesetz), it is also the task of schools to familiarize students with AI

within the framework of teaching and to experience together in a protected environment how AI-based text generators work, what potentials, but also what risks are associated with them. A ban on addressing and using AI as a teaching tool in the classroom cannot be a viable response in the context of an extremely dynamically developing world in which students live. It is therefore important for teaching and schools to open up further and to reflect on the developments together with the students. It is evident that AI is of great interest to students and can thus be strongly linked to their lives.

In what context should engagement with AI applications take place?

The Media Competence Framework for North-Rhine-Westphalia or the digital key skills that are used in vocational education provides the framework in which engagement with AI-based text generators is meaningful: students should know their functional scope and learn to use them creatively, reflectively, and purposefully (1.2 MCF NRW), evaluate the produced information (2.3 MCF NRW), as well as understand and consciously use the basic functionality (6.1 MCF NRW).

When engaging with text-generating AI applications, three essential aspects should be taken into consideration:

- the technological perspective: How does an AI text generator work?
- the sociocultural perspective: What are the societal impacts of using AI?
- the application-related perspective: What can the AI application be used for and what should be considered?

Implementation suggestion

To address the application-oriented perspective in teaching, it can be tried out together how to cleverly formulate prompts (= requests or instructions directed to the AI) to obtain the desired response. There are numerous instructions available online for this purpose. In addition, the outputs of the chatbot can be checked for errors, uncertainties, and gaps. Particularly in

the natural sciences and mathematics, misinformation or incorrect results can often be found, which can be used as a starting point for a discussion about the functionality and effective use of the chatbot. (Example experiment: Solving a simple quadratic equation: The formula used by ChatGPT is uncommon in schools, and the chatbot makes an error in the last step. All learners can understand this mistake, and as a differentiation strategy, students can verify that the given formula is valid.)

What legal and practical conditions must be observed when using AI in an educational context?

In the context of the use of AI applications in education, various legal aspects must be considered, particularly concerning the processing, evaluation, and possible transfer of personal data, which cannot yet be conclusively assessed. This is also because the respective application possibilities and usage conditions of AI applications may vary.

Regardless, the use of AI applications in education - like the use of an online platform or app - must only occur while adhering to the applicable and known data protection requirements. The responsibility for compliance with school data protection lies with the head of the respective school. It must be carefully assessed to what extent the personal data of users are collected, and if necessary, technical or organizational measures must be taken to protect such data.

Information on what data the provider of an AI application processes can be found, in particular, in the privacy policy and terms and conditions. These may include data that is fundamentally necessary for use, such as when creating an account (age limits may need to be considered), as well as data that arises during use (automatically). A provider must be transparent about its data processing and, in particular, about the right to deletion.

Regarding the various usage scenarios in schools and the existing knowledge about the usage conditions of an AI application, such as ChatGPT, the following assessments and recommendations can be made:

The use of ChatGPT in education with students' own devices or via their accounts/email addresses cannot be recommended given the current factual and legal situation (especially concerning data protection requirements).

If teachers have voluntary access to ChatGPT or other AI applications, they can use them to work with the application in class. As with all other applications, it must be ensured that no personal data of students is transmitted. This would be the case, for example, if prompts establish a connection to the class or individual students.

It is recommended to inform parents about the type of use of an AI application in class and the framework of legally permissible options in the context of educational partnerships. Providing information on AI applications in participation committees can also help to alleviate any existing uncertainty.

How the use by students with their accounts may be possible in the future depends on the design of the terms and conditions and data protection policies of the application(s) in each individual case.

For practical school purposes, initial assessments from the perspective of school data protection officers can be helpful, which can be found, for example, at <https://unterrichten.digital/2023/01/23/chatgpt-datenschutz-unterricht-schule>.

What impact can the text-generating AI application have on students' writing skills?

What consequences can the text-generating AI application have on the development of students' writing skills? So far, there is no scientific evidence about the effects of incorporating text-generating AI applications on the development of basic writing competencies. Although it is certainly true that writing skills are mainly developed through one's own efforts, a direct deduction of negative consequences is too simplistic. Especially

concerning students who may require more individual support, the possibilities of using a text-generating AI application must be further reflected upon (e.g., individual inquiries regarding text comprehension, scaffolding, and different levels of text difficulty in class).

Can AI be used to support the learning process?

A learning-friendly use of AI can contribute to expanding language, writing, and evaluation skills. It is necessary to design the use of AI didactically in terms of subject-specific and cross-disciplinary competencies (e.g., media competencies or digital key competencies of vocational colleges). For example, the texts generated by the AI can be examined for correctness, consistency, style, etc. Furthermore, working on factual accuracy or arguments is possible.

Scaffolding: In addition to supporting the processing of texts, AI can also provide individual assistance in other areas during the learning process. For example, it is possible to simplify or translate texts during internet research. In addition, learners can have concepts explained to them again and answer follow-up questions, and have appropriate examples or analogies searched for them.

Creation of exercise material: A text-generating AI can create self-tests to check whether students can correctly reproduce the content of a text. These gap-fill exercises, multiple-choice tests, or questions with answer choices can later be used for repetition and practice. The testing formats could also be integrated into learning management systems, such as Learning Management System for all public schools and recognized alternative schools in North Rhine-Westphalia, and made available as a practice opportunity for the class or course.

The defined evaluation criteria for a learning product can be used with the help of AI for any time available, timely formative feedback. It should be

clear to students that this is feedback to adjust the learning process and not a performance evaluation because the teacher is not directly involved in the feedback. The criteria must be not only transparent but also understandable for students to be able to work with the result meaningfully.

AI-generated texts can be used in the classroom in multiple ways as a basis for discussion or as a topic of discussion. On the one hand, they can serve as brainstorming, for initial orientation and structuring or for collecting arguments, but on the other hand, the creation of content must always be discussed. Since the texts are created based on an AI-based language model, correctness is not necessarily given. Learners must be guided to increasingly check the content against their prior knowledge independently. In addition, the texts created by AI must also be considered from a values perspective, as values, norms, and ethical aspects do not play a role in the creation of the text by AI.

How can teachers establish rules with students about how the use of AI in texts can be referenced?

If the students voluntarily use text-generating AI tools, they must, of course, indicate them as a source because only then can the teacher determine what performance the learner has achieved and evaluate it. Currently, citation rules that indicate AI as a source are being worked on in the scientific field. Based on this, a regulation for indicating the use in a school context could look like this:

"In producing this text [or image or programming code, etc.], X [=name of the AI-assisted tool] was used. I controlled the AI with the following prompts [=instructions or questions for the AI]: 1._____, 2._____"

The advantage of this type of indication is that the teacher can evaluate how extensive the use of AI was. It can also be assessed how competent the student controlled the use of the AI. The text generated by the AI should be included for evaluation.

How do I handle it if students use AI but do not disclose it?

The general principles for "scientific" work and performance evaluation still apply: If external aids are used, they must be fully disclosed. This also applies to the use of an AI application. Texts created using a text-generating AI are less easily identifiable as plagiarism than those copied and pasted from websites or texts because the AI-generated texts are not always identical. Strictly speaking, this is not plagiarism, but rather the use of an application to create a text. It is currently unclear whether it will be possible to develop reliable software that can detect AI-generated texts. Experts believe that such software would rather be used as a training opportunity for an AI system to improve.

Even if it is not plagiarism in the strict sense, not indicating that the text or parts of it were generated using AI is a deception about authorship. If the use of AI has been explicitly excluded in the assignment, it is also the use of an impermissible aid and an attempt to deceive. For tasks that are not carried out on-site in the school under supervision, teachers, as always, have the opportunity to check the degree of independent work of students: in this context, teachers have a high level of professional experience and can, among other things, recognize in classroom discussions whether students have completed products that they have created at home for the performance assessment independently or with impermissible assistance. Just as tasks in the home environment were not allowed to be created with the help of third parties, they must not be provided with a technical aid that is not adequately indicated. This also involves teaching students that it is in their own interest to independently complete learning tasks assigned to them (including homework, which might not be graded) and to correctly indicate any aids and sources used.

If there are acts of deception or other irregularities, the procedure is based on the principles of performance evaluation (§ 48 SchulG) and the relevant regulations of the examination regulations (in particular § 6 paragraph 7 APO-SI, § 13 paragraph 6 APO-GOST, § 20 APO-BK).

To avoid situations that are stressful for both students and teachers, the tasks should be designed preventively so that they cannot be completed exclusively with the help of AI.

How can I create tasks that are less susceptible to being exclusively completed by AI?

With sufficiently complex questions, it is apparent that texts created in the first step by a text-generating AI application do not necessarily meet the requirements. Often, multiple, very specific questions and corresponding subject matter knowledge are necessary to create appropriate texts for the assigned task. Training and mastering this type of control is a perspective that is important to convey to students. It is possible to motivate learners to work independently through challenging, interesting tasks and to show why their own, independent writing experiences are important to be able to reflectively evaluate the quality of both their own and others' texts. Addressing this perspective can be successful, especially in the cooperation between teachers and learners.

All of the mentioned aspects will contribute to the need to further develop learning and performance tasks. It is already advisable to combine existing formats for checking learning performance in such a way that they consider the work processes of the students and are less susceptible to being exclusively processed through the targeted use of AI.

For projects or term papers, for example, this could mean including their own, individual research task of an empirical nature or series of experiments that are tailored to the specific teaching situation. The guidance of the work process also becomes significantly more important. In conversation with the learner, it becomes clear to what extent he or she has dealt with the content. In addition, a final presentation with a discussion about the work can illustrate that the student independently completed the work or that passages of text generated with AI were not used without reflection.

In the classroom, it is necessary to motivate and guide learners to shape their learning process through creative, challenging, and real-world related tasks, opening up learning arrangements, and implementing peer approaches. Essential impulses for a learning-promoting education in the digital world can be found in impulse paper II.

Tools for the further development of tasks

If individual references are included in the learning and performance tasks, the AI cannot easily take them into account, so the individual contribution on the part of the learner is necessary. For example, if a survey in the class, an experiment, mapping, data collection, personal hobby, a self-selected focus about the place of residence or a comparison with the presentation of a classmate is included, AI text generators cannot fully take over the task. Additionally, working on a topic from an individual, self-chosen perspective can contribute to the motivation of learners.

A change in format or media in the task means that learners must independently implement their findings. While AI can provide ideas for creating a poster, an explanatory video, a podcast, a song, or a still image, suitable implementation and thus engagement with the content must be carried out by the learners. Additionally, this approach can promote argumentative competencies in learners, as they are increasingly asked to justify their own decisions, analyses, or formats.

Formative assessment - i.e., observation, feedback, or feedback accompanying the development process - can also contribute to ensuring the independence of the work submitted. It also promotes the motivation and acceptance of learners to write their own texts, as they can be further improved based on repeated personalized and timely feedback. Instead of writing several, ever-new texts for the class, fewer, but better, repeatedly revised texts are produced. Additionally, the continuous feedback contact between teacher and student makes it less likely to insert longer text passages that come from an AI.

What benefits could it have for me as a teacher?

In the discussion of the possible applications of text-generating AI, diverse opportunities also arise for teachers: Teachers can benefit from the use of AI text generators in preparing for class by generating differentiation options, diagnostic tests, ideas for implementing the class or initial corrections of products created by the learners. Thus, texts can be simplified and offered at different levels, analogies and examples can be found, or assessment formats can be created. Texts could be pre-corrected regarding grammar, spelling, and sentence structure. A media-competent use also means that data protection regulations must be followed, so no student data may be entered.

Outlook

Not only the school sector but also universities are currently intensively dealing with the consequences of text-generating AI applications. All discussions and publications by scientists are currently focussing on the 3 statements below:

1. A ban on the use of AI is unrealistic and unsustainable.
2. Instead, it is necessary to establish the potential and risks for teaching and learning.
3. Dealing with AI must be negotiated between teachers and learners and must be transparent and legally secured.

There is a high level of agreement that a competent handling of AI applications is essential for successfully dealing with future demands in education, study, work, and everyday life; for this reason, schools cannot turn a blind eye.

It requires a joint effort of schools, education authorities at all levels, and teacher training and development, to further develop the 3 aforementioned areas. Therefore, the Ministry of Education will continue to deal with these developments and provide information on the topic to all relevant bodies.