

Generative AI Systems in Education – Uses and Misuses

Generative AI systems, such as ChatGPT, DALL·E, and BlenderBot have been commanding news headlines and sparking [conversations](#) about the role of AI in [education](#), the workforce, and society in general. While these systems have the potential to be helpful tools, providing people with a new type of technological assistance, they also introduce a number of risks and challenges, and require careful introduction with clear guardrails and norms governing their use.



What is Generative AI?

Generative AI systems use machine learning to produce new content (e.g., text or images) based on large amounts of training data. That data is typically examples of the type of content the system will produce (such as [enormous amounts of text](#) for systems like ChatGPT that will produce text responses, or [hundreds of millions of images for DALL·E](#), which produces images in response to prompts). Using this data, these systems are trained in one of two ways:

- *Unsupervised*, meaning that the data that the system consumes in order to learn is not labeled or categorized by human experts, so the system does not know what data is good or high quality and what data is bad or poor quality; or
- *Semi-supervised*, meaning that most of the data the system consumes is unlabeled, but it may get some amount of labeled data.

The system uses all this training data to establish an understanding of what human-produced content looks like and aims to produce new content that mimics patterns it learned from the training data. The content can take a number of forms. For example, it might be language, in the case of systems like ChatGPT or BlenderBot, or art or imagery, in the case of DALL·E or [ThisPersonDoesNotExist](#). Importantly, these systems are largely aiming to produce content that feels “real” to human users, though what constitutes real depends on the type of content the system is producing. For text-producing systems, it typically means that the text produced mirrors that produced by humans, and a human could not tell the difference between human-generated content and content generated by the system. For image-producing systems, it might mean that the produced image looks like art that might have been made by a human, or it might mean that the image produced feels photorealistic.

While some generative AI systems produce content without any specific input or prompts from users (such as ThisPersonDoesNotExist, which presents a random photorealistic “human” face to website visitors), other systems provide content in response to specific queries or prompts from users. In order to respond to user prompts effectively, the system must be able to parse and “understand” what the user is asking for and how it will inform the generated output.



What are Uses for Text-Producing Generative AI in Education?

This technology has the potential for numerous applications, ranging from simple to complex, fanciful to tactical, delightful to deeply concerning. Significant attention has been paid to concerns like plagiarism, [sometimes resulting in blanket bans on the use of generative AI technologies](#), covering [not just students but teachers and administrators](#) as well. However, there are a number of potential constructive uses of generative AI in the education space, for both adults (such as teachers and school counselors) and students.

Adults

For adults who work in schools, like [teachers](#), [principals](#), and [school counselors](#), these uses may include [first drafts of lesson plans and rubrics](#), administrative tasks such as [drafting emails](#), using the system as a more responsive search engine, or [first-pass grading for essays and other assignments](#) (the system can provide information such as how well an essay follows a specific form, for instance). Using the AI for this sort of use may save teachers time, allowing them to focus their energies on other aspects of educating students. Some teachers are also incorporating generative AI into their classrooms to help students learn about [how these systems might be used in their adult lives](#), while [understanding their limitations and drawbacks](#).

Students

For students, there has been significant discussion of concerns like plagiarism, but some educators have noted that there are constructive uses of generative AI as well. These may include editing and improving a writing assignment draft (such as asking the system to identify areas where the text is unclear or too informal), [rephrasing complex topics in different ways for students who are struggling to understand a textbook explanation](#), or as a [more responsive search tool](#) that enables more complex queries than a traditional search engine. Additionally, the tool itself can provide a meta-lesson of sorts, as a way for students to explore concepts of media literacy and the source and value of different kinds of information and content, something teachers are beginning to [incorporate into their lessons](#).



What are the Risks and Challenges of Generative AI in Education?

While some of these uses have potential, there are certainly risks and drawbacks to generative AI systems being used in education as well.

Plagiarism

One of the most prevalent educational concerns around generative AI systems is their use for plagiarism, which in this context would mean students using the system to do work that they then present as something they created without AI assistance. This may mean producing [entire essays](#), or using the system along the way for tasks like outline generation or editing. While different

educators will disagree on at what point this use becomes problematic, part of the challenge lies in not being able to detect plagiarized work using traditional tools, as most current forms of plagiarism detection rely on the assumption that the offending text is drawn from content that already exists, which is not the case with AI-generated content.

Equity

As with most AI systems, equity is a concern. Generative AI systems are trained on data that will reflect the biases of the world that data stems from. This, in turn, [can lead to those biases and prejudices becoming embedded in the AI itself](#). While designers can take steps to limit this bias on both the input end (by curating the training data and trying to eliminate biases at that stage) and on the output end (trying to detect outputs that reflect a bias and stopping or modifying those outputs before they go to the user), [neither of these approaches will be completely effective](#). Because of the enormous volume of data needed to train generative AI systems, it is typically infeasible to have humans vet all the training data. Additionally, if the system is designed to continue learning from user queries and responses over time, those inputs will generally be outside the control of the system developers. This is particularly concerning in an education context where students may be using these tools to learn more about the world around them, meaning the tool may impart or reinforce biases in students' thinking.

Privacy

Another concern with generative AI systems is privacy. On one hand, there is the question of how training data is sourced. People may be surprised and upset to learn that information about them or content they created is being used to train or teach an AI system. If the generative AI system uses existing data corpuses with a clear use case of training AI systems this may be less of a concern. However, any system that gathers data from the public internet for novel purposes may be using data in ways that the data subjects did not anticipate and may not be comfortable with, even if they technically consented to it under a broad clause enabling unspecified future uses. Additionally, the data that is inputted or created during interaction with the system, whether that be the outputs from the data or search terms and queries provided to the system, may be sensitive as well. In some ways, this mirrors existing concerns around search engine privacy. A student asking for resources around gender and sexuality may be placed at risk if their teachers or school administrators get access to these queries and the student is outed to their family and community. Generative AI may exacerbate these considerations if those terms feed more heavily into the development and evolution of the system than they do for traditional search engines.

Efficacy

Another critical concern with generative AI systems is one of efficacy. Because of the unsupervised nature of their development, generative systems may "[hallucinate](#)," meaning they generate untrue responses. Of course, whether this is a problem depends on how the user is interacting with the system. If they asked the system to write a short fictional story, then untruth is not an issue; in fact it is expected. However, if the user was asking a factual question for research, a

hallucination is a failure case. Because of the multi-use nature of many generative AI systems, it can be hard for system developers to address this issue. This is both because they may not wish to restrict the system from hallucinating entirely, but even if they did desire to do so, it simply may not be possible. This means it would not necessarily be possible to make something like a research assistant that teachers could offer to their students with the assumption that it would only ever provide factual information. The system may not be able to understand ground truth in a meaningful way, because it is trying to “learn” for itself what data is more reliable than other data.

Detection

Partially in response to concerns about plagiarism, companies and people have begun building systems designed to detect content created by generative AI systems, and some developers have started watermarking the output of their systems. However, these are currently [largely ineffective](#), and are unlikely to ever be foolproof because they have to evolve along with the generative systems themselves, leading to what is often referred to as an “arms race.” As with the hallucination problem, the fallibility of detectors runs the risk that people will assume they are more effective than they actually are. Due to this, any solutions to the risk presented by generative AI are likely to have to be more robust than relying on detectors.

Appropriate Use

Generative AI has the potential to be incredibly useful, but societal norms for when and how it should be used are still very much in flux. Because of the human-imitation nature of these systems, there is a high potential for people to feel unsettled if they realize the systems have been [used in ways they find inappropriate](#). As these norms are developing, it is critical to engage in robust discussion with students and the broader school community about when and how they use the systems, the value and limitations they offer, and to set clear guidelines around their use in an academic context.

Authorship

A final concern with AI-generated content is one of authorship. This is closely related to the issue of plagiarism, but it raises broader considerations. It will not always be clear how much of a generative AI system’s output belongs to the user who prompted it, versus the developers of the system, versus the authors and creators of the system’s training data. The fact that content is often created in response to iterative prompts, or is used as a starting point for a piece of work that is then significantly altered or adapted by a user, makes this a complicated question. This may result in the need for things like clear guidelines around what appropriate uses are for generative AI when it comes to things like writing contests, school papers, and college essays.



Conclusion

Generative AI systems have the potential to be a remarkably adaptable and useful tool in education, both in and out of the classroom. As with almost all new technology, however, they raise risks and challenges. Reaping the benefits of these tools will require a careful and deliberate rollout and long term willingness to adjust and tune the tools themselves, as well as creating norms that govern how they are used by educators and students over time as new risks emerge and new mitigations are developed.

This is one in a series of information sheets designed by CDT's Equity in Civic Technology team to give practitioners inside public agencies clear, actionable guidance on how to most responsibly use technology in support of the communities they serve. More info: <https://cdt.org/civic-tech-inventory>.