

# Generative AI and Education: Adopting a Critical Approach

## Introduction

Research on Artificial Intelligence in Education (AIED) has been ongoing for more than 40 years, along with the promise that it will revolutionize education. However, rarely have the possibilities and challenges of AIED entered the public discussion. This all changed recently with the launch of ChatGPT and other Generative Artificial Intelligence (GenAI) tools, which are already the subject of heated debates in and beyond the classroom.

Commercial and open-sourced GenAI tools can generate text, images, audio, videos, simulations, and even software code. In the field of education, as a quick web search will reveal, GenAI is already being used for various purposes, such as suggesting ideas, editing texts, creating resources, providing feedback, tutoring, and assisting students with special needs, among many other things.

However, with their remarkable ability to generate human-like outputs within seconds, concerns have arisen that students might use these technologies for cheating; not only undermining the value of qualifications and the trust

covenant between teachers and their students but also leading some educators to question both the purpose and methods of assessment. In response, institutions worldwide have adopted various approaches, ranging from a [complete ban](#) on GenAI in schools to the [creation of guidelines](#) aimed at supporting staff members and students to use these technologies effectively, ethically, and transparently.

Given GenAI's fast-growing popularity (ChatGPT, for instance, gained [100 million users](#) in two months after the launch of its product) and the rapid emergence of new GenAI education tools, it has become clear that these technologies will increasingly impact education. Consequently, all educators and policymakers need to comprehend their limitations and potential negative effects.

## **Debunking the 'Intelligence' Myth**

On the [OpenAI website](#), we read that "while tools like ChatGPT can often generate answers that sound reasonable, they cannot be relied upon to be accurate." However, the fact that GenAI-generated text is becoming increasingly difficult to distinguish from human-produced text, is often leading to the common misperception that GenAI tools are intelligent.

In his influential publication, [Computing Machinery and Intelligence](#) (1950), Alan Turing describes what he called

the 'Imitation Game', what is now known as 'The Turing Test'. The purpose of the test is to determine whether a system is able to deceive a person by generating responses that are indistinguishable from those of a human. Turing considered a system 'unintelligent' if it was unable to achieve this level of performance. Today, many GenAI tools are capable of mimicking human responses to a wide range of questions and, therefore, passing the Turing Test. In the near future, these systems may be able to generate responses that are even more indistinguishable from those of humans. However, it remains incorrect to suggest that any GenAI tool is intelligent – as they lack any understanding of either the prompt or what they produce in response. In other words, GenAI cannot generate anything that it hasn't ingested; the production of text is solely based on statistical probability. Some have even called GenAI "[Stochastic Parrots](#)", as the systems simply repeat the text ingested in their training without comprehending or interpreting any meaning.

Because GenAI tools mislead us into thinking they are intelligent, it is also often assumed that they possess a certain level of agency; that they are autonomous and trustworthy. In reality, however, GenAI frequently generates factually inaccurate and unreliable information or made-up content, often referred to as "hallucinations" (although this is an unnecessary and misleading anthropomorphism). Further, the trust that people place in

GenAI opens up greater opportunities for disinformation, particularly targeting vulnerable groups or individuals lacking expertise in the field, such as students. This can have negative implications for education. Accordingly, a critical approach towards content generated by GenAI should be adopted by teachers and students alike.

The fact that GenAI-generated text is becoming increasingly difficult to distinguish from human-produced text, is often leading to the common misperception that GenAI tools are intelligent.

## **The Black Box**

GenAI is often referred to as a 'Black Box', as it lacks transparency both for users and for other interested parties. Even though we have a broad understanding of how the technology works, developers have not opened their tools to thorough scrutiny. We remain unaware of the specific algorithms they employ or how they generate particular responses. This lack of transparency poses a potential danger, as the inherent biases of these systems go unrecognized, leading to risks such as political manipulation, opinion influence, propaganda, disinformation, and more. This concern is magnified as the number of users continues to grow, and as more and more people rely on this technology for their daily tasks (the common verb to "Google it" is fast being replaced with "ChatGPT it"). Amid this growing popularity, teachers and students must be aware that it is impossible to

comprehend or explain why GenAI produces specific information and not others, and what should be trusted and what questioned.

## **More on Data**

GenAI relies on a massive amount of data collected from the Internet. However, access to these large datasets is restricted to big companies in the Global North that operate on a large scale. The associated capture and [processing costs](#) can amount to millions of dollars per month, preventing smaller players from competing effectively. This leads to market monopolization, as smaller companies, particularly those in developing countries, have limited access to such data or the necessary financial resources, leaving them without the opportunity to develop their own systems.

In addition, fewer people in developing countries have internet access, resulting in a lower production of local data compared to more developed countries. Language differences also play a role, further marginalizing the voices of linguistic minorities and those from subordinate backgrounds and reinforcing mainstream views and narratives. Policymakers, educators, and students all should be cautious about the content generated by GenAI, as it will likely not fairly represent the full variety of perspectives and will only reproduce dominant mainstream views. Furthermore, given that GenAI scrapes its data from the internet, it necessarily incorporates all of

the internet's biases, including hate speech, discrimination, racism, and homophobia.

The text generated by GenAI is already flooding the internet. This means that the new versions of these systems will be trained on data generated by GenAI, including all its errors and biases, leading to a recursive acceleration of inaccuracy.

While GenAI companies work to address these issues, by providing so-called 'guardrails' to avoid their tools generating offensive or harmful outputs, this often comes at the expense of workers in Global South countries who are employed on very low wages to identify and moderate the outputs, a task that can be particularly [traumatic](#) and even [life-changing](#). Since the production and use of data will continue to grow in the future, the need for ongoing training and addressing harmful outputs will be a never-ending task. Moreover, the text generated by GenAI is already flooding the internet. This means that the new versions of these systems will be trained on data generated by GenAI, including all its errors and biases, leading to a recursive acceleration of inaccuracy.

A final important consideration for students and teachers when using GenAI is that the content generated by these tools is often based on the intellectual property of others. The data that GenAI uses (text, audio, video, code, etc.) is scraped from the internet without any consent being sought or obtained. Several [lawsuits](#) have already been

filed on this [issue](#). This violation of Intellectual Property Rights raises a series of questions about the nonconsensual future use of personal data and the purposes for which it will be used.

## **Conclusion**

The GenAI genie is out of the bottle, and its impact is both irreversible and still to be fully worked out. While these technologies may bring about some yet-to-be-determined positive changes, it is essential to carefully weigh their benefits against their disadvantages. It is particularly important to explore deeper both their intended and possible unintended impacts on all spheres of life, as well as to address their limitations and their ethical, moral, and environmental implications (e.g., training GenAI models uses enormous amounts of energy). This is necessary to mitigate the drawbacks and to foster the development of more morally- and responsibly-driven GenAI for the future.

It is particularly important to explore deeper both their intended and possible unintended impacts on all spheres of life, as well as to address their limitations and their ethical, moral, and environmental implications.

In particular, the education sector should adopt a critical approach to the content generated by GenAI. Teachers and students alike should be aware that the outputs

generated through these technologies often exhibit errors, inaccuracies, and biases, favoring mainstream perspectives while overlooking those of marginalized communities. Moreover, caution should be exercised when incorporating content from these tools, given the Intellectual Property Rights and the limited understanding of the response generation processes. Finally, we also must take into consideration the human labor and financial resources required, as well as the negative environmental effects associated with these tools.