

# ARTIFICIAL INTELLIGENCE AND EQUITY IN EDUCATION: THE WAY AHEAD

ALDAN CREO

## ABSTRACT

How can we use AI to promote equity in education across cultural and socioeconomic divides? To answer this question, this article explores the old and new challenges that AI presents, including cost, implementation disparities, the closed nature of industry, inherent biases, and the impracticability of punitive approaches to AI use. I also discuss the way forward and how we can work to make AI a force for equity in education.

**Keywords:** Artificial Intelligence • Education • Equity • Bias • AI-Generated Text Detection



## THE PROMISE AND THE REALITY

Artificial Intelligence is a disruptive technology that promises to revolutionize the academic landscape. If we use it properly, we can promote equity at all levels of education.

Sounds cliché? Rightly so, we have been here before. I could have written almost the same line in the early 2000s. The internet was supposed to democratize access to information and make quality education a reality for everyone, regardless of location or socioeconomic status. While some progress has undoubtedly been made, inequity unfortunately remains a reality.

In this article, I hope to provide some insights into the challenges and opportunities that AI presents in the context of education, with a particular focus on the pathways to ensuring that AI helps to promote equity rather than exacerbate existing inequalities, and how Fulbrighters are uniquely positioned to play a crucial role in this transformation.

## OLD CHALLENGES THAT PERSIST

Several challenges that have persisted since the early days of the digital age continue to impede the realization of the promise of technology in education. Although AI was first conceptualized around the 1960s, its practical applications in education have only recently emerged, thanks to large language models. But first, we should define “equity” in this context. Here, I’ll use the term to mean that all students, regardless of background, have the same opportunities to succeed in school and in life, sometimes referred

to with the (similar but not identical) idea of “equality.” There are two sides to the intersection of equity and education systems: equal access to quality education and education in values shaped by equity. I believe both are equally important for societal progress, and AI can help us achieve both.

However, one reason often cited for why new technologies have not supported equity: their cost to less-developed nations. Less-developed nations have not been able to afford the same level of technology as their wealthier counterparts, thus creating a digital divide. While there may be some truth to this, we should not forget that market prices for technology are lower in those economies and that the growth of (mainly) Chinese exports in the 2000s and 2010s has significantly driven prices down. Open-source projects have also been instrumental in helping less-developed countries access technology at a lower cost, helping institutions save on software licenses that could otherwise become prohibitively expensive. Overall, cost has been a factor, but not the main reason the digital age’s promises have not been realized.

Indeed, if cost were the only factor, one would expect that advantaged countries would have been better able to use technology to reduce inequity. My experience of living in different European Union countries has taught me otherwise. For example, when I studied in a French lycée, I was surprised to see that students were not even allowed to use technology in the classroom; the use of smartphones was strictly forbidden. In Spain, students were given a personal tablet, but this had no clear educational purpose and lacked adequate teacher training, showing how poorly planned policies can remain ineffective. This is in stark contrast to Switzerland, where technology was an integral part of the curriculum. Thoughtful integration can reduce digital divides, but coordination becomes even more critical with AI. The same can be said on a broader global scale: the role of technology and AI in education is a policy choice that varies widely across countries, districts, and institutions. Students are given different opportunities to interact with systems that are becoming more pervasive in societies, which exacerbates existing inequalities. That said, other factors such as cost, infrastructure, and cultural attitudes certainly play their part in these disparities.

## **UNIQUE CHALLENGES OF AI**

While some challenges that AI poses to equity in the context of education are not new, there are other aspects that differ from previous technological revolutions.

One of the biggest challenges that AI poses for education is its cost. While the Internet has operated as a (mostly) open and free platform, AI comes at a high cost that we’re only beginning to understand. For example, the cost of training OpenAI’s GPT-5 model is estimated at \$500 million. That money is coming from investors looking for a return, and as the AI world begins

to show signs of strain, many of them will quickly start looking for ways to cash out. Are OpenAI's new ChatGPT pricing scheme (now up to \$200 per month) may hint at this, but this is likely just the beginning. Ultimately, users will bear the costs.

This is a problem for everyone, and especially for low-income countries. These countries used to be able to pay at least partially for the costs of their technological infrastructure out of their own pockets because their local markets have lower costs (for example, the median price of a month of internet in the US in 2024 was about \$60, while it was about \$10 in Egypt). But in the world of AI, the cost of developing and running a model will be the same regardless of where the user is located. ChatGPT Pro, for example, is priced at \$20 – equally, but not equitably (market-specific pricing). This is only natural in a capital-intensive market like AI, but it becomes a concern when these systems are what we base our education systems on. As costs are increasingly passed on to users regardless of their ability to pay, the digital divide can only widen.

Another challenge that is particularly pressing in the context of AI is the closed-source nature of most models. The open-source alternatives that have helped lower the cost of technology in less-developed countries are not as prevalent in the AI world. Some may point to new models like DeepSeek's V3 or R1 as examples of a new wave of open-source AI, but even these models are only open source in terms of their weights, not their training data. Researchers looking to develop more cost-effective solutions often need this data; without it, the range of possibilities is limited. For example, it is still possible to “extract knowledge” from open-source large language models (LLMs) to create smaller ones in a process known as distillation, where knowledge is transferred from a large AI model to a smaller, more efficient one. This process does not require retraining the model from scratch, but this is not always possible, and if researchers were to design an entirely new architecture, they would have to start from scratch. In general, the ability to develop cost-effective AI solutions is limited by the lack of open-source models or data, a challenge that disproportionately affects less developed countries, which have the greatest need for cost-effective solutions.

## **CONFRONTING BIAS IN AI SYSTEMS**

Beyond cost and access, another critical challenge that threatens AI's potential for educational equity is the presence of inherent biases within LLMs, an area that I am exploring in my own research.

While Meta's Llama-3.1 might predict that “John works as a freelance” (sic), it also infers that Mary is most likely to work as a nurse and Vivek as a software engineer. These seemingly innocuous predictions can reinforce gender and racial stereotypes. Consider a student using an AI-based tutoring

system: if the system consistently suggests different career paths or learning materials based on a student's name (which, as we just saw, is a source of bias), it could subtly limit their aspirations and opportunities. These are biases that are present in the data used to train the model and are not easily corrected.

There are some techniques, such as counterfactual role reversal, that “correct” the LLM by showing it examples that challenge stereotypes. For example, it could be shown many examples of both “Mary is an engineer” and “John is a nurse” to counteract preexisting biases. However, these methods are still in their infancy and may degrade the performance of the model; after all, biases play a large role in our understanding of the world. For example, if we removed the notion of gender, we would find that they might fail to distinguish that John is (often) a “he” rather than a “she,” or that breast cancer is more prevalent in females.

It might be useful to relax the definition of bias that I implicitly used above. Instead of defining “bias” based on what it is – as any difference in generated probability distributions when sensitive attributes like gender or race are changed – we can focus on what it causes and say that “bias” is any difference in generations that can be harmful to a particular group of people. This definition is more in line with the idea that bias is not inherently bad, but that it can be harmful; it is our responsibility to ensure that it is not.

The solution to this kind of bias is manifold. Public datasets exist that can be used to benchmark models for bias, and these results should be made public (model cards, similar to “nutrition facts” labels for LLMs, are a good example). This is an essential first step, although many users can be expected to ignore it – which emphasizes the need for awareness campaigns. We also need to empower researchers to develop alternatives that tackle bias at its root. One of the main hurdles is the lack of open-source training data; again, the solution is to advocate for more open-source models and data. Of course, researchers need to reciprocate by being mindful of the privacy concerns and sensitive nature of the data, such as personal information or medical records. This demands appropriate technical handling with rigorous data quality control and usage auditing.

## **BEYOND PUNITIVE APPROACHES**

Fearing the potential negative consequences of AI, many educational institutions have adopted a punitive approach to its use, much like they did with the internet. I am convinced that this is not the way forward. We talked earlier about the detrimental effects of exacerbating the digital divide with a divided approach to technology, but now I would like to shift the focus to the impracticality of this approach, since much of my work has focused on the recognition of AI-generated text.

While systems designed to distinguish between human- and AI-generated text are architecturally very diverse, they all come down to the same basic idea: they look for patterns that are characteristic of AI-generated text. This approach is fundamentally flawed. The reason why is surprisingly simple: the goal of large language models is to learn the best possible approximation of the distribution of human language - in other words, to become as human-like as possible. We can think of humans as language models themselves, black boxes that spend their childhood learning the rules of language and then output text based on what they have learned, like me writing this article. If we assume that the present trend of LLMs becoming increasingly better at approximating human language continues, it is only a matter of time before they become indistinguishable from humans. Once that happens, we will not be able to rely on any system, current or future, to detect if someone is using AI.

This insight leads to an important conclusion. Educational institutions need to move beyond punitive approaches to AI use. Not only is it impractical to try to detect AI-generated text, but it can be harmful to students if they are falsely accused of using these tools. Moreover, the harm can also come from its impact on the existing digital divide, as we have seen in the previous sections. The detection of AI-generated text is therefore yet another reason to advocate for the inclusion of AI as an integral part of curricula, harnessing its potential to open up new opportunities for students around the world.

### THE PATH FORWARD: BETTER AI FOR EDUCATIONAL EQUITY

If we want to promote equity in education, we need to ask for better AI. AI that is affordable, fair, and can be used by all.

To achieve this goal, I am convinced that it is essential to advocate for true open-source AI, where not only models but also training data are publicly available. It is also important to promote awareness and research on AI biases and its impact. To truly embrace AI as

a tool to empower students of all backgrounds to succeed in school and in life, AI needs to be accessible across cultural and socioeconomic divides, moving away from a punitive mindset that only serves to exacerbate existing inequalities.

That may sound compelling, but what exactly can Fulbrighters do? We are a diverse group of talented individuals united by the common goal of promoting cross-cultural understanding. Our global network makes us uniquely powerful to tackle AI biases and disseminate AI benefits across all countries in the program. To this end, I encourage you to consider:

*“If we want to promote equity in education, we need to ask for better AI, one that is not only affordable but also fair and that can be used by all.”*

**Open Source Advocacy:** Support or work with organizations like Apache, Hugging Face, Mozilla, or other local open source initiatives to advocate for the change you want to see. If you need a place to start, the Open Source Initiative has a list of partner organizations. Leverage the Fulbright network to coordinate advocacy efforts across multiple countries simultaneously.

**Research and Awareness:** Incorporate bias awareness into your teaching or research, even if it is not your primary area of expertise. Perhaps you can ask your students to compare how AI responds when they write in their native language versus English, develop educational materials on the topic, or collaborate with AI researchers globally through the program's connections. Studying bias requires cross-cultural understanding. Teams from a single country may fail to identify biases that could easily be detected by more diverse groups.

**Embrace AI in Education:** If you are an educator, lead the change by integrating AI literacy into your curriculum. For example, AI-powered translation tools for language learners, personalized learning platforms that adapt to individual student needs, or AI-powered chatbots that provide 24/7 student support, to name a few. Share the most and least effective practices in your experience with the broader community.

Of course, this is a long journey that requires the collaboration of educators, policymakers, students, and researchers. I will continue to work toward these goals, and I hope you can join me and other Fulbrighters in this important endeavor. Building a world where AI is a force for good in education is only possible if we work together to make it happen.

## FURTHER READING

1. Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, 15(15), Article 11452. <https://doi.org/10.3390/su151612451>
2. Naous, T., Ryan, M. J., & Xu, W. (2023). Having beer after prayer? Measuring cultural bias in large language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 1086–1099. <https://doi.org/10.18653/v1/2023.findings-acl.67>
3. Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can AI-generated text be reliably detected? *arXiv*. <https://doi.org/10.48550/arXiv.2303.11156>
4. Solatorio, A. V., Vicente, G. S., Krambeck, H., & Dupriez, O. (2024). Double jeopardy and climate impact in the use of large language models: Socio-economic disparities and reduced utility for non-English speakers. *arXiv*. <https://doi.org/10.48550/arXiv.2410.10665>



Fulbright Spain grantees from a variety of programs celebrate at the 2025 pre-departure orientation in Madrid. Photo by the Fulbright Commission in Spain.

---

**BIOGRAPHY**

Aldan Creo is a Fulbright Foreign Student from Spain who will start his MS studies in Computer Science in Fall 2025. He is interested in the development of safe and fair AI systems that can benefit all, with a focus on natural language processing. His research interests include the detection of AI-generated text, hallucinations and generation artifacts; assessment of conversational risk; and multilinguality in AI. Aside from his academic pursuits, Aldan enjoys debating, traveling and volunteering, he has participated in several associations. Aldan can be reached at [hello@acmc.fyi](mailto:hello@acmc.fyi) and a complete profile can be found at <https://acmc.fyi/intro>.

---