We do not have to accept AI (much less GenAI) as inevitable in education

Emily M. Bender

The systems being sold as 'AI' are not fit for purpose for educational applications and we should not treat it as a foregone conclusion that they represent the future of education. In brief, the technology underlying large language models amounts to little more than a parlour trick and only provides the illusion of 'intelligence'. Deploying such technology in the classroom, especially in resource-starved educational systems, is worse than nothing: on the one hand, large language models are designed to provide bespoke misinformation, and the way they are positioned constructs education as the accumulation of disembodied knowledge. On the other hand, any educational system purchasing them is misdirecting precious resources away from students and teachers, and instead towards the technology industry and the venture capitalists behind it. This is true even if the systems are allegedly provided for free: companies benefit from access to student data, as well as the reputational benefits of benevolently assisting in education.

The visions provided by the tech billionaires might sound appealing, especially to people and governments struggling to find resources to provide quality education to their populations. Sam Altman (2024) promises that the 'intelligence age', driven by his technology, will lead to a world in which 'our children will have virtual tutors who can provide personalized instruction in any subject, in any language, and at whatever pace they need'. Meanwhile, Bill Gates is predicting that advancements in 'AI' mean that within a decade 'great tutoring' will be free (Huddleston, 2025). These promises are empty, based on misapprehensions of both how the technology works and what education is. In this piece, I begin with a quick overview of why the technology can't do what is promised and then turn to the harms that follow if it is used anyway. The tech billionaires are seeking to disrupt, Silicon Valley-style, the communities built out of relationships between students, teachers and families, which are at the core of successful education. It is critical that educators and leaders of education systems bring a critical eye and skeptical attitude towards the sales pitches from AI companies and philanthropic organizations, so that they can protect the students in their care from exploitation and diminished educational services, in the name of 'progress'.

The trick of large language models

Large language models and chatbots built on top of them have been marketed as 'everything machines', nearly-there solutions to all that ails us: robo-doctors to step in when health-care systems are overstretched, robo-scientists to cure cancer and solve the climate crisis, and, of course, robo-teachers that can provide tireless personalized tutoring to every student. This is, in fact, all a con, as Alex Hanna and I write in *The AI Con* (Bender and Hanna, 2025).

What large language models are designed to do is to mimic the way that people use language. Based on very large input datasets, they can output text that takes the form of a medical diagnosis, a scientific paper, or a tutoring session. But the key thing to know here is that these models only ever have access to form: the spelling or sometimes the sounds of words. When we imagine them being 'trained on' or 'ingesting' enormous quantities of text, we

understand that text to be saying something (because we could make sense of it if we read it) and therefore imagine the systems to be 'learning' from knowledge represented in the text. But, in fact, all the systems ever have access to is the form of the linguistic activity that makes up the training data: literally the spellings of words.

The only reason they seem to be doing more is the way they play on how people make sense of language. Whereas it might seem like understanding a text is a matter of simply unpacking the meaning that the author packed into the words; in fact, psycholinguistics shows that the process is quite different and far more involved than that (Reddy, 1979; Clark, 1996). We use everything we know or imagine about the person who wrote the words, everything we imagine to be in our shared common ground with that person, and everything we imagine about what they know about our current beliefs (or those of their intended audience, at least). In that context, we then ask the question: what must they have been trying to communicate by choosing those words in that order? In other words, in order to make sense of language, we must imagine a mind behind the text, and furthermore we do this reflexively and instinctively. So, when we encounter the output of a large language model, we make sense of it by constructing an imagined mind that isn't there. Any 'intelligence' we perceive in these systems is purely a projection of our own cognition and linguistic competence.

Bespoke misinformation

The way that we make sense of language means we are quite vulnerable in the face of synthetic text extruding machines, especially ones designed to take an authoritative tone and sold as having access to all the world's information (Google, 2024). Furthermore, the process called 'reinforcement learning from human feedback' (Ouyang et al., 2022), used to reshape the probabilities associated with specific sequences of words such that the systems are less likely to produce particularly offensive outputs, has the effect of producing systems that tend to output strings that match what the user wants to see. All of this, combined with the fact that the outputs of synthetic text extruding machines are not controllable, means that students are going to get possibly subtly, possibly starkly, different 'information' out of the systems.

This 'information' will be presented authoritatively and convincingly, but without clear traces back to its provenance. Any given piece of 'information' might come from some specific underlying text, might be a decent summary of several texts, or might be simply a remix of words that, in fact, is not supported by any of the source texts. Where previously teachers might have had to deal with a variety of common misconceptions, now they are faced with bespoke misinformation being fed to each student, based on their interests and how they prompt the machine.

Critical engagement and communities of knowers

The lack of provenance information would be a problem even if the outputs were always strictly factual. Setting up a global 'answer machine' that can (seemingly) carry on conversations about any topic construes knowledge as something that exists separately from communities of knowers and education as the accumulation of that knowledge. But knowledge belongs to and is negotiated by communities of knowers (Hoffmann and Bloom, 2016). The source of even seemingly objectively knowable kinds of information matters immensely in making sense of it. Take, for example, the question of the length of the coastline of

some country. The particular value given will depend on how granular the measurement is (do you trace each and every inlet? each rock along the beach in each inlet?), and that, in turn, will depend on why the person measuring chose to do the measurement. The value will also depend on political facts, like where the boundaries of the country lie, boundaries that might well be contested, and so to make sense of the particular value, we need to know the political perspective of the measurer.

Especially when educational outcomes are measured via standardized testing, it is all too easy to see education as the accumulation of knowledge (including both 'knowing that' and 'knowing how'). But the deeper purpose of education, and one that could never be served by 'answer machines' involves knowing how to navigate an information ecosystem, how to understand ideas and positions and how they relate to each other and to the people who hold them, and how to articulate our own ideas and situate them in that broader landscape (Shah and Bender, 2024).

All kinds of synthetic media are problematic

I have been focusing on synthetic text, but all kinds of synthetic media are problematic and, in fact, poisonous to the educational context. None of these systems is built on consentfully collected datasets.[1] Modelling or encouraging their use teaches school kids to devalue the work of artists of all kinds, which was stolen to create them. It also devalues children's own creative expression, suggesting that their drawing, painting, writing, etc. aren't good enough and they should instead replace them with the more polished system outputs.

As a final example, consider the proposal by Alex Banks of The Signal (an organization whose stated purpose is to 'democratize AI education for everyone') to make students 'part of the story' in history lessons (Banks, 2025) by using synthetic video to depict historical events being studied. In contrast to an assignment that asks students to imaginatively depict an historical incident from various points of view, that is, one that provides structure for students to use their imagination to guide how they draw on primary sources, deploying synthetic media this way puts students in a passive role. Beyond that, it is guaranteed to misrepresent the people and events in question, mislead students by failing to demarcate what is known and what is imagined, and also mislead students in a more abstract way by suggesting that details of the past are knowable, which simply aren't.

Disrespectful to students, disrespectful to teachers

To suggest that synthetic text extruding machines are suitable for classroom use (to replace teachers, or just as an auxiliary aid) is to flatten the work of teaching and learning into just the words that are exchanged between student and teacher. While it is true that the words might be the most directly observable part of that activity, they are neither its heart nor where the value lies. Teaching and learning are intrinsically social activities, built around relationships in the classroom. When a teacher chooses words to say to their students, those words reflect not only the concept, technique and encouragement to critical thinking or other communicative goal that the teacher wishes to convey, but also their understanding of where their particular students are currently at, all of their expertise in pedagogical practice, and their care towards the students

1. See https://www.consentfultech.io

they are working with. And all of that is what makes the teacher's words effective. To say the value is in the words themselves is profoundly disrespectful to both teachers and students. It erases the work and expertise of teaching, on the one hand, and paints students as unworthy of caring mentors who help them grow their critical thinking, on the other hand.

Who really benefits? Whose interests should we be protecting?

We are often told that students must learn how to use chatbots or other 'AI' systems lest they get 'left behind'. But this argument presupposes that the future we must run towards is one where evermore aspects of our lives are mediated by technology. We can and should imagine futures where the development is towards other goals, such as more potential for self-actualization, better health outcomes, more environmental sustainability and stronger communities. Education has a role to play in all of those arenas, both through the learning that students do and through the connections that are strengthened among classmates, their families and their larger communities. Technologies of isolation (Gilliard, 2025) that discourage people from turning to each other for information, or to work together to understand information, run counter to these goals.

The world of education is currently besieged by marketing of so-called artificial intelligence solutions. This marketing positions the technology as beneficial, benevolent, magical and the way of the future. But we should always examine marketing with great skepticism. This means always asking: what is in it for them? Why are the tech companies and tech philanthropists so invested in shaping education? On top of lucrative contracts, there are also other incentives, such as access to data about students and further normalization and entrenchment of the power of tech oligarchies (Rhodes, 2025).

Educators and leaders of education systems must put the interests of students above all else. Any funds that are sent to tech companies, in the guise of being 'cheaper' solutions or 'better than nothing', are funds that could have been used for teacher salaries and other material support to education. If the tech is being adopted because it's 'better than nothing', it is always worth being skeptical of that claim, as the tech has documented potential for harm (Baker-White, 2025; Kosmyna et al., 2025). Furthermore, it is always worth asking why the alternative is 'nothing'.

It can be difficult to push back against the inevitability narrative and all of the associated glitzy, extremely well-funded marketing. I believe it can help education leaders to maintain a skeptical, even suspicious, stance, if we shift the conceptualization of education systems as perpetually in need of any resources available and understand them instead as being full of another kind of wealth: students' time and attention and their potential as individuals and as communities. These are extremely valuable and worth protecting. When tech companies or tech-funded philanthropic organizations come sniffing around with 'solutions' predicated on everyone using their software, these solutions are never cost-free in the bigger picture. If there isn't time to thoroughly examine the costs, then the default position can and should be, 'No, thank you'.

References

Altman, S. 2024. *The intelligence age*. Samaltman.com. https://ia.samaltman.com (Accessed 5 August 2025.)

Baker-White, E. 2025. *These AI tutors for kids gave fentanyl recipes and dangerous diet advice*. Jersey City, Forbes. https://www.forbes.com/sites/emilybaker-white/2025/05/12/these-ai-tutors-for-kids-gave-fentanyl-recipes-and-dangerous-diet-advice (Accessed 5 August 2025.)

Banks, A. 2025. *Learning will never be the same*. Sunnyvale, LinkedIn. https://www.linkedin.com/posts/a-banks_learning-will-never-be-the-same-people-activity-7340710510121611264-891j (Accessed 5 August 2025.)

Bender, E. M. and Hanna, A. 2025. *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want*. New York, HarperCollins.

Clark, H. H. 1996. *Using Language*. Cambridge, Cambridge University Press.

Gilliard, C. 2025. Surveillance scholar Chris Gilliard on Facebook's spy glasses. *Remarks on Techtonic with Mark Hurst*. Jersey City, WFMU [podcast]. https://www.wfmu.org/playlists/shows/152175 (Accessed 29 July 2025.)

Google. 2024. *Our approach to Search*. Mountain View, Alphabet. https://www.google.com/intl/en_us/search/howsearchworks/our-approach (Accessed 5 August 2025.)

Hoffmann, A. L. and Bloom, R. 2016. Digitizing books, obscuring women's work: Google books, librarians, and ideologies of access. *Journal of Gender, New Media, and Technology*, Vol. 9. Eugene, University of Oregon. https://hdl.handle.net/1794/26769 (Accessed 5 August 2025.)

Huddleston Jr., T. 2025. *Bill Gates: Within 10 years, AI will replace many doctors and teachers – Humans won't be needed 'for most things'*. Englewood Cliffs, CNBC. https://www.cnbc.com/2025/03/26/bill-gates-on-ai-humans-wont-be-needed-for-most-things.html (Accessed 5 August 2025.)

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I. and Maes, P. 2025. *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. Ithaca, arVix. https://doi.org/10.48550/arXiv.2506.08872 (Accessed 5 August 2025.)

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J. and Lowe, R. 2022. *Training language models to follow instructions with human feedback*. Ithaca, arVix. https://doi.org/10.48550/arXiv.2203.02155 (Accessed 5 August 2025.)

Reddy, M. J. 1979. The conduit metaphor: A case of frame conflict in our language about language. *Metaphor and Thought*. Cambridge, Cambridge University Press, pp. 164–201.

Rhodes, C. 2025. *Stinking Rich: The Four Myths of the Good Billionaire*. Bristol, Bristol University Press.

Shah, C. and Bender, E. M. 2024. Envisioning information access systems: What makes for good tools and a healthy web? *ACM Transactions on the Web*, Vol. 18, No. 3. New York, Association for Computing Machinery (ACM). https://doi.org/10.1145/3649468 (Accessed 5 August 2025.)