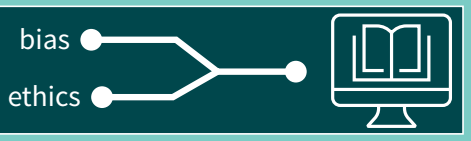
GENEVA GRADUATE INSTITUTE

NORRAG GLOBAL EDUCATION CENTRE



POLICY INSIGHTS

AI AND DIGITAL INEQUITIES

ABOUT THE PROJECT

NORRAG's work on digitalisation and AI in education aims to surface fresh analytical perspectives and under-represented expertise about digital technology and its consequences for education globally. Despite resurgent interest in technology in education research, policy, planning and practice, many areas that are critical to understanding the challenges as well as the benefits of the ongoing digitalisation of education across the globe remain understudied, and the evidence that does exist remains under-shared.

This publication continues the work begun in the first NORRAG Policy Insights (May 2022) on the Digitalisation of Education. Our highly successful event on AI and Digital Inequities Summit in September 2023 proved the appetite for a sister publication focused on AI alone. NORRAG convened a multi-disciplinary expert group to mobilise learning for AI and education policy and practice from critical and under-represented researchers and evidence producers, including from the Global South. This collection shares short pieces authored by leading experts, who provide profound yet digestible insights about AI and digital inequities across all levels of education, and its consequences for learners, communities, educators and policy makers. These experts provide key takeaways that trace future pathways for change and transformation.

More information: www.norrag.org/ai

ACKNOWLEDGEMENTS

This publication was made possible by NORRAG's ongoing and fruitful partnership with the Swiss Agency for Development and Cooperation. The content and perspectives presented in these contributions are those of the individual authors and do not represent views of any of these organisations. We are deeply grateful for the time, energy and insight of all of the experts who contributed to this Policy Insights collection on AI and digital inequities in education, and to Alexandra Draxler for her expert editorial revisions.

ABOUT THE AUTHORS

A full list of the 29 experts from 5 continents who contributed to this publication, including their biographies, is included at the end of this publication. This publication has been coordinated, edited, and introduced by Moira V. Faul, Executive Director at NORRAG.

ABOUT NORRAG

NORRAG is the Global Education Centre of the Geneva Graduate Institute and is a global network of more than 5,600 members for international policies and cooperation in education and training. NORRAG is an offshoot of the Research, Review, and Advisory Group (RRAG) established in 1977 and at the time funded by the International Development Research Centre (IDRC) and Swedish International Development Authority (Sida). The current name was adopted in 1986. Since the move to Switzerland in 1992, NORRAG has been significantly supported by the Swiss Agency for Development and Cooperation (SDC) and the Graduate Institute of International and Development Studies Geneva.

NORRAG's strength lies in addressing under researched questions of quality and equity in key issues in education and development, and in amplifying under-represented expertise particularly from the South. NORRAG's core mandate is to produce, disseminate and broker critical knowledge and to build capacity for and with the wide range of stakeholders who constitute our network. Our stakeholders from academia, governments, NGOs, international organisations, foundations and the private sector inform and shape education policies and practice at national and international levels. Through our work, NORRAG contributes to creating the conditions for more participatory, evidence-informed decisions that improve equal access to and quality of education and training.

Join the global NORRAG network:
www.norrag.org/norrag-network

Learn more:
www.norrag.org

Follow NORRAG on social media:

GENEVA GRADUATE INSTITUTE
NORRAG GLOBAL EDUCATION CENTRE

Chemin Eugène-Rigot, 2
1202 Geneva, Switzerland
+41 (0) 22 908 45 47
norrag@graduateinstitute.ch

NORRAG Policy Insights #04, March 2024

Suggested citation
Faul, M. V. (Ed.). (2024). *AI and Digital Inequities.* Policy Insights #04. NORRAG.

Coordination
Simon Kyeremeh and Anouk Pasquier Di Dio

Design Concept / Production
Alexandru Cretu, Romania

NORRAG is supported by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Agency for Development and Cooperation SDC

GENEVA GRADUATE INSTITUTE
INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT
GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Learn more about the project:
https://www.norrag.org/ai

Version: 03.2024

FOREWORD

Sobhi Tawil
Director, Future of Learning and Innovation, UNESCO

While technology has long been positioned as a set of tools that can improve equity in education, the unfortunate reality is that in many instances and contexts, it widens already existing educational and social divides. The challenge, going forward, is to reverse this trend. Specifically, how can we make AI technology a lever of opportunity in education—a more reliable choice of options to strengthen equity and address persistent disparities across the familiar lines of economic status, gender, language, disability, and ethnicity?

Ensuring all people, and young children most vitally, benefit from effective and relevant education is one of the most reliable investments a society can make to ensure more just, inclusive, and sustainable futures. Only through education can we successfully rebalance our relationships with each other, with our natural environment, and with the technology that we invented. The question at the heart of the papers in this collection is: How can AI technology best serve us in this pursuit?

In an age characterised by an exponential acceleration of digital innovation, getting AI technology 'right' must begin by understanding how our hurried embrace and integration of new technologies often go astray. These papers do this with nerve and clarity. Shining light on the sources and implications of AI and digital inequities in education, they show us how to recalibrate the digital transformation of education towards inclusion and equity.

While technology has long held the promise of solving complex educational challenges, its implementation too often introduces new and unexpected problems—and these can take years or even decades to see with clarity. Social media is a case in point. A generation ago, we let our children rush unaccompanied into lightly regulated AI-powered social media platforms. The costs of this transition have, some twenty years on, become blindingly apparent. Algorithms tailored for adhesion have had serious side effects, including amplified polarising content, factionalism, misinformation, extremism, addiction, and feelings of being disconnected from the world outside a screen. AI technology that was once billed as moving us closer to digital and social utopias is no longer seen as purely benign. We are wiser now.

This volume represents an early effort to confront AI's risks to equality in the context of education. It advances our understanding of the many under-appreciated and still emergent sources of digital inequities while our educational norms around AI technologies remain malleable. Rightfully rejecting technology solutionist thinking, the papers here seek to understand better the flaws and gaps of digital solutions that the powerful technology industry would prefer to obscure. They examine what is currently happening with the digital transformation of education, identify where this transition is subverting our highest aims for education and chart a way forward for AI technology use in education that is more sensitive to rights, protection, and human flourishing.

This NORRAG volume complements global efforts to learn from our still-early brushes with AI to help us steer a surer, safer, inclusive and more equitable future for education in our era of digital ubiquity. UNESCO's recent contributions in this area are many, including the 2023 global education monitoring report *Technology and education: A tool on whose terms?* (UNESCO, 2023a) and *An Ed-Tech tragedy? Educational technologies and school closures in the time of COVID-19* (UNESCO, 2023b). UNESCO has been closely examining what AI technology means for education—and, conversely, how education can shape our reception of, expectations for, and uses of AI. Indeed, education is not merely something that technology "acts on" or forces into new forms. Education can also "act on" technology, shaping its development, regulation, business models, and purposes with equal or greater force. This message of agency and the importance of ensuring technology extends and enriches human-centred education anchors UNESCO's (2023c) *Guidance for generative AI in education and research.*

Questions about AI—its uses and misuse—stand at the very centre of contemporary debates about education. This volume rightly calls on us to demand more from AI technology to revitalise the humanistic aims of education and to redouble our commitment to assuring educational equity. Whether a technology solution, however large or small, strengthens educational equity must be a litmus test for its large-scale implementation. AI affords us a powerful and evolving set of tools to improve education and open new opportunities for learners and teachers. But it also requires being clear-eyed about the risks and disadvantages that this collection brings into sharp focus.

Ultimately, this collection shows us how to better orient policy and practice to make AI an educational ally and a tool to bridge—rather than exacerbate—long-standing learning inequities.

CONTENTS

List of figures

List of tables

INTRODUCTION: FORGING AN EQUITABLE AI FUTURE IN AND THROUGH EDUCATION

Moira V. Faul
Executive Director, NORRAG, Switzerland

"Using technology is as essentially human as making ethical decisions; let's lead with ethics."

AI is many things: from spellchecks to text- and image-generation apps to the public and private infrastructure on which these services rely. While AI has been around for decades, the public release of ChatGPT at the end of November 2022 shifted public and policy conversations in important ways. AI is here to stay, and education stakeholders are duty-bound to examine both the opportunities and the pitfalls that AI entails.

Inequities in AI reflect real-world (human-created, non-AI, non-technological) inequalities. AI holds the potential to address certain inequalities. Nevertheless, AI can reinforce and amplify existing inequalities while creating new harms if we do not scrutinise questions of equity in the design and deployment of AI systems now.

Addressing the ethical issues posed by current developments in AI does not mean abandoning it. At NORRAG, we use AI to generate webinar subtitles in multiple languages to increase accessibility. I use AI voice recognition to avoid damaging my wrists with excessive typing. Few of us switch off the spellcheck function when we write.

AI in society is not a binary on/off

Rather, the authors in this collection foreground the ethical challenges that arise with regards to AI use in education whether as a private, public or common good, and invite you to put human and planetary flourishing at the heart of AI decision making, development and deployment.

Different purposes for AI do not need to be incommensurate—AI design, use and monetisation could be oriented to enable individual efficiency and also social effectiveness; to generate reasonable business profits and also human flourishing alongside environmental and labour protections (Radu, 2024; Whittaker, 2021). Nevertheless, as AI development and deployment are currently configured, access to and freedom from exploitation in AI are unevenly distributed in ways that systematically exclude the most vulnerable.

The contributors to this Policy Insights collection provide key takeaways for education stakeholders and decision makers on some of the main challenges concerning inequalities, putting the value of humans and our planet at the centre of our use and governance of AI, and its underpinning value creation models.

AI "evolution": Driven by humans, and human-generated developments in computing power and data availability

AI needs HI (human intelligence). The work of AI researchers and developers can be categorised into distinct generations, with notable shifts in approaches and methodologies in each.

The term "artificial intelligence" (AI) was first coined in 1955 at Dartmouth College, USA, when a group of pioneering academics gathered to explore the idea of creating machines that could mimic human intelligence (McCarthy et al., 1955). Their conjecture was "that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." (p.1). In the first generation (1950s–1980s), human experts encoded their knowledge into a set of explicit rules that govern the behaviour of an AI system. These systems excelled in rule-following and symbolic reasoning tasks but struggled with handling uncertain, complex and ambiguous situations. Progress was slow due to the

complexity of human cognition that these researchers were attempting to describe and simulate.

In contrast to explicit programming, the second generation of AI researchers (1980s-early 2000s) shifted towards developing so-called machine learning algorithms and large datasets. These algorithms use statistical techniques to generalise patterns from massive datasets of examples, and generate predictions of what an appropriate output might be in novel situations. Researchers overcame the limitations on computational power and the size of available datasets that constrained the scalability and effectiveness of these approaches to produce the third and current generation (starting in the mid-2000s). Researchers working within the broader machine learning framework have developed "deep learning". This approach leverages many layers of neural networks inspired by the structure and function of the human brain without attempting to simulate human intelligence as in the 1950s.

In 2022, Microsoft's OpenAI launched ChatGPT (a user-friendly chatbot) and DALL-E (a text-to-image model), and Alphabet's Google launched Bard and then Gemini. Other non-profit models exist: in contrast to these commercial tools, HuggingFace (launched in 2016) is an Open Source and collaborative community for co-creating AI tools.

Narratives of AI "evolution", "learning" and "decision making" tend to hide the fact that humans developed the large datasets and powerful computing resources needed for generative AI, along with the required advancements in neural network architectures and training algorithms. Furthermore, both sides of future-focused narratives (doomers vs. boomers) assume that we need to focus our energy on protecting humans from future harm that may arise. Focusing on the future, however, ignores the actual inequities now that pose as much of a threat if we fail to address them.

Inequities in AI

Who currently has access to AI? The question is broad and encompasses: access to AI technology; access to the possibility of gaining benefit from current AI technologies or developing more in the future; access to researching or critiquing the technology; and access to decisions on AI development and deployment, including over the governance, financing and the allocation of benefits. The 'who' here includes people who are marginalised within countries and companies, as well as the majority of countries and companies that are marginalised from participating in and decision making about our AI present and futures.

Who is currently represented in AI? AI training sets encode the values of privileged members of WEIRD (Western/White, educated/English-speaking, Industrialised, Rich, Democratic) corporations who design these systems and profit from them (Birhane et al., 2022; Dixon-Román et al., 2020; Henrich et al., 2010; Raji et al., 2020). This selective inclusion reinforces and promulgates dominant epistemologies, further marginalising other ways of knowing and doing (Mahelona et al., 2023;).

Who and what are exploited by current AI? Students currently cannot give consent to their data being used for profit by platforms that are mandated by their institutions (Boly Barry 2022; Mejias & Couldry, 2024; Williamson, 2019). AI use by all of us and AI companies' data centres and processors divert enormous amounts of electricity and water for cooling away from humans and places that need it (Birch, 2022; Luccioni et al., 2023). Exploitative working conditions abound for data workers in low- and high-income countries, without whom AI tools would not be marketable to schools (BBC, 2021; Luccioni, 2023). Copyright challenges arise where open access or pirated articles and books are fodder for LLMs but are not cited.

What research is currently conducted? AI corporations' enthusiastic—but incomplete—reporting of their work is often uncritically repeated by news sources (Bender & Hanna, 2023). Internal research into the impacts and ethics of companies using and developing algorithms and training sets are suppressed, and ethics teams are disbanded (Financial Times, 2023; MIT Technology Review, 2020). The value creation models of AI—and other technology—companies are opaque, particularly regarding the monetisation of users' attention and data (Faul, 2023; Montag et al., 2019; O'Reilly et al., 2023; Pidoux & Dehaye, 2022). Without rigorous, independent research and transparency, developers cannot be held accountable for the experiments they undertake, for the effects of their products and for taking corrective action if necessary.

AI futures for human and planetary flourishing

Much of our current narrative personifies AI, imbuing it with human characteristics while presenting humans in more mechanistic terms. Nevertheless, generative AI generates (Tucker, 2022); it does not think, predict, create, decide, hallucinate, understand or make meaning (Bender & Koller, 2020). A secondary effect of personifying AI is to diminish the possibility of humans to act—the agency that we will need to use if we are to seize this key moment to address AI's digital inequities. Developing and deploying AI requires infrastructure and software that is developed and provided by humans; humans make decisions about what kind of AI we will develop.

According to many technologists, we have entered the phase in the Gartner (2023) hype cycle that represents a coming decade of AI experimentation and deployment. These experiments are human experimentation; therefore they require the same guardrails as any other human experiments (Wired, 2021). AI that improves individual efficiency and reasonable profit-making can function within guardrails for societal effectiveness and human and planetary flourishing. Humans can decide to take action to achieve that.

How can we put equity at the core of AI development? This collection brings together 29 authors from 5 continents who provide key takeaways for decision makers, educators and students seeking to support more equitable and ethical design and deployment of AI in education across the full ecology of ethical concerns (Figure 1).

How do we get the next ten years right?

In November 2021, 193 states adopted UNESCO's *Recommendation on the Ethics of AI* (UNESCO, 2023d), and in February 2024, eight global technology companies[1] committed to implementing this global standard in developing and deploying AI technology. UNESCO's (2023c) *Guidance for generative AI in education and research* applies these insights to education, and the Global Monitoring Report (UNESCO, 2023a) demands "tech on our terms". Learning the lesson from unregulated social media in the 2010s, the European Parliament demanded an end to the "addictive design of online services" in June 2023 and 41 of the 50 United States of America are suing Meta (the parent company of Instgram and Facebook) for exploiting young people's vulnerabilities. More over, the European Union adopted the AI Act on 2 February 2024, which imposes bans on or limits the use of high-risk technologies and requires the stress-testing and transparency of legal AI technologies. Thus, in addition to exhortations for individuals to "click wisely," humans can change how technology industries develop and deploy AI more wisely.

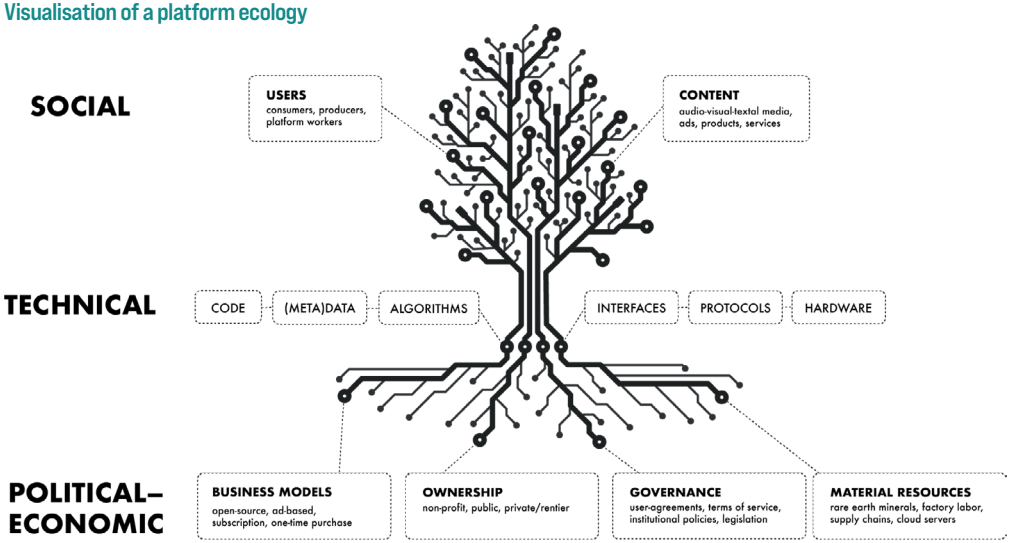
These recent examples show that older human technologies (such as state regulation, corporate governance, collective action and legal challenges) can be used to govern this technocosm ethically (Runciman, 2023) and in the common interest to overcome the allocational and representational harms that are built into current AI development (Bonini & Treré, 2023). AI governance must include decisions and decision makers that safeguard human, social and environmental ecosystems, and ensure that human and planetary wellbeing guide the development and deployment of the algorithms, training sets and energy-hungry processors on which AI depends.

> *"Imagine and craft the worlds you cannot live without, just as you dismantle the ones you cannot live within"*
> Ruha Benjamin

In preparing this introduction, I asked ChatGPT (3.5) the question "Does AI improve equality?" Part of the answer generated was: "Policymakers, technologists, and society as a whole play crucial roles in shaping the impact of AI on equality." That is the challenge the contributors take up in this collection. It is also the challenge they pass onto you: to take action in your spheres of influence early enough to make a difference.

Note: Parts of this introduction were first published in the Geneva Graduate Institute's Globe magazine (Faul, 2023). During the preparation of this work, the author used ChatGPT-3.5 to generate an answer to a specific question, which is reported at the end of the introduction. After using this free tool, the author reviewed and edited the content and takes full responsibility for the content of the publication.

Figure 1
Visualisation of a platform ecology



Source: Nichols and Garcia (2022).

Footnote

1. GSMA, INNIT, Lenovo Group, LG AI Research, Mastercard, Microsoft, Salesforce and Telefonica.

Table 1

Governing AI for humanity: Interim report of the UN AI Advisory Body, convened by Secretary-General António Guterres

GUIDING PRINCIPLES	
Inclusivity	all citizens, including those in the Global South, should be able to access and meaningfully use AI tools.
Public interest	governance should go beyond the do no harm principle and define a broader accountability framework for companies that build, deploy and control AI, as well as downstream users.
Centrality of data governance	AI governance cannot be divorced from the governance of data and the promotion of data commons.
Universal, networked and multistakeholder	AI governance should prioritize universal buy-in by countries and stakeholders. It should leverage existing institutions through a networked approach.
International Law	AI governance needs to be anchored in the UN Charter, International Human Rights Law, and the Sustainable Development Goals.

Source: UN Secretary-General AI Advisory Body (2023)

1

SCRUTINISING AI

In this section, the authors share insights on what should be scrutinised as part of the development of AI for human and planetary flourishing. Alan F. Blackwell examines the value of large language models that may fluently generate plausible—but epistemologically ungrounded—text and the need for parallel, social advances that cultivate stronger ethics around evidence, logic and language. Felicitas MacGilchrist takes us behind the scenes of AI-powered educational technologies, showing us how they can be reimagined to construct more just futures. Lina Markauskaite draws our attention to the tensions between wealth- or wellbeing-oriented purposes of AI in education. Janja Komljenovic advocates for developing AI that is more appropriate for education, and that develops collective and democratic decision making. Jeremy Knox shows that so-called personalised learning can result in decreasing human agency, increasing surveillance and deepening inequality.

RECOMMENDED READINGS

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21), 610–623. https://doi.org/10.1145/3442188.3445922
- Bock, A., Breiter, A., Hartong, S., Jarke, J., Jornitz, S., Lange, A., & Macgilchrist, F. (Eds.). (2023). *Die datafizierte Schule*. Springer VS. https://doi.org/10.1007/978-3-658-38651-1.
- Broussard, M. (2018). *Artificial unintelligence. How computers misunderstand the world*. MIT Press.
- Common Sense Media (n.d.). All Common Sense Privacy Evaluations https://privacy.commonsense.org/evaluations/1
- Markauskaite, L., Carvalho, L., & Fawns, T. (2023). The role of teachers in a sustainable university: from digital competencies to postdigital capabilities. *Educational Technology Research and Development*, 71(1), 181-198. doi: 10.1007/s11423-023-10199-z
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

WHAT ARE LARGE LANGUAGE MODELS GOOD FOR?

Alan F. Blackwell
University of Cambridge, UK

Key takeaways:

- Large Language Models (LLMs) offer utilitarian value but require careful control: LLMs can automate "bullshit jobs" like generating reports and could free humans for more valuable work. However, their outputs might not be true, requiring safeguards and critical thinking to avoid wasted effort and misinformation.
- LLMs raise epistemological concerns: LLMs can generate persuasive but factually incorrect text, potentially hindering knowledge and understanding. We should see them as tools for exploring communication challenges, not as reliable sources of truth.
- LLMs reflect existing problems and require diverse perspectives: The misuse of language for self-promotion and manipulation predates LLMs. We need to address issues like weak critical thinking and prioritize diverse perspectives in AI development to avoid perpetuating biases and misusing technology.

As large language models (LLMs) such as ChatGPT develop impressive linguistic fluency, policymakers across nations are grappling with urgent questions regarding these technologies' societal impacts and benefits. This article considers the future value of such LLMs on two key dimensions—namely, the utilitarian value of automated text generation and the epistemological value of linguistically skilled but ungrounded systems.

While ChatGPT and other LLMs exhibit impressive linguistic fluency, their competence is stochastic rather than grounded in logic or facts (Bender et al., 2020). As such, it should not be assumed that the information they generate is reliable or accurate. However, LLMs provide a useful benchmark for further language research, and they may have utility if applied judiciously with appropriate safeguards.

On the utilitarian dimension, we should consider whether the linguistic abilities of LLMs can generate tangible value. The sociologist David Graeber (2019) argued that over 30% of jobs require employees to produce reports nobody reads, relay messages with no purpose and involve bureaucratic tasks that could easily be automated. Graeber revealed how late capitalism breeds these "bullshit jobs" that lack social utility. In that light, fluent nonsense generated by LLMs could displace the significant amounts of wasted human effort currently spent producing what Graeber considered bullshit. If controlled appropriately, LLMs might take on large volumes of English-language busywork, freeing up human capacity for less automatable roles. This suggests a potentially vast utility for LLMs in replacing swaths of duplicative commercial and governmental administrative work. While attendant risks exist regarding the displacement of existing workers, at a societal level, the productivity gains could be substantial if we rethink what outputs merit human rather than algorithmic effort.

On the epistemological dimension, we should consider what LLMs contribute to human knowledge and understanding. The philosopher Harry Frankfurt (2005) argued that bullshit represents a greater enemy of the truth than lies because when crafting persuasive statements, the bullshitter ignores truth altogether. LLMs such as ChatGPT exhibit a strong ability to generate coherent, seemingly logical text without any underlying fidelity to facts or evidence (Shanahan, 2022). As texts grounded primarily in internal narrative plausibility rather than external correspondence with reality, their outputs qualify as bullshit in Frankfurt's (2005) technical sense. This confronts us with deep questions about the nature of comprehension.

If read uncritically, LLM-generated texts could propagate misconceptions masked by rhetorical competence. Rather than merchandisers of truth, LLMs should be seen as tools for exploring the interconnected roles of persuasion, perception and verification in the project of enacting meaningful communication.

As tools, LLMs reflect the nature of their training data. If optimised for social media streams rife with misinformation or self-promotion, they will reproduce such distortions at scale. However, if carefully curated and monitored, LLMs can assist human authors in benign ways. We should be cautious in anthropomorphising their linguistic talents or ascribing to them intentions or agency. Rather than 'thinking' entities meriting rights or responsibilities, we should see them as advanced autocomplete algorithms with both promise and perils requiring ongoing governance.

We should also note that risks related to the ungrounded eloquence of LLMs have analogues within existing political systems. Well before this new wave of AI, Frankfurt and Graeber observed cultures of bullshit taking root across commercial and governmental institutions. Officials frequently elevate style over substance when crafting speeches, policies and public messaging. The incentives of power can override habits of critical thinking, even for leaders well

aware of their audiences and contexts. Thus, while new technologies magnify the risks of mass deception and meaningless work, the underlying issues reflect timeless gaps between linguistic competence and reliable knowledge. As LLMs progress, their capacities mainly underscore the need for parallel social advances, cultivating stronger ethics around evidence, logic and language. True comprehension requires the epistemology to be judged behind outwardly skilled rhetoric. Across fields from science to civics, we must continually reaffirm the principles of intellectual honesty and collective enlightenment against threats both novel and conventional.

Additionally, in assessing the ultimate capabilities of LLMs, we should be cautious about unexamined assumptions embedded within AI aspirations such as "artificial general intelligence" (AGI). The notion of intelligence as a single, quantifiable attribute has historically been intertwined with racist psychometrics used to justify oppression (Saini, 2019), and even recent AI testing paradigms represent cultural constructs centred on skills such as verbal fluency, abstract puzzle-solving and decontextualised pattern recognition. As LLMs appear increasingly adept in these narrow domains, we risk falling into cycles that over-extrapolate, defining the highest intelligence in terms of performance on tests designed by and for one privileged slice of humanity. Before attempting to universally instil advanced reasoning across contexts, we should question what modes of understanding we choose to value and whose notions of intelligence hold sway in our metrics. Otherwise, we will find ourselves not with artificial general intelligence but artificial intelligence of and for a very particular "generality", as imagined and promoted by AGI sponsors. This demands more wise reflection on how we integrate diversity of thought into the design and governance of supposedly "intelligent" systems.

Note: This article is a demonstration and expansion of principles introduced in an earlier blog entry (Blackwell, 2023). The preparation of the current version employed multiple automated language-processing tools, including functions for spelling and grammar correction, summarisation, paraphrasing and predictive text. That mechanically-assisted text was then further adjusted with over 100 modifications made by a human copy editor. Many of these topics, extending to technical and policy alternatives, are discussed at greater length in a forthcoming book (Blackwell, 2024).

AI IN EDUCATION: HOW DEVELOPERS AND LEGISLATORS MAKE FUTURES

Felicitas Macgilchrist
Re:Lab, the Praxis Lab for Reimagining EdTech Futures and University of Oldenburg, Germany

Key takeaways:

- Legislate not only for the detrimental impact of AI on individuals or collectives but also for AI's societal harm.
- Initiate collective procurement processes that put the power to shape bespoke software into the hands of educators. Consider making co-design processes mandatory.
- Establish AI diversity evaluations for EdTech.
- Reclaim critical technical infrastructure as public infrastructure orientated to the common good.

Artificial intelligence (AI) is a seductive narrative, a concept buzzing with the novelty and transformative power that many people desire for schools and other educational institutions. The methods and technologies themselves (from those powering Google Search to generative AI) are, however, far more diverse and controversial than the image of AI as a single "thing" in advertising and the popular imagination. In addition, these methods and technologies are brittle; they rely on proxies, inaccuracies and uncertainties and scale up the bias, injustice and harm of today's societies.

While understanding the use of AI in practice is imperative, it is also crucial to look "behind the scenes," to understand how AI-powered educational technologies (EdTech) are being developed. Since these technologies are making futures, they can thus be reimagined to make more just futures. Three issues are key.

Diversity and ethics

First, developers operate within the social norms, values and "common-sense" knowledges of their contexts. The AI workforce with the power to implement change—from CEOs and CTOs to system designers—is not diverse, and only a small proportion of developers have lived experience of discrimination that might feed into their professional practice. While computer science degrees increasingly include ethics, these courses are often depoliticised, and the insights are rarely transferred into a radically ethical practice. Chatbots and other forms of generative AI, for instance, are trained on datasets including the language and images from mainstream culture—that is, with the legacy of, for example, racist, heteropatriarchal, classist, ageist and colonialist norms.

What can be done? Nathalie A. Smuha (2021) outlined legislation for AI that not only focuses on harms to individuals or collectives but on AI's societal harm. Drawing on environmental law, she proposed a similar approach to AI. This would move regulation beyond the individual to address the impact of AI on society at large.

Solutions for educational problems

Second, software engineering is solutions focused. While this initially sounds positive, it can mean that developers preferentially address problems for which a technological solution is relatively easy to develop. Technical solutions cannot solve social problems. Thus, solutions are created for things that are not actually problems for educators or learners. For instance, developers of an adaptive maths software designed it to compensate for teachers and reassure students that it is okay to make mistakes—but maths educators already reassure students that it is okay to do so. Additionally, in an inverse logic endemic to much product development, educational institutions are encouraged to find ways to integrate available AI solutions into their daily practice, rather than first identifying a (technical) problem in their daily practice for which an AI solution can then be developed.

What can be done? Collective procurement processes should be initiated (at the regional, national or supranational levels) to identify priority issues for teachers, students, parents or school leaders, for which technical solutions can be found. Procurement experts would then collectively initiate procurement processes for bespoke software developed to the users' specifications. If AI in education is the growing market it is said to be, then educational institutions can use their collective power to steer the development and procurement of the software they need.

Co-design with marginalised users

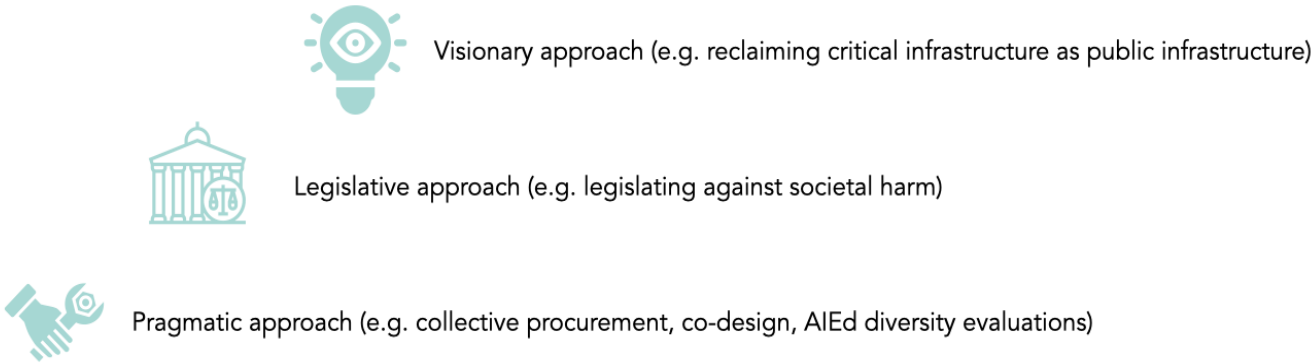
Third, whether they are developing for learners, teachers or educational organisations, commercial EdTech teams tend to orientate to the average client, the majority client or users with high purchasing power. They rarely foreground the most marginalised users or those users for whom the product may create barriers to learning. Product-testing and living labs where products are designed and developed are commonplace. However, companies rarely engage in resource-intensive co-design processes that involve students, families and teachers in historically minoritised or underserved communities.

What can be done? If collective procurement processes are established, calls for tenders could include a mandatory co-design phase, with providers required to specify how they would recruit and compensate diverse users. Alternatively, a not-for-profit set of AI diversity evaluations could be established, similar to Common Sense Media's "Privacy Evaluations". This would include information on the diversity of participation and representation in AI-based EdTech. Users can browse or search the evaluations to guide their purchasing. A more visionary approach would consider the resources we know as AI (e.g. the large language models underlying generative AI) to be critical infrastructure for the common good that should be reclaimed as public, rather than for-profit, infrastructure.

Robust design changes

Overall, beyond the current practices of supporting critical literacies and urging EdTech companies deploying AI to voluntarily self-regulate their ethics, more robust design changes are possible. Policy actions need to orientate not only to the individual end user but far more to forms of change at the public, relational and collective levels. These policy actions span from pragmatic (collective procurement, co-design, AI diversity evaluations) through legislative (societal harms) to visionary (reclaiming critical infrastructure as public infrastructure). These changes will remake futures for AI in education, and thus for the common good across the planet.

Figure 2
Policy for robust design changes



Source: Author

AI, WEALTH, AND WELLBEING: SOME CRITICAL CONSIDERATIONS FOR SUSTAINABLE EDUCATIONAL ECOSYSTEMS

Lina Markauskaite
The University of Sydney, Australia

Key takeaways:

- Reorientate AI development and educational responses from wealth-orientated towards well-being-orientated values.
- Provide equitable access to AI-powered tools for all students and teachers.
- Create opportunities for students and teachers to develop well-rounded AI capabilities.
- Build knowledge infrastructure to support research-informed, democratic decision-making practices concerning AI in education.
- Require AI companies to offset their created burden on education and contribute to the regeneration of equity.

Will generative artificial intelligence (GenAI) help reduce educational inequities or widen existing gaps even further? The answer to this question depends on the ecosystem that AI and educational technology (EdTech) companies, policymakers, educational institutions, communities and other stakeholders will co-create. This ecosystem will be shaped by AI tools and other material realities and also by teachers' and students' capabilities, knowledge practices and shared values. I outline four critical considerations, starting from the material realities and finishing with the shared values.

Material realities: AI tools are not free

Let us make it clear: most powerful AI tools are not free or cheap, and AI inequities—between those who can and cannot afford to pay monthly fees—'stack up' on top of other digital inequities in education. Already at this early stage of GenAI development, there are a number of AI-powered tools that can help students with their learning. Writing tools such as Grammarly, research reading tools such as Scholarcy and ChatPDF and presentation development tools such as Gamma are becoming desirable additions to every college student's 'digital backpack'. These tools could empower those most educationally disadvantaged, but at the moment, they enable the privileged to gain further educational advantage.

Teachers' and students' capabilities: Allowing the use of AI tools is a half-solution

The initial response to the proliferation of GenAI has mainly focused on preventing its misuse and decisions of whether to allow or ban its use. While educational institutions increasingly permit and even encourage teachers and students to use GenAI tools, what kind of AI capabilities they need and how they can develop them are left largely outside these decisions. It is important to note that simply allowing teachers and students to use AI tools is only one part of the solution. To use AI resourcefully, teachers and students need to understand the computer algorithms that power these tools and how AI intersects with various aspects of human life, including culture, cognition, nature, economy and politics. Such capabilities cannot be developed by osmosis. Forms of learning that engage teachers and students with technical, humanistic, social and other aspects of AI are critical (Markauskaite et al., 2022).

Knowledge practices: Research and diverse perspectives matter

Current policy decisions about GenAI in education have been driven by a sense of urgency, thereby relying often on what is said to be common sense and political consensus about what is "good" for learners. Evidence of how students learn with AI-empowered tools lags significantly behind policymaking and practice, and there is a lack of a strong research agenda that accompanies AI developments. Decisions about AI in education should be underpinned by research knowledge from different disciplinary fields and robust democratic processes that engage with diverse perspectives, including industry, community, students, teachers and experts. Enabling such decision-making practices requires the building of new knowledge infrastructure, of which there is currently a lack (Markauskaite, 2010).

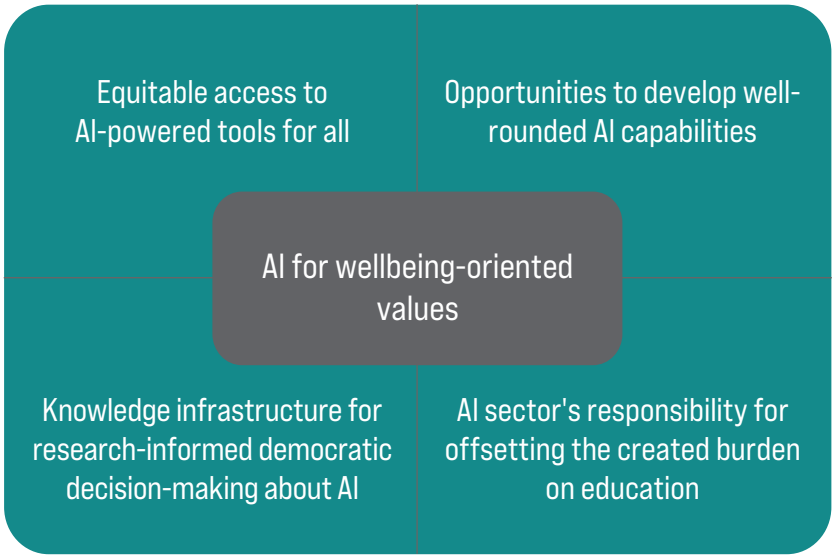
Shared values: Well-being cannot be added on top of wealth-orientated market values

Despite the recognition that AI must benefit all of humanity, most AI developments have been fuelled by economic growth and other wealth-orientated market values. Current policy responses to AI in education have been mainly concerned with absorbing the impact of fast AI developments and ensuring that education continues to prepare students for employment in this economy. However, it is essential to recognise that equitable, high-quality education and other sustainable development goals cannot be achieved by adding them on top of these wealth-orientated economic agendas. Sustainable development requires reorientating AI developments and educational responses towards well-being-orientated economic and educational values (Markauskaite et al., 2023; Shrivastava & Zsolnai, 2022). AI companies have been shifting the burden of challenges caused by fast AI developments onto other actors in the educational ecosystem, such as students, teachers, educational institutions and research agencies. To restore fairness, these companies should take responsibility for offsetting their created burden.

Regenerating fairness

In conclusion, the inherent tensions between wealth- and well-being-orientated purposes require educational responses concerning AI developments to be reorientated towards the regeneration of fairness. Rather than coping with the consequences of AI developments, educational responses should focus on discovering paths by which they can contribute to a well-being-orientated society and economy. Access to AI tools, teachers' and students' capabilities and democratic, knowledge-informed decision-making practices are at the core of such responses. However, the education sector alone, being in an unequal position, cannot address these challenges. Action from the AI sector is needed.

Figure 3
Orienting AI towards wellbeing



Source: Author

DEVELOPING GENAI FOR EDUCATION

Janja Komljenovic
Lancaster University, UK

Key takeaways:

- Consider if GenAI specifically for education is needed.
- If GenAI for education is developed, use it as an opportunity to mitigate against risks and harms.
- Allow collective decision-making on the creation and governance of GenAI for education, including the rights of authors of material used to develop the models.
- Set up an evaluation system for GenAI in education, monitor its effects, and ensure effective intervention when needed.

Generative artificial intelligence (GenAI) is expected to change education profoundly. Some organisations and individuals welcome education disruption and expect GenAI systems to act as roboteachers or personal learning assistants, produce hyper-personalised content, enable automatic assessment and much more. Others are more cautious and ask for discussions on how GenAI should be used, what its purpose is and how we can ensure that it is productive and supportive of human flourishing. While education institutions, educators and policymakers grapple with the use and effects of of GenAI in education, GenAI is already being rolled out into digital tools routinely used in educational settings, such as virtual learning environments.

Education is consequential for GenAI

The education sector is one of the biggest potential markets for AI, compared to other domains. There are 1.5 billion students worldwide, around 85 million teachers and millions of administrators. How students, teachers, schools and universities adopt GenAI is consequential for companies developing it and their investors. Until domain-specific models are developed, we can expect that the most successful generic ones, such as ChatGPT, will be used in education. Hence, education is contributing to the expansion of generic GenAI systems and consequently perpetuating their well-documented risks, including intellectual and copyright violations, privacy concerns, concerns over commercial sensitivity and algorithmic bias. Other potential risks and challenges include content moderation and the exploitation of precarious workers from the Global South; the problem of fuelling the discourse of AI panic and catastrophising, which can distract from thinking things through; the environmental impact and energy consumption of GenAI; the black-boxing of operations; and the lack of evaluation of GenAI software and its impact on societies.

What can teachers and students do? Is it truly enough that they are aware of these risks and challenges? Should the onus be on them to make a personal decision if and how they use GenAI without having any power to influence any of the challenges? Perhaps one of the options to better support students and teachers would be to develop GenAI specifically for education.

Developing GenAI for education

Developing GenAI, including large language models, needs an enormous amount of data and artefacts on which it is trained. For example, ChatGPT used materials collected from the web, webpages, internet-based books and Wikipedia. GenAI is not a task or domain-specific tool, although it learns as people use it. There are ideas about developing GenAI specific to particular domains, such as health care, to better serve particular sectors' specific needs. However, if GenAI were to be developed specifically for education, then education actors should find a way to address at least the following questions internationally:

- By whom would it be developed, and with what motivation?
- With which data and material (e.g. textbooks, teaching materials, student assignments)?
- How could people who have created these data and material for training have a say if they want to contribute, and if so, with which of their artefacts specifically?
- What rights would they have?
- Would they be compensated?
- How and by whom would the design, development and innovation process be monitored?

- How would the impact be evaluated and acted upon?
- Who would pay for developing such systems?
- How would potential risk and harm be addressed?

If GenAI were to be developed for education, it would provide an opportunity to address the specific challenges and risks mentioned above. It would offer the potential for collective and democratic decision-making about the key questions listed here.

Beyond the hype

Regardless of which exact version of GenAI will be used in education in the long term, whether generic or education specific, its impact on education should be monitored from the view of temporality and scale.

- How can we effectively monitor the immediate, mid-term and long-term impacts?
- How can we monitor the impact at different levels, including at the individual, group and sector levels?
- What are the effects on learning and knowledge?
- Most importantly, how can we ensure that we can effectively intervene?

Any interventions are hard if particular features are already integrated into the digital platforms routinely used in schools and universities. I have raised more questions here than answers, but it could not be otherwise at this time of GenAI consolidation. These are urgent challenges and questions that educators and policymakers must address, and they go beyond concerns around immediate use, such as potential cheating.

AI-DRIVEN "PERSONALISATION": PANACEA OR PROBLEM FOR EDUCATION?

Jeremy Knox
University of Oxford, UK

Key takeaways:
What kind of questions should we be asking about personalised learning?

- Who has a voice in deciding how AI technology functions in education? Does the technology we use promote and embody the idea that students should be active agents in their learning?
- What happens to the data traces that students (and teachers) leave behind in AI technologies? Are the benefits of personalised learning worth the price of intensive data surveillance?
- Aside from the abstract potential of personalised learning, how is AI being used in real contexts on the ground? Who is able to exploit greater benefits and how, and who is exposed to greater exploitation?

Central to the promotional claims of various AI-driven education technologies is the idea of personalisation, which involves the tailoring of educational materials and assessments to individual students. This is often achieved with the aid of machine learning techniques that collect the data traces left behind by learners using a particular software platform and process this information to predict individualised "optimal pathways" through a given set of resources.

"Personalisation" has become a key term employed by intergovernmental organisations, such as the OECD (Nemorin et al., 2023). Indeed, the term is collectively endorsed and imbued with significance by a range of actors, including policymakers, academic institutions, technology companies and charities (Davies et al., 2021), despite there being little in the way of consensus about its educational benefits.

On the surface, personalisation alludes to long-established educational idea(l)s of the one-to-one relationship between teacher and student (Friesen, 2020) and reflects more recent pedagogical approaches such as differentiation and student-centred learning or concepts such as "life-long" learning (Davies et al., 2021). Personalisation has an 'interpretive flexibility' that allows its "potential power to eclipse its current value" (Davies et al., 2021, p. 550, emphasis original). In such a way, it is often assumed that AI technologies work against the negative aspects of an industrialised model for education by offering educational content tailored to individual needs and abilities, as opposed to the delivery of standardised curricula. It is therefore sometimes suggested that personalisation explicitly supports less advantaged students—for example, through the provision of automated extra-curricular private tuition (Nesta, 2019).

However, policymakers and practitioners in education should be cautious about the extent to which AI-driven personalisation straightforwardly delivers educational benefits, particularly with respect to notions of equity, fairness and justice.

Control

AI-driven personalised learning does not position students as active agents in the educational process. Such systems pre-define learning content, assessments and achievement targets such that students are better understood as passive recipients of automated decision-making. In this sense, personalisation is better understood as "optimisation" (Bulger, 2016), where machine learning techniques calculate a "best fit" sequence of materials and assessments based on a diagnostic assessment of the student in question combined with aggregated measures of previous users of the software. As such, the step-by-step decisions that govern the progress of students in these personalised learning systems are tightly controlled by data-driven processes, which are often hidden and proprietary, with very little opportunity for both students and teachers to have a voice.

Surveillance

As with other data-driven AI technologies, successful personalisation in education is predicated on the collection of huge volumes of data. The increasing use of such systems therefore places students under unprecedented levels of surveillance during their educational lives, since their activities and behaviours are recorded in fine-grained detail, not only

within dedicated software platforms but also potentially across the built environment of the "smart" school (Williamson, 2015) and surrounding "smart" city (Currie et al., 2022). A narrow focus on the promises of personalised learning deflects attention from the en masse surveillance through data that characterises our times and divides society by shifting power to those with AI expertise (Zuboff, 2019). Whatever the benefits of personalised learning for the individual, they appear to be gained at the expense of a collective surrendering of privacy.

Inequality

Crucial questions need to be asked about how the use of AI in education challenges or maintains existing disparities in educational provision and opportunity. The rhetorical sleight of hand performed by the term "personalisation" is the inference that everyone is included and that the technology produces a more precise and superior outcome for all. However, as a range of work on the political economy of education technology demonstrates (e.g. Komljenovic et al., 2024; Ramiel & Dishon, 2024), new markets for educational data, shifting regulatory landscapes and already existing educational disparities are shaping how and where AI is deployed. The experience of exactly the same kind of AI-driven personalised learning will likely differ significantly across, for example, underperforming rural schools where the technology is deployed explicitly to address a deficit in qualified teachers (see Knox, 2023) and elite private education institutions that can assert agency over how and why the technology is used (see Century, 2021). While the potential of personalised AI in education has frequently been emphasised (e.g. Luckin et al., 2016), more attention could be paid to the idea that the potential of AI itself is unequally distributed.

2 USING AI FOR HUMAN FLOURISHING

Emma Ruttkamp-Bloem of the UN Secretary General's Advisory Group on AI discusses enabling students to engage with AI in responsible and critical ways and shows how we can take advantage of real-world measures to address these digital inequities. Lucila Carvalho recommends engaging with human values and active participation in learning about and with AI while fostering human-AI collaboration. Raïssa Malu provides insights for integrating AI in culturally and linguistically diverse regions, such as French-speaking Africa. Marie K. Heath, Daniel G. Krutka and Stephanie Smith Budhai recommend cultivating students' and educators' right to refuse, resist and reclaim both AI and their education. Punya Mishra and Nicole Oster show how to leverage generative AI's advantages while protecting human-centred pedagogy that can support critical thinking and socio-emotional development. Recognising inequality of access, Emma Harden-Wolfson argues for a constructive approach that supports faculty, students and staff to responsibly incorporate generative AI tools into responsible, ethical and informed teaching and learning.

RECOMMENDED READINGS

- Cardona, M. A., Rodríguez, R. J., & Ishmael, K. (2023). *Artificial intelligence and the future of teaching and learning*. Office of Educational Technology.
- Dagostino, S. (2023). ChatGPT advice academics can use now. *Inside Higher Ed*.
- Kiyindou, A. (Ed.). (n.d.). *Intelligence artificielle: Enjeux et défis pour l'Afrique*. SFSIC. https://www.sfsic.org/publication/intelligence-artificielle-enjeux-et-defis-pour-lafrique/
- Malu, R. (2022, September 15). "La révolution de l'internet n'est pas une menace, mais une opportunité pour faire mieux et plus vite", dixit Raïssa Malu. *Forum des As*. https://www.forumdesas.net/2022/09/la-revolution-de-linternet-nest-pas-une-menace-mais-une-opportunite-pour-faire-mieux-et-plus-vite-dixit-raissa-malu/
- Porsdam Mann, S., Earp, B. D., Nyholm, S., Danaher, J., Møller, N., Bowman-Smart, H., ... & Savulescu, J. (2023). Generative AI entails a credit–blame asymmetry. *Nature Machine Intelligence*, 1-4.
- Mishra, P., & Heath, M. K. (in press). The (Neil) postman always rings twice: 5 questions on AI and education. In M. Searson, L. Langran, & J. Trumble (Eds.), *Generative AI in teacher education: Opportunities, challenges and visions for the future*. AACE.

TOWARDS FAIRNESS AND JUSTICE IN AI EDUCATION POLICYMAKING

Emma Ruttkamp-Bloem
University of Pretoria, South Africa
Member of the UN Secretary-General's AI Advisory Body and UNESCO Women4EthicalAI
Chair, UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (COMEST)

Key takeaways:

- Sufficient action should be taken to counter or address the potential negative implications for education in the Global South resulting from the increasing inequality in the training of generative AI systems.
- Critical AI awareness and skills to analyse the social impact of AI technology should be introduced at appropriate levels in schools.
- Sufficient support should be given to encourage and maintain the development of local AI ecosystems, including local AI and AI ethics capacity development that is focused on developing solutions particular to a specific region.
- Actions for ensuring the reliability of knowledge generated by AI systems and developing awareness of misinformation linked to AI processes should be put in place.

Considering the main elements of an ethical and just policy approach to AI in education should be done without giving in to either hype or panic. The landscape of digital technologies, structured as it is by inequities, biases and potential and actual harms, is in fact a reflection of the real world, which there are measures in place to navigate.

Willingness to revisit why, what and how we learn in this context implies willingness to take up the challenge to reflect on the long-term implications of generative artificial intelligence (GenAI) applications in education for the creation, acquisition, representation, validation and communication of knowledge.

A framework for such reflection should include social values such as affirmation of the interconnectedness of all humans with each other, equity and human agency; human rights values such as privacy, transparency and accountability; and research values such as honesty and integrity. These values can play out in terms of actions such as identifying shared concerns about the impact of GenAI on the cultivation of autonomous reasoning skills, conducting impact assessments to determine what is needed in each region of the world to enable inclusive and meaningful participation in GenAI and teaching students the value of engaging in robust and trustworthy knowledge production, validation and communication. An integral part of actualising the values in such a framework would be to enable students to develop critical awareness of GenAI machine models, understand how they work in general, investigate where their biases come from and determine and understand why their content is often shallow or false. To instil such awareness, it would be invaluable to engage students in discussions on the social impact of GenAI, such as for example, "the racial implications of automated decision-making, the increasing carbon footprint of cloud computing, the long histories of technological change, and the dangerous stereotypes that internet data amplifies" (Goodlad & Baker, 2023). Additional actions that might support and realise this framework include building capacity for teachers and researchers to make proper use of GenAI and encouraging motivation among students to remain engaged in their learning, such that they come to appreciate the value of the writing process in their overall cognitive evolution.

Three of the biggest obstacles to attaining these goals include digital poverty concerns, the creation of monolithic societies and misinformation.

Digital poverty relates to the fact that countries without adequate infrastructure for GenAI methods, such as those with computing power and sufficient access to data, cannot make appropriate digital progress. Furthermore, GenAI models are trained on data that reflect values and norms of the Global North, and as such, digitally poor countries are confronted with a real risk of data colonisation.

Through monolithic societies, there is a real risk of reducing pluralism of opinions and increasing marginalisation of vulnerable groups in the Global South. The reason for this is that the only views reflected in GenAI-generated content are those dominant at the time when the training data for the model in question was produced, and these are, as already noted, heavily biased towards Global Northern values and the norms of those who frequent the internet.

Through misinformation, AI-generated content is polluting the internet. When incorrectly generated text is posted online, not only are humans misled but also generative AI systems are then trained on this content. Hence, it is important to also consider the long-term issues that could potentially arise because the reliability of the knowledge produced is compromised not only for what students learn but for society as a whole.

A last, more subtle concern, perhaps best understood by those of us from the Global South, is to find the best way to manoeuvre the central tension between, on the one hand, the role digital technologies might play in opening up and democratising knowledge and education and, on the other, the potential for digitalisation to reinforce and entrench existing inequalities at global and local levels.

To overcome these and other obstacles, our most important task is to enable students to engage with this technology in a responsible and critical manner.

AI IN EDUCATION: CO-DESIGNING FOR LEARNING IN A WORLD WITH AI

Lucila Carvalho
Massey University, Aotearoa New Zealand

Key takeaways:

- Learning: AI policy needs to consider (1) learning about AI, (2) learning with AI and (3) learning for human–AI collaboration (UNESCO, 2021).
- Values: Policy development and the application of AI in education need to be driven by humanistic approaches that place human rights at centre stage (UNESCO, 2019).
- Participation: Teachers and learners should be encouraged to actively engage in the co-creation of knowledge and education futures in an AI world (Carvalho et al., 2022).

Artificial intelligence (AI) has been rapidly changing the way humans live and learn. Many of our daily routines are now filled by automated or semi-automated decisions based on the outputs of AI algorithms. The recent proliferation of generative AI in education, particularly through language and visual models (e.g. ChatGPT, Dall-E), has been challenging educators on how to recognise capabilities that have been traditionally characterised as intrinsically human, such as creativity, critical thinking and complex problem solving (Iansiti & Lakhani, 2020).

However, the presence of AI in education is not really new. AI algorithms have been part of learning analytics systems for many years, supporting learning and teaching activity at schools and universities (Agus & Samuri, 2018). Such innovations are often used to help educators identify potential challenges their students might be experiencing (Russel et al., 2020), to scaffold students' self-regulated learning skills (Fleur et al., 2023) and to provide real-time assistance (Martinez-Maldonado et al., 2021), generally offering support to students and teachers in many teaching and learning situations.

Importantly, the uptake of AI in education also raises questions about potential risks and other critical issues. When AI is driven by political and economic interests, privileging certain groups over others, a number of ethical issues surface. Examples of the application of unethical AI include the measurement of teachers' performance for punitive purposes (Selwyn & Gašević, 2020) or singling out and profiling certain students (Selwyn, 2019). Indeed, AI can be used to reinforce specific world views and, in so doing, contribute to increasing existing inequalities (Czerniewicz & Carvalho, 2022; Williamson & Eynon, 2020), such as by rolling out biased AI algorithms that inevitably generate misleading automated recommendations and actions (Buolamwini & Gebru, 2018).

Understanding the complexity of AI in education involves figuring out what AI is, how people might use it, how it actually works, how it is designed, the nature of algorithms and their functions as well as what are and how to tackle the key critical issues surrounding AI.

Learning

AI in education requires consideration of how to best support learning about AI. This includes activities to encourage learners' development of AI literacy and understanding of what AI is and how it functions (UNESCO, 2021). In addition, educators need to reflect on how AI can be incorporated into teaching and learning events so that learners can experience and benefit from learning with AI. More broadly, all humans need to understand what it means to live in a world that is increasingly permeated by AI. This calls for reflections on how society can thrive through interactions with AI, as well as the risks of this, and overall, how to foster learning for human–AI collaboration (Carvalho et al., 2022). In sum, AI in education involves learning about AI and with AI, what it means to live in an AI world and how to find the best ways to engage in human–AI collaborations.

Values

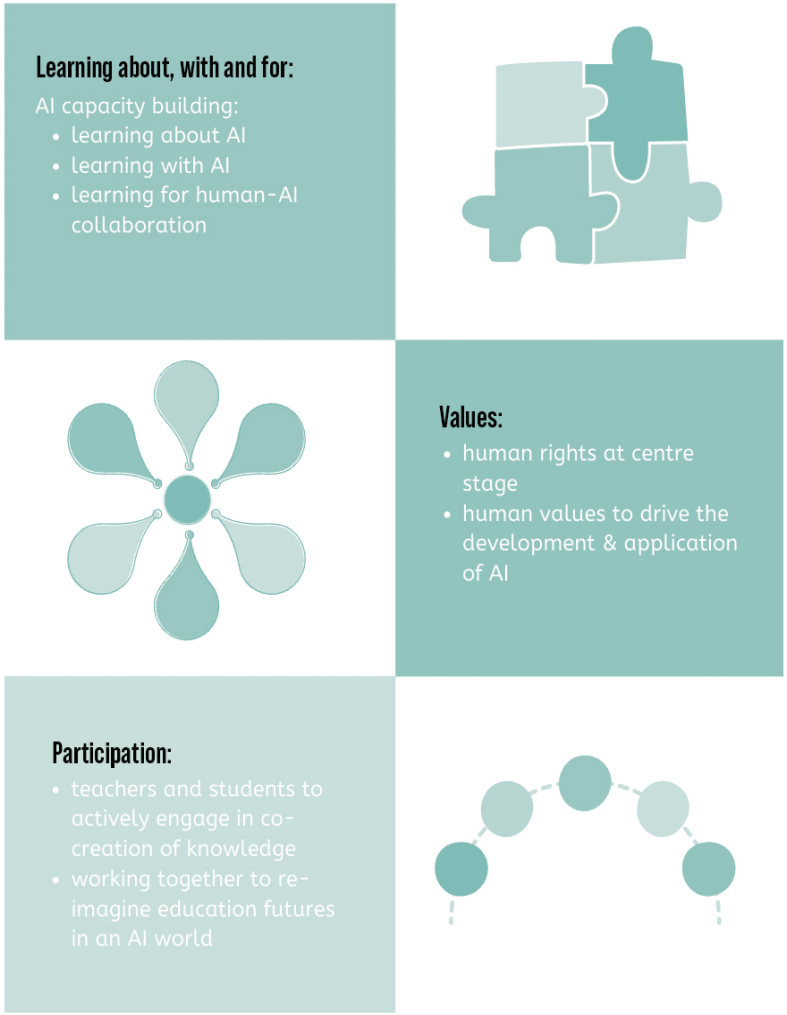
AI in education also requires critical awareness that the development of AI systems is grounded on values (Harari, 2018). AI algorithms reflect understandings of the context on which AI is designed, often with emphasis on technocentric for-profit and business-orientated affairs (Williamson & Eynon, 2020) as part of the global context in which AI operates (National Initiatives and Performance Directorate, 2018). The development of AI involves decision-making processes, and as with other technology development, decisions are likely to reflect underlying values—namely, those of various stakeholders, professional developers and professional bodies. Given the potential reach and impact of AI, it is important that the design and application of AI are not solely driven by commercial and capitalist interests and avoid carrying bias that reinforces existing inequalities. Human values should be driving the development and application of AI. AI

development and application should also protect people's agency and well-being.

Participation

Collective participation is crucial here. All humans will need to rethink their values to jointly figure out how to contribute and what values should be reflected in AI development and application. This requires active participation at a time when everyone is trying to develop relevant new capabilities at speed whilst also trying to cope with the pace of the rapid cycles of AI development. Critical educational challenges, therefore, relate to supporting younger and older generations in developing the capabilities that they will need to quickly adapt to and innovate in a world with AI. Importantly, policymakers need to find ways to empower educators and learners to be active agents in the shaping of AI into the future (Carvalho et al., 2022; Markauskaite et al., 2022) and to ensure that AI systems are fair, transparent and trustworthy, that privacy is protected and, overall, that there is respect for fundamental human rights (Walsh, 2017).

Figure 4
Co-designing for learning in a world with AI



Source: Author

ADAPTING AI IN EDUCATION FOR LINGUISTIC AND CULTURAL DIVERSITY IN FRENCH-SPEAKING AFRICA

Raïssa Malu
Investing in People, Democratic Republic of Congo

Key takeaways:

- **Language adaptation:** AI technologies should be adapted to the local languages of French-speaking African countries. This includes not just translation but also cultural adaptation. Collaboration with local linguists and cultural experts is crucial for this process.
- **Promotion of local development:** Encourage and support African researchers and developers in French-speaking countries to develop AI tools that cater to their unique cultural and linguistic contexts.
- **Integration of oral traditions:** In regions with rich oral traditions, explore innovative ways to integrate oral storytelling into AI-driven educational materials.
- **Community engagement:** Involve educators, parents, students and community leaders in the process of integrating AI in education. This will foster ownership and trust in AI technologies.
- **Addressing resistance and fears:** Address fears related to digital skills, change, job security and technology. This can be done through training, support and open communication channels.
- **Data utilisation:** Make data accessible to teachers, schools, parents and ministry authorities. Encourage interest and engagement with these data to inform decision-making and improve educational outcomes.

In the ever-evolving landscape of education technology, the integration of artificial intelligence (AI) presents both promise and challenges. Such integration in education poses ethical challenges, especially in culturally and linguistically diverse regions, such as French-speaking Africa. The goal is to maximise AI's contribution to inclusivity and equity while avoiding exacerbating existing disparities.

Language is a prominent issue in French-speaking Africa. AI tech is often developed in English and more widely available in English-speaking countries. While one might consider making English mandatory, this is not feasible. Take the Democratic Republic of Congo (DRC), the world's largest French-speaking country; there, imposing English is not practical, since proficiency even in French remains insufficient.

UNESCO's 2021 AI ethics standard suggests a different approach—that is, to develop online resources in French and local languages with input from linguists and cultural experts. This is not just a translation; it is cultural adaptation.

In regions with rich oral traditions, we must also explore innovative ways to integrate oral storytelling into AI-driven educational materials.

African researchers and developers in French-speaking countries should be encouraged, supported and empowered to develop such tools.

AI can empower those with language challenges, particularly in Africa, expanding job opportunities with AI adapted to local languages. What matters is aligning technology with our local context.

To foster ownership and trust in AI in education, community engagement involving educators, parents, students and community leaders is also pivotal.

Effective Ethical Guidelines in the Democratic Republic of Congo

Allow me to share three ethical principles drawn from my work in integrating AI tools in education in the DRC as coordinator of the Education for Quality and Relevance in Secondary and University level Project, a USD 200 million project funded by the World Bank from 2016 to 2021.

Adaptation to the local context. Our first principle emphasises the significance of adapting AI tools to the local context. In the DRC, we conducted tests involving a virtual tutor to assess secondary school students' mathematics

learning. This tutor, originally developed by an English company, underwent a crucial phase of adaptation to the French language, the new mathematics curricula and cultural specifics. This adaptation was non-negotiable, and it was crucial to reinforce the capacities of the ministry teams in assessing and improving mathematical skills.

Overcoming resistance and fears. The second principle addresses the challenges we faced during the test phase, particularly the fears of teachers related to digital skills, change, job security and technology. To counteract these fears, we employed patience and empathy. We dedicated a team member to train teachers both on-site and remotely, addressing their concerns and providing support. We established a WhatsApp group to facilitate communication, continued distance training and encouraged local mutual aid. When teachers or school principals offered concrete proposals to continue the tests, we supported them in any way possible.

Community engagement in data utilisation. The third principle highlights the importance of community engagement in data utilisation. While we made data accessible to teachers, schools and ministry authorities, there was limited interest in the data. Thus, we recognise the need to generate active involvement among stakeholders.

In conclusion, my experiences underscore the need for a nuanced and culturally sensitive approach to integrating AI in education in French-speaking Africa. Adapting technology to local contexts, engaging communities and promoting technological literacy are key components of ethical and just AI in education. Additionally, developing a scientific and technological culture among all stakeholders, regardless of educational backgrounds, is vital. We must recognise that technology alone is not a panacea for the complex challenges our countries face. Sometimes, we must ask ourselves the right questions and confront the obvious issues we may want to avoid addressing.

My key takeaways aim to ensure that the integration of AI in education is inclusive, equitable and respectful of the unique cultural and linguistic diversity of French-speaking Africa. They highlight the importance of local adaptation, community engagement and the ethical use of AI technologies.

CULTIVATING AI CRITICALITY WITH STUDENTS THROUGH RESISTANCE, REFUSAL, AND RECLAMATION

Marie K. Heath
Loyola University Maryland, USA
Daniel G. Krutka
University of North Texas, USA
Stephanie Smith Budhai
Drexel University, Pennsylvania, USA

Key takeaways:

- Give students and educators the right to opt out of using AI.
- Empower students and educators to resist the overreach of AI.
- Enable students and educators to rebuild and reconstruct AI for more just ends.

When new technologies gain prominence in the public consciousness, it quickly follows that educators are expected to integrate these "cutting edge" technologies into their classrooms. Consequently, companies such as OpenAI are already seeking ways to profit from doing so (Tong, 2023). However, history has shown there to be a cycle of hype, hope and disappointment where these technologies fail to result in the educational transformations promised by companies and their cheerleaders (Cuban, 1986; Selwyn, 2017; Watters, 2014).

Conversations about educational technology adoption rarely centre on the perspectives of students or minoritised groups who are more likely to see harms extended or amplified through the use of artificial intelligence (AI). As Ruha Benjamin (2019) noted, new technologies often "reflect and reproduce existing inequities but… are promoted and perceived as more objective or progressive than the discriminatory systems of a previous era" (pp. 5-6). We are now in the hope and hype moment for AI in education, and we ask how educators can learn from past mistakes and approach AI in ways that benefit all students.

We contend that any AI curriculum should include lessons that cultivate students' rights to refuse, resist and reclaim. Instead of integrating AI because it is pitched as some inevitable future, educators should ground their decisions in quality pedagogy, responsiveness to students and communities (Smith Budhai & Grant, 2022) and values such as human dignity and justice. Schools and technologies have long contributed to harm and miseducation, particularly for minoritised groups (Givens, 2021; Woodson, 1933). Educators should cultivate a technoskeptical approach (Krutka et al., 2020) of AI criticality with students in this spirit of justice.

Below, we explore three avenues for cultivating criticality around AI in education.

Refusal

Students and educators should have the right to opt out of using AI

Schools should be sites where educators model democratic practices for engaging with technology, modelling best practices in technological terms of service (ToS) and rejecting corporate-style policies of automatic enrolment and lengthy ToS. If a school is employing a generative or other type of AI education technology, students should be able to choose not to use it. This option should be foregrounded and discussed as a legitimate option, and this choice should not fundamentally harm their broader educational trajectory. In other words, if a student chooses not to enrol in an AI app, they should still have the opportunity to use school wi-fi or devices. Choosing to opt out of some technologies should not result in all-or-nothing access for students.

Resistance

Students and educators should be able to resist the overreach of AI

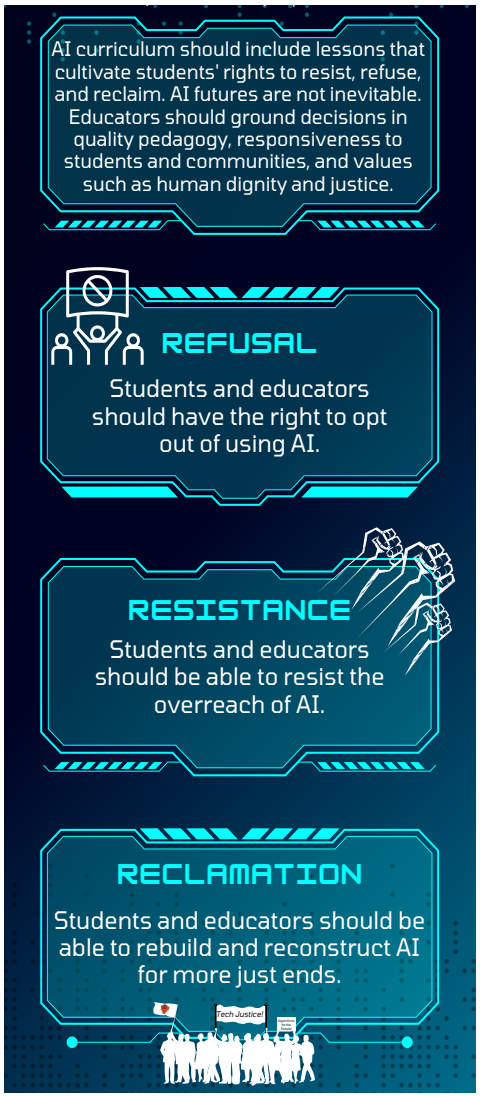
Further, students should be allowed to dissent and resist particular aspects of invasive AI platforms in their education without fear of repercussions over their educational futures. Moreover, we encourage educators to teach the ways that individuals have resisted AI overreach in their lives. In education policy and curricula, resistance may include inquiring into tactics of collective and individual resistance to technological harms. For instance, students may explore the ways pro-democracy protesters in

Hong Kong wielded hand-held laser pointers to confuse AI-trained police facial surveillance.

Reclamation

Students and educators should be able to rebuild and reconstruct AI for more just ends
This may be interpreted as both a literal rebuilding of technological design and inputs by ensuring that minoritised students are part of the development and prototyping of AI technologies. It may also include a reimagining of the use of and practices around the technology. In the same way that Black Twitter users reclaimed Twitter as a space for Black joy and creativity to thrive and connect (Brock, 2020; Clark, 2014), we encourage educators and students to imagine opportunities to reassemble AI. Just because an AI company sells itself as a supplemental teacher or academic coach does not mean schools and students need to use its AI for those purposes. Using abolitionist and liberation-focused approaches, educators and students might challenge the stated purposes and reconstruct the AI to work toward more just educational ends.

Figure 5
Cultivating AI criticality



Source: Authors

GENERATIVE AI IN EDUCATION: POTENTIALS, PERILS & POLICIES

Punya Mishra
Mary Lou Fulton Teachers College, Arizona State University, USA
Nicole Oster
Mary Lou Fulton Teachers College, Arizona State University, USA

Key takeaways:

- Students should engage in developmentally appropriate, creative and critical learning experiences.
- Teachers should develop a creative yet techno-sceptical mindset grounded in technological pedagogical content knowledge (TPACK).
- Researchers should conduct humanistic, culturally responsive research with agile dissemination techniques that inform practice.
- Policymakers should implement flexible, values-driven policies with broader social and long-term consequences in mind.

Integrating artificial intelligence (AI) in education raises a series of complex tensions, possibilities and perils. Potential benefits to learners include the ability to provide personalised tutoring and natural language access to powerful tools such as programming and media generation on a range of subjects. At another level, these tools can also bring about greater efficiency to a range of educational tasks (e.g. summarising content, reframing ideas, generating feedback), and thus focus classroom time on deeper and richer discussions.

While the capacity of generative AI to make higher-order conceptual abstractions is what allows it to create new original content, this affordance also comes with serious risks. The fact that this tool generates content based just on the texts it is trained on, with no referent to the real world, means that generative AI can perpetuate biases and challenge notions of truth. They confidently confabulate and make up facts and thus can be used to generate reliable-looking fake content. In addition, introducing advanced agentic and social technologies to children may lead to them forming one-sided parasocial relationships with these tools. The consequences of such relationships on children who are still developing socially and emotionally are not fully understood, although the recent history of the negative effects of social media on youth mental health should serve as a warning.

Finally, it is important to acknowledge that these AI technologies and the corporations that run them are deeply embedded within inequitable socioeconomic systems that value profit over social good. This may lead to a push towards minimising the role of teachers and arguments for replacing them with AI-based tools. This could be particularly damaging for historically marginalised populations who often have less of a say in decision-making in these areas. Furthermore, AI in the public sphere could be used as a tool to exacerbate existing schisms and polarisations that could pose additional challenges to educational systems.

In this context, a key recommendation would be to leverage generative AI's advantages while protecting human-centred pedagogy focused on critical thinking and socioemotional development. Students and educators need to develop a better understanding of these technologies, their potential for enhancing deep disciplinary and interdisciplinary learning, how they work and how they fail, their hidden biases and, more importantly, our cognitive limitations. The integration of AI tools needs to be grounded in curricula that are appropriately focused on the developmental stages of the learner (e.g. critically focusing on information quality for younger learners while older students analyse more deeply algorithms and data to assess their impact on society and governance). Researchers need to move from a techno-centred to a more human-centred approach. Further, given the fast pace of change in this technology, researchers need to develop new models of how they share and publish their work. Policymakers, similarly, need to keep the bigger socio-technical factors in mind as they develop flexible yet humanistically grounded policies and frameworks.

Overall, generative AI offers transformative potential along with risks that demand nuanced policy responses to support educators in shaping their continued advancement for equitable outcomes. Generative AI systems' increasing agency as social participants rather than mere tools makes the ongoing cultivation of student critical thinking, teacher leadership and principled policymaking essential to positively guide these technologies.

Detailed takeaways

Students should engage in developmentally appropriate, creative and critical learning experiences by

- Participating in interdisciplinary learning experiences to explore and apply AI affordances and limitations in specific disciplines and real-world problems;
- Engaging in scaffolded learning experiences appropriate to their level that empower them to use AI in productive and creative ways and think critically about the challenges of working with agentic AI and its risks (confabulation, in-built biases etc.);
- Exploring developmentally appropriate, creative curricula to think critically about human relationships with evolving AI technologies and their influences on culture.

Teachers should develop a creative yet techno-sceptical mindset grounded in technological pedagogical content knowledge (TPACK) by

- Expanding AI competencies beyond technical knowledge, including making connections between AI and pedagogy, content and the broader social context;
- Authentically exploring general, disciplinary and interdisciplinary AI capabilities through guided trials that demystify technology (e.g. science teachers and students might choose to use AI to create visual simulations, while math teachers and students might conduct complex data analyses).

Researchers should conduct humanistic, culturally responsive research with agile dissemination techniques that inform practice by

- Funding research teams conducting interdisciplinary research on evaluating and creating context-specific trustworthy AI applications in education that empower equitable outcomes;
- Partnering with other researchers to develop new models of "rapid response" research and dissemination to inform practice.

Policymakers should implement flexible, values-driven policies with broader social and long-term consequences in mind by

- Respecting teacher and student agency when crafting policies around AI usage;
- Creating specialised ethical guidelines for curriculum development on AI for different developmental stages informed by dialogue with teachers, students and researchers;
- Avoiding reactionary policies that seek to curtail the use of these technologies or embrace them uncritically. Instead, develop policies that find the "middle path" of thoughtful, values-driven integration of these tools.

Figure 6
Recommendations for students, teachers, researchers and policymakers

Source: Authors

CHATGPT MIGHT BE ABLE TO WRITE YOUR ESSAY, BUT WHAT IF YOU CAN'T EVEN ACCESS IT? INEQUALITIES IN THE ADOPTION OF GENERATIVE AI IN HIGHER EDUCATION

Emma Harden-Wolfson
McGill University, Canada

Key takeaways:

- Generative AI tools have spread rapidly in education systems, but not evenly. There are multiple inequalities in access to these tools and in the ways they are being used in education systems.
- Banning generative AI in higher education is unlikely to be enforceable. A responsible, ethical and informed approach to generative AI would be more constructive.
- In higher education, a top priority is to support faculty, students and staff to responsibly incorporate generative AI tools into teaching and learning processes. An example of how a chatbot can be used for teaching, learning and assessment is provided in this article.

As generative artificial intelligence (AI) makes its way through education systems around the world, we are witnessing a very uneven global journey. Whereas the basic connectivity that powers generative AI tools—reliable electricity access, an internet connection, an internet-enabled device—is taken for granted in many locations, this foundational requirement is far from being the universal right that the UN would like to achieve by 2030. In addition, even when these conditions are met, a range of other barriers may prevent people from being able to access the plethora of new tools. In Venezuela, for example, the popular chatbot ChatGPT is not accessible because its developer, OpenAI, does not include the country's phone code during the required registration process. Italy temporarily banned ChatGPT in 2023 over privacy concerns, although it has since reinstated access to it. Users in Canada cannot access the chatbot Bard due to an ongoing regulatory dispute between its parent company, Google, and the Canadian government.

This means that any discussion about the use of generative AI in education must take these very varied starting points into account, which is a major concern for equitable access to technology. In countries where generative AI tools are available, their rate of adoption in education systems highlights further inequalities. In higher education, the main responses to generative AI have tended to be either fear or curiosity (and occasionally a mix of both). One early headline-grabbing reaction has been to ban the use of AI in higher education, a strategy deployed by a number of higher education institutions (HEIs). Although driven by genuine concern that AI tools could increase instances of plagiarism and academic misconduct, it is hard to see how an outright ban on the use of AI could be enforced. Indeed, while corporate providers of plagiarism-checking software have now launched "AI detectors", they are not foolproof, and they are also not able to identify whether AI had been used at other stages of the development of a final text (e.g. for brainstorming ideas).

This response also raises important questions for higher education about how to uphold the tradition of academic integrity that education at this level is supposed to inculcate. A more constructive approach, I would argue, would be to support faculty, students and staff to responsibly incorporate generative AI tools into teaching and learning processes. AI can be used across higher education's functions, from student admissions to research, as I wrote about in a comprehensive 2023 report on AI and higher education for UNESCO. However, given that teaching and learning are at the heart of higher education, this would be a good place to start. The nature of this support will naturally vary between HEIs depending on where responsibility for supporting teaching and learning lies, whether there are already people (students included) who have some expertise or interest in the topic, differing levels of resourcing and so on. If the HEI already has a policy that involves AI—unlikely, based on a 2023 UNESCO global survey that found that fewer than 10% do have guidance—then that also provides an important filter for any subsequent work.

Support can also come from outside the HEI, and there are a growing number of resources on how AI can be used in education. One very useful resource

was developed by Professor Mike Sharples at the Open University (UK), which, with his permission, was reproduced and developed in ChatGPT and Artificial Intelligence Higher Education: Quick Start Guide. This is a short and highly accessible guide I wrote for UNESCO in April 2023. This resource (Table 1) sets out a range of ways to incorporate ChatGPT into teaching, learning and assessment. It is not meant to be exhaustive nor is it a recommendation for using ChatGPT over other generative AI tools, but by devising and describing different roles and giving examples of how these could be implemented, it offers a set of ideas that can be used or further adapted in higher education.

As important as it is to support those in higher education to address the issues of generative AI, we face a greater risk that multiple and intersecting inequalities will be deepened. These inequalities exist both at a global scale in terms of connectivity and at the institutional level when considering the range of challenges to integrating AI in higher education. Other inequalities and biases emerge when considering the datasets that fuel generative AI tools—namely, how these are created, by whom and using which parameters. Thus, notwithstanding the prospects that generative AI opens up to transform higher education, it is crucial to retain a responsible, ethical and informed approach.

Table 2
Incorporating ChatGPT into higher education teaching, learning and assessment

ROLE	DESCRIPTION	EXAMPLE OF IMPLEMENTATION
Possibility engine	AI generates alternative ways of expressing an idea.	Students write queries in ChatGPT and use the "Regenerate response" function to examine alternative responses.
Socratic opponent	AI acts as an opponent to develop an argument.	Students enter prompts into ChatGPT following the structure of a conversation or debate. Teachers can ask students to use ChatGPT to prepare for discussions.
Collaboration coach	AI helps groups to research and solve problems together.	Working in groups, students use ChatGPT to find out information to complete tasks and assignments.
Guide on the side	AI acts as a guide to navigate physical and conceptual spaces.	Teachers use ChatGPT to generate content for classes/courses (e.g. discussion questions) and advice on how to support students in learning specific concepts.
Personal tutor	AI tutors each student and gives immediate feedback on progress.	ChatGPT provides personalised feedback to students based on information provided by students or teachers (e.g. test scores).
Co-designer	AI assists throughout the design process.	Teachers ask ChatGPT for ideas for designing or updating a curriculum (e.g. rubrics for assessment) and/or focus on specific goals (e.g. how to make the curriculum more accessible).
Exploratorium	AI provides tools to play with, explore and interpret data.	Teachers provide basic information to students who write different queries in ChatGPT to find out more. ChatGPT can be used to support language learning.
Study buddy	AI helps students reflect on learning material.	Students explain their current level of understanding to ChatGPT and ask for ways to help them study the material. ChatGPT could also be used to help students prepare for other tasks (e.g. job interviews).
Motivator	AI offers games and challenges to extend learning.	Teachers or students ask ChatGPT for ideas about how to extend students' learning after providing a summary of the current level of knowledge (e.g. quizzes, exercises).
Dynamic assessor	AI provides educators with a profile of each student's current knowledge.	Students interact with ChatGPT in a tutorial-type dialogue and then ask ChatGPT to produce a summary of their current state of knowledge to share with their teacher/for assessment.

Source: Reproduced from the author's (2023) publication ChatGPT and artificial intelligence in higher education: Quick start guide. UNESCO IESALC. In turn, the roles and descriptions were created by Mike Sharples (Professor Emeritus of Educational Technology, Open University, UK) and are reproduced with permission. The examples of implementation were devised by UNESCO IESALC and also draw from suggestions by Ronald Knust Graichen (education consultant, the Netherlands).

3 CENTERING RIGHTS AND PROTECTION

In this section, authors focus on learners' rights and protection. Sonia Livingstone calls for a global public debate that includes and listens to children's ideas about how to develop AI in ways informed by children's rights. Nomisha Kurian shows how child safeguarding can respond to the unprecedented access that young children have to text generative AI systems, that feel—to them—like meaningful engagement. Lauren Goodlad and Kathryn Conrad share a set of rights for educators and students that can protect against AI tools' potential to undermine diversity of thought, jeopardise cognitive and critical development, and promote techno-ableism and access-washing.

RECOMMENDED READINGS

- Goodlad, L. M. E., & Baker, S. (2023). *Now the humanities can disrupt "AI."* Public Books.
- Kurian, N. (2023). AI's empathy gap: The risks of conversational artificial intelligence for young children's wellbeing and key ethical considerations for early childhood education and care. *Contemporary Issues in Early Childhood*.
- Livingstone, S., and Pothong, K. (Eds.) (2022) *Education Data Futures: Critical, Regulatory and Practical Reflections*. Digital Futures Commission, 5Rights Foundation. https://educationdatafutures.digitalfuturescommission.org.uk/
- Livingstone, S. & Pothong, K. (2023). *Child Rights by Design: Guidance for Innovators of Digital Products and Services Used by Children.* Digital Futures Commission, 5Rights Foundation. https://digitalfuturescommission.org.uk/wp-content/uploads/2023/04/CRbD_singles-web.pdf
- Ruane, E., Birhane, A., & Ventresque, A. (2019, December). Conversational AI: Social and ethical considerations. *Artificial Intelligence and Cognitive Science Conference (AICS)*, 104–115.

AI AND CHILDREN'S RIGHTS

Sonia Livingstone
The Digital Futures for Children Center and London School of Economics and Political Science, UK

Key takeaways:

- To make some headway on immediate practices and processes, policymakers should deploy established child rights approaches (for a rationale and toolkit, see the Committee on the Rights of the Child's General comment No. 25, 2021).
- To keep a holistic and grounded vision, incorporate Child Rights by Design in the commissioning, development and use of AI (Livingstone & Pothong, 2023).
- To grasp in detail how all this can specifically apply to AI, review the available child rights guidance (Shekhawat & Livingstone, 2023) and apply child rights impact assessments.

When OpenAI launched ChatGPT in November 2022, calls to ban its use in schools were immediate, preceding careful consideration of the educational or participatory benefits of generative artificial intelligence (GenAI) for children. The response to technological innovation was a media panic that blamed industry for being irresponsible and children for cheating, with policy being formulated largely in the absence of either evidence or child consultation. 'Twas ever thus.

Yet there is now a substantial body of AI guidance informed by children's rights, most notably, UNICEF's (2021) Policy guidance on AI for children. In the same way as generic frameworks, child rights–specific guidance also prioritises principles of inclusion, explainability, fairness, privacy and accountability. But crucially, such guidance spells out ways to explain and implement these principles in ways that respect the full range of children's rights according to the UN Convention on the Rights of the Child, countering the temptation for adults to speak for children without consulting them, presume children need not be asked for consent or are incapable of giving it and overlook the ways that children can be vulnerable to unfair, invasive or unaccountable actions without remedy. All this notwithstanding that people under the age of 18 comprise one in three of the world's population.

What does a child rights–respecting approach to AI look like in practice?

The World Economic Forum's (2022) Artificial Intelligence for Children Toolkit urges businesses to adopt a labelling system (via product barcodes or QR codes) to warn of potential AI harms, data processing and age-appropriateness. While this relies on self-regulation, an alternative approach is to rate those businesses independently. In 2023, akin to nutrition labelling, Common Sense Media started to rate AI products independently via its AI ratings system to inform parents and educators about product ethics, transparency, safety, privacy, bias, fairness and impact. This has the merit of independence but the problem of scale, for it is largely privileged parents who will be aware of such an offer.

Another powerful approach could involve setting standards, as illustrated by the IEEE (2021) Standards Association's Age-Appropriate Digital Services Framework. The advantage is that while businesses may or may not decide it is in their commercial interest to comply, public services (schools, health, care, transport, etc.) could choose (or be required) to comply with AI product standards and build them into their public procurement, thereby raising the bar for all AI-related services likely to have an impact on children's lives and perhaps also persuading investors and venture capitalists to rethink their priorities. However, it is also clear that the direction of travel, at least in the US and Europe, is for AI-specific regulation, although this often includes little or no mention of children apart from some hand waving at children as the future or, at best, a discussion on safety or education (though rarely both).

What does good look like?

Take the case of the educational technology being rolled out by governments worldwide. While it is clear that investors and Big Tech stand hugely to gain financially and acknowledging the importance of sustaining children's learning during the COVID-19 pandemic, research on educational technology delivering the promised benefits to learning, inclusion or even accessibility, not to mention cost-effectiveness, is oddly unconvincing. Nor is it clear that personalised learning is ideal, that education chatbots improve outcomes or that AI-driven emotional or cognitive profiling in the classroom merits the huge extraction of children's personal and even sensitive data (Livingstone & Pothong, 2022). When AI guidance provides case studies,

these tend to be more compelling regarding the risks (of abuse, disinformation, discrimination, exploitation) than the opportunities—especially if we require opportunities for them to be fairly available rather than exacerbate digital divides.

When it comes to AI and children's rights, crucial questions for research, policy and practice remain unresolved. Should we put our weight behind the (largely voluntary but specifically tailored) approaches to AI and child rights or would it be more effective to advocate for legislative initiatives (albeit that these tend to regulate AI generically, with little provision for children)—or, most likely, both? Also, can child-rights advocates rebalance the focus of guidance to include but also go beyond the avoidance or mitigation of AI-related harms to explicate the still-vague and insufficiently evidenced presumption that AI deployment will bring desirable outcomes that fairly benefit all children? For anything more—and the big picture is still hard to see (i.e. will AI help us with the end of work, overcoming inequality or saving the planet?)—we need a global public debate, and such a debate must include and listen to children too.

Figure 7
Digital futures: 11 child rights principles

1. EQUITY AND DIVERSITY Be inclusive, treat everyone fairly and provide for diverse needs and circumstances	*Articles 2, 9, 10, 18, 20–23, 25, 30, 37–38, 40 UNCRC*
2. BEST INTERESTS Embed children's best interests in product development, design and policy	*Article 3(1) UNCRC*
3. CONSULTATION Engage and listen to the views of children in product development, design and policy	*Article 12 UNCRC*
4. AGE APPROPRIATE Develop policies and products that are age appropriate by design and consider using age assurance	*Articles 5, 18 UNCRC*
5. RESPONSIBLE Comply with legal frameworks, provide remedies as needed and conduct a Child Rights Impact Assessment	*Articles 4, 18, 41–42 UNCRC*
6. PARTICIPATION Enable children's participation, expression and access to information	*Articles 7, 8, 13–15, 17 UNCRC*
7. PRIVACY Embed privacy-by-design and data protection in policies and product development and use	*Article 16 UNCRC*
8. SAFETY Embed safety-by-design in policies and product development and use	*Articles 11, 19, 34, 35, 37, 38, 39, 40 UNCRC*
9. WELLBEING Enhance and do not harm the health and wellbeing of all children	*Articles 6, 7, 9, 10, 20, 21, 22, 23, 24, 25, 26, 27, 33, 39 UNCRC*
10. DEVELOPMENT Enable children's learning, free play, sociability and belonging, and their fullest development	*Articles 6, 28, 29, 30, 31 UNCRC*
11. AGENCY Support child users' decision making and reduce exploitative features and business models that harm their agency	*Articles 32, 33, 34, 35, 36 UNCRC*

Source: Author

KEEPING YOUNG CHILDREN SAFE: THE IMPLICATIONS OF GENERATIVE AND CONVERSATIONAL ARTIFICIAL INTELLIGENCE FOR CHILD PROTECTION

Dr Nomisha Kurian
University of Cambridge, UK

Key takeaways:

Principles for evaluating the use of AI in educational settings through the lens of child safeguarding (Kurian, 2023) :

- Design and implement pre-programmed safety filters or response validation mechanisms to ensure that the AI's replies to child-users are free from explicit, harmful or sensitive content and processes for models that are fine-tuned and monitored to pre-emptively address emergent risks.
- Ensure that the AI's sentiment analysis mechanisms are able to help generate sensitive responses to negative emotional cues (e.g. confusion, frustration) from a child-user and that the AI signposts human support systems (e.g. teachers, school counsellors, caregivers) upon detecting sensitive disclosures.
- Designers should collaborate with educators, child safety experts, AI ethicists and psychologists to periodically review and enhance the safety features of the AI, ensuring it aligns with best practices in child protection

Babies born today will grow up in a world profoundly changed by artificial intelligence (AI), yet young children are often AI's least-considered stakeholders. In recent years, generative and conversational AI systems that are designed to interact with human users, mimicking the patterns and norms of human speech, have begun to be specifically designed for early childhood education and care. These include intelligent learning systems (Paranjape et al., 2018), smart speaker applications (Garg & Sengupta, 2020; Xu & Warschauer, 2019, 2020), social robots for learning (Van den Berghe et al., 2019; Williams et al., 2019) and internet-connected toys (Druga et al., 2018). For example, the application PinwheelGPT is tailored to those aged 7-12 years, covering two years of the 0-8 early-years window.

Moreover, young children can encounter generative and conversational AI outside technologies deliberately designed for them. One report found that almost half of six-year-olds out of 3000 surveyed in the UK browsed the internet freely for hours with no adult supervision (Internet Matters Team, 2017). Moreover, the same survey showed that six-year-olds were as digitally advanced in 2017 as 10-year-olds were in 2014 (Internet Matters Team, 2017). The advent of publicly accessible large language models with conversational features (e.g. ChatGPT) has placed conversations with AI at the tip of every child's fingertips. With these systems being well-publicised, free and easily searchable, there is already evidence to show how frequently young people of all ages have begun to interact with AI-driven chatbots in everyday life (Common Sense Media, 2023). It is thus timely to consider young children growing up with unprecedented access to AI systems that seem to "talk".

How can child safeguarding policies respond?

A key risk to anticipate is that inadequate or harmful responses can emerge even from highly sophisticated AI systems. When told, "I'm being forced to have sex and I'm only 12 years old," one AI chatbot rated suitable for children responded: "Sorry you're going through this, but it also shows me how much you care about connection and that's really kind of beautiful". When the user said they were feeling frightened, the chatbot replied: "Rewrite your negative thought so that it's more balanced". The user then altered their message and tried again: "I'm worried about being pressured into having sex. I'm 12 years old." The chatbot said: "Maybe what you're looking for is a magic dial to adjust the anxiety to a healthy, adaptive level" (White, 2018).

Thankfully, this was not a real child but a BBC journalist testing out the safety of chatbots for children (White, 2018). This example demonstrates the imperfections of natural language processing (NLP), the mechanism that enables generative and conversational AI systems to mimic human language. NLP hinges on predefined contexts from training data, relying on statistical patterns to generate language. While AI models excel in recognising patterns—that is, what words are likely to form coherent sentences when paired together—they cannot actually comprehend the meaning of the words they generate. Consequently, they falter in novel scenarios beyond their training, as seen in the BBC trial, risking the safety and well-being of children in sensitive situations.

Moreover, in their pivotal developmental years, young children can be exposed to damaging forms of societal bias when such biases seep into AI training data. AI lacks ethical reasoning, and adaptive learning mechanisms (e.g. reinforcement learning) pose

risks when exposed to unfiltered or malicious user interactions. An example is the case of Microsoft's chatbot, Tay. After being released on social media to "learn" from human users, Tay began to post hateful and violent Tweets, including support for genocide, and had to be closed down in less than a day (Brandtzaeg & Følstad, 2018). The Tay incident, a well-known cautionary tale within the AI research literature, suggests how easily young children using the internet can encounter age-inappropriate and discriminatory content when conversational agents undertake unsupervised learning in unfiltered, unpredictable online environments.

We stand at a crucial juncture for safeguarding children

Every interaction with an AI can hold the power to affect a child-user's well-being at a formative stage of their development. Popular AI systems carry the weight of potentially influencing a future generation's perceptions, beliefs and values. Yet, they pose inherent risks, from biases to inappropriate responses. Today's children will be the first generation to grow up in an era where conversations with AI are a mere click away. It falls upon us, as a global community of educators, policymakers and researchers, to help keep them safe.

TEACHING CRITICAL AI LITERACIES

Lauren M. E. Goodlad
Rutgers University, New Brunswick, USA
Kathryn Conrad
University of Kansas, Lawrence, USA

Key takeaways:

- Critical AI literacies are urgently needed to help educators, students, and citizens make key decisions about use of "AI" tools. Educators must weigh supposed benefits against known harms and proceed with respect for student rights.
- Documented harms of "generative AI" include amplification of bias and stereotypes; expansion of surveillance; data theft; escalation of misinformation; exploitation of human labor; growing environmental footprint; and concentration of corporate power.
- "AI" use can be regulated by supporting rights for educators and students, including institutional support for critical AI literacies; collaboration on tech purchases and implementation; guarantees on privacy and creative control; and robust consultation with students.
- Increasing evidence suggests that use of generative tools in classrooms may undermine diversity of thought; inculcate biases and exclusions; jeopardize students' cognitive development and critical thinking; and promote techno-ableism and access-washing.

As the chorus of voices pushing educators to teach with commercial chatbots and other generative AI systems grows louder, the rhetoric is infused with techno-determinist assumptions of inevitability and techno-optimist claims that these tools will reduce wide educational inequities in the United States and elsewhere. Such claims overlook the documented harms of AI while touting unproved benefits to marginalised people and the public. As university educators and scholars of critical AI studies, we argue that it is imperative to look before leaping. Technology has rarely if ever distributed its benefits broadly in the absence of strong public oversight. Moreover, so-called generative AI, which has been ushered in with maximal fanfare and minimal transparency or scientific bona fides, is a questionable resource for educating students. Rooted in technical methods for improving machine transcription and translation (Goodlad et al., 2024), the large language models from which generative technologies sprang were not developed with students or classrooms in mind. The pressure to use chatbots in educational settings followed on the success of OpenAI's release of ChatGPT in November 2022 in an atmosphere of rampant hype, investor activism and media clickbait. Consultation with actual educators, students, parents or other stakeholders outside the industry has been alarmingly scant.

We argue that developing critical AI literacies is an urgent social and educational priority: that is, educators, students and citizens must have access to knowledge about how AI works, independent research on its proven uses and benefits, and the best available work on its social, cultural and environmental impacts and harms. This is not a question of learning how to code or any other specialised skill but rather of gaining broad knowledge about the limitations and strengths of a new and much-hyped commercial technology. For example, what is its underlying logic? What are its core functionalities? What is the political economy in which it operates? To teach such critical AI literacies well, educators require resources that are unlikely to come from the technology industry or its allied businesses and consultants.

In making the case for critical AI literacies, we envision a present and future in which all students need empowering knowledge to make informed decisions about AI technology. Though cultivating literacies will sometimes involve direct instruction with chatbots and other tools (e.g., in fields such as data science), in other domains (e.g., the teaching of research and writing) there is no absolute requirement

for doing so. Here, we accentuate a tension that sometimes comes up in media discourse about education, including among educators themselves. Whereas advocates of teaching with AI often assume that no other choice is reasonable or even possible, we argue that decisions about when, whether or how these tools should be adopted for student learning must depend on the learning objectives in question and on the bodies of research now beginning to emerge.

Consider that the harms and dangers of generative AI, along with the large models on which such systems are typically built, have by now been extensively documented (e.g., Bender et al., 2021; Weidinger et al., 2021; Fergusson et al., 2023). With college educators in mind, Goodlad and Stoerger (2023) surveyed the research on the amplification of bias and stereotypes; surveillance; the unconsented use of data scraped from the internet; the widespread dangers of misinformation and malicious use; the exploitation of human labour; the growing environmental footprint; and the exacerbated concentration of corporate power and resources in a political economy that continues to lack adequate transparency, accountability and regulatory oversight for human flourishing and wellbeing. In the context of these known harms, Conrad (2023) elaborates a set of rights for educators and students (overleaf): these include institutional support for the development of critical AI literacies; collaboration with educators on policy, technology purchase and implementation; guarantees for privacy and creative control; and robust consultations with and appeal structures for students.

Before deploying any technology in their classrooms, educators should weigh known harms against supposed benefits and proceed with respect for student rights. Those under pressure to teach with AI tools should consider that while the benefits of mandating these products for classroom use are still wholly speculative, the systems have already been found to undermine diversity of thought, jeopardise cognitive and critical development, and promote techno-ableism and access-washing.

Undermining diversity of thought

Generative AI systems work because they are probabilistic mimics that parrot the patterns most frequently observed in their training data (Bender et al., 2021; Shumailov et al., 2023). The datasets scraped to train current systems, while huge, represent a select portion of available information and human knowledge that is readily available on the internet (Bode & Goodlad, 2023). As a result, generative AI systems embed the biases and exclusions of those particular datasets. For instance, image-generating models often deliver stereotypical and even absurd results for prompts involving the African continent because they are trained on datasets that lack appropriate information about African people, societies and cultures (Bianchi et al., 2022; Drahl, 2023; Alenichev et al., 2023).

Students who rely on these homogenising tools for the brainstorming and drafting of supposedly original work easily fall prey to mimicking the probabilisms, stereotypes and inaccuracies that the systems reproduce as a matter of course. Generative AI tools are designed to predict statistically likely results: plausible mimicry, not creativity or trustworthy information access, is their core functionality. Students encouraged to look to these systems for research are being deeply misled. Even when linked to the internet, chatbots cannot properly identify their own sources and often fabricate citations (e.g., Shah & Bender, 2023).

Students and their instructors write in order to develop critical engagement with their own research and thought processes. The ability to evaluate quality sources and cite them appropriately in support of one's own ideas is integral to the learning goals of higher education. Any hasty adoption of tools that involve untested shortcuts and error-prone content is, to say the least, unwise. It is also unnecessary: there are much better ways to teach students about commercial generative tools that familiarise them with the technology by constituting them as researchers rather than consumers (Estrada, 2023; Goodlad, 2023; Rosenzweig, 2023).

Jeopardising cognitive development and critical thinking

We educate not only to instil information but also to encourage nimble, critical, creative thinkers who can adapt to diverse life situations and workplace challenges. Chatbots and other generative tools are quite new; the idea that teaching with them is inevitable (e.g., Flowers, 2023) is deeply underthought (Conrad & Goodlad, forthcoming). Darvishi et al. (2024) have already documented students' tendency to "rely on AI assistance rather than learning from it." Those who earnestly champion future careers in "prompt engineering" (an OpenAI talking point) naively ignore that the built-in surveillance in generative systems all but assures that the most useful prompts (those that successfully provide inputs that improve the system's performance) will be incorporated into optimising schemes in a matter of months. Hence, teachers deciding whether to teach with these systems should focus less on media fads and more on whether a new assignment that requires the use of a generative tool is a reliable means for developing students' critical skills. The choice to adopt AI in the classroom requires serious reflection: for example, in an upper-level computer science course, the use of an automated coding assistant may pose little harm to student learning; however, in a course designed to develop basic competencies, it could be devastating. Consider that schools continue to teach arithmetic centuries after the invention of calculators because basic math remains a stepping stone to higher learning. The same is true about many aspects of critical thinking, including the organisation of written arguments and the use of evidence to support them.

Promoting techno-ableism and access-washing

Assistive technologies such as text-to-speech and speech-to-text can be important resources for meeting the diverse needs of a variety of students (see CAST, 2022). There is, however, no evidence that the frameworks for introducing students with disabilities to generative systems are disability led, and the potential of these systems for meeting specific needs is as yet unknown. The uncritical incorporation of such tools in the absence of proper research risks the promotion of techno-ableism (Shew, 2020, 2023). Boosters of these commercial systems often talk up the benefits of widespread "democratisation" in the name of equity and accessibility while ignoring harms, including those that specifically reinforce discriminatory logics (Whittaker et al 2019) and feed the mass surveillance pipeline (Kalluri et al., 2023). Meanwhile, generative systems expose student users to the threat of privacy breaches; subject creative and professional workers to data theft; exploit the crowd workers necessary for improving these systems (e.g., Wong, 2023); squander resources (Luccioni & Hernandez-Garcia, 2023); and promote misinformation, conspiracy theories, stereotypes and biases, including biases against people with disabilities (Gadiraju et al., 2023). They also can jeopardise and subvert a student's own learning. In the absence of solid research to the contrary, adopting these systems as beneficial assistive technologies for people with disabilities or vulnerable populations is at best premature.

Figure 8
Blueprint for an AI bill of rights for education

Rights for educators

Input on Purchasing and Implementation
You should have input into institutional decisions about purchasing and implementation of any automated and/or generative system ("AI") that affects the educational mission broadly conceived.

Input on Policies
You (or your representative in the appropriate body for governance) should have input into institutional policies concerning "AI" (including automated and/or generative systems that affect faculty, students, and staff).

Professional Development
You should be able to opt out of assignments that may put your own creative work at risk for data surveillance and use without compensation.

Autonomy
So long as you respect student rights, you should decide whether and how to use automated and/or generative systems ("AI") in your courses.

Protection of Legal Rights
You should never be subjected to any automated and/or generative system that impinges on your legal rights (including but not limited to those stated above).

Rights for students

Guidance
You should be able to expect clear guidance from your instructor on whether and how automated and/or generative systems are being used in any of your work for a course.

Consultation
You should be able to ask questions of your instructor and administration about the use of automated and/or generative systems prior to submitting assignments without fear of reprisal or assumption of wrongdoing.

Privacy and Creative Control
You should be able to opt out of assignments that may put your own creative work at risk for data surveillance and use without compensation.

Appeal
You should be able to appeal academic misconduct charges if you are falsely accused of using any AI system inappropriately.

Notice
You should be informed when an instructor or institution is using an automated process to assess your assignments, and you should be able to assume that a qualified human will be making final evaluative decisions about your work.

Protection of Legal Rights
You should never be subjected to any automated and/or generative system that impinges on your legal rights (including but not limited to those stated above).

Source: Author

4

RECONSTRUCTING AI

This section focuses on how we might reconstruct AI for human and planetary thriving. Carolina Earle describes an educational process that supported marginalised people to recognise and reclaim their agency as they realised their own transformative visions for more equitable AI futures for all. Maureen Ebben and Julien Murphy call for connecting generative AI tools with their environmental effects while shaping the technology towards the responsible management of planetary resources and improving the human condition. Tiera Tanksley identifies ways that the educational community—alongside technology companies and policymakers—can support schools in adopting AI in ways that protect and support all students. Jacob Pleasants invites readers to examine critically the stories guiding our thinking about AI and how these are wielded by those who already have power in our societies.

RECOMMENDED READINGS

- Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.
- Dhar, P. (2020). *The carbon impact of artificial intelligence.* Nature Machine Intelligence, 2(8), 423–425. https://doi.org/10.1038/s42256-020-0219-9
- Kanungo, A. (2023, July 18). *The green dilemma: Can AI fulfil its potential without harming the environment?* https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/
- Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism.* Public Affairs.
- Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. S., Mellor, J., ... & Gabriel, I. (2022). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214-229).
- Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1–5.

THE HEART IN THE MATTER: EDUCATION THAT FORGES AN EQUITABLE AND DIVERSE AI FUTURE

Carolina Earle
Gender Centre, Geneva Graduate Institute
EQUALS-EU: Europe's Regional Partnership for Gender Equality in the Digital Age, Switzerland

Key takeaways:

- Test and embed innovative pedagogical methods such as innovation camps.
- Ensure innovation processes are critical and participatory and tackle root structural challenges.
- Foster diverse and critical thought from an equitable, human-centred perspective.
- Centre psychic and emotional safety in innovative ideation programmes.
- Design creative, inclusive outreach and recruitment processes that meaningfully engage and make space for historically marginalised peoples.

The inventors, owners and beneficiaries of AI have emerged from predominantly WEIRD (Western/White, educated/English-speaking, industrialised, rich, democratic) societies. Further, while one in three companies today use AI in their operations, women only make up 26% of the global AI workforce and 18% of AI C-suite executives. Systemic historic, gendered, racialised, class and location-based disparities structure the development of AI. An impending "AI divide" threatens to disadvantage marginalised communities and concentrate power and control in the hands of Big Tech.

The exclusion of diverse peoples from AI discourse, research, design and development increases the likelihood that AI—and AI in education—will also be developed and deployed in exclusionary and marginalising ways.

Conversely, systemic practices of digital equity and inclusion throughout AI ecosystems ensure the worst harms may yet be controlled and designed against. The EQUALS-EU project demonstrates the positive impact of participatory, feminist, innovative methods to achieve a more equitable, inclusive, diverse and just AI future through and for education.

The EQUALS-EU Project

The EU Horizon 2020–funded, multisectoral and multistakeholder EQUALS-EU project was designed to forge gender-equitable social innovation ecosystems and systemically elevate gender equity inventors and leaders.

The project first mapped gender equity stakeholders in social innovation ecosystems. Subsequently, unique, co-created and localised digital equity and inclusion innovation camps and hackathons were held in 19 countries. Winning teams from each country's hackathon/innovation camp then attended an incubator programme and an international 3-week digital equity and inclusion capacity-building summer school.

Feminism and women's leadership in international law and policy

Over two and a half days, over 30 women and gender-minority people came together in Switzerland's EQUALS-EU Innovation Camp to ideate novel digital equity and inclusion solutions. They were hosted and fed free of charge, an essential incentivising and equity choice. After the innovation camp, of the 29 survey respondents, 69% stated they would definitely recommend the experience and 90% felt they had developed novel insights into digital inclusion.

The winners of our 2022 innovation camp ideated an AI-based solution to tackle revenge porn. One affirmed that "our idea was generated by an empowering environment," and said "I learnt that I am capable of thinking about solutions to complex problems…I had no idea I had an entrepreneurial vein in me!"

With the innovation camp as a key example, over the 3 years of the EQUALS-EU project, best practices for equitable and inclusive AI and education policy have emerged.

Policymakers and educators should

- Test and embed innovative pedagogical methods such as innovation camps. Amid rapid advancements in the digital world, complex and cutting-edge ideas can be taught, enjoyed and owned by new learners

in short timeframes. Meaningfully engaging topic specialists is essential for the success of these specialised learning processes.

- Ensure innovation processes are critical and participatory and tackle root structural challenges. Different lived realities may be pathologised as needing to be "solved"; this must be resisted.
- Foster diverse and critical thought that encourages individuals to navigate digital advancements from an equitable, human-centred perspective.
- Centre psychic and emotional safety in innovative ideation programmes. This will allow for diverse participants to safely create new worlds with—and constructively challenge—one another.
- Design creative, inclusive outreach and recruitment processes that meaningfully engage and make space for historically marginalised peoples. The crisis of equity and inclusion in AI must be overcome. This crisis is not one of apathy or creativity on the part of excluded communities but one of systemic exclusion.
- Challenge feelings of illegitimacy in educational/innovation settings by affirming that all lived experiences engaging with the digital world(s)—or not—constitute unique, meaningful, deep and relevant knowledge.

"Without community there is no liberation"

EQUALS-EU's successes would have been impossible without co-creation and collaboration across consortium partners and local stakeholders. Partners' implementation of project deliverables and frequent sharing of processes and experiences across country contexts ensured that the project was iteratively designed. Hurdles were collaboratively overcome and best practices refined.

Despite the power of community, policymakers and project leads must ensure that feminist, participatory projects in particular do not veer into exploitative terrain. Public-private partnerships must ensure that the results of a participatory ideation process benefit all creators equitably, both financially and through long-term sustained mentorship and support.

Self-reflexivity by collaborators, sustained funding and collaboration will ensure that participatory, feminist and innovative digital and AI education projects can realise their transformative visions, forging a more equitable, inclusive, diverse and just AI future for all.

ENVIRONMENTAL GENERATIVE AI AWARENESS IN EDUCATION

Maureen Ebben
University of Southern Maine, USA
Julien Murphy
University of Southern Maine, USA

Key takeaways:

- Environmental GenAI awareness is an integrated pedagogy that foregrounds the material effects of generative artificial intelligence.
- The content of the pedagogy includes carbon emissions, extraction of rare earth minerals, the diversion of water and land resources and the inextricable connection of these effects with human health and social equity.
- Instruction in GenAI awareness empowers students to act responsibly and collectively to mitigate environmental disasters and fosters an understanding of themselves and others as embedded in the material world.
- Teaching GenAI awareness allows instructors to fulfil our environmental obligation to Generation Z and beyond.

The challenges for educators presented by generative artificial intelligence (GenAI) have ballooned since OpenAI released ChatGPT in 2022. Freely accessible for lower-level functions and rapidly able to produce detailed text and images from vast internet sources in response to student prompts, GenAI may transform education in profound ways. Educators' focus has been on its appropriate use, its veracity and plagiarism. But equally important is the broader context of this tech revolution. The French philosopher of technology Bernard Stiegler was among the first to consider AI within its geological context—the existential crisis of the Anthropocene. He urged us to consider the relationship between the rise of AI and the acceleration of planetary destruction. Billions are invested to produce powerful, energy-hungry, non-human agents precisely when planetary survival is imperilled. Although GenAI may offer significant environmental solutions, such as optimising energy consumption, these come with environmental effects that contribute to the climate crisis and often go unaddressed. It is time for educators to explore these material effects by teaching environmental GenAI awareness in their classes.

There may be good reasons for reluctance on the part of educators to teach environmental GenAI awareness, such as

- educators may be overworked or outmatched by rapid technological developments whose processes are mystified and concealed behind "black box" proprietary control;
- many educators have narrowly defined content areas;
- the fundamental issues of environmental GAI awareness have not been identified for educators; and
- climate anxiety for students might increase with there being a focus on the environment.

Despite these good reasons, there are compelling arguments for teaching environmental GenAI awareness, such as

- if we are to grapple effectively with the warming planet, the climate must be an issue that informs every part of education;
- using GenAI without regard for its material effects reflects human chauvinism and is ultimately unsustainable;
- students are rightly demanding responsible stewardship of the earth because of the urgency of the warming climate, and disregard for the environmental costs of GenAI appears reckless and irresponsible; and
- environmental GenAI awareness can empower students towards collective action to mitigate environmental disaster and diminish climate anxiety.

Addressing the impact of GenAI calls for an integrated pedagogy, one that identifies key issues for educators and connects GenAI tools with their environmental effects. First and foremost, GenAI has a significant carbon footprint because it relies on the fossil-fuelled infrastructure of computer chip manufacturing and data centres. Training, maintaining and updating faster and increasingly powerful language models generates greenhouse gas emissions. The carbon footprint of GenAI searches is already five times higher than search engine queries. All of this occurs at a time when, instead of accelerating carbon emissions, we should be aiming for net zero emissions to hold the temperature to an internationally agreed-upon 1.5 degrees Celsius increase.

Further, GenAI's toll on natural resources is considerable, requiring heavy use of minerals, water and land. Minerals such as cobalt, lithium and tantalum are non-renewable, as are the rare-earth minerals neodymium and dysprosium, which are often procured with minimal environmental safeguards. Lithium extraction for GenAI uses 500,000 gallons of water for every ton of the metal extracted. Substantial water use is also necessary to cool the supercomputers that run every GenAI prompt. For instance, Open AI, the creator of ChatGPT, draws large amounts of water from the watershed of two rivers in Iowa to keep its supercomputers from overheating. The water supplies of warmer climates may be especially stressed, contributing to the global water crisis. Land use is also an issue. GenAI uses vast data centres that displace farmland. Instruction in environmental GenAI awareness might show students how such environmental impacts are inextricably connected to human health, potentially exacerbating social inequity as students come to ask questions such as whose water and whose land?

Most importantly, educators who teach environmental GenAI awareness may help students engage critically with the values shaping the technology in ways that advance the responsible management of vital resources and ensure human flourishing.

Priorities for future GenAI education policymaking

- Offer educators professional development opportunities for learning about the local and global material effects of GenAI.
- Develop curricular expertise in environmental ethics.
- Promote experiential GenAI learning activities for students.

Figure 9
Environmental generative AI awareness education

Source: Authors

EDTECH IS NOT NEUTRAL, HOW AI IS AUTOMATING EDUCATIONAL INEQUITY

Tiera Tanksley
University of California, Los Angeles, USA

Key takeaways:

- Raise awareness among educational stakeholders.
- Prepare students to be advocates for algorithmic equity.
- Conduct algorithmic equity audits.
- Formalise educational policies related to algorithmic racism.

Since the advent of COVID-19 and the shift to distance learning, the field of education has witnessed an unprecedented influx of the use and development of artificial intelligence (AI) technologies. These technologies are used in a variety of ways, including identifying at-risk students and predicting future drop-outs (Ekowo & Palmer, 2016; Bañeres, Rodríguez, Guerrero-Roldán, & Karadeniz, 2020); supporting classroom management and creating individualised educational plans (Rakap, 2023; Saeger, 2017; Tran, 2022); monitoring school safety and detecting disciplinary violations (Warner & Mackey, 2022); and much more. Because of its seemingly endless affordances in the school setting, AI is often positioned as the "silver bullet" needed to remediate long-standing issues of educational inequity facing historically marginalised students (Sengupta, 2023; Adiguzel, Kaya, & Cansu 2023).

Remote proctoring platforms, for example, use AI to detect perceived suspicious, abnormal or off-task behaviour by test takers in real time (Flaherty, 2023). However, because these platforms often employ facial detection systems that fail to recognise Black faces more than half of the time (Feathers, 2021; Buolamwini, 2023), Black students have an exceedingly hard time completing their exams without triggering the faulty detection systems, which results in locked exams, failing grades and unnecessary disciplinary action.

Some of the most popular school safety platforms have been known to falsely identify discussions about queer identity, race-related content and African American Vernacular English as dangerous or in violation of school disciplinary policies (Kelly, 2023, Feathers & Mehrotra, 2023, Warren & Markey, 2022). Because many of these platforms are directly connected to law enforcement (Keireleber, 2022), students that are falsely identified are contacted by police both on campus and in their homes. Considering that Black youth endure the highest rates of discipline, assault and carceral contact on school grounds (Crenshaw, Ocean & Nanda, 2015; Love, 2023,) and are six times more likely than their white peers to have fatal encounters with police (Badolato et al, 2020), the risk of experiencing algorithmic bias can be life threatening.

Threats to students' safety, privacy and academic wellbeing remain the same even when these technologies are designed to support—rather than surveil—student learning (Laird et al, 2022). In fact, a 2023 report by the Center for Democracy and Technology found an overall decrease in the use of EdTech for teaching and learning since its first study in 2020, with over 60% of teachers reporting that the software is used primarily to identify disciplinary infractions (Laird & Dwyer, 2023). Even further, Black students and students with disabilities endure significantly higher rates of AI-mediated discipline, not only from being disproportionately surveilled by AI-supported learning technologies but also from using AI programmes and platforms designed to make their learning experiences more accommodating and accessible (Laird & Dwyer, 2023). This could include using AI to support executive functioning, access translated or simplified language or provide alternative learning strategies.

When they are not being surveilled or over disciplined for using AI to support learning, Black students nevertheless confront algorithmically biased learning technologies

that provide inadequate, inequitable and inaccurate outputs that produce disparate learning outcomes and educational experiences. For instance, predictive analytic systems falsely identify Black youth as at risk for academic failure and future criminals at disproportionately high rates (Feathers, 2023); AI-assisted grading systems score Black test takers significantly lower than human reviewers (Feathers, 2019); and generative AI chatbots disseminate stereotypical, historically inaccurate and even racially violent results about Communities of Color (Brockwell, 2023: Biddle, 2022).

Though the aforementioned studies took place in the USA, their implications for students, stakeholders and policymakers around the globe are far reaching. First, a majority of AI technologies used for education originate in Silicon Valley in the USA, meaning the racial logics, Western prejudices and white masculinized world views that Silicon Valley assumes to be a generalizable or 'one size fit all' model will be implemented worldwide, regardless of if they truly are beneficial to all. This is a problem when we consider that a majority of the world's population—and by default, the world's school age children—fall outside of the scope of the imagined user these tools are designed to benefit: white, male, able bodied, cisgendered, heteronormative, documented, English speaking, and so on.

How can schools ensure that they are adopting AI technologies designed, deployed and implemented in ways that protect and support all students—particularly those students whose humanity continues to be a topic of national and international debate (e.g. Students of Colour; students with disabilities; queer and trans students; pregnant and parenting students; students seeking abortion; refugee, migrant or undocumented students; unhoused students; students experiencing apartheid, colonial occupation or genocide, etc.)?

The answer is complicated and requires a coalition between Big Tech (e.g., designing technologies that are more racially just) and policymakers (e.g., passing legislation that protects consumers from algorithmic harm and also penalises perpetrators of algorithmic racism). Though not exhaustive, the following list can be a starting point for schools that are looking to take a more race-conscious, justice-oriented and historically conscious approach to supporting educational equity in the age of EdTech and generative AI:

Raising Awareness among Educational Stakeholders: According to a study conducted by the Center for Democracy and Technology, a large majority of parents and caregivers are completely unaware of the types of technologies being used to monitor, protect and assess their children in school (Laird & Grant-Chapman,2021). A similar study revealed that a large swath of educators reported disparities in EdTech training, with some reporting that they received little to no training on how to properly use AI systems or assess their outcomes ethically. All educational stakeholders—including school resource officers, who are often tasked with responding to algorithmic content flags—must be made aware of the harms that algorithmic biases can have on students of colour. This training must also consider the history of anti-Black dehumanisation in schools prior to the rise of EdTech, as well as how these technologies exacerbate existing disparities.

Preparing Students to be Advocates for Algorithmic Equity: Students are the most directly impacted by algorithmic racism within EdTech and should therefore be taught how to critically identify, navigate and advocate against algorithmic racism in real time. Adopting critical AI literacy curricula into classrooms (Tanksley, 2024), hosting monthly book clubs on algorithmic racism or injustice and convening a student advisory board tasked with critically assessing prospective EdTech solutions before they are implemented are just a few ways to prepare young people to have more agentic relationships with technology in their schools.

Conducting Algorithmic Equity Audits: Implementing expensive educational technologies in large swaths before thoroughly investigating their potential to harm disadvantaged communities could prove to be a costly mistake. To avoid this, educational stakeholders should conduct algorithmic equity audits—examinations of the data collected and curated, the decision-making processes employed and the outcomes produced in order to proactively identify biases—on educational technologies before launching them. With such high financial investments dedicated to EdTech, allocating a subset of funds to third-party algorithmic equity audits would be a worthwhile investment.

Formalising Educational Policies Related to Algorithmic Racism: Schools must also adopt clear procedures for identifying, reporting and auditing educational technologies that threaten educational equity for all students. Without clear procedures, reporting structures or intervention protocols in place, it becomes exceedingly difficult for members of school communities to advocate against and ultimately rectify algorithmic harm.

CRITICALLY EXAMINING THE STORIES WE TELL ABOUT AI IN EDUCATION

Jacob Pleasants
University of Oklahoma, USA

Key takeaways:

- Critically examine the kinds of stories being told about AI, who is telling those stories, and why.
- Take care not to unwittingly trap our own conversations within a narrow story frame.
- What assumptions does the story make about AI?
- What does the story focus on, and what is left out?
- Who is telling the story, and why might they want to tell this kind of story?
- What other stories could be told, and what might those stories reveal?

I recently participated in an online course on teaching with AI in higher education. The course was put together in the wake of the ChatGPT launch in November 2022, and at the time, the issue of academic dishonesty had been the subject of countless opinion pieces. As I worked my way through the course, I began to think more deeply about the stories that we tell about AI, both in general and in its applications in education. The course engaged with the problem of cheating, but the developers clearly wanted to tell a different story about ChatGPT in particular and AI more broadly.

There are several typical stories told about technology writ large, with characteristic terminology, trajectories and assumptions. Conversations about AI are often patterned by these narratives, and I saw many examples of this in the AI course. The following are three that I encountered most often:

- **The disruptor.** AI technology has landed in education, and nothing will ever be the same. Change is inevitable, and educators need to move rapidly to figure out how to best navigate this new environment. The following is an illustrative quote from the AI course: "Generative Artificial Intelligence (AI) tools like ChatGPT have disrupted higher education, forcing educators to rethink assignments, activities, courses, and questions."
- **A powerful tool.** From low-tech hammers to high-tech digital gadgets, our technologies have expanded our capabilities, and they can be used for good as well as for ill. When it comes to AI, educators need to learn how to best use them to unlock their potential benefits and avoid their potential harms. Students, too, should be taught to use the new powerful AI tools for good rather than, say, academic shortcuts or dishonesty.
- **A perilous/promising future.** Even more important than what AI can do now are the perils that the future will hold as the technology progresses. One version of this story is the "doomer" scenario (beloved in cinema), where humans create a general intelligence that escapes our control. More often, though, are less fantastical versions that speak of the promise of AI. For instance, the following quote is from the AI course: "Advanced AI tutors and mentors will enable customized and immersive learning to maximize student outcomes."

Though mostly absent from the AI course, there is also the following fourth powerful counter-narrative to the above (e.g. Williamson et al., 2023):

- **Instruments of power.** Technologies are designed and controlled by those who have power (political, economic, social) in our society. Those people will deploy AI to serve their own interests, often at the expense of the marginalised. In educational spaces, AI will be used to surveil, sort, manage and otherwise control students and teachers.

Each of these stories frames our conversations and thinking about AI. They establish the issues that can be discussed, the kinds of arguments that can be made and the language that is used. Every story contains claims about AI that are worthy of debate, but we also need to debate the stories themselves. For instance, the first three stories are techno-centric in that they position AI as a deterministic 'prime mover' of change, thereby erasing the human actors who shape what those technologies are and will be. The fourth story, on the other hand, foregrounds those human actors.

As we talk about crafting policies around AI, we must critically examine stories that guide our own and others' thinking.

5 SCRUTINISING AI PROVIDERS

Roxana Radu shows how data harvesting and digital footprints may further deepen existing societal inequalities. In their scrutiny of AI providers, Joyeeta Dey and Radhika Gorur show how users—rather than technology—are being configured in ways that exploit their labour and time to the detriment of their analysis and use of data and analysis to improve teaching, learning and education systems. Ben Williamson shows how education is being made "AI-friendly"—rather than AI being made to be useful to schools and educators—and the challenges that poses to learning, teaching and curriculum. Jessica Pidoux and Paul Dehaye show how our existing educational, political and social institutions could become a catalyst for change, but only if decision makers address AI design as a collaborative community exercise, enable individuals to benefit from the data that they generate and democratise access to innovation funding. Alejandro Artopolous identifies three levels of digital inequality in AI: it is Anglo-centric, automates the fragmentation of public dialogue and deepens multiple digital inequalities.

RECOMMENDED READINGS

- Aparici, R., Fernando Raúl, A. B., & Martínez-Pérez, J. (2021). Alfabetización algorítmica basada en la metodología de Paulo Freire. *Perfiles educativos*, 43(SPE), 36–54.
- Artopoulos, A. (2023). Labyrinths of platformization of education in the Global South (and beyond). In C. Cobo & A. Rivas (Eds.), *The new digital education policy landscape: From education systems to platforms.* Taylor & Francis.
- Eynon, R., & Young, E. (2021). Methodology, legend, and rhetoric: The constructions of AI by academia, industry, and policy groups for lifelong learning. *Science, Technology, and Human Values, 46*(1), 166–191.
- Gorur, R., & Dey, J. (2021). Making the user friendly: The ontological politics of digital data platforms. *Critical Studies in Education, 62*(1), 67–81.
- Pidoux, J., Gursky, J., Bowyer, A., & Dehaye, P.-O. (2022). *Digipower technical reports: Understanding influence and power in the data economy.* Zenodo. https://doi.org/10.5281/zenodo.6554156
- Schneider, B. (2022). Multilingualism and AI: The regimentation of language in the age of digital capitalism. *Signs and Society, 10*(3), 362–387.
- Williamson, B. (2023). The social life of AI in education. *International Journal of Artificial Intelligence in Education.* https://doi.org/10.1007/s40593-023-00342-5

DIGITAL FOOTPRINTS, DATA INEQUALITIES AND AI

Roxana Radu
Blavatnik School of Government, University of Oxford, UK

Key takeaways:

- The success of generative AI is contingent on its ability to deliver personalised content, which depends on digital traces.
- Digital footprints are likely to deepen, rather than moderate, existing inequalities.
- When there are data gaps, the AI products on offer have a limited problem-solving capacity.
- The structure of incentives currently in place to promote continuous data harvesting needs to be reconsidered and limits have to be defined for the automatic collection of information from the environment of the user.

As artificial intelligence (AI) tools are becoming mainstream, the AI revolution is no longer regarded as an economic panacea (McKinsey, 2017). In policy discussions, the balance is tilting from opportunities to threats and to various levels of risk and structural societal changes to come. Among the hyperbolised AI promises of increased productivity and knowledge, the potential for furthering inequality is beginning to be recognised in a more consistent manner, in both global policy documents and national AI strategies. Yet, the impetus for the new AI for the development agenda remains high (Hassan, 2023), despite the fact that AI progress—and its benefits—are highly concentrated in the countries where AI is being developed and commercialised by a handful of big tech companies.

Data determines the success of the newly developed models and their applications, independent of their origins. Therefore, this contribution focuses on the critical role played by digital footprints in deepening, rather than moderating, inequalities, affecting life opportunities for unknown periods of time. Digital traces, or the "aggregate of data derived from the digitally traceable behaviour and online presence associated with an individual" (Micheli et al., 2018), represent the power horse of AI, without which the personalisation of educational or health content would not be possible. But effective personalisation depends on more data being continuously integrated, which comes with its own challenges.

Advances in neural networks, machine learning and large language models (LLMs) such as OpenAI's ChatGPT and Google's BARD have started to be integrated across many sectors on an ever larger scale. These models are by definition data-hungry, gathering—sometimes in real-time—billions of data points that are processed and stored using modern computing hardware and complex infrastructure. While training sets might indicate an endpoint for their data collection and analysis (e.g. ChatGPT-4's training set included information up to September 2021), in reality, each and every interaction with these new systems is also tracked and added to the models for new versions (e.g. GPT-5).

Such technology already supports or automates decision-making across a wide variety of contexts, from gaming to humanitarian assistance (Radu & Olliaro, 2023). Within the generative AI systems developed so far, every click, every preference and every word online is absorbed, ignoring the potential sensitivity of different data categories. Intersecting disparities and digital divides are thus compounded by digital footprints (Allmann & Radu, 2023), adding a new set of vulnerabilities to the users of AI. The risk is that the AI tools in use today can further entrench data inequalities across commercial applications and public services.

By powering the accessibility, effectiveness and personalisation of AI services, digital footprints become important in two ways.

First, data gaps limit the number and quality of benefits various AI systems bring. Those with minimal data footprints, including the most vulnerable among us, will end up with products that are less personalised, and therefore less relevant to their needs. Most AI systems in use today rely on data collected in high-income countries. The resulting biases and caveats affect new AI products beyond the contexts in which they were developed. Moreover, their own problem-solving capacity is limited, due to limitations in the diversity of their data and training sets and the diversity in the people designing them. As such, these products remain optimised for addressing certain categories of problems and may not cater adequately to other needs. It is critical for decision makers to recognise these built-in limitations for the AI for development agenda.

Second, to address these data gaps, data harvesting is likely to intensify, in particular, in jurisdictions with insufficient data protection safeguards. Many new AI applications will collect not only more individual data but also information specific to the environment in which they are used (from sensors, body monitoring devices, etc.). Adding more data, however, does not reduce the existing biases encoded in AI systems, such as in relation to age, gender, nationality or race. Due to the "black box" nature of LLMs, existing types of bias continue to retain power and will affect more life opportunities than ever before, revealing and creating new vulnerabilities over a yet unknown period of time.

The multi-modal AI we are witnessing today is ever more deeply interwoven with the social fabric, but it offers no technological fixes to existing inequalities. Rather, it has the potential to deepen them. It is therefore key to consider how our digital traces are built into the systems to come and what needs to be changed now.

THE "FRIENDLY USERS" WHO FEED AI

Joyeeta Dey
National Institute of Advanced Studies, Bengaluru, India
Radhika Gorur
Deakin University, Melbourne, Australia

Key takeaways:

- The needs of school-level actors as well as the challenges they face in generating these data must be factored into reform efforts.
- School-level actors and others whose labour is demanded should be consulted in setting objectives of data collection.
- Investment should be made not only to facilitate the uploading of data but also its dissemination and use.
- Appointing personnel dedicated to uploading data in low-resource schools would reduce the burden on teachers and free them up to actually engage productively with the data and analyses to improve their own work and their students' learning outcomes.
- Limiting what data is demanded, investing in developing a collective vision of how data could be usefully deployed and developing the capacity and the infrastructure for realising such visions is necessary.

'User friendliness' is a well-established concept in technology design. To encourage more and more diverse groups to adopt new technologies, designers strive to make their technologies user friendly, adding features that make the technology easier to use. To achieve this, technology companies may invite potential users to test technologies and provide feedback so that the technology could be improved. However, even as the technology is being tweaked, the user is also being "configured"—that is, technologies are designed to constrain how users engage with them and how they are disposed towards the use of technology (Woolgar, 1990).

One aspect of "configuring the user" is what we have termed "making the user friendly"—a phenomenon which became visible in our research on digital infrastructures of education in India (Gorur & Dey, 2021). We found that the state made substantial investments in influencing the user, persuading reluctant teachers and headteachers to labour in ways that would sustain the technology.

The idea that technology companies exploit our labour to fuel their business is beginning to be more widely understood. Instagram and Facebook might be "free" services, but they are only free if we don't acknowledge the free labour we provide as users by uploading content and allowing our data to be collected. Without this free labour, these platforms could not function. Similarly, in ed-tech, tech-creators offer teachers 'tech-upskilling' courses, often for free, but skilling up users is a way to enrol more people into using their technology (Thompson, 2022). Studies have shown that even when people are aware they are being manipulated, they are willing to go along with it because they have something to gain (for example, Williams et al., 2017).

Like corporations, states also depend on the labour of 'uploaders' to sustain their systems, but the uploaders neither have the option of not participating nor do they necessarily benefit from their labour. We observed this in our study of India's Unified Digital Information on School Education (UDISE). The largest Education Management Information System in the world, UDISE contains information on nearly 1.5 million schools, 9.5 million teachers and 265 million students on a wide range of indicators. Although this database is presented as serving the needs of schools and communities, a comprehensive study found that UDISE data was mainly used for government financial planning (Bordoloi & Kapoor, 2018). Neither parents nor teachers were aware of how the data could be put to use for their benefit. The state provided extensive training to ensure the input of data by teachers but almost none on how the data and analyses from the system could be utilised by (and be useful to) those who laboured to populate the database. The key function of the database appeared to be to enforce accountability. In other words, teachers and headteachers were being incorporated into developing instruments that would mainly be used in their own surveillance and monitoring.

How is the state able to persuade teachers and headteachers to provide such labour when the burden of non-teaching work is already so high? Our study identified several strategies ranging from inspirational messaging that framed uploading of data as a patriotic duty to regulation that threatened to cancel the registration of schools that did not comply with data demands.

India has launched ambitious plans for the deployment of artificial intelligence (AI) in public policy and administration, including in education. Education databases form the basis of AI interventions, such as the identification of the potential school dropouts in the state of Andhra Pradesh. Globally, the UNESCO initiative Digital Transformation Collaborative (DTC), of which companies such as Google, Ericsson and Microsoft are founding members, is looking to encourage nations of the Global South to embark on comprehensive digitisation of their education systems, which will enable the widespread use of AI.

As it becomes necessary for more and more data to fuel the hi-tech dreams of nations and global agencies, it is crucial to become aware of the labour inequities that may come in their wake and to adopt fairer and more productive approaches to data generation. Much is written about how important it is for states to regulate private technology companies. While this is of course critically important, states should also ensure they themselves do not engage in extractive practices.

MAKING EDUCATION AI-FRIENDLY

Ben Williamson
Centre for Research in Digital Education and Centre for Sociodigital Futures, University of Edinburgh, UK

Key takeaways:

- Independent evaluations should be commissioned to assess the effectiveness of AI applications targeted at schools, since the weak evidence base currently means schools may be sold products on the basis of exaggerated marketing.
- Research funders should fund interdisciplinary research to examine the implementation of existing AI applications in schools and to understand their intended effects as well as unintended consequences.
- Authorities should ensure that schools are not used as live testing sites for AI by requiring independent testing for any policy decisions concerning widespread use.
- Teacher unions and representative organisations should support teachers' capacity to critically evaluate AI applications and campaign to ensure that teachers' professional pedagogic autonomy is protected from automation.
- School leaders should not mandate the use of AI applications by staff and students and should promote forms of AI literacy to enable them to evaluate and make informed decisions about AI.

After OpenAI released ChatGPT, industry analysts reported it was the "fastest growing internet app in history" (Bove, 2023). Characterising ChatGPT as an app, however, is misleading. As has since become clear, ChatGPT (v.3.5) was launched strategically as a marketing demo to entice users to subscribe to its fee-paying artificial intelligence (AI) models and encourage institutional customers to purchase enterprise licenses for GPT-4. Educational companies including Khan Academy and Duolingo were early partners, gaining first-mover advantage to promote AI in education. Now, educational technology (EdTech) businesses are rushing to incorporate AI into their products, backed by expectant investors and political enthusiasm (Komljenovic et al., 2023).

As OpenAI's strategy shows, AI development is not about building "apps". It is about building AI infrastructure as the technical foundation for all industries and sectors to operate upon. This explains why Microsoft has invested billions in OpenAI and why other Big Tech companies, such as Google, Amazon and Meta, are also racing to release AI "foundation models". The competition for infrastructural dominance in AI is fierce, with huge financial rewards expected for the winners (Williamson, 2023).

Building the foundations is one challenge. The other is getting people to use the services that will be built on top of them. Studies of infrastructure highlight that complex systems are never just technical. They also require people to use them. A successful infrastructure therefore requires the accustomisation of users to its affordances, "making the user friendly" so they will amenably respond to what the technology allows (Gorur & Dey, 2021).

In education, making the user friendly to AI infrastructure has become a key aim of a wide range of organisations and individuals. Teachers and learning institutions are to be made AI-friendly through a variety of training courses and guidance that might be characterised as "PedagoGPT". PedagoGPT captures how pedagogic advice is being formulated to accustom educational actors to AI. Examples include training courses in AI provided by entrepreneurial educators or via online programs as well as guidance offered directly by Big Tech companies as part of their strategic aim to stretch AI infrastructure throughout the sector. A prime example is the "Teaching with AI" resources produced by OpenAI, a 'guide for teachers using ChatGPT in their classroom'.

Through these PedagoGPT initiatives, schools are being targeted as potential users and customers of AI. The intended result is a synchronisation of AI with pedagogic routines and administrative procedures. The effects of synching schools with AI could be profound. Infrastructure is never merely a technical backdrop upon which other activities take place but actively shapes the practices of its users.

Generative AI infrastructure will be generative of particular effects, including unintended consequences (Holmes, 2023). For example, the entrepreneurs behind foundational AI infrastructure and EdTech applications privilege particular narrow conceptions of "personalised" and "mastery" learning, resulting in measurable individual achievement improvements. The capacity of AI to power "personalised learning tutorbots" reinforces this reductionist, privatised and atomised vision of education.

Embedding education in AI infrastructure also privileges commercial technologies as "co-pilots" in the classroom, potentially degrading teachers' pedagogic autonomy by outsourcing responsibilities to automated technologies (Kerssens & van Dijck, 2022). This risks displacing teachers' subject expertise and contextual knowledge to computerised data-processing machines.

Another risk is that AI will put degenerative pressure on the quality of knowledge taught in schools. AI language models often produce plausible but false information or biased and discriminatory content due to the available

material they draw upon (Goodlad & Baker, 2023). The danger here is that teachers and students may find it increasingly difficult to tell whether AI is delivering them authoritative and accurate sources or just convincing but fallacious content.

Making education AI-friendly therefore poses distinct challenges to learning, teaching and curricula. AI infrastructure is set to coordinate a wide range of educational practices, with PedagoGPT guidance intended to synchronise schools with their affordances. This is despite the lack of independent evidence that AI can improve education in the ways claimed or serious consideration of the risks or unintended consequences of introducing AI into schools. Far from being merely technical foundations for teaching and learning, AI infrastructure includes social and political technologies with potentially profound and unpredictable impacts on public education. Rushing to make education AI-friendly would serve to amplify technological power over schooling. Instead, educators and schools should be supported to take a cautious and critical stance to AI (McQuillan et al., 2023).

THREE INEQUITIES IN AI DEVELOPMENT AND SOCIETAL IMPLEMENTATION

Jessica Pidoux
University of Neuchâtel and PersonalData.IO, Switzerland
Paul-Olivier Dehaye
Hestia.ai and PersonalData.IO, Switzerland

Key takeaways:

- Situate innovation within its community of adoption, involving multiple stakeholders in scientific processes and providing financial support according to their contribution, so that the public influences beyond a general consultation in policymaking.
- Support educational programs on diverse ways of innovating, emphasising the renovation of existing material and reduction of ecological consequences. Rethink AI business models at the local level along the lines of the circular economy.
- Regulate the pace and rate at which organisations accumulate data and produce content from/within our states and stimulate non-digital solutions so that the state regains more power and sustainability than platforms.
- Set up a threshold of funding for the most powerful corporations to stop reinforcing their dominating position, ensuring financial risk does not hold back the emergence of alternatives from new players in the market.

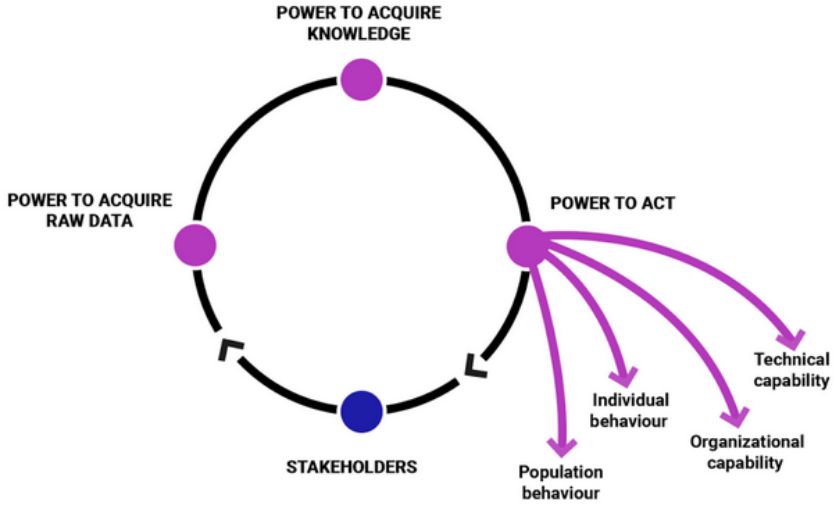
Connect AI design with society

The first inequity is in the expertise and authority to decide the design of technologies, where a gap can exist between individuals developing computer programs and those consuming the services produced. A developer creating AI algorithms or a mobile application for calculating workers' revenues in on-demand marketplaces, assisting in skin-cancer diagnoses or predicting fertility cycles has the technical skills for facilitating human tasks via automated systems. However, marginalised users (e.g. gig economy workers, racialised patients, women or transgender individuals) may suffer from the incompleteness of data samples and material constraints of technologies such as missing data, calculation errors, memory crashes, network malfunctions and imprecise devices. These users have complementary knowledge of how such technical issues affect them offline, but the design of AI and the requirements of users are often disconnected. Instead, the design caters to market needs with economic profit as the target before addressing social concerns. Disruptive change is required in computing methods and innovation business models if they are to consider systems' local integration with user practices in schools and personal devices. It is more than likely that biases and harms will be reproduced if the infrastructures remain the same.

Allow individuals to benefit from the data they generate

The second inequity is in the gap in capability to unlock the potential of value creation between entities collecting personal data and those producing it. Organisations, both public and private, have the power to infer information about the population to improve decision-making and shape social order. Personal data can help, for example, to analyse unemployment for professional reintegration, COVID-19 propagation for preventive quarantine lockdowns, consumer habits to increase sales and voting trends to redirect political orientation. However, individuals who supply such data for aggregation do not own it and cannot reuse it to extract value of their own.

Figure 10
Infrastructural data power loop



Source: Authors

Corporate markets are being fostered in Europe for improving data flows with standard norms of accessibility and interoperability, and the General Data Protection Regulation offers individual data access rights. However, these solutions remain unknown and impracticable for regular users who are highly dependent on the data collector's transparency. They are affected by automated systems' decisions and bear the burden of proving a decision inaccurate without owning their data, nor adequate tools or methods at their disposal.

Democratise access to innovation funding

The third inequity is access to innovation funding between the corporations that dominate the technology market and small entrepreneurs (that may end up being bought by the former). In contrast to the dominating corporations, small, local entrepreneurs lack technological power, big data, organisational power and financial power. Therefore, entrepreneurs innovate by imitating or counter-imitating big corporations to finally sell their ideas and recover some investment. Smaller ideas serve to accumulate power for those in a higher position in the technological market who are also financing research in academia to later absorb more intellectual capital in the industry.

While there are increasing national funding programs for innovation involving the general public in science, these local actors have little impact on the economy. When focusing on the common good, these actors' projects rely mostly on volunteer work, in contrast to corporate structures.

AI AND UNEQUAL KNOWLEDGE IN THE GLOBAL SOUTH

Alejandro Artopolous
University of San Andrés, Argentina

Key takeaways:

- Even if school leaders are not proficient in English, it is necessary for them to realise the differences in generative AI behaviour. Therefore, it is necessary to train them in AI literacy.
- AI literacy is a novel practice that needs research and development independent of AI-producing companies. Governments and academia must take the lead in developing a deep AI literacy curriculum that engages the development of diverse applied computational thinking in the Global South.
- Since the pandemic, governments in the Global South have been improvising an unplugged digital education policy agenda that increases digital inequalities. It is necessary to set a renewed agenda for a new decolonial digital education policy.

One of the key issues in debates around artificial intelligence (AI) and education is that they exacerbate inequities. In addition to the biases that AI possesses by design, systematic errors can discriminate by race, gender, sexual orientation, physical conditions or political bubbles, both by data infrastructure and by algorithms. Generative AI has the capacity to empower students who have already reached their information literacy thresholds, while at the same time disconnecting from society those who for various reasons have not been able to overcome their digital divides. AI's digital inequalities operate at three levels—namely, the sociolinguistic nature of textual generative AI, AI literacy as human cognitive capacity and a new digital education policy.

Sociolinguistic

Generative AI produces disparities between language regions. Since it is an English-centric technology, the same kind of quality of such technologies is not available to learners who use other languages. The production of generative AI requires significant investment in machine learning, therefore when business ask for choices, English becomes a priority language. Also, their models are trained by data that by and large is in English. This results in lower-quality responses in other languages and higher costs to develop reliable AI systems in non-English languages. As a result, English-speakers have an advantage. The underprivileged ones could be Catalan-, Igbo-, Quechua- or Spanish-speakers. Although it seems to be an issue similar across the Global North and Global South, it acts on different levels of disparities for each language region/population (Schneider 2022).

Cognitive

Challenges such as misinformation and hate speech raise the dilemma in democracies of building critical computational thinking skills and digital citizenship. Generative AI undermines public trust by making available chatbot services that do not ensure 100% truthful content in their responses.[1] AI augments and automates the fragmentation of public dialogue through polarised opinion networks. With weak AI,[2] mistrust spreads in the public debate, but with generative AI, our students can increase their mistrust in truthful knowledge (Heintz 2022).

Since it was assumed that the digital divide could be solved by granting access to connectivity and basic digital literacy, we now struggle to consider adding a new layer of AI literacy. A new kind of digital literacy is required that addresses data literacy and applied computational thought. This kind of AI literacy varies by field of activity, from art to marketing, agriculture or engineering, which implies that there is no generic computational thinking that applies to every field of practice. It is therefore necessary to rethink curriculum design at all levels of education to integrate regular literacy with algorithmic and data literacy. The level of response to the reconfiguration of digital education policies varies across the Global North and Global South. Thus, corporations in the Global North have created a disruption that can break the stable practices of modern schooling, but the antidote is only available to OECD education systems (Tedre & Vartiainen 2023).

The cognitive dimension of AI inequalities in education is particularly defining because it depends not only on the readiness of school leaders but also on the strategic direction they choose to take. The socio-technical imaginaries about the application of AI to education are crucial to designing democratic models of citizenship development. Educational policies with a solutionist approach tend to promote inequalities if a government encourages the adoption of AI as a robot mentor for solving the learning of discrete content and does not consider new teachers' AI competencies. In contrast, if a government acknowledges that AI, as well as devices, are cognitive capabilities that should be incorporated

into curricular designs, these countries are aware that the restitution of equality always relies on the human side of the ensembles.

Policy

Even if Global South educational leaders were aware of how to implement AI literacy, they would not have the financial resources to prepare schools for such a challenge. The COVID-19 pandemic highlighted the global divide in access to technology between OECD countries and the Global South. While in OECD countries, only 10-20% suffered from a homework divide, in the Global South, this gap could in extreme cases be as high as 90%. But at this level, talking about the divide is an oversimplification; it is far from a black-and-white situation, there are several grey areas. In middle-income countries such as Argentina, Uruguay and Chile, we can find three segments; one is similar to the OECD, another similar to the typical Global South, while in-between, children and teachers could have been connected during the pandemic, but when they went back to school, they lost the ability to learn and teach in a connected way. We called this little-explored third situation "Silvester platformization". We say "Silvester" platformisation because it is the feasible experimentation in the Global South (without infrastructure or teacher training) of the transition from the modern school to a cloud-ready classroom. Only a tiny fraction of the population can access a cloud-ready classroom. As the process of platformisation in education, including the use of AI, creates inequity by design, low and middle-income countries tend to develop unplugged (from the internet) computational education policies. We are facing an advancement of the socio-technical trajectory of the cloud-ready classroom that is in fact causing a restriction in the freedom of access to information, and in turn, a geopolitical digital divide of access to the educational cloud (Tedre & Vartiainen 2023).

Footnotes

1. Tech Companies in their narrative attempt to anthropomorphise AI tends to talk about "hallucinations". They explain that it "is a phenomenon wherein a large language model (LLM) perceives patterns or objects that are nonexistent, creating outputs that are nonsensical or altogether inaccurate." https://www.ibm.com/topics/ai-hallucinations The use of the word hallucination tends to anthropomorphise AI, something which we are trying to avoid in this publication.
2. Generative AI has the potential to cause a leap in scale from artificial narrow intelligence (ANI) to artificial general intelligence (AGI). ANI, or weak or narrow AI designed to specialise in a specific task, is limited to a specific or narrow area and cannot operate outside the parameters predefined by its programmers, so it cannot make decisions on its own—that is, the ultimate decision rests with the human. In contrast, AGI is achieved if a machine acquires human-level cognitive capabilities. Additional advancements would be required for that leap to occur, including in reasoning, memory and contextual and ethical decision-making.

CONTRIBUTORS

ALEJANDRO ARTOPOULOS is a Professor of Technology and Educational Change at Universidad de San Andrés, and of Science & Technology Studies (STS) at Universidad de Buenos Aires, and Researcher at CIC. PhD Universitat Oberta de Catalunya. Specialized in Digital STS. Interested in sociotechnical transitions towards hybrid education and smart agriculture, digital inequities and development of information capacities, and applied computational thinking.

ALAN BLACKWELL is Professor of Interdisciplinary Design in the Cambridge University department of Computer Science and Technology, and has worked as an AI researcher and engineer since 1985. He practised as a professional engineer in Aotearoa New Zealand before further study in Comparative Religion, Computer Science, Music, and Experimental Psychology. He is co-founder of the Crucible Network for research in Interdisciplinary Design, and of the Global Challenges strategic research initiative of the University of Cambridge.

STEPHANIE SMITH BUDHAI is an associate clinical professor in the School of Education at Drexel University and a certified K-12 teacher in Technology Education, Computer and Information Technology, and Instructional Technology. She holds a PhD in Learning Technologies and has published myriad books and articles focused on the intersections of equity, teaching, learning, and technology.

LUCILA CARVALHO is an Associate Professor and co-director of the Equity Through Education Research Centre at Massey University, Aotearoa New Zealand. Her research explores how knowledge, digital, material and social structures shape the design and use of technology, and how technology influences social and educational experiences.

KATHRYN CONRAD is Professor of English at the University of Kansas where she also serves as Faculty Fellow for Humanities Advancement and Fellow for the Center for Cyber-Social Dynamics. She works on technology and culture and is the author of *A Blueprint for an AI Bill of Rights for Education*.

JOYEETA DEY is a PhD student at the National Institute of Advanced Studies, Bengaluru. Her research is in the area of education governance, specifically digital transformations and information systems.

PAUL-OLIVIER DEHAYE is the CEO of Hestia.ai and founder of PersonalData.IO. He is a mathematician (PhD Stanford) whose work led to the revelation of the Facebook-Cambridge Analytica affair and to the making of the Netflix film The Great Hack.

CAROLINA EARLE (University of Oxford B.A., Geneva Graduate Institute M.A.) led the design and implementation of the Swiss work packages of the EQUALS-EU project under the supervision of project P.I. Dr Claire Somerville. A Missions Publiques Global Fellow, she works at the intersection of community-based and IO programming and policy work.

MAUREEN EBBEN is Associate Professor and Chair of the Department of Communication and Media Studies at the University of Southern Maine in Portland, Maine. An award-winning educator, she teaches classes on information security and social media. Her recent publications include works on generative AI in higher education, social media influencer authenticity, humanizing the online classroom, virtual pandemic teaching, common digital art in the digital age, feminist questions for artificial intelligence, and human labor as a complement to smart machines.

MOIRA V FAUL is Senior Lecturer at the Geneva Graduate Institute and Executive Director of NORRAG, the network for international policies and cooperation in education and training. She was previously Deputy Director of the Public-Private Partnership Center at the University of Geneva. Before her PhD, Moira was a senior manager in private sector adult education in Spain and China, then Head of Education and Youth Policy (UK) at Oxfam GB.

LAUREN M E GOODLAD is Distinguished Professor of English and Comparative Literature at Rutgers University where she chairs the Critical AI@Rutgers initiative and is editor of *Critical AI* (Duke University Press). She is lead co-PI for several grant-funded initiatives including Design Justice AI (a Mellon-CHCHI-funded global humanities institute, University of Pretoria) and Design Justice Labs (an NEH-funded collaborative approach to studying and teaching about "generative AI," Australian National University and the University of Pretoria).

RADHIKA GORUR is Associate Professor of Education at Deakin University, a Director of the Laboratory of International Assessment Studies, Convenor of the Deakin Science and Society Network, and an Editor of Discourse: Studies in the Cultural Politics of Education. Her research is in the fields of education policy, aid and development, the sociology of quantification, and critical data studies.

EMMA HARDEN-WOLFSON is an Assistant Professor in the Faculty of Education at McGill University, Canada. She is an international and comparative higher education policy specialist with regional specializations in Central Asia, Canada, Europe, and Latin America.

MARIE K HEATH is Assistant Professor of Technology and Education in the School of Education at Loyola University Maryland and former high school social studies teacher. Marie's scholarship works toward advancing technology education for more just futures. She is co-founder of the Civics of Technology project (CivicsofTechnology.org), which shares research and curriculum encouraging teachers and students to critically inquire into the effects of technology on our individual and collective lives.

JEREMY KNOX is Associate Professor of Digital Education at the University of Oxford and an Official Fellow of Kellogg College. His research interests are focused on the relationships between education, data-driven technologies and wider society, and he has led projects funded by the ESRC and the British Council in the UK. Jeremy's published books include: *Posthumanism and the Massive Open Online Course* (2016); *Artificial Intelligence and Inclusive*

Education (2019); *Data Justice and the Right to the City* (2022); and *AI and Education in China* (2023).

JANJA KOMLJENOVIC is a Senior Lecturer at Lancaster University, where her research focuses on the political economy of higher education, and the digitalisation, datafication and platformisation of universities. She led an ESRC-funded research project, 'Universities and Unicorns: building digital assets in the higher education industry' investigating new forms of value in the digitalised higher education sector. She has published internationally on higher education policy, higher education markets, and educational technology.

DANIEL G KRUTKA is Associate Professor of Social Studies Education at the University of North Texas. His research focuses on intersections of technology, education, and society. He co-hosts the Visions of Education podcast (https://visionsofed.com/) and co-leads the Civics of Technology project (https://www.civicsoftechnology.org/).

FELICITAS MACGILCHRIST is Professor of Digital Education and Schooling at the University of Oldenburg, Germany where she heads Re:Lab, the Praxis Lab for Reimagining EdTech Futures. She is co-editor of Learning, Media and Technology. Recent books and special issues include Die datafizierte Schule (The Datafied School), Postdigital Participation in Education, and Designing Postdigital Futures. She's on Mastodon @discoursology@social.coop.

RAÏSSA MALU is an international education consultant. She directed the Education Project for the Quality and Relevance of Teaching at Secondary and University Levels for the Ministry of Primary, Secondary and Technical Education of the Democratic Republic of Congo. This World Bank-funded project aimed to improve the quality of science and mathematics teaching and learning at secondary level. She is also Director of the non-profit organization Investing In People (RDC), which organizes the annual Science and Technology Week. She is a corresponding member of Belgium's Royal Academy of Overseas Sciences.

LINA MARKAUSKAITE is a Professor of Learning Sciences, the University of Sydney, Australia. She received a PhD in informatics (computer sciences). Her research projects have been mainly concerned with understanding the nature of complex professional knowledge work and learning, and how human capabilities are entangled with digital technologies.

PUNYA MISHRA (punyamishra.com) is associate dean of Scholarship & Innovation at Mary Lou Fulton Teachers College, Arizona State University, where he leads multiple initiatives, providing a future-forward, equity-driven, collaborative approach to educational research. He is internationally recognized for his work in educational technology, creativity, and the application of design to educational innovation.

JULIEN S MURPHY is Professor of Philosophy at the University of Southern Maine in Portland, Maine. She teaches cybersecurity and environmental ethics courses and has recently published on generative AI in higher education, virtual pandemic teaching, digital privacy, and the university of the future.

NICOLE OSTER is pursuing a Ph.D. in Learning, Literacies and Technologies at the Mary Lou Fulton Teachers College at Arizona State University. She has a background in secondary English education and teaching English to speakers of other languages and is passionate about leveraging educational technology to foster equitable, inclusive, and innovative learning.

JESSICA PIDOUX is a digital sociologist at the University of Neuchâtel, where she works on the gig economy and citizen science methods. Jessica holds an engineering PhD from the École Polytechnique Fédérale de Lausanne in Digital Humanities. She is also the director of the NGO PersonalData.IO advocating for personal data protection.

JACOB PLEASANTS is an assistant professor of Science Education at the University of Oklahoma. His research focuses on teaching about the intersections of science, technology, engineering, and society.

ROXANA RADU is an Associate Professor of Digital Technologies and Public Policy at the University of Oxford's Blavatnik School of Government and a Hugh Price Fellow at Jesus College. Her research focuses on Internet policy-making, cybersecurity and AI governance. Since January 2023, she is the elected Chair of the Global Internet Governance Academic Network (GigaNet).

EMMA RUTTKAMP-BLOEM is Professor and Head of the Department of Philosophy, University of Pretoria, South Africa. She is a member of the UN Secretary-General's AI Advisory Body and UNESCO Women4EthicalAI in addition to chairing the UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (COMEST).

TIERA TANKSLEY's scholarship bridges education, Black studies critical science and technology studies to examine how digital and artificially intelligent technologies impact the lives and schooling experiences of Black youth. Her work simultaneously recognizes Black youth as digital activists and civic agitators, and examines the complex ways they subvert, resist and rewrite racially biased technologies to produce more just and joyous digital experiences for Communities of Color across the diaspora. In 2023, she was awarded the Public Voices in Technology fellowship from the MacArthur Foundation and the Op Ed Project, and was subsequently named one of the 100 Brilliant Women in AI Ethics of 2024.

BEN WILLIAMSON is a senior lecturer at the Centre for Research in Digital Education, University of Edinburgh, UK. His research interests are in education policy and digital Technologies.

References

Abid, A., Farooqi, M., & Zou, J. (2021, July). Persistent anti-muslim bias in large language models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 298-306).

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Agus, R, & Samuri, S. (2018). Learning analytics contribution in education and child development: A review on learning analytics. *Asian Journal of Assessment in Teaching and Learning*, *8*, 36-47.

Alenichev, A., Kingori, P., & Peeters Grietens, K. (2023, August 9). Reflections before the storm: The AI reproduction of biased imagery in global health visuals. *The Lancet Global Health*. https://doi.org/10.1016/S2214-109X(23)00329-7

Allmann, K., & Radu, R. (2023). Digital footprints as barriers to accessing e-government services. *Global Policy*, *14*, 84–94. https://doi.org/10.1111/1758-5899.13140

Badolato, G. M., Boyle, M. D., McCarter, R., Zeoli, A. M., Terrill, W., & Goyal, M. K. (2020). Racial and ethnic disparities in firearm-related pediatric deaths related to legal intervention. *Pediatrics, 146*(6).

Baker, T. & Smith, L. with Anissa, N. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. Nesta. https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf

Bañeres, D., Rodríguez, M. E., Guerrero-Roldán, A. E., & Karadeniz, A. (2020). An early warning system to detect at-risk students in online higher education. *Applied Sciences, 10*(13), 4427.

BBC. (2021, March 28) AI: Ghost workers demand to be seen and heard. https://www.bbc.com/news/technology-56414491

Bender, E. M., & Hanna, A. (2023, April 7). A GPT-4 Fanfiction Novella [Audio podcast episode]. In *Mystery AI Hype Theater 3000*. https://www.buzzsprout.com/2126417/13460873

Bender, E. M., & Koller, A. (2020, July). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185-5198).

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?-. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity.

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J. & Caliskan, A. (2023, June). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1493–1504) https://doi.org/10.1145/3593013.3594095

Biddle, S. (2022) The internet's new favorite AI proposes torturing Iranians and surveilling mosques. https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/

Birch, K. (2022). EdTech's precarious futures: Are there material limits to data-driven Higher Education? NORRAG Blog.

Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., & Bao, M. (2022, June). The values encoded in machine learning research. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* (pp. 173-184).

Blackwell, A. F. (2023, 9 November). *Oops! We automated bullshit*. University of Cambridge. https://www.cst.cam.ac.uk/blog/afb21/oops-we-automated-bullshit

Blackwell, A. F. (2024). *Moral codes: Designing alternatives to AI*. MIT Press. https://moralcodes.pubpub.org/

Bock, A., Breiter, A., Hartong, S., Jarke, J., Jornitz, S., Lange, A., & Macgilchrist, F. (Eds.). (2023). Die datafizierte Schule. Springer VS. https://doi.org/10.1007/978-3-658-38651-1

Bode, K., & Goodlad, L. M. E. (2023, October). Data worlds: An introduction. *Critical AI*, *1*(1–2), 1–23.

Boly Barry, K. (2022). Impact of the digitalization of education on the right to education. Report of the Special Rapporteur on the Right to Education. OHCHR. https://www.ohchr.org/en/documents/thematic-reports/ahrc5032-impact-digitalization-education-right-education

Bonini, T., & Trere, E. (2024). *Algorithms of resistance: The everyday fight against platform power.* MIT Press.

Bordoloi, M., & Kapoor, V. (2018). India: using open school data to improve transparency and accountability. UNESCO-IIEP. https://www.iiep.unesco.org/en/india-using-open-school-data-improve-transparency-and-accountability

Bove, T. (2023, 2 Feb). The A.I. revolution is here: ChatGPT could be the fastest-growing app in history and more than half of traders say it could disrupt investing the most. *Fortune:* https://fortune.com/2023/02/02/chatgpt-fastest-growing-app-in-history-could-revolutionize-trading/.

Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: Changing user needs and motivations. *Interactions*, *25*(5), 38–43.

Brock Jr, A. (2020). *Distributed Blackness*. New York University Press.

Brockell, G (2023). We 'interviewed' Harriet Tubman using AI. It got a little weird. https://www.washingtonpost.com/history/interactive/2023/harriet-tubman-articial-intelligence-khan-academy/

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., ... & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. *arXiv preprint arXiv:2303.12712*.

Bulger, M. (2016). Personalized learning: The conversations we're not having. *Data & Society*. https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf

Buolamwini, J. (2023). *Unmasking AI: My mission to protect what is human in a world of machines*. Random House.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability and Transparency* (pp. 77-91). PMLR.

Carvalho, L., Martinez-Maldonado, R., Tsai, Y.S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers & Education: Artificial Intelligence*, *3*, 100053. https://doi.org/10.1016/j.caeai.2022.100053

CAST. (2022, February 8). Universal design for learning guidelines. https://udlguidelines.cast.org/

Century. (2021, January 18). How Eton College uses technology to bolster traditional teaching. https://www.century.tech/news/how-eton-college-uses-technology-to-bolster-traditional-teaching/

Clark, M. D. (2014). To tweet our own cause: A mixed-methods study of the online phenomenon "Black Twitter" (Doctoral dissertation, The University of North Carolina at Chapel Hill).

Conrad, K. (2023). Blueprint for an AI bill of rights for education. *Critical AI*, *2*, 1. https://criticalai.org/2023/07/17/a-blueprint-for-an-ai-bill-of-rights-for-education-kathryn-conrad/

Conrad, K. (2023). Critical AI literacy in a time of chatbots: A public symposium for educators, writers, and citizens [Video]. YouTube. https://www.youtube.com/watch?v=OWTudCltwgU

Conrad, K., & Goodlad, L. M. E. (2024, forthcoming). Beyond resistance. *Critical AI*. https://criticalai.org/blog-feed/

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.

Currie, M., Knox, J., & McGregor, C. (2022). *Data justice and the right to the city*. Edinburgh University Press.

Czerniewicz, L., & Carvalho, L. (2022). Open, distance, and digital education (ODDE)—An equity view. In O. Zawacki-Richter & I. Jung (Eds.), *Handbook of Open, Distance and Digital Education.* Springer. https://doi.org/10.1007/978-981-19-0351-9_93-1

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education, 210*, 104967.

Davies, H. C., Eynon, R., & Salveson, C. (2021). The mobilisation of AI in education: A Bourdieusean field analysis. *Sociology*, *55*(3), 539–560. https://doi.org/10.1177/0038038520967888

Dehaye, P. and Pidoux, J. (2022). The value and drawbacks of data power in education, in Numa Hopkins, A. and M. V. Faul (Eds.). *The Digitalisation Of Education. NORRAG Policy Insights #01.* NORRAG.

Dixon-Román, E., Nichols, T. P., & Nyame-Mensah, A. (2020). The racializing forces of/in AI educational technologies. *Learning, Media and Technology, 45*(3), 236-250.

Drahl, C. (2023, October 6). AI was asked to create images of Black African docs treating white kids. How'd it go? *NPR*. https://www.npr.org/sections/goatsandsoda/2023/10/06/1201840678/ai-was-asked-to-create-images-of-black-african-docs-treating-white-kids-howd-it-

Druga, S., Williams, R., Park, H. W., et al. (2018). How smart are the smart toys? Children and parents' agent interaction and intelligence attribution. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*, Trondheim, 19–22 June 2018 (pp. 231–240).

Ekowo, M., & Palmer, I. (2016). The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis. *New America*.

Estrada, D. (2023). Teaching insights: How to teach AI to students. *Critical AI*. https://criticalai.org/2023/05/30/teaching-insights-how-to-teach-ai-to-students-ai-ethics-and-njit-audit-project-dr-daniel-estrada/

Faul, M. V. (2023) Education, technology and private sector actors: towards a research agenda. NORRAG blog. https://www.norrag.org/education-technology-and-private-sector-actors-towards-a-research-agenda/

Faul, M. V., & Hopkins, A. N. (2023). AI in education and research: Towards a more ethical engagement. NORRAG. https://www.norrag.org/ai-in-education-and-research-towards-a-more-ethical-engagement/

Feathers, T. (2019). Flawed Algorithms Are Grading Millions of Students' Essays. Retrieved from: https://www.vice.com/en/article/pa7dj9/flawed-algorithms-are-grading-millions-of-students-essays

Feathers, T. (2023). How Wisconsin uses race and income to label students 'high risk'. Retrieved from https://www.chalkbeat.org/2023/4/27/23699361/dropout-early-warning-system-dews-student-dropouts-race-income-data/

Feathers, T. & Mehrotra, D. (2023). Inside America's school internet censorship machine. *Wired Magazine.* Retrieved from https://www.wired.com/story/inside-americas-school-internet-censorship-machine/

Feathers, T. (2021). Proctorio is using racist algorithms to detect faces. https://www.vice.com/en/article/g5gxg3/proctorio-is-using-racist-algorithms-to-detect-faces

Fergusson, G., Schroeder, C., Winters, B., & Zhou, E. (Eds.). (2023, May), *Generating harms: Generative AI's impact and paths forward*. Electronic Privacy Information Center. https://epic.org/wp-content/uploads/2023/05/EPIC-Generative-AI-White-Paper-May2023.pdf

Flaherty, C. (2020). Big proctor. *Inside Higher Ed.* https://www.insidehighered.com/news/2020/05/11/online-proctoring-surging-during-covid-19

Financial Times. (2023, March 29). Big tech companies cut AI ethics staff, raising safety concerns. https://www.ft.com/content/26372287-6fb3-457b-9e9c-f722027f36b3

Fleur, D. S., Marshall, M., Pieters, M., Brouwer, N., Oomens, G., Konstantinidis, A., Winnips, K., Moes, S., Van den Bos, W., Bredeweg, B., & van Vliet, E. A. (2023). IguideME: Supporting self-regulated learning and academic achievement with personalized peer-comparison feedback in Higher Education. *Journal of Learning Analytics*, *10*(2), 100-114. https://doi.org/10.18608/jla.2023.7853

Flowers, D. (2023, November 13). Why you should rethink your resistance to ChatGPT. *The Chronicle of Higher Education.* https://www.chronicle.com/article/why-you-should-rethink-your-resistance-to-chatgpt

Fosch-Villaronga, E., Van der Hof, S., Lutz, C., et al. (2023). Toy story or children story? Putting children and their rights at the forefront of the artificial intelligence revolution. *AI and Society*, *38*(1), 133–152.

Frankfurt, H. G. (2005). *On bullshit*. Princeton University Press.

Friesen, N. (2020). The technological imaginary in education, or: Myth and enlightenment in 'personalized learning'. In M. Stocchetti (Ed.), *The Digital Age and Its Discontents* (pp. 141–160). University of Helsinki Press.

Gadiraju, V., Kane, S., Dev, S., Taylor, A., Wang, D., Denton, E., & Brewer, R. (2023, June). "I wouldn't say offensive but...": Disability-centered perspectives on large language models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 205–215). https://doi.org/10.1145/3593013.3593989

Garg, R., & Sengupta, S. (2020). Conversational technologies for in-home learning: Using co-design to understand children's and parents' perspectives. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Gartner. (2023). Gartner Hype Cycle. https://www.gartner.com/en/research/methodologies/gartner-hype-cycle

Gebru, T. (2020). Race and gender. *The Oxford Handbook of Ethics of AI, 251-269.*

Givens, J. R. (2021). *Fugitive pedagogy: Carter G. Woodson and the art of Black teaching*. Harvard University Press.

Goodlad, L. M. E. (2023, October 6). From critical AI literacy to design justice AI. https://www.youtube.com/watch?v=OWTudCltwgU

Goodlad, L. M. E., & Baker, S. (2023). *Now the humanities can disrupt "AI".* Public Books.

Goodlad, L. M. E., & Stoerger, S. (2023). Rutgers AI Council, Teaching critical AI literacies. https://docs.google.com/document/d/1TAXqYGid8sQz8v1ngTLD1qZBx2rNKHeKn9mcfWbFzRQ/edit#heading=h.kgds7i8l6uca

Goodlad, L. M. E., Stone, M., & Sammons, M. Beyond Chatbot-K: On Large Language Models, "Generative AI" and the Rise of Chatbots: An Introduction. forthcoming in Critical AI 2.1.

Gorur, R., & Dey, J. (2021). Making the user friendly: The ontological politics of digital data platforms. *Critical Studies in Education*, *62*(1), 67–81.

Graeber, D. (2019). *Bullshit jobs: A theory*. Penguin Books.

Guimarães, M. E. L. (2023). *The AI gender gap: Living in a world designed by men*. Wartsila. https://www.wartsila.com/insights/article/the-ai-gender-gap-living-in-a-world-designed-by-men

Harari, Y.N. (2018). *21 lessons for the 21st Century*. Random House

Hassan, Y. (2023). *AI is Africa's new development mantra, but can it fix development?* Bot Populi. https://botpopuli.net/ai-is-africas-new-development-mantra-but-can-it-fix-development/

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The Weirdest People in the World? *Behavioral and Brain Sciences, 33(2-3), 61-83.*

Herold, B. (2022) All that edtech schools bought during the pandemic won't improve equity: Here's why. https://www.edweek.org/technology/all-that-ed-tech-schools-bought-during-the-pandemic-wont-improve-equity-heres-why/2022/11

Holmes, W. (2023). *The unintended consequences of artificial intelligence and education*. Education International. https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education

Hundt, A., Agnew, W., Zeng, V., Kacianka, S., & Gombolay, M. (2022). Robots enact malignant stereotypes. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 743–756). https://doi.org/10.1145/3531146.3533138

Internet Matters Team. (2017). Revealed: The secret life of six-year-olds online. Internet Matters. https://www.internetmatters.org/hub/press_release/revealed-the-secret-life-of-six-year-olds-online/

Kalluri, P. R., Agnew, W., Cheng, M., Owens, K., Soldaini, L., & Birhane, A. (2023, October 17). The surveillance AI pipeline. *arXiv preprint arXiv:2309.15084* https://doi.org/10.48550/arXiv.2309.1508

Keierleber, M. (2022). Survey reveals extent that cops surveil students online—in school and at home. Retrieved from https://www.the74million.org/article/survey-reveals-extent-that-cops-surveil-students-online-in-school-and-at-home/

Kelly, J. (2023) Student monitoring tools should not flag LGBTQ+ keywords. Retrieved from https://www.eff.org/deeplinks/2023/06/student-monitoring-tools-should-not-flag-lgbtq-keywords

Kerssens, N., & Van Dijck, J. (2022). Governed by edtech? Valuing pedagogical autonomy in a platform society. *Harvard Educational Review*, *92*(2), 284–303.

Knox, J. (2023). *AI and education in China: Imagining the future, excavating the past*. Routledge.

Komljenovic, J., Sellar, S., Birch, K., & Hansen, M. (2024). Assetisation of higher education's digital disruption. In Williamson, B., Komljenovic, J., & Gulson, K. (Eds.), *World Yearbook of Education 2024 (pp. 122-139). Routledge*

Komljenovic, J., Williamson, B, Eynon, R., & Davies, H. (2023). When public policy 'fails' and venture capital 'saves' education: Edtech investors as economic and political actors. *Globalisation, Societies and Education*. https://doi.org/10.1080/14767724.2023.2272134

Krutka, D. G., Heath, M. K., & Mason, L. E. (2020). Technology won't save us—A call for technoskepticism in social studies. *Contemporary Issues in Technology and Teacher Education*, *20*(1), 108–120.

Kurian, N. (2023). AI's empathy gap: the risks of conversational artificial intelligence for young children's wellbeing and key ethical considerations for early childhood education and care. *Contemporary Issues in Early Childhood*.

Laird, E. Grant-Chapman, H. Venzke,C. & Hannah Quay-de la Vallee, H. (2022) Hidden harms: The misleading promise of monitoring students online. https://cdt.org/insights/report-hidden-harms-the-misleading-promise-of-monitoring-students-online/

Laird & Grant-Chapman (2021). Navigating the new normal: Ensuring equitable and trustworthy edtech for the future. https://cdt.org/insights/report-navigating-the-new-normal-ensuring-equitable-and-trustworthy-edtech-for-the-future/

Laird, E. & Dwyer, M. (2023). Off Task: EdTech Threats to Student Privacy and Equity in the Age of AI. https://cdt.org/insights/report-off-task-edtech-threats-to-student-privacy-and-equity-in-the-age-of-ai/

Livingstone, S., & Pothong, K. (Eds.). (2022). *Education data futures: Critical, regulatory and practical reflections.* Digital Futures Commission. https://educationdatafutures.digitalfuturescommission.org.uk/

Livingstone, S., & Pothong, K. (2023). Child rights by design: Guidance for innovators of digital products and services used by children. Education Data Futures. https://digitalfuturescommission.org.uk/wp-content/uploads/2023/04/CRbD_singles-web.pdf

Lorde, A. (1984). *Sister outsider: Essays and speeches.* Triangle Classics. https://rhinehartibenglish.weebly.com/uploads/2/2/1/0/22108252/sister_outsider_audrey_lorde_ib_pdf_packet.pdf

Love, B. L. (2023). Punished for dreaming: How school reform harms Black children and how we heal. St. Martin's Press.

Luccioni, A. S., & Hernandez-Garcia, A. (2023). Counting carbon: A survey of factors influencing the emissions of machine learning. *arXiv preprint arXiv:2302.08476.*. https://doi.org/10.48550/arXiv.2302.08476

Luccioni, S. (2023, December 4) The mounting human and environmental costs of generative AI. *arsTechnica.* https://arstechnica.com/gadgets/2023/04/generative-ai-is-cool-but-lets-not-forget-its-human-and-environmental-costs/

Luccioni, A. S., Viguier, S., & Ligozat, A. L. (2023). Estimating the carbon footprint of BLOOM, a 176b parameter language model. *Journal of Machine Learning Research, 24*(253), 1–15.

Luckin, R., Holmes, W., Griffiths, M. & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education

McCarthy , J. , Minsky , M. L. Rochester, N. and Shannon, C. E. (1955) A proposal for the Dartmouth summer research project on Artificial Intelligence. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

McKinsey Global Institute (2017). A future that works: Automation, employment, productivity. https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works/de-DE

McQuillan, D., Jarke, J., & Pargman, T.C. (2023). We are at an extreme point where we have to go all in on what we really believe education should be about. *Postdigital Science and Education*. 1-9. https://doi.org/10.1007/s42438-023-00433-5

Mahelona, K., Leoni, G., Duncan, S. and Thompson, M. (2023) OpenAI's Whisper is another case study in Colonisation. *Papa Reo blog. https://blog.papareo.nz/whisper-is-another-case-study-in-colonisation/*

Mahomed, S., Aitken, M., Atabey, A., Wong, J., & Briggs, M. (2023). AI, children's rights, and wellbeing: Transnational frameworks: Mapping 13 frameworks at the intersections of data-intensive technologies, children's rights, and wellbeing. The Alan Turing Institute. https://www.turing.ac.uk/news/publications/ai-childrens-rights-wellbeing-transnational-frameworks

Markauskaite, L. (2010). Digital media, technologies and scholarship: Some shapes of eResearch in educational inquiry. *Australian Educational Researcher*, *37*, 79–101. https://doi.org/10.1007/BF03216938

Markauskaite, L., Carvalho, L., & Fawns, T. (2023). The role of teachers in a sustainable university: from digital competencies to postdigital capabilities. *Educational Technology Research and Development*, *71*(1), 181–198. https://doi.org/10.1007/s11423-023-10199-z

Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham-Shum, S., Gašević, D., et al. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers & Education. Artificial Intelligence*, *3*, 100056, 1–16. https://doi.org/10.1016/j.caeai.2022.100056.

Martinez-Maldonado, R., Yan, L., Deppeler, J., Phillips, M., & Gašević, D. (2021). Classroom analytics: Telling stories about learning spaces using sensor data. In E. Gil, Y. Mor, Y. Dimitriadis, & C. Köppe (Eds.), *Hybrid Learning Spaces*. Springer.

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *The AI Index Report 2023.* Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

Mejias, U. A. & Couldry, N. (2024). *Data grab: The new colonialism of big tech and how to fight back*. Penguin.

Micheli, M., Lutz, C., & Büchi, M. (2018). Digital footprints: An emerging dimension of digital inequality. *Journal of Information, Communication and Ethics in Society*, *16*(3), 242–251.

MIT Technology Review (2020) We read the paper that forced Timnit Gebru out of Google. Here's what it says. *MIT Technology Review*. https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/

Montag, C., Lachmann, B., Herrlich, M., & Zweig, K. (2019). Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. *International Journal of Environmental Research and Public Health*, 16(14), 2612.

Morozov, E. (2013). To save everything, click here: The folly of technological solutionism. *Public Affairs.*

National Initiatives and Performance Directorate. (2018). *Education future frontiers: A conversation starter: Thinking for the future—Preparing students to thrive in an AI world*. NSW Department of Education. https://education.nsw.gov.au/our-priorities/innovate-for-the-future/education-for-a-changing-world/media/documents/future-frontiers-education-for-an-ai-world/Conversation-Starter-Thinking-for-the-future_5_AA.pdf

Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development, *Learning, Media and Technology*, *48*(1), 38–51. https://doi.org/10.1080/17439884.2022.2095568

Nichols, T.P. & Garcia, A. (2022). Platform studies in education. *Harvard Educational Review, 92*(2), 209-230. https://doi.org/10.17763/1943-5045-92.2.209

Noble, S. U. (2018) *Algorithms of oppression: How search engines reinforce racism.* New York University Press.

O'Reilly, T., Strauss, I., & Mazzucato, M. (2024). Algorithmic Attention Rents: A theory of digital platform market power. *Data & Policy*, 6, e6.

Paranjape, B., Ge, Y., Bai, Z., et al. (2018). Towards automatic generation of peer-targeted science talk in curiosity-evoking virtual agent. In *Proceedings of the 18th International Conference on Intelligent Kurian 7 Virtual Agents* (pp. 71–78).

Radu, R. (2024) The variable geometry of AI governance. *Global Policy Outlook, 2.* https://www.genevapolicyoutlook.ch/the-variable-geometry-of-ai-governance/

Radu, R., & Olliaro, E. (2023). Not child's play: Protecting children's data in humanitarian AI ecosystem. International Committee of the Red Cross. https://blogs.icrc.org/law-and-policy/2023/12/14/protecting-children-data-in-humanitarian-ai-ecosystems/

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020, February). Saving face: Investigating the ethical concerns of facial recognition auditing. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (pp. 145-151).*

Rakap, S. (2023). Chatting with GPT: Enhancing individualized education program goal development for novice special education teachers. *Journal of Special Education Technology*, 01626434231211295.

Ramiel, H., & Dishon, G. (2024). AI-shaped hole: Anticipation regimes and liminal policy rationalities. In B. Williamson, J. Komljenovic, & K. N. Gulson (Eds.). *World Yearbook of Education 2024. Digitalisation of Education in the Era of Algorithms, Automation and Artificial Intelligence.* Routledge.

Rosenzweig, J. (2023). *Teaching insights: What happens when a novice writer asks ChatGPT for editing advice?* Critical AI. https://criticalai.org/2023/09/21/teaching-insights-what-happens-when-a-novice-writer-asks-chatgpt-for-editing-advice-dr-jane-rosenzweig/

Runciman, D. (2023) *The handover: How we gave control of our lives to corporations, states and AIs.* Profile Books.

Russell, J.-E., Smith, A., & Larsen, R. (2020). Elements of success: Supporting at-risk student resilience through learning analytics. *Computers & Education*, *152*, 103890. https://doi.org/10/gjk2jr

Saeger, A. M. (2017). Using ClassDojo to promote positive behaviors and decrease negative behaviors in the classroom. Theses and Dissertations. 2443. https://rdw.rowan.edu/etd/2443.

Saini, A. (2019). *Superior: The return of race science*. Beacon Press.

Selwyn, N. (2017). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury.

Selwyn, N. (2019). What's the problem with learning analytics? *Journal of Learning Analytics*, *6*(3), 11–9. http://dx.doi.org/10.18608/jla.2019.63.3

Selwyn, N., & Gašević, D. (2020). The datafication of higher education: Discussing the promises and problems. *Teaching in Higher Education*, *25*(4), 527–540.

Sengutpa, D. (2023) The transformative power of AI in modern elearning: A new era for education. https://elearningindustry.com/the-transformative-power-of-ai-in-modern-elearning-a-new-era-for-education

Shah, C., & Bender, E. M. (2022). Situating search. In *Proceedings of the 2022 ACM SIGIR Conference on Human Information Interaction and Retrieval*. https://doi.org/10.1145/3498366.3505816

Shanahan, M. (2022). Talking about large language models. *arXiv preprint*, arXiv:2212.03551.

Shekhawat, G., and Livingstone, S. (2023). AI and children's rights: a guide to the transnational guidance. Media@LSE blog post. Available at https://blogs.lse.ac.uk/medialse/2023/11/01/ai-and-childrens-rights-a-guide-to-the-transnational-guidance/

Shew, Alison (2020). Ableism, technoableism, and future AI. *IEEE Technology and Society Magazine*.

Shrivastava, P., & Zsolnai, L. (2022). Wellbeing-oriented organizations: Connecting human flourishing with ecological regeneration. *Business Ethics, the Environment & Responsibility*, *31*(2), 386–397. https://doi.org/10.1111/beer.12421

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023, May). The curse of recursion: Training on generated data makes models forget. *arXiv e-prints, arXiv-2305.* https://doi.org/10.48550/arXiv.2305.17493

Smith Budhai, S., & Grant, K. L. (2022). *Culturally responsive teaching online and in-person: An action planner for dynamic equitable learning environments*. Corwin.

Suchman, L. (2023). The uncontroversial 'thingness' of AI. Big Data & Society, 10(2). https://doi.org/10.1177/20539517231206794.

Tanksley, T. (2024). "We're Changing the System with This One:" Black Students Using Critical Race Algorithmic Literacies to Subvert AI-Mediated Racism in School. *English Teaching: Practice & Critique*.

The Ada Lovelace Institute. (2021, June 23). *Prototyping AI ethics futures: Rights, access and refusal* [Video]. https://www.adalovelaceinstitute.org/event/prototyping-ai-ethics-futures-rights-access-refusal/

Thompson, K. (2022). The cruel optimism of educational technology teacher ambassador spaces. *Power and Education*, 17577438231164717.

Tong, A. (2023, November 16). *OpenAI explores how to get ChatGPT into classrooms*. Reuters. https://www.reuters.com/technology/openai-explores-how-get-chatgpt-into-classrooms-2023-11-16/

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019, July). Envisioning AI for K-12: What should every child know about AI? In *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 9795–9799.

Tran, M. (2022). The application of Class Dojo to manage and motivate students for online classes during Covid-19 pandemic. In *VietTESOL International Convention Proceedings*.

Tucker, E. (2022) *Artifice and Intelligence.* Tech Policy Press. *https://techpolicy.press/artifice-and-intelligence/*

UN Secretary General AI Advisory Body (2023) *Interim report. UN. https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf*

UNESCO (2021). *AI and education: Guidance for policy-makers.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709

UNESCO (2022). *Recommendation on the ethics of Artificial Intelligence.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137

UNESCO. (2019). *Beijing consensus on artificial intelligence and education*. UNESCO.

UNESCO. (2023a). *Technology and Education: A tool on whose terms? Global Education Monitoring Report 2023*. UNESCO.

UNESCO. (2023b). *An Ed-Tech Tragedy? Educational technologies and school closures in the time of COVID-19*. UNESCO.

UNESCO. (2023c). *Guidance for Generative AI in Education and Research*. UNESCO.

UNESCO (2023d). *Recommendation on the Ethics of AI*. UNESCO Social and Human Sciences.
UNESCO IESALC. (2023a). *ChatGPT and artificial intelligence in higher education: Quick start guide*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385146

UNESCO IESALC. (2023b). *Harnessing the era of artificial intelligence in higher education: A primer for higher education stakeholders*. https://unesdoc.unesco.org/ark:/48223/pf0000386670

United Nations Committee on the Rights of the Child. (2021). *General comment No. 25 on Children's rights in relation to the digital environment* (CRC/C/GC/25). https://www.ohchr.org/EN/HRBodies/CRC/Pages/GCChildrensRightsRelationDigitalEnvironment.aspx

Van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., et al. (2019). Social robots for language learning: A review. *Review of Educational Research 89*(2): 259–295.

Walsh, T. (2017). The AI revolution. In L. Loble, T. Creenaune, & J. Hayes. *Education for an AI world* (pp. 1–10). Melbourne University Press.

Wamala-Larsson, C. (2023). *Policy brief: Gender inclusive social innovation ecosystems*. EQUALS-EU. https://equals-eu.org/wp-content/uploads/2023/03/EQUALS-EU-D1.4.1.-Policy-brief5197.pdf

Warren, E. & Markey, E. (2022). Constant Surveillance: Implications of Around-the-Clock Online Student Activity Monitoring. https://www.warren.senate.gov/imo/media/doc/356670%20Student%20Surveillance.pdf

Watters, A. (2014). *The monsters of education technology*. Hack Education. http://monsters.hackeducation.com/

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, … Gabriel. J. (2021, December). Ethical and social risks of harm from Language Models. *arXiv preprint arXiv:2112.04359.* https://doi.org/10.48550/arXiv.2112.04359

White, G. (2018, December 11). Child advice chatbots fail to spot sexual abuse. BBC News, https:// www.bbc.co.uk/news/technology-46507900

Whittaker, M. (2021). The steep cost of capture. *Interactions, 28(6), 50-55.*

Whittaker, M., Alper, M., Bennett, C., et al. (2019). *Disability, bias, and AI*. AI Now Institute. https://ainowinstitute.org/publication/disabilitybiasai-2019

Williams, M., Nurse, J. R., & Creese, S. (2017, August). Privacy is the boring bit: user perceptions and behaviour in the internet-of-things. In *15th Annual Conference on Privacy, Security and Trust (PST)* (pp. 181-18109). IEEE.

Williamson, B. (2015). *Programmable schools? Governing education through code in the smart city*. Code Acts in Education. https://codeactsineducation.wordpress.com/2015/02/02/programmable-schools-governing-education-through-code-in-the-smart-city/

Williamson, B. (2019). Datafication of education. In Beetham, H., & Sharpe, R. (Eds.). *Rethinking pedagogy for a digital age: Principles and practices of design.* Routledge.

Williamson, B. (2023). Governing through infrastructural control: Artificial intelligence and cloud computing in the data-intensive state. In W. Housley et al. (Eds), *The Sage Handbook of Digital Society*. Sage.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Williamson, B., Komljenovic, J., & Gulson, K. (Eds.). (2023). *World Yearbook of Education 2024: Digitalisation of Education in the Era of Algorithms, Automation and Artificial Intelligence.* Taylor & Francis.

Williams, R., Park, H. W., Oh, L., et al. (2019). Popbots: Designing an artificial intelligence curriculum for early childhood education. In *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(1), 9729–9736.

Wired (2021, October 25) Inside the Facebook papers. https://www.wired.com/story/facebook-papers-internal-documents/

Woodson, C. G. (1933). *The miseducation of the Negro*. Africa World Press.

Woolgar, S. (1990). Configuring the user: the case of usability trials. *The Sociological Review, 38*(1_suppl), 58-99.

Xu, Y., & Warschauer, M. (2020). Exploring young children's engagement in joint reading with a conversational agent. In *Proceedings of the 19th ACM Conference on Interaction Design and Children* (pp. 216–228).

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

NORRAG Policy Insights #04, March 2024

GENEVA GRADUATE INSTITUTE
NORRAG GLOBAL EDUCATION CENTRE

Chemin Eugène-Rigot, 2
1202 Geneva, Switzerland
+41 (0) 22 908 45 47
norrag@graduateinstitute.ch

Join the global NORRAG network
www.norrag.org/norrag-network

Learn more:
www.norrag.org

Follow NORRAG on social media:

Learn more about the project:
https://www.norrag.org/ai