Nordisk tidsskrift for pedagogikk og kritikk
Volume 10 | 2024 | pp. 3–14

Essay

On the Limits of Artificial Intelligence (AI) in Education

Neil Selwyn
Monash University, Australia

ABSTRACT
The recent hyperbole around artificial intelligence (AI) has impacted on our ability to properly consider the lasting educational implications of this technology. This paper outlines a number critical issues and concerns that need to feature more prominently in future educational discussions around AI. These include: (i) the limited ways in which educational processes and practices can be statistically modelled and calculated; (ii) the ways in which AI technologies risk perpetuating social harms for minoritized students; (iii) the losses incurred through reorganising education to be more 'machine readable'; and (iv) the ecological and environmental costs of data-intensive and device-intensive forms of AI. The paper concludes with a call for slowing down and recalibrating current discussions around AI and education – paying more attention to issues of power, resistance and the possibility of re-imagining education AI along more equitable and educationally beneficial lines.

Keywords: *artificial intelligence; automation; digital; education; harms*

Received: October, 2023; Accepted: October, 2023; Published: January, 2024

Introduction

The past twelve months have seen artificial intelligence (AI) attract heightened levels of popular and political interest rarely seen before in the sixty-year history of the field. Much of this has been fuelled by financiers chasing quick profits, policymakers keen to appear supportive of national innovation, and Big Tech corporations scrambling to catch-up with more agile specialist start-ups. One consequence of this furore is the difficulty of now engaging in balanced and reasoned discussions about the societal implications and challenges of AI. For example, we have reached a point where the majority of US adults are now prepared to accept that "the swift growth of artificial intelligence technology could put the future of humanity at risk" (Reuters, 2023). This special issue of the *Nordic Journal of Pedagogy & Critique* therefore comes at a moment when a lot is being said about AI, albeit little of which is likely to hold up to scrutiny a few years hence.

Correspondance: Neil Selwyn, e-mail: neil.selwyn@monash.edu

Citation: Selwyn, N. (2024). On the Limits of Artificial Intelligence (AI) in Education. Nordisk tidsskrift for pedagogikk og kritikk: Special Issue on Artificial Intelligence in Education, 10, 3–14. http://doi.org/10.23865/ntpk.v10.6062

While not suffering the extreme peaks and troughs of general public discussions around AI, the education sector has also been experiencing its own version of AI-fever. This has perhaps been most obvious in educational reactions to Chat-GPT and other 'generative AI' writing tools capable of producing pages of plausible-sounding text in response to short written prompts. At the beginning of 2023, initial publicity around this particular form of AI raised widespread concerns over the likelihood of students using such tools to fraudulently produce written assignments. This triggered a succession of university and school-wide 'bans,' hasty reformulations of assessment tasks, and the rapid marketing of new AI counter-measures claiming to be capable of detecting algorithmically-generated writing. Observing this from the outside, it seemed alarming how quickly the educational debate around Chat GPT spiralled out of control, with many otherwise sober commentators reaching extreme conclusions over the transformative connotations of this technology.

This paper calls for more reasoned responses to the educational possibilities of AI. While educators should not *completely* ignore recent developments around machine learning, large-language models and the like, there is certainly a need to resist the more extreme hopes and fears that the idea of AI technology continues to provoke. At the same time, there is also a need to better engage with complex issues and concerns that have so far tended to remain sidelined in educational discussions around AI. This requires sustained, ongoing and open dialogue that brings in perspectives not usually given space in conversations around digital innovation and education futures. In particular, this requires paying closer attention to the experiences and standpoints of those groups likely to gain least (and likely to lose most) from the unfettered implementation of AI technology in education. As a start, this brief paper sets out some pertinent starting-points from which such discussions can progress in earnest.

AI and education – some basic points of definition

It is perhaps helpful to first set out the nature and form of the technology under discussion. While many teachers and students understandably might feel that they are yet to encounter this technology, tangible applications of AI in education are fast emerging. For example, government authorities and agencies are beginning to adopt various forms of 'automated education governance' where AI tools are used to process big data sets from entire school systems in order to model 'business decisions' ranging from future school building priorities through to teacher recruitment. Conversely, individual schools are now beginning to assign all manner of tasks to AI that would previously have been delegated to teachers. These include automated grading and online exam proctoring systems, chat bots that automate general interactions between teachers and students, and surveillance tools which judge the extent to which a class is diligently working or not. At the same time, AI tools and diagnostics are also regularly part of how students are supported in their studies. This includes the use of AI-driven search, natural language processing to provide automated writing support, and the use of personalized learning systems to curate online learning content and activities for different students on the basis of their prior performance.

Crucially, while these applications might seem incredibly sophisticated in comparison to the educational technologies of the 2000s and 2010s, such examples all constitute what is termed 'narrow artificial intelligence.' In other words, these AI systems are designed to address one specific task (such as grading essays or predicting student behaviours). These AI tools are refined using training data relating to this specific area of education, and then operate within pre-defined boundaries to recognise patterns in a limited range of input data. Thus, the forms of AI currently entering our schools and classrooms are far-removed (if not totally distinct) from the speculative forms of AI that often feature in popular discussions of how 'sentient' forms of AI might soon replace teachers, render schools obsolete, and even do away with the need for humans to learn things for themselves. Thus, in contrast to the fears and hopes that have fast grown up around ideas of 'general AI,' 'digital minds,' 'superintelligence' and so-called 'singularity,' the first step in establishing a healthy response to the emergence of AI technologies into schools is to foreground what Divya Siddarth and colleagues (2021) term 'Actually Existing AI' – i.e. the computational limitations of this technology alongside the IT firms and flows of funding that are promoting it.

In particular, the idea of actually existing AI pushes us to frame educational AI in terms of maths, statistics and computation. As Hilary Mason (2018, n.p.) puts it, "AI is not inscrutable magic – it is math and data and computer programming, made by regular humans." Indeed, some elements of the computer science community have recently begun to deliberately distance themselves from the term 'AI' and revert to using labels that better describe the types of machine learning and algorithmic developments that underpin their work (see Jordon in Pretz, 2021). Elsewhere, policymakers and industry actors are also beginning to turn to alternate terms, such as 'automated decision making' and 'algorithmic forecasting.' Such linguistic turns reinforce Emily Tucker's (2022, n.p.) assertation that "whatever the merit of the scientific aspirations originally encompassed by the term 'artificial intelligence,' it [has become] a phrase that now functions in the vernacular primarily to obfuscate, alienate and glamorize."

Recognising AI as a sophisticated form of statistical processing quickly raises questions over what can (and what cannot) actually be accomplished with these technologies in education. For example, from this perspective, a seemingly sentient AI tool such as Chat GPT is more accurately understood as assembling and re-arranging pre-existing scraps of text taken from the internet in ways that are statistically likely to resemble larger pieces of pre-existing text. Generative AI – as with any AI tool – does not 'know' or 'understand' what it is doing any more than any other non-human object. Even if it is producing apparently plausible reams of text, a generative AI language tool has no 'understanding' or 'knowledge' of what its output might mean. Instead, just as a parrot can mimic human speech without any reference to meaning, so too will a large language model – albeit using sophisticated probabilistic information about how text has previously been put together by human authors (Bender et al., 2021). At best, then, these are statistical simulations, or more accurately, replications of human-produced text with none of the human ingenuity, imagination or insight that was used to produce the original source materials.

AI and education – some things to be concerned about

Understanding AI technology as a complex statistical procedure (based on enormous computational power and data processing) therefore pushes education debates on AI to reflect on some of the obvious limitations of this technology not usually acknowledged. For example, as with any computational process, AI technologies are reliant on the quality of the data they are working with. As with any computational process, AI technologies operate through iteration and optimisation, the use of approximations and correlations, the production of errors and false matches. All this makes the application and outputs of any AI system incredibly context-specific and inherently limited. As the computer scientist Melanie Mitchell (2019, n.p.) puts it: "People have been trying to get machines to reason since the beginning of the field [...] but they're what people call 'brittle' – meaning you can easily make them make mistakes, and reason incorrectly."

It is well worth thinking further about how this statistically-derived 'brittleness' might be evident in educational AI – in particular, taking time to consider how the statistical limitations of AI might bump up against educational contexts and educational ambitions. At its heart, the ontological premise of educational AI is that the social world of any student or classroom is broadly quantifiable and subject to statistical control. Key here is the idea that the social world can be reduced, represented and modelled in an abstract form. In other words, it is presumed that all of the key features of any social context can be represented, ordered and rendered calculable – what Wajcman (2019) describes as an 'engineering' mindset. From this perspective, a social system (such as a classroom) can be unproblematically modelled as a set of variables than can be manipulated in order to achieve optimal efficiency.

In this sense, educational AI applications are dependent on the input of data relating to education phenomena. This might take the form of data generated from students' uses of devices and software, data collected in classrooms through sensors and/or pre-existing contextual data generated offline (such as assessment results, demographic details, and so on). In this sense, most AI technologies currently being used in schools and universities are dependent on various 'proxy' variables – easily extractable data points that can substitute for direct measures of a particular aspect of education. For example, the time that a student spends watching an online instructional video might be used as a proxy for their levels of 'engagement' with the content of that video. If large sets of such data can be collated and analysed, then algorithmic models can be constructed to anticipate what might happen in similar future events. Key here is the capacity of these systems to adjust and 'learn' from mismatches. Indeed, in simple terms, machine learning involves a computer autonomously developing a mathematical model and refining it each time an error occurs.

All told, the delegation of key educational decisions and actions to these statistical logics certainly marks a radical shift in the provision, organisation and governance of education. While many people seem willing to presume that the AI technologies just described are capable of increased efficiency, precision, standardisation and consistency of outcomes when compared to traditional human-centred approaches, concerns are growing that this might not be the case. The following sections briefly outline four such areas of uncertainty and push-back.

Problems of representation and reduction

First, is the extent to which education can be adequately represented, modelled and manipulated in data form. A strong argument can be made that many of the basic aspects of teaching and learning *cannot* be captured reliably in data form. This is even more true for capturing and representing the complexities of a classroom or a student's social circumstances. While all data-driven processes are compromised by issues of representativeness, reductiveness, and explainability, these constraints are especially pertinent to uses of AI to model 'real world' issues that are embedded in social contexts such as classrooms. To paraphrase Murray Goulden (2018), even the most 'technologically smart' innovation is likely to be 'socially stupid' when deployed in a real-life context such as a school. As Meredith Broussard (2019, p. 61) argues: "Math works beautifully on well-defined problems in well-defined situations with well-defined parameters. School is the opposite of well-defined. School is one of the most gorgeously complex systems humankind has built."

Thus, however sophisticated AI becomes, any efforts at statistically modelling the contextual layers implicit in any educational episode or moment will continue to result in blunt computational approximations of the real-life complexities purportedly being captured. This phenomenon was illustrated in a Princeton University study which provided teams of statisticians, data scientists, AI and machine learning researchers with comprehensive data-sets covering over 4,000 families. Even with this wealth of data, stretching back over 15 years and boasting nearly 13,000 data points per child, all these expert teams failed to develop even moderately successful statistical models for children's life outcomes relating to school grades and competencies. As Karen Hao (2020, n.p.) reported at the time: "AI can't predict how a child's life will turn out even with a ton of data."

The social harms of AI

Second, then, are the social consequences of these statistical frailties – the gaps, omissions, and false errors that arise from the conflation of complex social phenomena into numbers. Recently, we have a trend to acknowledge such issues in loosely-defined terms of 'AI ethics' and 'AI safety.' However, there is now growing recognition of the real-life harms and violence that occur as a result of AI technologies being deployed in a social setting – what Shelby et al. (2022, p. 2) define as "adverse lived experiences resulting from a system's deployment and operation in the world." In terms of the ongoing educational application of AI, then, one set of concerns relates to what Shelby refers to as 'allocative harms' – i.e. how AI systems are proving prone to reaching decisions that result in the uneven – and sometimes unfair – distribution of information, resources and/or opportunities. This is reflected in various recent reports of 'algorithmic discrimination' in education – such as automated grading systems awarding higher grades for privileged students who fit the profile of those who historically have been awarded high grades, or voice recognition systems repeatedly making false judgements of cheating on language tests against students with non-native accents (NAO, 2019).

Also of concern are 'quality-of-service harms' – i.e. instances where AI systems fail systematically to perform consistently and to the same standards regardless of a person's background or circumstances. This has already come to the fore in instances where US schools have deployed facial recognition systems that regularly fail to recognise students of colour (Feathers, 2020), or systems developed to detect AI-generated writing that discriminate against non-native English speakers, whose work is more likely to be written formulaically and use common words in predictable ways (Sample, 2023). Of particular concern is the emergence of educational AI systems that rely on processes unsuited to disabled and neuro-diverse students – for example, eye-tracking technologies that take a steady gaze as a proxy for student engagement (Shew, 2020).

Alongside these concerns are what Shelby terms 'representational harms' – i.e. the ways in which AI systems rely on statistical categorisations of social characteristics and social phenomenon that often do not split into neatly bounded categories. This can lead to mis-representations of who students are, their backgrounds and behaviours in ways that can perpetuate unjust hierarchies and socially-constructed beliefs about social groups. Finally, are concerns over AI technologies adversely impacting on social relations within education settings – what Shelby terms 'interpersonal harms.' These include AI-driven 'student activity monitoring systems' now being marketed to allow teachers to surveil students' laptop uses at home, or school authorities using students' online activities as the basis of algorithmically-profiling students who might be deemed 'at risk' of course non-completion.

Running throughout all these examples is the underpinning concern that even the most 'benign' uses of AI in a school or classroom setting is likely to exacerbate and entrench pre-existing institutional forms of control. Schools *and* AI technologies are similarly built around processes of monitoring, categorising, standardising, synchronising and sorting. All told, while such exclusionary glitches might not be a deliberate design feature, AI technologies are proving prone to replicating and reinforcing oppressions that minoritized students are likely to regularly encounter during their educational careers. In this sense, one of the most important conversations we should now be having around the coming-together of education and AI relates to how AI is imbued with "a tendency to punch down: that is, the collateral damage that comes from its statistical fragility ends up hurting the less privileged" (McQuillan, 2022, p. 35).

Fitting education around the needs of AI

Third, is the concern that approaching students, teachers, classrooms and schools primarily in terms of what can be captured in data implies a number of fundamental rearrangements and reorganisations of education – what might be described as a recursive standardisation, homogenisation and narrowing of education. This relates to the question of what AI technologies expect of education (and, more pointedly, what AI technologies expect of the people involved in education). As Tennant and Stilgoe (2021, p. 846) remind us, "technological promises, if they succeed, end up making demands on the world." Here, then, we are already seeing an increased imperative to arrange education settings in 'machine readable' ways that will produce data that can be recognised and captured by AI technologies. This chimes with the phenomenon of what Langdon Winner (1978) termed reverse adaptation – i.e. rather than expecting technology to adapt to the social world, most people prove remarkably willing to adapt their social worlds to technologies.

In this respect, one immediate concern is that teachers and students are now beginning to be compelled to do *different* things because of AI technologies. For example, we are seeing reports of students now having to act in ways that are machine-readable – what might be described as 'adapting to the algorithm' (see Høvsgaard, 2019). This might involve a student having to write or speak in a manner that can be easily recognised by the computer, or to act in ways to produce data that an AI system can easily process. Similarly, teachers might have to develop 'parseable pedagogies' – i.e. easily codified ways of teaching that result in outcomes that can be inputted into the system. Perhaps less obvious, is the concern that teachers and students end up engaging in empty performative acts in order to trigger appropriate algorithmic responses. For example, this is already being seen in reports of call centre workers repeatedly saying 'sorry' during their interactions with callers in order to meet their automated 'empathy' metrics – regardless of whether saying 'sorry' is appropriate or not (Christl, 2023).

AI as environmental burden

Finally, is the underpinning concern that the data-intensive and device-intensive forms of AI currently being taken up in education incur unsustainable ecological and environmental costs. For example, MIT Technology Review reported in 2019 that the carbon emissions associated with training one AI model had been estimated to exceed 626,000 pounds of carbon dioxide (equivalent emissions to driving 62 petrol-powered passenger vehicles for twelve months). Similarly, conducting a 'conversation' with Chat GPT of between 20 to 50 prompts is estimated to consume 500 ml of water (Li et al., 2023). Thus in terms of natural resource consumption and energy drain alone, as Thompson et al. (2021, n.p.) understatedly puts it, “the cost of [AI] improvement is becoming unsustainable.”

It is therefore beginning to be argued that educators need to temper any enthusiasms for the increased take-up of AI with the growing environmental and ecological harms associated with the production, consumption and disposal of digital technologies. In this sense, AI should not be seen as an immaterial, other-worldly technology – somehow weightless, ephemeral and wholly ‘in the cloud.’ In reality, AI is reliant on a chain of extractive processes that are resource-intensive and with deleterious planetary consequences. In short, the growing use of AI technologies in education comes at considerable environmental cost – implicated in the depletion of scarce minerals and metals required to manufacture digital technologies, massive amounts of energy and water required to support data processing and storage, *and* fast-accumulating levels of toxic waste and pollution arising from the disposal of digital technology (see Brevini, 2021).

Given all the above, any enthusiasms for the increased use of AI in education must address the growing concerns among ecologically-concerned commentators that it might not be desirable (and perhaps even impossible) to justify the development and use of AI technologies in the medium to long-term. On the one hand, this necessitates proponents of educational AI to explore how the continued use of AI in schools and universities might be aligned with ‘green-tech’ principles and perhaps make a positive contribution to forms of eco-growth. In this sense, there is certainly a pressing need to explore the extent to which educational AI might be oriented toward emerging developments in areas such as ‘carbon-responsive computing’ and ‘green’ forms of machine learning. This implies, for example, developing different forms of AI built around small datasets and refined processing techniques, and moving beyond ‘brute force’ computational approaches (Nafus et al., 2021).

On the other hand, however, we also need to give serious consideration to the idea that AI is ultimately an irredeemable addition to education, and needs to be rejected outright. Strong arguments are being made that the environmental and ecological harms arising from AI use *cannot* be offset by efforts to instigate ‘greener’ forms of carbon-neutral digital technology and ‘cleaner’ forms of renewable energy. As such, educationalists would do well to be open to the possibility that most – if not all – forms of AI technology “are intrinsically incompatible with a habitable earth” (Crary, 2022, n.p.). If this is the case, then it makes little sense to continue to push for the reframing of education in an era of climate crisis and environmental breakdown around these technologies. From this perspective, then, AI is nothing more than a dangerous distraction from much more pressing and threatening planetary issues.

AI and education – some ways forward

The main challenge now facing educators is to avoid getting mired in the considerable hype that will continue to surround AI in the months (and perhaps years) ahead. At the moment, the emergence of AI is prompting a familiar response that has regularly accompanied educational discussions of previous 'new' technologies over the past 40 years or so. In short, this has involved the sudden appearance of 'common-sense' arguments that: (i) the increased incursion of AI tools into classrooms is inevitable; (ii) that teachers quickly need to upskill (become 'AI literate') in order to make best use of these technologies, and (iii) that we need to seriously rethink how traditional educational forms and practices might need to change and adapt to the affordances of AI. In short, educators are positioned as having little control over the nature, pace and direction of this technological change. Existing forms of schools and schooling are positioned as providing impediments and barriers to the smooth use of the technology, and teachers are positioned in deficit. The underpinning logic here is simple – education needs to change quickly in order to 'catch up' with this seismic technological change that has the position to radically transform all aspects of what it means to educate and be educated.

In contrast, this paper has attempted to recast the imperatives of AI and education in a substantially different light. Above all, it has stressed the need for educators to take control and work to proactively shape the agendas that are continuing to form around what AI might mean for schools, and how we might see AI playing a constructive role (if at all) in the future classroom. This means getting actively involved in the conversations and debates that are currently swirling around the topic of AI and education, led largely by voices with little or no direct expertise in schooling and education. Education experts need more confidence in speaking up and leading these debates. One key area of discussion are questions over exactly what 'added-value' AI technology can be said to offer. Here, educators are in a key position to push back against vague claims of AI radically relieving teachers' workloads or acting as a 'one-to-one tutor for the world.' More immediately, perhaps, educators are also in a key position to demonstrate the limited outcomes that result from limited educational AI technologies. At the same time, it is also important for educators to speak up about the *other* forms of AI technology that we might collectively believe as capable of being of genuine education benefit.

In all these ways, then, education communities should be looking to play a key role in providing a collective counter-balance to the hyperbole that has engulfed recent debates around AI and education. This requires challenging IT industry-led visions of how education might be best reorganised and/or dissembled, as well as the associated surrender of public education interests to the economic and political interests that continue to push AI into education. This also requires pointing to the disadvantages and harms that are now being noted as key aspects of education become increasingly reliant on AI technologies – from concerns over AI-led administrative violence and algorithmic discrimination through to the diminished quality of educational provision and support. Above all, this requires moving away from portraying AI in education as a technical object, and instead framing AI as a system that is bound up with the messy realities of education systems, economic systems, political systems and other social systems. Finally, amidst these clarifications, counter-arguments and critique there is also a need for educators to talk more about possible alternate forms of AI that might better fit education – i.e. ways in which AI might be genuinely useful in being a part of a response to educational needs. As Nick Couldry reasons, making criticisms of the recent AI turn does not necessarily denote a wholesale rejection of AI technology altogether:

> We are not objecting to the use of AI tools to solve specific problems within clear parameters that are set and monitored by actual social communities. We are objecting to the rhetoric and expansionist practice of offering AI as the solution for everything. (Couldry, 2023, n.p.)

In this spirit then, it falls to the education community to now begin to work out how to shape a new wave of discussions around AI in education that are framed in more emancipatory, fair, or perhaps simply kinder ways than the brut(ish) forms of corporate algorithmic control currently on offer. Indeed, there are some burgeoning examples of how this might be done. On the one hand, we are beginning to see some radical calls for feminist, queer, decolonialised and indigenous reimaging of what AI might be (e.g. Adams, 2021; Klipphahn-Karge et al., 2023; Munn, 2023; Toupin, 2023). On the other hand, a few mainstream public education agencies and organisations are also beginning to make a decent start in calling for new forms of AI that emphasize human elements of learning and teaching, that are sympathetic to education contexts, that involve educators in their conception, development and implementation, and are based around values of trust, care and that align with shared education visions. For example, as the US Office of Educational Technology (2023, p. 10) recently contended:

> Use of AI systems and tools must be safe and effective for students. They must include algorithmic discrimination protections, protect data privacy, provide notice and explanation, and provide a recourse to humans when problems arise. The people most affected by the use of AI in education must be part of the development of the AI model, system, or tool, even if this slows the pace of adoption.

Conclusions

All told, this paper has begun to outline the case for slowing down, scaling back and recalibrating current discussions around AI and education. While this might not feel like an easy task, the urgency of current conversations around AI and education is clearly unproductive in the long run. It makes good sense for educators to try to disconnect themselves from the apparent imperatives of AI-driven educational 'transformation,' and instead work to slow down discussions around AI and education, and introduce an element of reflection and nuance. Given the technical *and* social complexity of AI, it behoves us to try to develop forms of public debate that engage with these complexities rather than descend to overly-simplistic caricatures and fears. Given the clear inequalities and injustices already arising from AI technologies it also behoves us to pay closer attention to "the oppressive use of AI technology against vulnerable groups in society" (Birhane & Van Dijk, 2020, n.p.). Moreover, all of the concerns raised in this paper all point to key questions of power – i.e. who gets to decide what AI tools are implemented in education will inevitably wield considerable influence over what goes on in that education setting. As Dan McQuillan (2023, n.p.) argues:

> From this perspective, AI is not a way of representing the world but an intervention that helps to produce the world that it claims to represent. Setting it up in one way or another changes what becomes naturalised and what becomes problematised. Who gets to set up the AI becomes a crucial question of power.

Seen in this light, then, it seems crucial that educators and the wider education community become more involved in debates and decision-making around who gets to 'set up' AI and education. The future of AI and education is *not* a foregone conclusion that we simply need to adapt to. Instead, the incursion of AI into education is definitely something that can be resisted and reimagined.

Acknowledgements

This paper arises from research supported by funding from the Australian Research Council (DP240100111).

Author biography

Neil Selwyn has been researching and writing about digital education since the mid-1990s. He is currently a professor at Monash University, Melbourne. Recent books include: *Should Robots Replace Teachers? AI and the Future of education* (Polity 2019), *Critical Data Literacies* (MIT Press 2023, with Luci Pangrazio), and the third edition of *Education and Technology: Key Issues and Debates* (Bloomsbury 2021).

References

Adams, R. (2021). Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews*, *46*(1–2), 176–197.

Bender, E., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the dangers of stochastic parrots. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623).

Birhane, A., & van Dijk, J. (2020, February). Robot rights? Let's talk about human welfare instead. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 207–213). Association for Computing Machinery.

Brevini, B. (2021). *Is AI good for the planet?* Polity.

Broussard, M. (2021, 22 April). [Tweet]. Twitter. https://twitter.com/merbroussard/status/1384934004030418945

Caird, S., Lane, A., Swithenby, E., Roy, R., & Potter, S. (2015). Design of higher education teaching models and carbon impacts. *International Journal of Sustainability in Higher Education*.

Christl, W. (2023). *Surveillance and algorithmic control in the call centre*. CrackedLabs. https://crackedlabs.org/en/data-work/publications/callcenter/

Couldry, N. (2023, 11 April). *AI as colonial knowledge production*. University World News. https://www.universityworldnews.com/post.php?story=2023041014520289

Crary, J. (2022). *Scorched earth*. Verso.

Feathers, T. (2020, 2 December). *Facial recognition company lied to school district about its racist tech*. Vice Motherboard. https://www.vice.com/en/article/qjpkmx/fac-recognition-company-lied-to-school-district-about-its-racist-tech
Giannini, S. (2023). *Generative AI and the future of education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385877
Goulden, M. (2018). [Tweet]. Twitter. https://twitter.com/murraygoulden/status/1038338924270297094
Hao, K. (2020, 2 April). AI can't predict how a child's life will turn out even with a ton of data. *MIT Technology Review*.
Høvsgaard, L. (2019). Adapting to the test. *Discourse: Studies in the Cultural Politics of Education*, *40*(1), 78–92.
Klipphahn-Karge, M., Koster, A., & Bruss, S. (Eds.). (2023). *Queer reflections on AI*. Routledge.
Li, P., Yang, J., Islam, M., & Ren, S. (2023). *Making AI less 'thirsty': uncovering and addressing the secret water footprint of AI models*. arXiv. https://doi.org/10.48550/arXiv.2304.03271
Mason, H. (2018, 3 July). [Tweet]. Twitter. https://twitter.com/hmason/status/1014180606496968704.
McQuillan, D. (2022). *Resisting AI*. Policy Press.
McQuillan, D. (2023, 6 June). Predicted benefits, proven harms. *The Sociological Review: Magazine*. https://thesociologicalreview.org/magazine/june-2023/artificial-intelligence/predicted-benefits-proven-harms
Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.
Munn, L. (2023). The five tests: Designing and evaluating AI according to Indigenous Māori principles. *AI & Society*. https://doi.org/10.1007/s00146-023-01636-x
Nafus, D., Schooler, E., & Burch, K.. (2021). Carbon-responsive computing. *Energies*, *14*(21), 6917.
NAO. (2019) *Investigation into the response to cheating in English language tests*. National Audit Office. https://www.nao.org.uk/wp-content/uploads/2019/05/Investigation-into-the-response-to-cheating-in-English-language-tests.pdf
Pretz, K. (2021, 31 March). *Stop calling everything AI, machine-learning pioneer says*. IEEE Spectrum. https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says
Reuters. (2023, 18 May). *AI threatens humanity's future, 61% of Americans say: Reuters/Ipsos poll*. www.reuters.com/technology/ai-threatens-humanitys-future-61-americans-say-reutersipsos-2023-05-17/
Salganik, M., Lundberg, I., Kindel, A., Ahearn, C., Al-Ghoneim, K., Almaatouq, A., & Altschul, D. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences*. www.pnas.org/content/117/15/8398
Sample, I. (2023, 10 July). Programs to detect AI discriminate against non-native English speakers, shows study. *Guardian*. www.theguardian.com/technology/2023/jul/10/programs-to-detect-ai-discriminate-against-non-native-english-speakers-shows-study
Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla, N., Gallegos, J., Smart, A., Garcia, E., & Virk, G. (2022). *Sociotechnical harms: Scoping a taxonomy for harm reduction*. arXiv. https://doi.org/10.48550/arXiv.2210.05791
Shew, A. (2020). Ableism, technoableism, and future AI. *IEEE Technology and Society Magazine*, *39*(1), 40–85.
Siddarth, D., Acemoglu, D., Allen, D., Crawford K., Evans, J., Jordan, M., & Weyl, G. (2021, 1 December). *How AI fails us*. https://ethics.harvard.edu/files/center-for-ethics/files/howai_fails_us_2.pdf?m=1638369605
Tennant, C., & Stilgoe, J. (2021). The attachments of 'autonomous' vehicles. *Social Studies of Science*, *51*(6), 846–870.
Thompson, N. Greenewald, K., Lee, K., & Manso, G. (2021, 24 September). Deep learning's diminishing returns. *IEEE Spectrum*. https://spectrum.ieee.org/deep-learning-computational-cost
Toupin, S. (2023). Shaping feminist artificial intelligence. *New Media & Society*, 14614448221150776.
Tucker, E. (2022, 17 March). *Artifice and intelligence*. Tech Policy Press. https://techpolicy.press/artifice-and-intelligence/
U.S. Office of Educational Technology. (2023). *Artificial intelligence and future of teaching and learning*. Washington DC, U.S. Department of Education.
Versteijlen, M., Salgado, F., Groesbeek, M., & Counotte, A. (2017). Pros and cons of online education as a measure to reduce carbon emissions in higher education in the Netherlands. *Current Opinion in Environmental Sustainability*, *28*, 80–89.
Wajcman, J. (2019). How Silicon Valley sets time. *New Media & Society*, *21*(6), 1272–1289.
Winner, L. (1978). *Autonomous technology*. MIT Press.