Out of the laboratory and into the classroom: the future of artificial intelligence in education

Daniel Schiff[1]

Abstract

Like previous educational technologies, artificial intelligence in education (AIEd) threatens to disrupt the status quo, with proponents highlighting the potential for efficiency and democratization, and skeptics warning of industrialization and alienation. However, unlike frequently discussed applications of AI in autonomous vehicles, military and cybersecurity concerns, and healthcare, AI's impacts on education policy and practice have not yet captured the public's attention. This paper, therefore, evaluates the status of AIEd, with special attention to intelligent tutoring systems and anthropomorphized artificial educational agents. I discuss AIEd's purported capacities, including the abilities to simulate teachers, provide robust student differentiation, and even foster socio-emotional engagement. Next, to situate developmental pathways for AIEd going forward, I contrast sociotechnical possibilities and risks through two idealized futures. Finally, I consider a recent proposal to use peer review as a gatekeeping strategy to prevent harmful research. This proposal serves as a jumping off point for recommendations to AIEd stakeholders towards improving their engagement with socially responsible research and implementation of AI in educational systems.

Keywords Artificial intelligence · Education technology · Social implications of technology · Educational agents · Responsible research and innovation

1 Introduction

AI has received substantial attention in the twenty-first century by powerful business interests, researchers, governments, and the public. Along with efforts to implement AI in organizational settings and economize on its value, the ethical, legal, and social implications have come into the forefront of discussion. Much of the focus of public, private, and academic discourse surrounding AI has addressed labor shifts and displacement, military and cybersecurity concerns, ethical issues such as bias, and economic benefits.

However, less public attention has been given to other areas of AI in social policy, especially AI in education (AIEd). This paper attempts to address that gap by characterizing and then exploring the implications of AIEd as well as providing recommendations to AI researchers and developers towards a responsible research agenda. Moreover,

✉ Daniel Schiff
schiff@gatech.edu

1 School of Public Policy, Georgia Institute of Technology, Atlanta, GA, USA

this analysis brings a forward-looking policy lens that can help situate the work of AI developers and researchers in the larger picture.

In Sect. 2, I contextualize the role of AIEd historically and discuss the excitement, fears, and often faulty predictions that have accompanied new forms of education technology. Digital distance education serves as a historical example from which lessons can be learned regarding AIE's capacities and implications. Section 3 reviews three reasons why a belief in the technological inevitability of distance education fared poorly against social realities: a failure to consider implementation of education technology, a mistaken assumption of technological linearity, and the inability of distance education to reproduce teacher–student interaction. I highlight the role of teachers in quality education to explain why massive open online education has been unable to achieve its vision of high-quality, universal education.

Section 4 offers an analysis of AIEd and argues that AIEd is a candidate to overcome this barrier through the simulation of teachers, including complex social behaviors previously thought non-reproducible. I provide an introduction to AIEd and its most prominent tool, intelligent tutoring

systems and then pay special attention to anthropomorphized educational agents, which most commonly simulate teachers. Section 5 discusses possible futures as AIEd begins to penetrate educational systems. I explore, both optimistically and critically, implications for pedagogy, curricula, international development and equity, flexibility and ownership of educational systems, the role and possible automation of teachers, and the use of nudging/manipulation by AIEd.

In light of the risks posed, Sect. 6 reviews recommendations to promote responsible research in AIEd and AI generally. I consider a recent proposal in computing to use the peer review process to prevent harmful research. I suggest revisions and expansions of this proposal towards a responsible approach to AIEd research and implementation. Section 7 concludes.

2 Education technology and sensational history

With new technology come utopian reformers and dystopian cynics. This is true in the realm of education, where the stakes are believed to be nothing short of the future of society. Here, reformers and cynics posit, respectively, that young learners will be radically empowered to reshape the world or will be placed into the service of unfeeling, technocratic forces (Feenberg and Hamilton 2012). Similar competing progressive and dire predictions have accompanied the advent of media forms that affect education, including radio, television, the Internet, and even writing. For instance, Plato, speaking through Socrates in Phaedrus, feared that writing would lead to forgetfulness and the appearance of wisdom without the reality. Like a painting, writing was static and distant and would "go on telling you just the same thing forever." Ultimately, Plato worried that writing was a poor replacement for 'true education' which instead required a teacher capable of instructing and facilitating, of defining and defending ideas, and knowing "when to speak and when to be silent" (Jowett 1901). Not dissimilar critiques around educational implications of technology have accompanied radio (Siepmann 1941), television (Emery 1972; Schwartz 1974), the Internet (Iseke-Barnes 1996), and most recently, virtual reality (Evans 2003), and social media (Sunstein 2018).

In the realm of education technology, a stereotypical progressive reformer is a technologist, entrepreneur, or philanthropist who believes that the status quo is failing us and so invests time or money to promote a new vision.[1] The reformer's counterpart is then a ground-level practitioner, philosopher, or sociologist who has witnessed or studied how so-called progressive ideas are, in fact, counterproductive.[2] Yet, advocates and skeptics of educational technology often shared a key proposition. Regardless of their support or opposition, many agreed that the pace of change would be rapid, the scale of disruption substantial, and the transformation inevitable (Agre 1999; Hamilton and Feenberg 2005). Ironically, this lone area of agreement often missed the mark, with its adherents failing to anticipate social and technical constraints and ultimately make faulty predictions. The mode of writing has been dynamic and productive, contrary to Plato's expectations, and video has been successfully integrated into educational settings. More recently, while brick-and-mortar universities have not yet been supplanted by online ones, the COVID-19 virus has accelerated such a transformation (LeBlanc 2020). Simply, predicting the limits of technology—and its eventual social implications—is difficult.

While prediction is imperfect, this familiar cycle suggests that history is a useful lens with which to understand AIEd's possible future. One such antecedent for AIEd is distance education or (digital/online) learning, which is not typically AI-based. Emerging into prominence along with the early Internet in the 1990s, distance education promised to automate, streamline, and universalize education (Feenberg 2002; Li 2002). AIEd is, in some ways, a continuation of this vision of educational technology: both are software-based, challenge the role of teachers, and potentially upset traditional understandings of classroom size, pedagogy, and more. Most importantly, both distance education and AIEd advocates imagine technology's role in providing the highest quality instruction for learners around the world. However, distance education typically replicates existing classroom structures such as textbooks and lectures, often putting them into a video format, and lacks 'intelligent' features that characterize AIEd. In contrast, the intelligent tutoring systems that are central to AIEd seek to reproduce one-on-one human instruction, with personalized and real-time feedback (Self 1998). Section 4 reviews AIEd's characteristics in more detail.

By the late 1990s, the majority of universities had an Internet-based educational option in place. Yet with the early potentials of distance education recognized, criticism followed the more optimistic aspirations. David Noble argued vehemently that distance education was being gleefully serviced towards commodification of education, deskilling and displacing professionals, and harming students, all in the name of economic efficiency (Noble 1998a, b). He feared

---

[1] See for example Nicholas Negroponte, founder of MIT's media Lab, whose ambitious One Laptop per Child program did not achieve its goal of reaching hundreds of millions of students and transforming global education (Warschauer and Ames 2010).

[2] See the idea of "threat rigidity" when teachers face a constant stream of externally imposed reforms (Olsen and Sexton 2009).

that educational technology would turn schools into "digital diploma mills." These criticisms echoed those of industrial and postmodern critics like Andrew Ure and Jean-François Lyotard, who saw the automating forces of their times as similarly reductive of human capacities (Lyotard 1984; Ure 1836).

A particularly salient case in distance education was the cascade of excitement about the universalizing potential of Massive Open Online Courses (MOOCs) in the 2010s, an online educational technology that is not typically AI-based (Bond et al. 2019). Unlike AIEd, MOOCs have little capacity for real-time or differentiated instruction, and instead aim at mass delivery of lecture content (Scagnoli et al. 2019). Even so, MOOCs were imagined to play a key role in 'flattening' education (Friedman 2005), opening up the best courses and instructors to anyone with an Internet connection. This access would rapidly shift educational infrastructure towards skill-based online education and away from long-term, credential-based education in physical spaces. Universities were among those most concerned about losing their clientele, and those most eager to adapt MOOC-style education into their approaches to evolve with the new market. While these questions have yet to play out in full, the perception of inevitability was tempered rapidly as it became apparent that dropout rates were staggeringly high and that participants tended to be well-educated and from wealthier countries (Christensen et al. 2013; Rivard 2013). MOOCs seemed to reach students who had already been reached, and employers and students were not ready to abandon the traditional degree and educational experience for a series of online videos. The predictions of MOOC's inevitability went unrealized.

Not dissimilarly, some artificial intelligence commentators have blundered into this context, seemingly forgetting AI's own history of failed predictions including "AI winters" in the 1970s, 80s, and 90s wherein slowdowns of technological progress and funding impeded the sense of inevitability (Hendler 2008).[3] Famously, Hubert Dreyfus' (1979) discussion of the limits of AI had to be walked back two decades later (Dreyfus 1992) after his assumptions about AI's core focus on symbolic methods and subsequent limitations were undermined. Nevertheless, general AI-related predictions today continue to be just as, if not more extreme. Surveys of experts demonstrated substantial consensus that AI will achieve 'superintelligence' (surpassing humans) by the end of the century. Leading scholars, scientists, and technologists predict variously a golden age wherein all major civilizational challenges have been solved and humans are entirely free to pursue creative interests, or alternatively, the literal end of the humanity (Armstrong et al. 2014; Müller and Bostrom 2016).

Within the realm of AI and education, predictions are less sweeping but no less bold. It is widely believed by AIEd advocates that "AI will be a game changer in education," wherein global classrooms will enable "individuals and groups to learn remarkably better than if they are taught by a single human teacher" (Woolf et al. 2013). The Department of Education's STEM 2026 vision states that intelligent tutoring systems, the primary AI-based educational tool currently, "may play a key role in the future of education" (Tanenbaum 2016). Founder of Squirrel AI Derek Li argues that only AIEd's adaptive systems can produce "miracles" and that entrusting teachers may risk "damaging geniuses" (Hao 2019).

The above examples do not entail that prediction is a foolish enterprise. Though it is difficult to place bounds on AI's capacities or social implications, technological forecasting is a rich and valuable toolkit with qualitative and quantitative aspects. It can help us to envision possible (extreme) scenarios that may serve as idealistic goals or cautionary tales and can help us chart a more considered course. For example, not all commentators of distance education took polarizing positions; some offered moderating principles and strategies to address possible challenges, noting the ways in which developmental trends are positive or negative and work that remains (Beldarrain 2006; Chickering and Ehrmann 1996). The value of future-oriented thinking is also recognized in the AIEd community, including in a special edition of the International Journal of Artificial Intelligence in Education (Lane et al. 2016).[4] It is in this moderating spirit that I hope to contribute to the conversation on AIEd.

3 Technological inevitability and evitability

What then is the gap between the ostensible technological inevitability (or fatalism) and the much clunkier reality in the historical case of education technology? I highlight three candidates which are neither mutually exclusive nor exhaustive:

---

[3] Marvin Minsky predicted that before 1980, "we will have a machine with the general intelligence of an average human being." Herbert Simon offered that by 1985, "machines will be capable of doing any work a man can do."

[4] Woolf et al. also follow in this more sober tradition for predictions and challenges facing AIEd (Woolf et al. 2013) as does the US Department of Education working with American Institutes for Research. For example: "As with all education technology tools, however, the benefits of intelligent tutoring systems will be realized only if they are equitably accessed and effectively integrated into and coordinated with complementary learning activities such as group discussions and project-based learning" (Tanenbaum 2016).

- The failure to consider implementation;
- The assumption of technological linearity; and
- The need for human interaction, especially teacher–student interaction

3.1 Implementation barriers

I will not overly focus on the first in this paper, as the topic has been discussed thoroughly.[5] I observe, however, that this remains a perennial problem. Both national and international education technology initiatives have failed to sufficiently consider cost/funding, teacher training and ongoing professional development, timely rollout of complementary curricular and pedagogical supports and strategies, necessary school and system-level structural changes, and student and family context among other issues (Argueta et al. 2011; Warschauer and Ames 2010; Nye 2015).

Design-based implementation research (DBIR) is one framework that has emerged in response to these failures, focusing on collaborative multi-stakeholder design and development of sustainable capacity to address previous shortcomings in implementation (Fishman et al. 2013). Practitioners in DBIR as well as user-centered design, human–computer interaction, and other fields are increasingly expounding on contextual, regional, and demographic factors necessary for successful design and implementation. However, the fact that major initiatives like purchases of millions of tablets/laptops still fail to take account of the lessons of education technology implementation provides a cautionary note.

3.2 Technological linearity

As for the second point, the nonlinearity of technological development was pointedly observed by Pinch and Bijker in their seminal work on the social construction of technology. Technology could not be said to be linearly entailed by the nature of its underlying science. Instead, prospective technologies are open to a kind of interpretive flexibility, wherein the contextualized relevance to certain social groups and a bit of historical accident leads to closure and a stable technological form that we later accept (mistakenly) as obvious and inevitable (Pinch and Bijker 1984). Feenberg has carefully demonstrated the ways in which the relationship between technical capacities of a new educational technology and its applications is neither linear nor deterministic (Feenberg 2002). Writing has long since emerged as a powerful educational tool capable of displaying complex traits that would surprise Plato, and the Internet has developed

social-interactional mechanisms that early theorists could hardly speculate. Why should we expect AIEd to be any different?

Even this criticism of retrospective closure and stability in light of sociotechnical mediation is incomplete. Following Ihde, Rosenberger argues that education technology is additionally subject to prospective reinterpretation/reinstrumentalization and thus multi-stability (Ihde 1986). Education technologists who believe their products will be used in a certain way are bound to be disappointed or at least surprised as teachers and students modify and innovate through a process of strategic adaptation.[6] A tool imagined for one purpose becomes repurposed for another; closure may never actually obtain. However, these potential usages are still circumscribed by the technical nature of the tools and the social context, leading to at least temporary stabilities, hence multi-stability (Rosenberger 2017). Education technology, then, is deterministically realized neither in the past nor in the future.

A final point is relevant to the assumption of technological linearity. This is that technologies interact not only with a local social context (e.g., a classroom) or with consumer behavior, but also more pointedly with our politics and our values. AIEd, like other education technologies, is inherently political and value-laden. It implicitly assumes much about the structural norms of education—the nature of curricula, pedagogy, classrooms, educational trajectories and national goals, the local culture, and so on. Subsequently, then, educational technology reinforces these norms, reifies structures, and even preserves authorities (Winner 1980). These structural assumptions as embedded in technology are critically important to how the technology is eventually mediated. Unpacking and challenging these assumptions in technology design can reveal nonlinear possibilities, a point to which we will return.

3.3 Teacher–student interaction

Third, the nature of human interaction becomes especially salient for educational technologies that radically change the teacher–student relationship, e.g., distance education, and as we will see, AIEd. Distance education explicitly re-envisions the role of teachers and the teacher–student ratio, and less intentionally, the phenomenology of student experience. The critique is this: while distance education advocates have realized the potential of near-zero marginal cost scaling and access to 'high-quality' education, they failed to sufficiently appreciate how much educational quality depends on human

---

[5] For a few excellent overviews, see Ely (1999); Leggett and Persichitte (1998), Romiszowski (2004), Nye (2015).

[6] A calculator can be used for mathematics, storing information, playing games, sharing silly messages, and even making music, but it is not particularly helpful for viewing microscopic organisms.

interaction and teachers. Content was imagined above form, and educational 'quality' as measured by the best instructors and best schools was abstracted away from the original context which enabled that quality.

This account helps to explain why distance education has failed to universalize equitably or without substantial dropout rates. As educational philosophers, developers, and evaluators have noted, simulated, impersonal digital learning environments just do not seem to be able to replace the classroom (Mackness et al. 2010). Will AIEd be different, if it can simulate student–teacher interaction or teachers themselves? Or are there other abstractions involved in the design of AIEd which leave more to be desired?

Predictive failures, then, are explained in part by an overestimation of the quality of the novel educational technology compared with the norm it purports to supplant. This overestimation can result from a formalized abstraction of the previously successful setting (the classroom) which leaves out critical elements of quality (teachers). The gap between hypothesized and actual educational effectiveness explains the gap in implementation success, as education stakeholders are forced to revert back to previous methods when facing suboptimal results, or are simply savvy enough to be skeptical in the first place (Norris et al. 2012).

Explaining this quality gap further, Dreyfus focuses specifically on the nature of teaching and learning. He argues that basic competence can be achieved with online education, but to achieve true proficiency, expertise, and mastery, students must be exposed to risk and failure in a way that is only achievable in person. Students must be able to propose and defend ideas, make fools of themselves, and face social approbation and reckoning from the teacher. An online 'like' or comment does not engender this kind of emotional investment and thus learning. Moreover, Dreyfus suggests that the highest levels of expertise and mastery require apprenticeship, not only to learn complex skills like perspective- taking, but also to be socialized into professional norms and cultures. Students cannot be encouraged, motivated, challenged, or ultimately acclimated without the teacher, and the Internet provides a weak approximation (Dreyfus 2002, 2008). It may be unsurprising then that MOOCs often have limited student engagement (Cohen et al. 2019).

There are, of course, various pedagogical and design strategies being tested to promote motivation and even social interaction in online education. Notifications and quizzes during lectures, reminder emails, and suggestions on planning the course workload are all efforts towards increasing attention, motivation, and successful course completion. These efforts have been able to draw on big data, tracking individual student clicks and attention to videos, to optimize MOOC design (Simonite 2013). Small discussion forums, peer review, and even in-person meetups have been attempted to improve elements of social interaction (Goel

et al. 2016). To Dreyfus, these laudable steps are nevertheless tweaks on a fundamentally flawed model. Without the teacher, digital education simply does not have the tour de force to achieve its ambitious aims.

More optimistically, Feenberg asserts that online education can achieve the quality of in-person education, but only provided it incorporates teachers more robustly. What online education needs to match the quality of traditional education is facilitation, moderation, synthesis, and leadership: "raising topics, recognizing participants' contributions, and summarizing discussion at key points" (Feenberg 2008). Similar critiques are manifested in the Distributed Open Collaborative Course (DOCC) movement, a MOOC alternative based on feminist pedagogy. DOCC's pedagogy emphasizes distributed rather than central expertise, collaboration and community rather than isolation, and diversity in participants and perspectives. A clear expression of this philosophy is the idea of digital office hours, a personalized interaction that is incompatible with typical MOOC structures (Koh 2017).

3.4 Differentiation

It should also be pointed out that these conceptualizations of the critical role of teachers offered above imply an additional need: individualization or differentiation. To provide effective guidance, motivation, facilitation, and emotional investment, a teacher must know their students. They must understand the students' social, emotional, and educational past and present. Only with this knowledge and skillful practice can educators inspire students with unique backgrounds, learning styles, and interests. This vision is clearly not embedded in MOOC-style distance education, as asynchronous posting, passive reception of video content, and even the use of pseudonyms (usernames) exacerbates the issue of impossible student–teacher ratios.

Importantly, while in-person meetups and digital office hours are plausible at small scale and do start to address the concerns of Dreyfus and Feenberg, these teacher-centered approaches to digital education contradict the raison d'être of massive online education—large scale and reduction of human capital costs. In the face of this tension of quality versus scale, scale dominates. The result of this model is that scale-oriented MOOC education model often leaves students untethered in the face of "the lack of structure, support and moderation" (Mackness et al. 2010).

In contrast to MOOCs and other distance education tools which do not traditionally involve AI, AIEd technologies focus heavily on the perspective of learners (Lane et al. 2016), on pedagogy, and on differentiated, one-to-one instruction. AIEd tools are becoming capable of social–emotional engagement and can even be integrated into scale-focused distance education technologies. Moreover, while AIEd has been used in applied settings for decades, its

prospects (Nye 2016) and usage (Hao 2019) are now greater than ever before, thanks to research advances and especially increased access to personal computers and the internet. In short, AIEd is leaving the laboratory and entering the classroom. In light of this, the next section reviews the argument that AIEd is a candidate to bridge the quality gaps surfaced by the previous round of education technology.

4 AI and educational agents

Distilled to a single term, the argument for the quality gap presented above is teachers. Without teachers (or teaching), we have good reasons to think that the quality gap between traditional and digital education will not be so easily closed and hence implementation will fall utterly short of the hopes of educational technology's strongest proponents. This is why AIEd is so enticing: it offers to close the quality gap by tackling the most insoluble barrier—teaching.[7] I argue that the central ambition of modern AIEd is to simulate teachers.

A most useful definition of artificial intelligence[8] in the case of education is "the study of how to make computers do things at which, at the moment, people are better" (Rich and Knight 1991). The 'thing' referred to in this paper is teaching, and those people that AI hopes to emulate and surpass are most often teachers (otherwise tutors, mentors, or even educational administrators). Below, I provide a brief overview of AIEd technologies (both extant and emerging), focusing on the predominant class of tools known as intelligent tutoring systems.[9] I characterize their basic design, functional capacities, and expected utility, demonstrating how they increasingly encroach on the traditional niche of human teachers. I then highlight a subset of intelligent tutoring systems known as educational agents, effectively anthropomorphized software-based characters. Educational agents represent the attempt to tackle some of the hardest challenges in simulated teaching, namely perceiving and portraying emotion in a way that engenders authentic student response.

AI is used in a wide variety of educational applications, including: adaptive assessments (Heffernan and Heffernan 2014), educational and social robots (Timms 2016; Wang and Johnson 2016), lifelong intelligent mentors (Woolf et al.

2013), and more (McArthur et al. 2005; Czopek and Pietrzak 2016). AI is also being applied to traditional educational tools, with intelligent features being incorporated into 'adaptive' learning management systems (Nenkov et al. 2016), 'adaptive' MOOCs and distance education tools (Sonwalkar 2013; Xiaohong 2017; Rosé and Ferschke 2016), and predictive analytic systems in higher education, for example, aimed at improving graduation rates and career outcomes (Dimeo 2017; Tastimur et al. 2016). For a use case to qualify as an example of AIEd, it must involve some 'intelligent' component. Finally, AI is important to related educational fields such as educational data mining (Romero and Ventura 2010) and computer-supported collaborative learning (Dillenbourg et al. 2009).

However, the most prominent form of AIEd is intelligent tutoring systems (ITS) (Niculescu 2016; Roll and Wylie 2016). ITS are defined as "computer-based learning systems which attempt to adapt to the needs of learners" with the "scientific goal to make computationally precise and explicit forms of educational, psychological and social knowledge which are often left implicit" (Self 1998). ITS are, therefore, aimed at opening the black box of teaching and learning and simulating it computationally. While this paper focuses on ITS, ITS are often connected with other AIEd tools, such as educational robots and adaptive assessments, creating broader AIEd ecosystems (Nye 2016). An ITS is typically constituted by some variation of the four following software components (Samuelis 2007):

- An expert domain, which holds the relevant subject knowledge, e.g., contains all the information related to high school algebra that will be employed by the ITS.
- A learner domain, which keeps track of how the student has engaged with the ITS, e.g., what questions the student has answered correctly and incorrectly and, via prediction, what the student (probabilistically) understands.
- A pedagogical domain, which structures strategies for providing targeted feedback and material to students, addressing mistakes and questions as would a teacher.
- An interface through which the student interacts with the software, potentially including various forms of multimedia or different representations of the digital learning environment. An educational agent, for example, would manifest as a quite different kind of interface, one in which the student interacts with an anthropomorphized character, even a robot.

The resemblance of this four-prong structure to our model of formal education is no coincidence. Indeed, the goal is to unpack—and, therefore, eventually simulate—what it is that happens within the black box of teaching and learning. Importantly, our interpretation of that black box is not a straightforward or unambiguous representation; how we

---

[7] More skeptically, when Dede argued that "even given the optimistic technological assumptions," AIEd would "have no significant potential for standalone 'education'" (Dede 1988). This very much remains to be seen, as much has already changed in the capabilities of AI.

[8] This operational definition avoids semantic and technical debates about what defines AI. Readers can refer to Krafft et al. (2020) for alternative definitions.

[9] Replace "intelligent tutoring systems" with "intelligent teaching systems" and the impetus is obvious.

Table 1 Ideal-type educational agents

Role	Teacher, professor	Tutor, coach, teaching assistant	Learning companion
Content expertise	Expert, full domain knowledge	Quasi-expert, substantial domain knowledge	Peer, similar to student within some range
Emotional positionality	Strict, neutral, authoritative, older	Supportive, motivating, semi-authoritative, older peer	Student-like, engaging, non-authoritative, same-age peer
Activity	Instruct, challenge, hold accountable	Guide, facilitate, encourage	Excite, relate, foster learning-by-teaching

conceive of teaching and learning and what they unpack into is shaped by the social and political context of educational settings at the micro-level and our educational systems writ large at the macro-level.

Among the most evident influences of our social context on AIEd is the fundamental emphasis on scale. In theory, an effective ITS can serve any number of students all at the same time, assessing student answers and differentiating instruction instantly. It can store a student's entire history of mistakes and growth, their patterns of learning, and even micro data about engagement, motivation, and more. The capacity for differentiation is especially notable in contrast with MOOCs. MOOCs cannot at present differentiate at all; all students watch the same videos and complete the same assignments with no attention paid to student characteristics, educational history, or learning style. ITS can differentiate readily and en masse in ways that arguably surpass the capacities of human teachers. In sum, ITS promise impressive capacities, filling the very gaps that plague traditional non-intelligent distance education.

Furthermore, ITS are becoming richer and more robust, increasingly attempting to simulate what were once thought to be 'uniquely' human capacities, namely emotional/sentiment measurement, analysis, and representation (Woolf et al. 2010; Arroyo et al. 2014). ITS are attempting to track student attention (D'Mello 2016), detect emotions such as frustration, and even teach physical motor skills (DeFalco et al. 2017; Harley et al. 2017). Along with improving the ability of ITS to measure and interpret emotional data through affective computing, efforts are increasing to portray emotion, encourage motivation, and "care" about students (Du Boulay 2011; Du Boulay et al. 2010; San Pedro et al. 2013).

Among the most interesting and cutting-edge ITS tools which often use social–emotional capacities are advanced versions of educational or pedagogical agents. These agents had early precursors in the 1990s but are now reaching far higher levels of sophistication (Dinçer and Doğanay 2017). Educational agents can be represented visually, with sound or with text (Baylor et al. 2003). They may appear as human faces, full human bodies, cartoons/animations, or may even be represented in embodied robots (Dunsworth and Atkinson 2007). Text-based educational agents must respond to students in a convincing human-like manner with words, while

auditory agents and embodied robots must master inflection and facial expressions. In light of these features, educational agents are theorized to foster socio-emotional connection, interest, and engagement, and thus produce superior learning outcomes. Evidence even in these early stages supports the claim that that humans do bond with nonhuman entities (Tanaka et al. 2015) and that educational agents have positive impacts on motivation and academic outcomes (Bickmore, Pfeifer, and Paasche-Orlow 2009; Domagk 2010; Woolf et al. 2010).[10]

On top of the different interfaces through which an educational agent can manifest, the agent can also take on different roles. I adapt work from Chou et al. (2003) and Payr (2003) on offering a simple ideal type depiction of educational agents (Table 1). Educational agents can be parsed by role (teacher, teacher's assistant) learning companion, peer tutor), by content expertise (expert to beginner), by emotional stance (from strict/authoritative to neutral to highly emotional), and by activity (from direct instruction to guidance to co-learning). Example ideal-type agents are the authoritative teacher, who contains all the relevant domain knowledge but portrays limited emotionality; the coach or tutor, who provides more mentorship, guidance, and support; and the learning companion, who has more restricted expertise and serves as a co-learner and peer.[11]

This is just a sample of the forms that an educational agent might take, as professor-like agents need not be emotionless nor learning companions juvenile, and indeed thoughtful design might suggest a hybrid of these kinds of characteristics. Nevertheless, these efforts to simulate humans as educational agents are revealing in terms of how they are encroaching on the traditional niche of human teachers and what that might mean for the future of education.

---

[10] Subsequently though, the use of explicitly emotional educational agents raises a cascade of ethical and design issues related to student privacy and manipulation. Do we want educational software to shame students for not working hard enough if it is an effective strategy, or to use gamification-style rewards (bright colors and addictive structures) to engender engagement? I consider this in the next section.

[11] Payr offers the idea of agents as actors, used in simulations for healthcare, management, etc., towards educational purposes. I restrict my attention to more classic instructional agents.

At this point, it is worth reviewing the nature of AIEd and ITS in light of our prior discussion of distance education:

- There are strong analogies between distance education and AIEd as forms of digital education. Both purport to achieve massive scale at low marginal cost. Both promise to challenge traditional educational structures, especially with regard to the role of teachers. And of course, both are subject to hype and skepticism.
- However, the disanalogies are also important. While MOOCs traditionally lack intelligent features and provide effectively no individualization or differentiation, AIEd is built on the AI paradigm of individualized, immediate responses. AIEd thus explicitly tackles one of the great barriers of distance education by simulating teachers, promising something closer to what Feenberg desired in a small classroom distance education setting. In essence, it is a candidate for the best of both (but not necessarily all) worlds.
- Moreover, while MOOCs offer little in the way of socio-emotional engagement and support, especially with respect to the student–teacher relationship, AIEd is advancing unafraid towards this goal, especially via educational agents. AIEd is thus explicitly addressing the very gaps proposed by Feenberg and Dreyfus regarding why distance education just can not replace traditional human teachers.

Feenberg (2002) writes that "whenever a new educational technology is introduced, arguments emerge for substituting interaction with the technology for the process of intellectual exchange. But there is something about dialogue, and the active involvement of the teacher, that is fundamental to the educational process and that should be woven into the design of every new instructional tool." Is it possible that something about AIEd technology is different, that it breaks with the typical cycle of substitution, overreach, retreat, and eventual synthesis?

If it has become accepted that computers can store more information and perform calculations more quickly, it has certainly not been accepted that computers can replace teachers as mentors and motivators. And yet, this perennial expectation is being challenged. While it is yet to be seen how robustly, convincingly, and effectively emotionality can be assessed and portrayed by AI, those who assume the irreplaceability of teacher interaction, differentiation, facilitation, and emotionality ought to revisit these assumptions. I discuss a related issue of replacing or otherwise transforming the work of teachers in Sect. 5.3.

Even if AIEd's simulation of teachers is less-than-perfect and clear quality gaps remain, as they will initially, advocates for technological substitution will have increasingly strong arguments for highlighting the advantages of AIEd—scalability, differentiation, instant adaptability, and lower cost. For this reason, it is important to review possible models of AIEd in the future. What sociotechnical considerations will impact the success of AIEd? What are the risks if the technology is pushed too far too soon? What ethical and design questions need to be asked and answered before major implementation efforts are undertaken? In the next section, I consider these questions.

5 Possible futures for AIEd

I review possible futures for AIEd and educational agents in particular with respect to two competing visions. These resemble what Feenberg calls the factory and city models. The factory model represents "a society reflecting in all its institutions the logic of modern production, obsessed by efficiency achieved through mechanization and management." In contrast, "the city is the place of cosmopolitan interactions and enhanced communication. Its god is not efficiency but freedom. It is not dedicated to rigid reproduction of the same, 'one best way', but to the flexible testing of possibilities" (Feenberg 2002). If the factory model is a technocratic dystopia, then the city model is its progressive counterpart.

My approach emphasizes the factory or dystopian model and, therefore, appears closer to the viewpoint of the skeptic. The purpose is to emphasize social and ethical risks by making them especially salient (a risk-averse approach). As a result, there are many benefits from AIEd that are not fully explored here. Even some of the negative scenarios imagined here could be viewed more charitably.[12] I encourage the reader to keep this lens in mind even where I have not.

I discuss social and ethical implications of AIEd regarding bias, curricula, international development, flexibility and ownership of educational systems, the role and potential replacement of teachers, and the use of nudging/manipulation by AIEd. In each case, I consider the optimistic 'city' perspective alongside a more critical 'factory' analysis.

5.1 Pedagogy and bias

Fully realized, an optimistic AIEd vision of pedagogy under the city model is one of robust differentiation. AIEd systems will be able to deeply understand students, including their cognitive development and history, emotional reactions, level of engagement, and even their future educational pathways. While differentiation remains a major challenge in classrooms today, with the flexible and responsive

---

[12] We might greatly be concerned about displacing and disempowering teachers, but lowering marginal costs is not necessarily bad, and technological destruction and creation can be for the ultimate good.

technology applied by AIEd, students would learn optimally, supported cognitively, structurally, and socio-emotionally with unprecedented levels of sophistication (Tomlinson 2000).

In contrast, an emerging fear in artificial intelligence is that of algorithmic bias, which may favor or disfavor certain subgroups by gender, sex, race, ethnicity, socioeconomic grouping, and other more subtle factors (Friedman and Nissenbaum 1996; Howard and Borenstein 2017; Bolukbasi et al. 2016; Buolamwini and Gebru 2018). AIEd systems are developed through training data, and those data will be extracted from dominant and early user bases. AIEd, therefore, learns about student engagement, learning, and cognitive styles from a subset of potential future users. Plausibly, these users will come from more highly-resourced school systems that can afford to purchase and experiment with new tools. These users may also come from wealthier countries and especially from countries where AIEd development is predominant, likely in high-income countries. This issue of intercultural use and bias is also considered by Pinkwart's (2016) dystopian scenario and likely to be a real concern. This is unfortunate given that low-income countries are the places that could be most positively impacted by AIEd.

What sorts of assumptions are implicit in the educational settings of wealthier nations, in the structure of their classrooms, and in the backgrounds of their students? ITS and educational agents built on training data from one part of the world may not speak to or effectively educate students from another (Anderson and Adams 1992; Fejerskov 2017; Kak 2020).[13] Indeed, challenges with technological access and literacy affect a huge portion of learners around the world (Nye 2015). Thus, when AIEd favors certain pedagogies, learning styles, and educational systems, it ultimately dis/advantages certain students and their communities. Finally, implicit and explicit biases in AIEd development also reflect the slanted demographics of its researchers—their own backgrounds and assumptions about education are already shaping AIEd (Blanchard 2015).

5.2 Curricular emphasis on STEM subjects

A major promise of online education has been its potential to take any subject material, perhaps taught by the 'best instructors' in the field, and make that available to anyone at any time. This complete universalization of curricula, coupled with the powers of AIEd, could ostensibly enable the original promise of MOOCs but with critical educational quality issues such as the role of teachers/teaching largely solved. This would represent the true flattening of education.

However, a similar risk obtains as with pedagogy. The AIEd development ecosystem will make strategic choices about what content is developed first, and as AIEd moves out of the laboratory and into classrooms, corporate developers and large purchasers will have an increasing say. Education technology companies have a critical role in translating 'laboratory' research into practice, but their financial motives may favor development decisions that neglect nuances important to AIEd researchers.

Moreover, the development of AIEd systems is itself conditioned by the nature of the technology. In particular, some subjects are easier to model computationally than others. Algebra is quite easy to create an ITS around; history, not so much. Mathematics relies on well-established equations and symbols, and even standard problem-solving approaches and student misconceptions are well understood and easily modeled by AIEd. Understanding meaning and context via natural language processing is far more challenging than simple mathematical representations. Evaluating the quality of a student's personal reflection on current issues or providing feedback on poetry are far more ambitious goals for AIEd. This is true both with respect to the expert domain (the content itself) and the pedagogical domain (how to teach the content). For example, data used to train an ITS can be obtained from both student answers and teacher feedback, and these data are much more straightforward to obtain from simple question–answer pairs in mathematics than from a winding and complex teacher–student conversation about historical perspectives or literary criticism. Indeed, it is no coincidence that current AIEd reflects this bias towards modeling STEM content (Roll and Wylie 2016).[14]

There is a further corollary of this logic related to subject bias having to do with the popularity of the subject material. More training data can be obtained from larger, more common courses as compared to smaller and more esoteric ones. Large introductory courses will provide robust and consistent data sets; esoteric and upper-level courses will be too rare, inconsistent, and expensive to capture data from. Many AIEd introductory algebra courses already exist, while courses on philosophy and art history are unavailable. Given educational and developmental pressures to produce content, developers may not wait to produce a full suite of English, social studies, science, and math courses, much less anthropology or gender studies. They will produce what can be done more easily and cheaply and marketed in turn.

---

[13] Research about the failure of MOOCs in India contrasted against the success of local technical institutes reveals that Indian students have different learning styles, a need for career relevance, and technological challenges with navigating digital educational tools as compared to the presumed (Western) consumers of MOOCs (Joshi et al. 2018).

[14] Even within STEM, some content and some courses are more easily modeled than others.

The upshot of these technical realities and developmental pressures is that large, introductory courses in STEM will be favored while courses in arts, social sciences, humanities, and other areas may be further relegated to second-tier importance (Olmos-Peñuela et al. 2015). This outcome, also discussed by Pinkwart (2016), is highly reminiscent of complaints that labor force, economic, and even military motives have pushed well-rounded, liberal education out of the way in the pursuit of STEM. This educational emphasis has further implications for student moral and civic development, diversity, equity, and more (George et al. 2001). The technical nature of AIEd may provide even more momentum towards this imbalance.

5.3 Teachers' role and displacement

One more well-established concern with AI is that of technology-induced role change and displacement, with subsequent effects for the exacerbation of inequality and economic and political instability (Autor 2015; Makridakis 2017; Manyika et al. 2017). The AIEd community is often careful to argue that its projects are tools to support and augment rather than displace teachers (Corbett et al. 1997; Pinkwart 2016). Many AIEd systems are designed to be used in coordination with teachers and cannot function without teachers' active facilitation and complementary instruction. For example, the ASSISTments project (Heffernan and Heffernan 2014) recognizes the mantra "put the teacher in charge, not the computer." This view is one of positive complementarity (Hannafin and Savenye 1993).

Under this paradigm, AIEd will serve as a tool to empower teachers. It will provide detailed information about student learning, enable unprecedented levels of differentiation, and enhance the status of the profession. Teachers will be freed up from the most rote tasks, and afforded the flexibility, creativity, and time to pursue advanced professional development or other educational goals in one's school and career.

There are a few reasons to be hopeful about this trajectory. One is the strong intentions of many AIEd researchers themselves, an important force shaping design. A second reason is that teachers do have some agency over whether and how tools are used; if they are not supportive of AIEd, the tools are unlikely to flourish in classrooms. Finally, some evidence suggests that teaching is among the hardest professions to automate (Frey and Osborne 2017), particularly because of its focus on social–emotional skills.

However, some empirical and socio-political considerations cut in the other direction. First, the intentions of AIEd researchers may not be enough. As discussed, designers may have only partial control over the multi-stable usages of a design. This is especially true when commercial interests develop tools based on economic logic. Education technology companies and governments may find it profitable and efficient to emphasize labor-saving versions of AIEd, after initial versions are transferred by AIEd researchers.

The aspiration to support, not replace, teachers is also undercut by research efforts to learn from and model human teachers, to function in environments without teachers, and to be evaluated against the performance of teachers (Boulay and Luckin 2016; Cumming and McDougall 2000). Comparative studies have found that ITS are rapidly converging on the effectiveness of human teachers.[15] Van Lehn's (2011) meta-review of 28 comparison studies found that ITS were essentially equal to human tutors in driving student learning gains (VanLehn 2011).[16] Kulik and Fletcher (2016) find similarly in a meta-analysis of 50 studies that students taught by ITS outperformed control students who received no ITS in 92% of studies, stronger than typical human tutors. The very fact that the AIEd community is proactively aware of this possibility of replacing teachers in at least some settings speaks to its plausibility.

For example, amongst the largest uses of AIEd out of the laboratory is Squirrel AI (or Yixue), an education technology company in China (Cui et al. 2018). Squirrel has registered over a million students in more than 200 cities, and is growing. For both students and instructors, the laptop is the main vehicle of learning. Squirrel's founder, Derek Li, argues that when AIEd technologies reach their ultimate vision, "human teachers will be like a pilot," playing a passive role, monitoring computer dashboards until a student is flagged, and even then focusing on emotional (not academic) communication (Hao 2019). Thus, while AIEd researchers may believe that their tools require the involvement of teachers in some way, it is not merely wholesale replacement of teachers that is important. Squirrel AI still involves a role for teachers, but it makes it easier to increase student–teacher ratios, displacing or removing a need to hire teachers in the first place. Moreover, it puts teachers into an arguably lower-level role of facilitating computer systems and managing student social and emotional needs.

Second, teachers may not have as much agency in making choices about the usage of AIEd as imagined. The history of education technology and its intersection with teachers reveals a more pained relationship (Cuban 1986). Teachers tend to be (perhaps rightly) resistant to the emphatic insistence of technology reformers, and implementation, use, and eventual outcomes fall far short of the optimistic education

---

[15] The validity of grades and test scores as measures of effective teaching/learning in contrast with various social–emotional/cognitive/twenty-first-century skills and long-term outcomes is beyond the scope here. Even these latter outcomes for ITS are within the realm of possibility now.

[16] The relative effect sizes were calculated as a Cohen's d of .76 for ITS and .79 for human tutors.

technology predictions. Some teachers thrive, leveraging new tools and innovating beyond even the creator's original intentions, but other teachers may lack comfort with the tools or reject the implied pedagogical and structural changes. This resistance itself feeds justifications for resentment and blame of teachers, and arguments for more aggressive enforcement and deprofessionalization of teachers. Moreover, increasing public expenditures in education, concerns about economic competitiveness and inequality, and actual or perceived stagnation of educational outcomes all increase pressure on teachers.[17] While it is true that teachers have some agency over whether AIEd tools are used, they are often overruled by administrators, policymakers, and other powerful interests.

Third, while some projections of automation suggest that teachers are relatively insulated, this insulation is granted by the focus of teachers on social–emotional tasks, which are currently hard to automate. Yet, even these projections are based on only the presumed near-future capacities of machine learning (Frey and Osborne 2017). This suggests caution, as other tasks previously thought non-automatable, like driving, can now be computerized. It is impossible to predict the limits of AI. Moreover, as we have seen, the AIEd community is undertaking very deliberate efforts to close the social–emotional gap, and to computerize the very tasks that currently insulate teachers.

If the complementarity thesis fails then, the alternatives of displacement, deprofessionalization, and deskilling may dominate. The eventual outcomes of the substitution paradigm are larger class sizes, job displacement, and even relegation of the teacher role to a mere facilitator job, where human decision-making is replaced by algorithms (Guilherme 2019; Croy 1989). In this role, low-skill instructors walk up and down the aisles of a large computer lab, policing student behavior and ensuring logistical smoothness and basic student engagement. This fear, dating back to Nobel's concern of digital diploma mills, it is not so far-fetched. It is evidenced by Squirrel AI's rapid growth, as well as in classrooms in low-income countries which currently employ prepackaged curricula coupled with low-skill teaching guides (Rangan and Lee 2011). Even success stories like Georgia State University, with increased graduation rates and closure of achievement gaps, have relied on large computer-based courses with high student–teacher ratios to achieve educational goals inexpensively (Dimeo 2017).

5.4 International development and premature automation

The above analyses can be extended to consideration of education across different regions, socioeconomic settings, and countries (Nye 2015). The utopian view is, again, universalized access to top-quality educational materials and now AIEd-supported instruction (Pinkwart 2016). While digital education and MOOCs were previously unable to address issues of differentiation, local context, and socio-emotional engagement, AIEd promises to improve on these challenges. However, even a successful long-term complementarity might be preceded by short-term sufficiency. Perhaps, the best we can do given established challenges with the educational systems and teacher workforce in many countries is to provide highly packaged instructional tools (Birdsall et al. 2005).[18]

On the other hand, investing in technological solutions and accepting packaged education facilitated by low-skill classroom guides may in fact be the short-term good destroying the long-term better. While educational aid has recently stagnated, the promise of cheap, effective educational tools could lead to a revolution in educational foreign aid (UNESCO 2017). If AIEd technologies are viewed as relatively superior to the status quo without demanding major infrastructural spending and change, the expansion of AIEd may come at the cost of long-term educational development. We might call this the premature automation of education.[19] That is, a country's investment in its educational infrastructure, systems, and perhaps most importantly preparation and ongoing professional development of high-quality teachers may suffer, because AIEd is viewed as a sufficient short-term solution.

This risks precluding long-term educational effectiveness and international equity, locking certain communities or countries into a kind of permanent second-tier educational system, particularly if low-income countries are viewed as a laboratory for AIEd and commercial interests (Fejerskov 2017). Therefore, even if we accept deprofessionalization of teachers and increased teacher–student ratios as possibly 'better' short-term strategies endorsed by the likes of USAID, we ought to be very careful with long-term implications for educators and educational systems, especially when our proposed solutions are imposed from the outside.

---

[17] Teachers are simultaneously some of the most respected and least empowered individuals in society (Gallup 2010; Ingersoll and Perda 2008). Teachers are expected to be able to solve educational problems even without necessary support and resources and are nevertheless blamed when they fail to do so.

[18] See the XSEED program which tries to leverage and improve human capacity while maintaining scalability through highly packaged education (Nayak et al. 2010).

[19] This is somewhat analogous to the concept of premature deindustrialization (Rodrik 2016).

5.5 Flexibility and ownership of educational systems

Accompanying the vision of finely individualized pedagogy and curricula is a view that local communities, families, and students themselves will have unprecedented control over their own educational destinies. Students will have significant flexibility in choosing the courses they desire at the pace they want towards their anticipated academic and professional trajectory. Schools will be able to experiment under a supportive regulatory environment, tailoring educational experiences to the needs of individuals and their communities. Under this view, educational ownership becomes an increasingly local project, engaging and inspiring teachers, families, and students towards better educational, social, and civic outcomes.

The contrasting view reflects challenges already presented in the pedagogy and curriculum sections above. Under this scenario, curricula will be packaged with aforementioned biases toward certain content, cultures, and learning styles. Official entities at higher levels of aggregation, such as school districts, states, and countries, will determine educational priorities for progression and graduation. These priorities may reflect national interests, rather than individual or local ones, and may additionally reflect unbalanced private sector market influence and lobbying.[20] The irony here is that a technology lauded for its ability to individualize could be used instead for mass standardization. If AIEd enters the classroom in such a way that it decenters and supplants teachers, it could detract from the creativity and diversity made possible by the esoterica of individual teachers, and even decrease educational freedoms of students and families.

5.6 Nudging and manipulation

One last element worthy of review is the idea of nudging, or framed less positively, manipulation. If we anticipate increased abilities of AIEd systems to predict, measure, and respond to emotional data, we raise questions about the form those responses will take. Ideally, these emotional systems are used to support and improve students educationally as well as socio-emotionally, taking into account the kinds of broader goals we might have for students beyond test scores and graduation.

However, the possibility of nudging raises a suite of challenging ethical issues, presented in a slightly different context by Borenstein and Arkin (2017), which I expand on here. For example, do we expect or want students to be consciously aware of, or be able to decide whether a nudge occurs? Do we want nudges that harm, even in service of calculated long-term benefit, such as an AI system shaming a student for laziness or mistakes?[21] More generally, is it appropriate to treat student emotion as a means to an end? Do we want AIEd systems to manipulate student cognition with flashy colors and potentially addictive gamification techniques?[22] Are there particularly vulnerable students, "due to age, physical or mental characteristics, socioeconomic status, or other relevant factors" toward which we would want to pay special attention (Borenstein and Arkin 2017)? A further issue compounding these nuanced challenges is the so-called black box problem, wherein some AI systems built on big data rely on complicated and fundamentally opaque decisions (Castelvecchi 2016). In such cases, even the programmers of AIEd systems may not understand where and why students are being nudged, rendering legitimacy even more problematic.

5.7 Final notes on the city and factory

Readers interested in other depictions of possible futures of AIEd should see similar work by Rummel et al. (2016) and Pinkwart (2016). Both present utopian and dystopian scenarios, with the latter serving as a cautionary tale. Rummel et al. imagine a future of computer-supported collaborative learning where AIEd systems lack flexibility, are opaque and intrusive, and ultimately undermine student collaboration and motivation. Pinkwart imagines an ITS which is culturally biased towards US learners, pedagogically favors subjects like math rather than ethics, and where companies commercialize student data and even profile students as dangerous. In response, both papers offer suggestions for research. Pinkwart proposes increased attention to issues such as intercultural education, achieving practical impact in classrooms, and student privacy. Rummel et al. emphasize a need for more transparent and flexible AIEd tools. In the next section, I move to recommendations to address some of the ethical and social risks raised in this paper.

---

[20] I would not be the first person to observe that STEM has captured education policy with talks of gaps, pipelines, and crises (Metcalf 2010).

[21] Dreyfus suggested that social functions like shame are common and indeed useful in education. Does the same apply when a computer is doing the shaming?

[22] Of note, we accept many of these nudging practices in children's books, television shows, in classrooms, and on cereal boxes. It is important to uncover what the ethical lines are and what makes AI applications different, if anything. Consider the recent energy behind gamification and related criticisms (Rughinis 2013).

6 What's an AIEd researcher to do?

These possible futures represented as the factory and city model may ultimately be as inaccurate as some of the predictions cited previously, but they minimally help to situate AIEd in the present and reveal novel ways in which AIEd as a technical tool can interact with social and political realities. Awareness is a good starting point, but how is this information and speculation useful to us beyond mere consciousness-raising? The proximal audience for this paper is those with some interest in shaping the future of AIEd systems, and AI more broadly as it impacts social and policy outcomes. As such, the reader may have some influence over the possible futures of AIEd and AI generally. How might this influence be exerted in a responsible and prudent manner?

One promising and growing framework in this vein is Responsible Research and Innovation (RRI), which "asks researchers to take on a practice of critical reflection considering the potential societal impacts of their research outputs, and second to include the general public in a dialogue around the development of research goals and strategy" (Schomberg 2013). RRI has been applied to a number of domains and regions, and can be applied to computer science research and AI in particular.

Aligned with the sentiment of RRI, a proposal from the ACM Future of Computing Academy (FCA) argues that in light of increasing recognition of possible harms of computing, it is incumbent on researchers to "work much harder to address the downsides of our innovations" and that the failure to do so has represented "a serious and embarrassing intellectual lapse" and arguably an ethical one (Hecht et al. 2018). The recommended response is to leverage the gatekeeping function of the peer review process, requiring papers to include robust broader impact discussions that consider (1) potential harms as well as potential benefits of the research, and (2) technical and policy options that would mitigate those harms. The proposal also suggests that research that entails net harms and researchers who repeatedly pursue such research should face strong resistance and rejection.

While this use of the peer review process represents an interesting and potentially viable mechanism to promote socially responsible research, there are also significant limitations as well as reasons to favor other approaches. Tapan Parikh rebuts that pursuit of harmful research may be important for prophylactic reasons, and points out that researchers are not wholly responsible for how the technologies they develop are used politically.[23] Even seemingly beneficial or harmless research can be misused, and these misuses cannot be controlled easily by researchers (Waddell 2018b). Moreover, the ability of both authors and reviewers to accurately and unbiasedly capture, predict, and review risks in such uncertain domains and potentially far outside of their field of expertise is highly questionable (Parikh 2018). Is it reasonable to expect an AI researcher to provide accurate models of labor displacement, produce a cost–benefit analysis, or to engage with critiques in education implementation or sociology?

Consequently, attempting to assess net harms may not be practical and thus is unlikely to be a reliable rubric for critical gatekeeping decisions. I would add that the focus on net harms versus benefits itself misses the point in an important sense. A net harm/benefit discussion emanates from the tradition of optimization and efficiency and is often orthogonal to critiques regarding distribution and equity (Callahan 1964). Ironically, these are often the very intellectual and ethical lapses that computing research has been accused of in the first place. For example, it may very well be that a certain job-displacing automated technology leads to net benefits, measured by economic growth, GDP, and perhaps even net jobs. But concerns about automation were never purely about GDP in the first place; the concern is that many people in certain job types, sectors, or demographic groups will lose out, irrespective of society's net benefit. Similar points can be made about racially biased facial recognition or criminal sentencing AI systems, which are opposed not primarily on aggregate efficiency grounds, but on ethical ones related to fairness, transparency, and accountability (Buolamwini and Gebru 2018). Therefore, again, even if peer review is a viable gate, net harm and benefit are not appropriate rubrics for gatekeeping.

Another challenging question is the scope of implications that a researcher is responsible for. Must a researcher be accountable to the implications of their entire field or sub-field (the large-scale implications of AIEd discussed previously)? Must every AI paper about image processing include a boilerplate discussion of the risks of facial recognition technology? This seems both unfair and wasteful. However, equally problematic is the risk that few individual papers (e.g., review papers, meta-analyses) synthesize and encompass the set of developments and risks inherent in the broader field. As such, most single papers will not and perhaps should not discuss these issues in depth.[24] Instead, collective gatekeeping of a domain of research or new technology seems more effective and appropriate than gatekeeping

---

[23] Nuclear technology is a relatively easier case, for example. In general, knowing where to draw the lines of responsibility can be very challenging, especially given the problem of many hands (Thompson 1980).

[24] An exception is the case where lines of code included in the research are readily corruptible towards harmful purposes. In this case, disclosure to developers or authorities or withholding code is an established approach in computer security (Waddell 2018a).

of individual papers. Along these lines, one of Parkih's helpful proposals is to leverage existing research fora outside of the peer review process. As is done in some cases, conferences can host interdisciplinary panels, and journals can encourage integration of extradisciplinary ideas even if they do not advance 'core research.'[25]

Ultimately, the FCA proposal raises an important issue, and is certainly correct that research should not be presented through solely rose-tinted glasses. Minimum consideration of risks is a good idea, and while thoughtful expansion of broader impact consideration is a great start, the research community needs even more robust ways to engage with these complex technoethical issues and their associated uncertainties. Instead of individual gatekeeping then, responsible research and implementation suggests broader thinking of impacts at all stages of the research process, from brainstorming to publication to implementation and evaluation.

A key element of this responsible research is the inclusion of the subjects, targets, and decision makers who surround the research/technologies in question into the research process, implementation, and associated fora. Gatekeeping should not be done solely by academic reviewers. While incorporating broader voices is complex and can be especially challenging for technical communities (Shilton 2018), work on participatory methodologies and DBIR has demonstrated that inclusivity is perfectly feasible (Schuler and Namioka 1993). For example, the healthcare research community is currently working through the micro- and macro-challenges of incorporating patients as co-equal researchers, and has valuable lessons and models from which the computing and AI communities could learn (Domecq et al. 2014).

In light of all of this, what might improving our capacity to conduct thoughtful and socially responsible research look like in the case of AIEd?

- Currently, teacher collaboration in AIEd research is only partial, with teachers involved most often as accessories during implementation (Roll and Wylie 2016). Instead, teachers and other educational stakeholders (students, parents, administrators, education researchers, and policymakers) should take a much greater role in the development and implementation of AIEd research. Fortunately, there is extensive work on design in educational technology that provides methods and frameworks to better incorporate teachers and other stakeholders in design

(Luckin et al. 2013). The incorporation of diverse perspectives is beneficial both for researchers and developers who want their products to be truly useful, while also encouraging development of technologies along socially responsible pathways.[26]

- AIEd research publications and platforms should actively discuss how exactly they interface with larger scale educational systems and affect social and policy considerations, not merely how the technology works in abstract or within a single classroom. This suggests an emphasis on implementation studies, perhaps along the lines of DBIR (Fishman et al. 2013), as well as increased work with social science and policy scholars outside of education. Academic and corporate developers should articulate clearly how exactly their platforms interact with existing (and future) settings, including the projected degree of collaboration with and the anticipated role of teachers and in what kinds of environments the tool is appropriate and inappropriate.

- Importantly, parallel considerations also hold for those funding and procuring AIEd technology. School administrators, school districts, regulators, etc., should be transparent about how exactly the technologies they procure are meant to interface with their educational systems and stakeholders. This, at a minimum, should open these decisions to transparency and accountability. Thus, if a developer's tool deprofessionalizes teachers, or a school district uses technology to justify cutting class sizes, these decisions can more easily face public scrutiny and accountability.

- Finally, journals, conferences, private and public funders, graduate programs, and commercial parties interested in economizing AIEd tools should emphasize social responsibility in research, development, and implementation. Such emphasis should be deeply integrated into processes and cultures, not offered superficially or nominally, as can be the case in computing and engineering education (Cech 2013; Leydens and Lucena 2018). For example, research fora should discuss the 'corruptible potential' of their technologies, mitigative strategies, and how they might engage positively with subjects, targets, and policymakers connected with the technology. And then, they should do so.

---

[25] A great exemplar of this is the International Journal of Artificial Intelligence in Education, which published a series of interdisciplinary discussions on the future of AIEd in a special edition (Lane et al. 2016).

[26] A promising example is that of authoring tools, which allowed nonexperts to have some say in creating their own AIEd platforms (Dermeval et al. 2017). However, authoring tools have a long way to go before they are truly accessible to teachers and other curriculum developers.

7 Conclusion

In Dede's prognostications about the future of artificial intelligence in education 30 years ago, he offered that "the two most common errors in technology assessment are overestimating the speed of diffusion of an innovation and underestimating its eventual consequences and side effects" (Dede 1988). While predicting the future of educational technology often lends itself to uncertainty and extremes, a socially and historically grounded consideration of educational technology helps to reveal possible trajectories for AIEd. In particular, AIEd and ITS in particular promise (or threaten) to encroach on the roles of teachers and close perhaps the most prominent barriers keeping distance digital education from achieving its lofty goals—differentiation, and social–emotional engagement.

In this light, it is important to evaluate the implications of AIEd for our educational systems, including considerations of pedagogy, curricula, the role and possible automation of teachers, international development, ownership of educational choices, and behavioral manipulation. By reviewing these potential futures, I hope to promote critical reflection on possible developmental pathways of AIEd as it comes closer to mainstream implementation. Beyond that, I have offered preliminary strategies towards a more thoughtful engagement with AIEd's future and how the AIEd community can exercise the social responsibility incumbent on it. Moreover, these considerations and strategies are relevant to other research and practice communities, such as those in AI, computing, and technology design generally. It is my hope that through sustained interdisciplinary dialogue and research, and continued reflection on our progress, we can realize the best of promising technologies like AIEd while avoiding the worst harms.

References

Agre PE (1999) Information technology in higher education: the "Global Academic Village" and intellectual standardization. Horizon 7(5):8–11

Anderson JA, Adams M (1992) Acknowledging the learning styles of diverse student populations: implications for instructional design. New Dir Teaching Learn 1992(49):19–33

Argueta R, Jessica Huff D, Tingen J, Corn JO (2011) Laptop initiatives: summary of research across six states. Friday Institute for Educational Innovation, North Carolina State University, Raleigh, pp 1–20

Armstrong S, Sotala K, Éigeartaigh SSÓ (2014) The errors, insights and lessons of famous AI predictions—and what they mean for the future. J Exp Theor Artif Intell 26(3):317–342. https://doi.org/10.1080/0952813X.2014.895105

Arroyo I, Woolf BP, Burelson W, Muldner K, Rai D, Tai M (2014) A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. Int J Artif Intell Educ 24(4):387–426. https://doi.org/10.1007/s40593-014-0023-y

Autor DH (2015) Why are there still so many jobs? The history and future of workplace automation. J Econ Perspect 29(3):3–30. https://doi.org/10.1257/jep.29.3.3

Baylor A, Ryu J, Shen E (2003) The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. EdMedia: world conference on educational media and technology. Association for the Advancement of Computing in Education (AACE), Chesapeake, pp 452–458

Beldarrain Y (2006) Distance education trends: integrating new technologies to foster student interaction and collaboration. Distance Educ 27(2):139–153

Bickmore TW, Pfeifer LM, Paasche-Orlow MK (2009) Using computer agents to explain medical documents to patients with low health literacy. Patient Educ Couns 75(3):315–320

Birdsall N, Levine R, Ibrahim A (2005) Towards universal primary education: investments, incentives, and institutions. Eur J Educ 40(3):337–349

Blanchard EG (2015) Socio-Cultural imbalances in AIED research: investigations, implications and opportunities. Int J Artif Intell Educ 25(2):204–228

Bolukbasi T, Chang K-W, Zou JY, Saligrama V, Kalai AT (2016) Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R (eds) Advances in neural information processing systems 29. Curran Associates Inc, New York, pp 4349–4357

Bond M, Zawacki-Richter O, Nichols M (2019) Revisiting five decades of educational technology research: a content and authorship analysis. Br J Edu Technol 50(1):12–63. https://doi.org/10.1111/bjet.12730

Borenstein J, Arkin RC (2017) Nudging for good: robots and the ethical appropriateness of nurturing empathy and charitable behavior. AI Soc 32(4):499–507

Boulay Du, Benedict KA, Luckin R, Martínez-Mirón E, Méndez GR, Carr A (2010) Towards systems that care: a conceptual framework based on motivation, metacognition and affect. Int J Artif Intell Educ 20(3):197–229

Buolamwini J, Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency, pp 77–91

Callahan RE (1964) Education and the cult of efficiency. University of Chicago Press, Chicago

Castelvecchi D (2016) Can we open the black box of AI? Nat News 538(7623):20. https://doi.org/10.1038/538020a

Cech EA (2013) Culture of disengagement in engineering education? Sci Technol Human Values 39(1):42–72. https://doi.org/10.1177/0162243913504305

Chickering AW, Ehrmann SC (1996) Implementing the seven principles: technology as lever. AAHE Bull 49:3–6

Chou C-Y, Chan T-W, Lin C-J (2003) Redefining the learning companion: the past, present, and future of educational agents. Comput Educ 40(3):255–269

Christensen G, Steinmetz A, Alcorn B, Bennett A, Woods D, Emanuel EJ (2013) The MOOC phenomenon: who takes massive open online courses and why? https://doi.org/10.2139/ssrn.2350964

Cohen A, Shimony U, Nachmias R, Soffer T (2019) Active learners' characterization in MOOC forums and their generated knowledge. Br J Edu Technol 50(1):177–198. https://doi.org/10.1111/bjet.12670

Corbett AT, Koedinger KR, Anderson JR (1997) Intelligent tutoring systems. Handb Hum Comput Interact 5:849–874

Croy MJ (1989) Ethical issues concerning expert systems' applications in education. AI Soc 3(3):209–219. https://doi.org/10.1007/BF01891431

Cuban L (1986) Teachers and machines: the classroom use of technology since 1920. Teachers College Press, New York

Cui W, Xue Z, Thai K-P (2018) Performance comparison of an AI-based adaptive learning system in China. In: 2018 Chinese automation congress (CAC), pp 3170–75. https://doi.org/10.1109/CAC.2018.8623327

Cumming G, McDougall A (2000) Mainstreaming AIED into education? Int J Artif Intell Educ (IJAIED) 11:197–207

Czopek A, Pietrzak P (2016) Unlocking the potential of technology in education. E-Mentor 3:78–82

D'Mello SK (2016) Giving eyesight to the blind: towards attention-aware AIED. Int J Artif Intell Educ 26(2):645–659. https://doi.org/10.1007/s40593-016-0104-1

Dede CJ (1988) Probable evolution of artificial-intelligence-based educational devices. Technol Forecast Soc Chang 34(2):115–133

DeFalco JA, Rowe JP, Paquette L, Georgoulas-Sherry V, Brawner K, Mott BW, Baker RS, Lester JC (2017) Detecting and addressing frustration in a serious game for military training. Int J Artif Intell Educ. https://doi.org/10.1007/s40593-017-0152-1

Dermeval D, Paiva R, Ibertbittencourt I, Vassileva J, Borges D (2017) Authoring tools for designing intelligent tutoring systems: a systematic review of the literature. Int J Artif Intell Educ. https://doi.org/10.1007/s40593-017-0157-9

Dillenbourg P, Järvelä S, Fischer F (2009) The evolution of research on computer-supported collaborative learning. In: Balacheff N, Ludvigsen S, de Jong T, Lazonder A, Barnes S (eds) Technology-enhanced learning: principles and products. Springer, Dordrecht, pp 3–19. https://doi.org/10.1007/978-1-4020-9827-7_1

Dimeo J (2017) Georgia state improves student outcomes with data. Inside Higher Edc

Dinçer S, Doğanay A (2017) The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. Comput Educ 111:74–100

Domagk S (2010) Do pedagogical agents facilitate learner motivation and learning outcomes? J Media Psychol 22:84

Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, Brito JP, Boehmer K, Hasan R, Firwana B (2014) Patient engagement in research: a systematic review. BMC Health Serv Res 14(1):89

Dreyfus HL (1979) What computers can't do: the limits of artificial intelligence. Harper Colophon Books 613, Rev edn. Harper & Row, New York

Dreyfus HL (1992) What computers still can't do: a critique of artificial reason. MIT press, New York

Dreyfus HL (2002) Anonymity versus commitment: the dangers of education on the internet. Educ Philos Theory 34(4):369–378

Dreyfus HL (2008) On the internet. Routledge, London

du Boulay B (2011) Towards a motivationally intelligent pedagogy: how should an intelligent tutor respond to the unmotivated or the demotivated? New perspectives on affect and learning technologies. Springer, New York, pp 41–52

du Boulay B, Luckin R (2016) Modelling human teaching tactics and strategies for tutoring systems: 14 years on. Int J Artif Intell Educ 26(1):393–404. https://doi.org/10.1007/s40593-015-0053-0

Dunsworth Qi, Atkinson RK (2007) Fostering multimedia learning of science: exploring the role of an animated agent's image. Comput Educ 49(3):677–690

Ely DP (1999) Conditions that facilitate the implementation of educational technology innovations. Educ Technol 39(6):23–27

Emery E (1972) The press and America: an interpretative history of the mass media. Pearson, London

Evans A (2003) This virtual life: escapism and simulation in our media world. Vision, London

Feenberg A (2002) Transforming technology: a critical theory revisited. Oxford University Press, Oxford

Feenberg A (2008) The online education controversy. In: Conference on technology for learning, teaching and the institution

Feenberg A, Hamilton E (2012) Alternative rationalisations and ambivalent futures: a critical history of online education. (Re)inventing the internet: critical case studies. Sense Publishers, Boston, pp 43–70

Fejerskov AM (2017) The new technopolitics of development and the global south as a laboratory of technological experimentation. Sci Technol Human Values 42(5):947–968. https://doi.org/10.1177/0162243917709934

Fishman BJ, Penuel WR, Allen A-R, Cheng BH, Sabelli N (2013) Design-based implementation research: an emerging model for transforming the relationship of research and practice. Natl Soc Study Educ 112:136–156

Frey CB, Osborne MA (2017) The future of employment: how susceptible are jobs to computerisation? Technol Forecast Soc Chang 114:254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Friedman T (2005) The world is flat: a brief history of the twenty-first century. Macmillan, New York

Friedman B, Nissenbaum H (1996) Bias in computer systems. ACM Trans Inf Syst (TOIS) 14(3):330–347

Gallup (2010) Americans' views of public schools still far worse than parents'. Gallup.Com. https://news.gallup.com/poll/142658/Americans-Views-Public-Schools-Far-Worse-Parents.aspx. Accessed 25 Aug 2010

George YS, Neale DS, Van Horne V, Malcolm SM (2001) In pursuit of a diverse science, technology, engineering, and mathematics workforce. In: American association for the advancement of science, pp 1–24

Goel A, Anderson T, Belknap J, Creeden B, Hancock W, Kumble M, Salunke S, Sheneman B, Shetty A, Wiltgen B (2016) Using watson for constructing cognitive assistants. Adv Cognit Syst 4:1–6

Guilherme A (2019) AI and education: the importance of teacher and student relations. AI Soc 34(1):47–54. https://doi.org/10.1007/s00146-017-0693-8

Hamilton E, Feenberg A (2005) The technical codes of online education. E-Learn Digit Media 2(2):104–121

Hannafin RD, Savenye WC (1993) Technology in the classroom: the teacher's new role and resistance to It. Educ Technol 33(6):26–31

Hao K (2019) China has started a grand experiment in AI education. it could reshape how the world learns. MIT Technology review

Harley JM, Lajoie SP, Frasson C, Hall NC (2017) Developing emotion-aware, advanced learning technologies: a taxonomy of approaches and features. Int J Artif Intell Educ 27(2):268–297. https://doi.org/10.1007/s40593-016-0126-8

Hecht B, Wilcox L, Bigham JP, Schöning J, Hoque E, De Russis L, Yarosh L, Anjum B, Contractor D, Wu C (2018) It's time to do something: mitigating the negative impacts of computing through a change to the peer review process. ACM FCA (blog)

Heffernan NT, Heffernan CL (2014) The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. Int J Artif Intell Educ 24(4):470–497. https://doi.org/10.1007/s40593-014-0024-x

Hendler J (2008) Avoiding another AI winter. IEEE Intell Syst 23(2):2–4

Howard A, Borenstein J (2017) The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. Sci Eng Ethics. https://doi.org/10.1007/s11948-017-9975-2

Ihde D (1986) Experimental phenomenology: an introduction. SUNY Press, Albany

Ingersoll RM, Perda D (2008) The status of teaching as a profession. In: Schools and society: a sociological approach to education, pp 107–118

Iseke-Barnes JM (1996) Issues of educational uses of the internet: power and criticism in communications and searching.

J Educ Comput Res 15(1):1–23. https://doi.org/10.2190/FLYP-YNQC-9T55-MKB5

Joshi M, Joshi T, Rangaswamy N (2018) Scaling classroom IT skill tutoring: a case study from India. In: Proceedings of the 2018 CHI conference on human factors in computing systems, vol 630, pp 1–630:12. CHI '18. ACM, New York, NY, USA. https://doi.org/10.1145/3173574.3174204

Jowett B (1901) The dialogues of plato, translated into english with analyses and introductions, 3rd edn. Clarendon Press, Oxford

Kak A (2020) 'The global south is everywhere, but also always somewhere': national policy narratives and AI justice. In: Proceedings of the AAAI/ACM conference on AI, ethics, and society. AIES '20. Association for Computing Machinery, New York, NY, USA, pp 307–312. https://doi.org/10.1145/3375627.3375859

Koh A (2017) Feminist pedagogy in the digital age: experimenting between MOOCs and DOCCs. In: Losh E (ed) MOOCs and their afterlives: experiments in scale and access in higher education. University of Chicago Press, Chicago, pp 123–134

Krafft PM, Young M, Katell M, Huang K, Bugingo G (2020) Defining AI in policy versus practice. In: Proceedings of the AAAI/ACM conference on AI, ethics, and society. AIES '20. Association for computing machinery, New York, NY, USA, pp 72–78. https://doi.org/10.1145/3375627.3375835.

Kulik JA, Fletcher JD (2016) Effectiveness of intelligent tutoring systems: a meta-analytic review. Rev Educ Res 86(1):42–78. https://doi.org/10.3102/0034654315581420

Lane HC, McCalla G, Looi C-K, Bull S (2016) Preface to the IJAIED 25th anniversary issue, part 2. Int J Artif Intell Educ 26(2):539–543

LeBlanc P (2020) COVID-19 has thrust universities into online learning—how should they adapt? Brookings (blog)

Leggett WP, Persichitte KA (1998) Blood, sweat, and tears: 50 years of technology implementation obstacles. TechTrends 43(3):33–36

Leydens JA, Lucena JC (2018) Engineering justice: transforming engineering education and practice. In: IEEE PCS professional engineering communication series. Wiley, Hoboken. IEEE Press

Li H (2002) Distance education: pros, cons, and the future. Presented at Western States Communication Association annual meeting 2002

Luckin R, Puntambekar S, Goodyear P, Grabowski BL, Underwood J, Winters N (2013) Handbook of design in educational technology. Routledge, London

Lyotard J-F (1984) The postmodern condition: a report on knowledge, vol 10. University of Minnesota Press, Minneapolis

Mackness J, Mak S, Williams R (2010) The ideals and reality of participating in a MOOC. In: Proceedings of the 7th international conference on networked learning 2010. University of Lancaster, Lancashire

Makridakis S (2017) The forthcoming artificial intelligence (AI) revolution: its impact on society and firms. Futures 90:46–60

Manyika J, Lund S, Chui M, Bughin J, Woetzel J, Batra P, Ko R, Sanghvi S (2017) What the future of work will mean for jobs, skills, and wages: jobs lost, jobs gained. McKinsey Global Institute. https://www.mckinsey.com/featured-insights/future-of-organizations-and-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages. Accessed 23 Aug 2018

McArthur D, Lewis M, Bishary M (2005) The roles of artificial intelligence in education: current progress and future prospects. J Educ Technol 1(4):42–80

Metcalf H (2010) Stuck in the pipeline: a critical review of STEM workforce literature. InterActions UCLA J Educ Inf Stud 6(2):27–47

Müller VC, Bostrom N (2016) Future progress in artificial intelligence: a survey of expert opinion. Fundamental issues of artificial intelligence. Synthese library. Springer, Cham, pp 555–572. https://doi.org/10.1007/978-3-319-26485-1_33

Nayak A, Arora S, Joshi H (2010) A Million children now! Transforming school education in India through the IDiscoveri XSEED living knowledge system (innovations case narrative: IDiscoveri). Innov Technol Governance Glob 5(2):31–51

Nenkov N, Dyachenko Y, Dimitrov G, Koeva K (2016) Artificial intelligence technologies for personnel learning management systems. In: Yager R, Sgurev V, Hadjiski M, Jotsov V (eds) 2016 IEEE 8th international conference on intelligent systems (Is), pp 189–194

Niculescu C (2016) Intelligent tutoring systems—trends on design, development and deployment. In: Roceanu I, Holotescu C, Colibaba AC, Blaga M, Beligan D, Adascalitei A, Miron C, Herman C, Grigore V, Stanescu M (eds) Elearning Vision 2020!, vol Iii. Carol I Natl Defence Univ Publishing House, Bucharest, pp 280–285

Noble DF (1998a) Digital diploma mills, part II: the coming battle over online instruction. Sociol Perspect 41(4):815–825

Noble DF (1998b) Digital diploma mills: the automation of higher education. Sci Culture 7(3):355–368

Norris C, Hossain A, Soloway E (2012) Under what conditions does computer use positively impact student achievement? Supplemental vs. essential use. In: 2021–28. Association for the advancement of computing in education (AACE). https://www.learntechlib.org/primary/p/39886/. Accessed 24 Aug 2018

Nye BD (2015) Intelligent tutoring systems by and for the developing world: a review of trends and approaches for educational technology in a global context. Int J Artif Intell Educ 25(2):177–203. https://doi.org/10.1007/s40593-014-0028-6

Nye BD (2016) ITS, the end of the world as we know it: transitioning AIED into a service-oriented ecosystem. Int J Artif Intell Educ 26(2):756–770. https://doi.org/10.1007/s40593-016-0098-8

Olmos-Peñuela J, Benneworth P, Castro-Martínez E (2015) Are sciences essential and humanities elective? Disentangling competing claims for humanities' research public value. Arts Hum Higher Educ 14(1):61–78. https://doi.org/10.1177/1474022214534081

Olsen B, Sexton D (2009) Threat rigidity, school reform, and how teachers view their work inside current education policy contexts. Am Educ Res J 46(1):9–44

Parikh T (2018) Mitigating the negative implications of computing: making space for debate. Medium (blog)

Payr S (2003) The virtual university's faculty: an overview of educational agents. Appl Artif Intell 17(1):1–19

Pinch TJ, Bijker WE (1984) The social construction of facts and artefacts: or how the sociology of science and the sociology of technology might benefit each other. Soc Stud Sci 14(3):399–441. https://doi.org/10.2307/285355

Pinkwart N (2016) Another 25 years of AIED? challenges and opportunities for intelligent educational technologies of the future. Int J Artif Intell Educ 26(2):771–783. https://doi.org/10.1007/s40593-016-0099-7

Rangan VK, Lee KL (2011) Bridge International Academies: a school in a box. SSRN scholarly paper ID 2017973. Social Science Research Network, Rochester

Rich E, Knight K (1991) Artificial intelligence. McGraw-Hill, New York

Rivard Ry (2013) Measuring the MOOC dropout rate. Inside Higher Educ 8:2013

Rodrik D (2016) Premature deindustrialization. J Econ Growth 21(1):1–33. https://doi.org/10.1007/s10887-015-9122-3

Roll I, Wylie R (2016) Evolution and revolution in artificial intelligence in education. Int J Artif Intell Educ 26(2):582–599. https://doi.org/10.1007/s40593-016-0110-3

Romero C, Ventura S (2010) Educational data mining: a review of the state of the art. IEEE Trans Syst Man Cybern Part C (Appl Rev) 40(6):601–618. https://doi.org/10.1109/TSMCC.2010.2053532

Romiszowski AJ (2004) How's the e-learning baby? factors leading to success or failure of an educational technology innovation. Educ Technol 44(1):5–27

Rosé CP, Ferschke O (2016) Technology support for discussion based learning: from computer supported collaborative learning to the future of massive open online courses. Int J Artif Intell Educ 26(2):660–678. https://doi.org/10.1007/s40593-016-0107-y

Rosenberger R (2017) The ICT educator's fallacy. Found Sci 22(2):395–399

Rughinis R (2013) Gamification for productive interaction: reading and working with the Gamification debate in education. In: 2013 8th Iberian conference on information systems and technologies (CISTI), pp 1–5. IEEE

Rummel N, Walker E, Aleven V (2016) Different futures of adaptive collaborative learning support. Int J Artif Intell Educ 26(2):784–795. https://doi.org/10.1007/s40593-016-0102-3

Samuelis L (2007) Notes on the components for intelligent tutoring systems. Acta Polytechnica Hungarica 4(2):77–85

San Pedro MO, Baker RS, Gowda SM, Heffernan NT (2013) Towards an understanding of affect and knowledge from student interaction with an intelligent tutoring system. International conference on artificial intelligence in education. Springer, New York, pp 41–50

Scagnoli NI, Choo J, Tian J (2019) Students' insights on the use of video lectures in online classes. Br J Edu Technol 50(1):399–414. https://doi.org/10.1111/bjet.12572

Schuler D, Namioka A (1993) Participatory design: principles and practices. CRC Press, New York

Schwartz T (1974) The responsive chord: how radio and TV manipulate you… Who you vote for… What you buy… and how you think, vol 126. Doubleday, New York

Self J (1998) The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. Int J Artif Intell Educ (IJAIED) 10:350–364

Shilton K (2018) Engaging values despite neutrality: challenges and approaches to values reflection during the design of internet infrastructure. Sci Technol Human Values 43(2):247–269. https://doi.org/10.1177/0162243917714869

Siepmann CA (1941) Further thoughts on radio criticism. Public Opinion Quarterly 5(2):308–312

Simonite T (2013) Search under way for gold in online education data trove. MIT Technology Review

Sonwalkar N (2013) The first adaptive MOOC: a case study on pedagogy framework and scalable cloud architecture—part I. In: MOOCs forum, vol 1. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, pp 22–29. https://doi.org/10.1089/mooc.2013.0007

Sunstein CR (2018) #Republic: divided democracy in the age of social media. Princeton University Press, Princeton

Tanaka F, Isshiki K, Takahashi F, Uekusa M, Sei R, Hayashi K (2015) Pepper learns together with children: development of an educational application. In: 2015 IEEE-RAS 15th international conference on humanoid robots (Humanoids), pp 270–275. IEEE

Tanenbaum C (2016) STEM 2026: a vision for innovation in stem education. Office of Innovation and Improvement US Department of Education, Washington, District of Columbia

Tastimur C, Karakose M, Akin E (2016) Improvement of relative accreditation methods based on data mining and artificial intelligence for higher education. In: 2016 15th international conference on information technology based higher education and training (Ithet). IEEE, New York

Thompson DF (1980) Moral responsibility of public officials: the problem of many hands. Am Polit Sci Rev 74(4):905–916. https://doi.org/10.2307/1954312

Timms MJ (2016) Letting artificial intelligence in education out of the box: educational cobots and smart classrooms. Int J Artif Intell Educ 26(2):701–712. https://doi.org/10.1007/s40593-016-0095-y

Tomlinson CA (2000) Reconcilable differences: standards-based teaching and differentiation. Educ Leadersh 58(1):6–13

UNESCO (2017) Aid to Education Is Stagnating and Not Going to Countries Most in Need (EFA Global Monitoring Report Policy Paper No. 31). In: Global education monitoring report. UNESCO, Paris

Ure A (1836) The Philosophy of Manifactures, or an exposition of the scientific, moral and commercial economy of the factory system of Great Britain by Andrew Ure. Charles Knight

VanLehn K (2011) The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educ Psychol 46(4):197–221

von Schomberg R (2013) A vision of responsible research and innovation. Responsible innovation. Wiley-Blackwell, New York, pp 51–74. https://doi.org/10.1002/9781118551424.ch3

Waddell K (2018a) Computer Science's public safety question. Axios

Waddell K (2018b) Confronting demons of the computer age. Axios

Wang N, Johnson WL (2016) Pilot study with RALL-E: robot-assisted language learning in education. Intelligent tutoring systems. Springer, Croatia, p 514

Warschauer M, Ames M (2010) Can one laptop per child save the world's poor? J Int Affairs 1:33–51

Winner L (1980) Do artifacts have politics? Daedalus 1:121–136

Woolf BP, Arroyo I, Muldner K, Burleson W, Cooper DG, Dolan R, Christopherson RM (2010) The effect of motivational learning companions on low achieving students and students with disabilities. International conference on intelligent tutoring systems. Springer, New York, pp 327–337

Woolf BP, Chad Lane H, Chaudhri VK, Kolodner JL (2013) AI grand challenges for education. AI Mag 34(4):9

Xiaohong Z (2017) Application of distance education combined with artificial intelligence. Agro Food Industry Hi-Tech 28(1):555–559