11-24-2019

Augmenting Education: Ethical Considerations for Incorporating Artificial Intelligence in Education

Dana Remian

A final project presented to the faculty of the
Instructional Design Master's Degree Program
University of Massachusetts at Boston

Augmenting Education:
Ethical Considerations for Incorporating Artificial Intelligence in Education

Submitted by
Dana Remian

In partial fulfillment for the requirement of the degree

MASTER OF EDUCATION

November 24, 2019

Dr. Carol Ann Sharicz

Approved by Dr. Carol Ann Sharicz, Faculty

Abstract

Artificial intelligence (AI) has existed in theory and practice for decades, but applications have been relatively limited in most domains. Recent developments in AI and computing have placed AI-enhanced applications in various industries and a growing number of consumer products. AI platforms and services aimed at enhancing educational outcomes and taking over administrative tasks are becoming more prevalent and appearing in more and more classrooms and offices. Conversations about the disruption and ethical concerns created by AI are occurring in many fields. The development of the technology threatens to outpace academic discussion of its utility and pitfalls in education, however. Conversations about the disruption and ethical concerns created by AI are occurring in many fields. To ensure that AI in education serves learners and educators and that ethical concerns are answered or mitigated, the field must first clarify what those concerns are. This paper surveys academic and trade literature and draws upon a parallel questionnaire deployed to define existing and emerging ethical concerns of AI in education.

Keywords: Artificial intelligence, machine learning, education, ethics

Augmenting Education: Ethical Considerations for Incorporating Artificial Intelligence in Education

Artificial intelligence (AI) has existed in theory and practice for decades. Until recent years, applications for AI could generally be described as exploratory, with limited practical relevance or utility. The earliest AI included programs that could play checkers. Arthur Samuels, based in part on the work of his contemporary Christopher Strachey, worked on a program that could apply both rote learning and generalization to the game of checkers. In 1962, after ten years of periodic expansion and improvement, the program was sophisticated enough to win a game against a human checkers champion ("Artificial Intelligence," 2019a). Competency in the game of chess was proposed as an early milestone of AI development in the 1950s, but the definitive proof of this capacity for many, the performance of IBM's Deep Blue computer against chess champion Gary Kasparov, came four decades later in 1997 (Bloomfield & Vurdubakis, 2018). The theoretical underpinnings of how such an AI chess champion could operate existed from the earliest days of AI, but the computational power required to make it work was not yet present ("Artificial Intelligence," 2019a). In the interim, AI theory and computational power made advances, and applications continued to be developed and used in practice, including in education.

Anxieties about the role of the technology in education, and discussions of the ethics involved, have occurred across these decades as well. However, with the emergence of a growing number of practical applications for AI, these anxieties have transitioned from speculative to immediate. Artificial intelligence, long the topic of both dystopian and utopian science fiction, is now pervasive in developed countries. In public, industrial, and academic spheres, talk of the impending disruption which AI may bring has assumed a new urgency.

Scholarly work devoted to examining the ethical concerns raised by AI is flourishing, especially in the computer sciences. The significance of these concerns in the domain of education and concerns which may be specific to education must be defined and explored. Thus far, private-sector research and development, along with relevant partners in academia, have been the primary steering force in artificial intelligence (Cath, Wachter, Mittelstadt, Taddeo, & Floridi, 2018). The need for interrogation of AI by other schools within the academy, and from practitioners and the public at large, must inform future development. This paper surveys academic and trade literature and incorporates responses from a parallel questionnaire to define existing and emerging ethical concerns of AI in education.

Background

Artificial Intelligence

The term “artificial intelligence” was coined by John McCarthy in the 1950s, but the idea of a machine that could replicate human behavior is significantly older (“Artificial Intelligence,” 2019a; “Artificial Intelligence,” 2019b). There is no single definition for what constitutes AI, and this is only further complicated by revisions to the definition based on continuous advancements in computer science (Hoeschl, Bueno, & Hoeschl, 2017). Artificial intelligence can be described in a broad sense as “the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings” (“Artificial Intelligence,” 2019a). A slightly more refined definition of AI can be stated as the “ability of computers to learn from data, as opposed to being explicitly programmed” (Sullivan, 2018, as cited in Heller, 2019). These two definitions, taken together, can serve as a useful basis for understanding. Like the issue of defining AI, there is no universally accepted definition or set of parameters that define intelligence (“Artificial Intelligence,” 2010a; Heller, 2019). When discussing intelligence in this context, however, it is worth incorporating the understanding that intelligence may be best understood as an aggregate of a diversity of traits, abilities, and qualities ("Artificial Intelligence," 2019a).

The creation of the digital computer allowed for the development of so-called "weak" or "narrow" AI. Weak AI can be described as "algorithms able to reproduce or supplement human intelligence in very specific areas" (de Saint Laurent, 2018). Under this definition, many existing pieces of software can be described as weak AI. "Strong" AI, what is sometimes referred to as "General AI," on the other hand, involves the creation of machine intelligence which is "capable of equaling or surpassing human intelligence in general" (de Saint Laurent, 2018; Heller, 2019). While weak AI has been a reality for decades, strong AI remains a matter for speculation and theory. Weak tasks like mastering games and mathematical skills are early examples of AI development, and recent years have seen AI overcome tasks with increasingly complex, knowledge-based problems and real-world perceptive challenges once believed to be impossible for non-human intelligence (Schalow, 2016).

The example of self-driving cars demonstrates the strides in AI in recent years, but this is still an example of weak AI in action. A system as sophisticated as a vehicle capable of analyzing an enormous amount data from a number of instruments and then using that data to respond within a set of parameters like traffic laws and safety rules may be considered impressive by today's standards and past examples, but it is still only a handful of very narrow tasks being performed and refined by a computer. Representations of AI in popular culture tend toward sensational depictions of strong AI run amuck, but even weak AI still has great potential to impact how we live our lives as it becomes less and less narrow in application (Bloomfield & Vurdubakis, 2008; Schalow, 2016).

Machine learning. Beyond the definition of weak and strong AI, a few other terms deserve explanation before examining the application of AI and the implications which flow from it. Machine learning denotes the ability of a system to learn on its own from data (de Saint Laurent, 2018; Heller, 2019). As with artificial intelligence, of which machine learning is a sub-field, machine learning dates to the 1950s, and Strachey's checkers-playing AI demonstrated machine learning by playing against itself and updating its strategy ("Artificial Intelligence," 2019a; Brynjolfsson & Mitchell, 2017). When exposing a machine learning model to a sufficient amount of quality data, including recognizable features, this data is referred to as a training sample (Mohri, Rostamizadeh, & Talwalkar, 2012). With this training sample, a successful result occurs when the model can construct generalized rules which it can then use to interpret further samples (Mohri et al., 2012).

There are a variety of models, or carriers, presently available, and a variety of approaches to exposing the system to data (Zimmerman, 2018). The approaches to providing data to a given instance of machine learning model include supervised learning, in which the model is provided training data with labeled features, unsupervised learning, in which the model receives unlabeled training data, and variations wherein the model receives updated data to emphasize desired results or reinforcement upon achieving desired results (Kose, 2018; Mohri et al., 2012). Major classes of problems to which these models can be applied include classification, regression (or prediction based on statistical evidence), ranking, and clustering (Mohri et al., 2012). One overview of this process, featured in Michelle Zimmerman's book Teaching AI, begins with gathering data, wherein a sufficient amount of quality data is gathered, and continues with preparing that data into a uniform format (Zimmerman, 2018). Next, a model is selected, and various models excel for different tasks, such as binary classification, understanding and generating text, or interpreting digital images (Zimmerman, 2018). Once a model has been selected, the aforementioned training can take place, and once the system has been trained, it must be evaluated, exposing the trained model to more data to determine whether or not it can produce the desired results without human intervention (Zimmerman, 2018). Depending on the results of the evaluation, the model can then be improved by tuning its parameters, and a determination can be made regarding the overall success of the model (Zimmerman, 2018).

Neural networks and deep learning. Where these machine learning models seek inspiration from animal neurology, the term neural network becomes essential to note. A neural network operates by processing information through many interconnected processing nodes (“Artificial Intelligence,” 2019b). These neural networks have allowed for more complex applications, such as handwriting interpretation and financial predictions (“Neural Network,” 2019). Neural networks are, in turn, necessary for defining one further, related term. Deep learning is a subset of machine learning in which several layers of neural networks are applied toward tasks (Zimmerman, 2018). Deep learning can be used for even more complex learning tasks and outcomes, but are larger, slower, and more challenging to understand as a result (Zimmerman, 2018). Neural networks are responsible for many recent developments in reproducing human-level capabilities in computers, but since most of the information processing takes place in the network, not in the human-created program itself, explaining this “hidden” layer of processing is difficult (“Neural Network,” 2019). Where these layered networks demonstrate expertise in a given domain or task, they are referred to as “expert systems” or “knowledge-based systems” in the context of a model (Arruda, 2017; Crowe, Lapierre, & Kebritchi, 2017; McArthur, Lewis, & Bishary, 2005). IBM’s Watson platform operates in a multitude of disciplines by setting its numerous expert systems to work in coordination (Crowe et al., 2017) (See *Figure 1*).

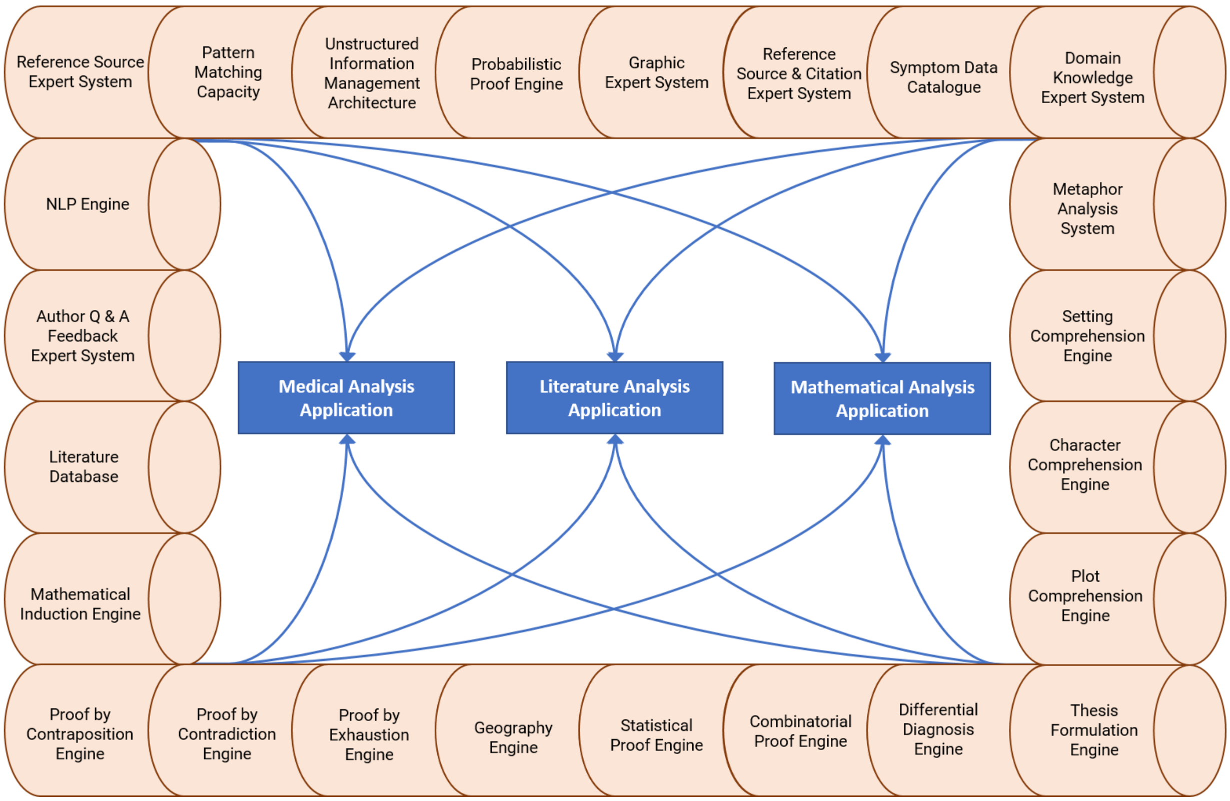

Figure 1. A diagram mapping IBM Watson expert systems (adapted from Crowe et al., 2017).

Natural language processing & sensory inputs. One of the areas of development further enabled by neural networks and deep learning is in natural language processing. Natural language processing indicates methods whereby a computer can understand spoken or written language for its meaning, not just the mechanical rules (“Artificial Intelligence,” 2019a). Human beings use language full of idioms, symbols, and tone requiring interpretation. The capacity to interpret language beyond its mechanical rules becomes necessary for AI to interact with humans outside of highly regimented settings. Natural language processing is key for expanding the potential for AI and human interaction and was foundational to allowing IBM’s Watson platform to defeat human players on the television game show Jeopardy, a game where the interpretation of wordplay and hints, not just the rote recall of facts, is essential (Arruda, 2017; Crowe et al., 2017). Utilizing natural language processing to interact with subsequent knowledge-based systems, platforms like Watson can perform in a responsive manner to interpret and respond to input calling on a multitude of domains of knowledge. Further, this language processing allows for enhanced translation, allowing for inputs and outputs in a variety of languages, even in real-time (Zimmerman, 2018).

Beyond the processing of written or auditory language, methods of processing images, biometrics, and other perceptive sources of data allow for more significant applications. Machine learning has enabled computers to recognize objects and features in images, allowing for some of the advances related to self-driving vehicles, but also allowing for user discovery and investigation ("Artificial Intelligence," 2019a; Kompella, 2018; Zimmerman, 2018). Applications available on mobile devices can tap into this potential to identify objects in real-time (Vanitha, Jeeva, & Shriman, 2019). Facial recognition and eye-gaze systems can provide additional data for applications to learn from, as well as open the door to new forms of interaction (Zimmerman, 2018). These perceptive tasks, formerly outside of the capabilities of machines, are now viable ("Artificial Intelligence," 2019a).

Current applications of AI in Ed

Artificial intelligence in educational software is not a new development, but the new capacities of AI are allowing for more meaningful applications. Objective and rigid subject matter, like mathematics and computer sciences, are well-suited for computer-based education, and their compatibility with a "drill-and-practice" format makes them well suited for classroom or homework activities where learners may have access to a computer (McArthur et al., 2005).

As the capacity and variety of AI expert systems increases, and the natural language processing capabilities of AI expands, new possibilities for AI in education will continue to emerge. With a multitude of commercially available AI-enhanced software and service options, organizational implementation of AI has grown by 270 percent in the past four years, and AI adoption in education in the United States is anticipated to grow by nearly 48 percent in the next three years (Bonderud, 2019). The applications for AI in education are interlinked, as well. Although several different applications can be discussed, they may be used in concert and share data. The potential benefits and risks of AI must consider the aggregate ability and power of these systems, not only the individual.

ITS and ILE. Intelligent Tutoring Systems (ITS) and Intelligent Learning Environments (ILE) are among the primary expressions of AI in education, with growing adoption in a variety of institutional learning environments. Intelligent tutoring systems like WEST and SOPHIE emerged beginning in the 1970s, allowing learners to be exposed to different content according to their behavior and achievement (McArthur et al., 2005). An ITS replicates the role of a human tutor, monitoring learner progress and adjusting the focus and difficulty of instructional content based on the learner's interest and aptitude, as well as providing timely information and guidance. Intelligent learning environments operate in a similar manner, by incorporating aspects of ITS, and adaptive educational hypermedia systems (AEH), which deliver interactive lessons in multiple mediums and in an adaptive manner (Herder, Sosnovsky & Dimitrova, 2017). Some ILEs are more open-ended and exploratory, as in microworlds, wherein learners can approach given materials in a Constructivist manner and test their ideas in a trial-and-error fashion (du Boulay, 2019; McArthur et al., 2005).

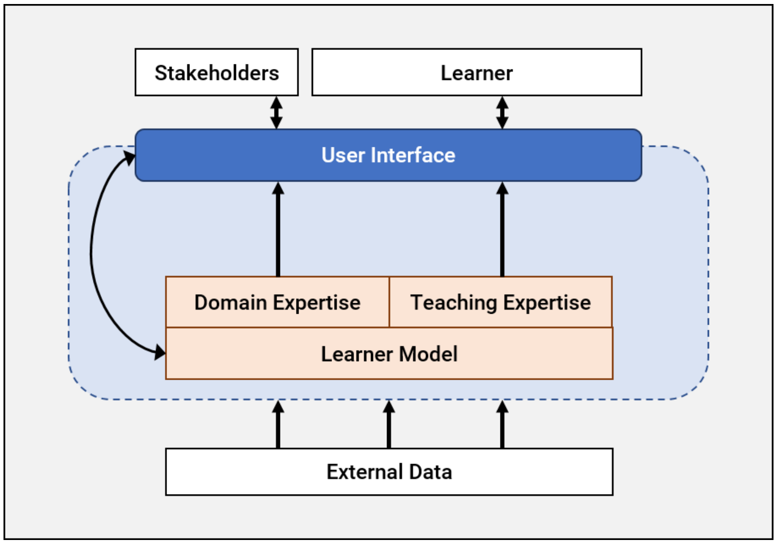

Figure 2. Data transfer in an ITS. Adapted from Kay (Kay, 2012).

These systems develop models of learners on an ongoing basis, including the modeling of learner knowledge, cognition, and metacognition, learning about them as they learn from the system (Herder et al., 2017). This learner model is used in conjunction with domain expertise and teaching expertise derived from the platform's models and external data to deliver content to the learner via the user interface (Kay, 2012). The user interface provides a variety of platform and domain-appropriate learning resources for the learner according to the system's model for that learner (Kay, 2012). Additionally, the user interface can deliver feedback about the learner model it has developed for stakeholders, such as parents or teachers, as well as for student awareness and reflection (du Boulay, 2019; Kay, 2012) (See figure 2). Social, emotional, motivational, and environmental models are additional aspects proposed for inclusion in AI learning platforms but are not yet common (Desmarais & Baker, 2012; Herder et al., 2017). While the efficacy of these systems cannot be overly generalized, a significant number of studies have determined that ITS and ILE can produce measurable improvement in learning outcomes, and even improvements beyond those anticipated from interaction with human tutors in some cases (du Boulay, 2019; McArthur et al., 2005; McCarthy, Rosenblum, Johnson, Dittel, &

Kearns, 2016; Krueger, 2017; Pierce & Hathaway, 2018; Wijekumar, Meyer, & Lei, 2012; Xu, Wijeukumar, Ramirez, Hu, & Irey, 2019).

Intelligent tutoring systems and ILE often fit into existing classroom and distance learning practices but may also reshape classroom practices in the future. Existing applications can take the place of traditional at-desk work and homework, as well as forming the basis for a flipped-learning approach in some cases (du Boulay, 2019; Krueger, 2017). Rather than practicing math on a worksheet, for example, a student might interact with an AI-augmented classroom application. Rather than opening a physical textbook, students can open an interactive and adaptive multimedia version of course content. Within some of these uses, there is also the ability for the instructor to track individual and group progress. In the case of one Arizona State University course on biology for nonmajors, adaptive courseware allowed students to review textbook and multimedia content, take a quiz, and receive feedback and review guidance; in turn, the courseware also informed the instructor of progress and allowed for updated in-class approaches and considerations (McMurtrie, 2018). One proposed use for these systems for use in classroom settings is to enable teachers to customize not only their lessons but also how they use their time and whom they need to devote attention to (du Boulay, 2019; Kokku, 2018). Fitting into the grander scheme of classroom education, there has been some suggestion that these systems might be fit with classroom sensors capable of observing student behavior and biometrics to create an entirely integrated “Smart classroom” (Kim, Soyata, & Behnagh, 2018; Timms, 2016).

Chatbots and assistants. Chatbots, also known as conversational agents or dialogue systems, are interfaces that allow people to have a conversation, simulated or otherwise, with a machine (Henderson et al., 2018). Whereas early chatbots typically operated via typed text and featured limited response capabilities, new virtual assistants, such as those who populate many smartphones and other "smart" home devices, now integrate voice input and output as well and can draw upon machine learning and AI systems for increasingly responsive and accurate replies and inquiries. In education, as in banking and other customer-service industries, one of the apparent applications is in fielding questions from students or parents. Georgia State University has employed a text-based chatbot called Pounce to follow students through admissions and registration, financial aid, and other administrative tasks, answering common student questions (Bendici, 2018). The use of Pounce at Georgia State was credited with a 21 percent reduction in accepted freshman students failing to finish matriculation while allowing university personnel to focus on more involved problems and other tasks (Bendici, 2018). At Winston-Salem State University, the university's administrative chatbot Winston nudges students to complete required paperwork and tasks with far greater efficacy than traditional mail or email prompts (Bendici, 2018).

Chatbots have been implemented in classrooms and online learning environments, as well. A notable example is that of Jill Watson, an AI chatbot powered by IBM's Watson platform and designed for use as a teaching assistant to support Professor Ashok K. Goel's online computer science class at the Georgia Institute of Technology (Gose, 2016). Chatbots, in this sense, can serve as valuable classroom aides, responding to discussion board posts and emails regardless of the availability of a human representative. Because of this, this form of the chatbot has been advanced as potential improvement in the practice of delivering massive open online courses (MOOCs) as it could mitigate the absence of feedback and guidance in the format as presently employed (Gose, 2016). Even in primary education with a teacher present, virtual assistants have been promoted for their ability to encourage exploratory learning (Krueger, 2017). Likewise, while a chatbot could be a source of information, other uses include the promotion of interaction between students by linking discussions or by prompting intermediate learners to share their learning with beginners (Gose, 2016; McMurtrie, 2018). In one example, a bot named Question has been used to track learner questions and flag them for response by a peer-tutor, or to direct the learner to available lecture transcripts using natural language processing (Zimmerman, 2018).

Grading and feedback. One of the roles often taken by classroom assistants is in relieving teachers of time-consuming basic tasks, such as grading. Multiple-choice questions have long been favored due to their ability to be quickly graded by a human or a machine, but multiple-choice questions can be more difficult and time-consuming to construct while being less suited to higher-order learning assessment than other formats. Software powered by AI has been designed to alleviate some of this tension. A tool called Gradescope has been employed by over 550 colleges in order to streamline and improve assessment and grading (Blumenstyk, 2018). By marking correct answers and flagging incorrect answers, the tool saves instructors from most of the work involved in manual grading while affording them more time to scrutinize where students are struggling, including by highlighting patterns that may indicate the need for an improved assessment (Blumenstyk, 2018). Other tools allow for qualitative analysis of writing, and natural language processing can enable the automation of grading far beyond multiple-choice questions. The ability to assign essay questions in large lecture courses and MOOC environments has been restricted by the time and effort required to grade essay responses, but AI tools can now facilitate their grading or crowdsource essay review by aggregating peer reviews (Schauffhauser, 2019b; McMurtrie, 2018). One example of these concepts at work is that of engineering lecturer Daniel Kellerman at the University of New South Wales, who has used machine learning to facilitate grading for himself and his human assistants (Zimmerman, 2018). The AI systems he has worked with allow for the grading of materials for his 500-student courses, including hand-written test responses and thousands of pages of essays (Zimmerman, 2018). If the system finds incorrect or unclear answers, it flags those items and provides them to Zimmerman and his assistants by category, allowing for easier grading but also improved detection of poor assessment items (Zimmerman, 2018).

Student monitoring and support. Artificial intelligence can also augment teachers' ability to monitor and support their students. This includes tools like GoGuardian, which can replace traditional keyword-based and URL-level security monitoring software in school settings (Pierce & Hathaway, 2018). Using AI to analyze page content and learner text entry within their contexts, systems like this can alert administrators to students who might be accessing inappropriate materials or flagged behaviors, like content which might be indicators of suicidality (Friedman, 2019b; Pierce & Hathaway, 2018). The increased monitoring and analysis of student data can also increase the ability of teachers and administrators to discover when and how students might need additional support. Using predictive analytics reported by software, teachers may be able to more easily identify at-risk students and intervene to ensure that they get the supports that they need (Herodotou, Rienties, Boroowa, Zdrahal, & Hlosta, 2019). The University of Arizona is using data derived from student ID-use and other activities in order to predict which students may be at risk of dropping out with 73% accuracy on the day they start classes, and with increasing accuracy as time passes (Gilliard, 2018).

Further, learner skill, behavior, and attitudes can be measured and used to suggest approaches for maximized engagement and outcomes (Desmarais & Baker, 2012). McGraw-Hill Education's ALEKS platform has integrated a feature called Insight, which will do just that, by alerting teachers to a variety of issues related to student behavior in the ALEKS learning environment (Schaffhauser, 2019a). Beyond basic monitoring for compliance and retention, these systems may have potential for flagging students who demonstrate characteristics of undiagnosed learning, developmental, or mental health disorders (Friedman, 2019a; Friedman, 2019b; Lee, Maenner, & Hellig, 2019; Schaffhauser, 2019a). As predictive models are improved, potential applications in surveillance and predictive policing emerge (Asaro, 2019).

Accessibility. Because AI can work with input and out across multiple modes and in multiple languages, AI could transform accessibility in the classroom. The potential for language learners to receive instructional material in their primary language and their new language may increase the learner's success with both languages and the topic of instruction (Laere & van Braak, 2017). Live captioning with multi-language support and text-to-speech can also allow for increased classroom integration and access to learning materials for students with a variety of learning needs (Bonderud, 2019; Laere & van Braak, 2017). Because AI can work with multiple sensors and input tools, this may also reduce the barriers between learners and learning materials and activities. Work has been done to create an ITS to teach braille to educators and visually-impaired learners, and adaptable output can be delivered via braille display with input through a braille keyboard (McCarthy et al., 2016).

Non-classroom applications. Education is not the only field grappling with the inclusion of AI. Artificial intelligence has been used to create tools for medicine, for law, and for general business use. Customer service and relations and general communication applications have the potential for wide use across industries, including by administrators of education. Likewise, general productivity tools and modeling of individual job roles could allow for greater efficiency and reduced losses during absences or due to job turnover (Heller, 2019). Tools created to support customers and professionals are constructing and filling basic forms and documents according to user needs (Heller, 2019; Medianik, 2018). Companies like HireVue are using AI to screen job applicants, and it is conceivable that interactions with AI could serve as the first interview for professionals and applying students alike (Burke, 2019). In the legal profession, a Watson-based platform named ROSS is serving as a research assistant, combing through case law and statutes using natural language processing to deliver superior results to traditional Boolean searches (Medianik, 2018; Nunez, 2017). Such a platform could be tremendously useful for librarians, researchers, educators, and students in all fields, and Microsoft is offerings its semantic Academic Search engine to offer similar services (Ojala, 2018).

Methods

This paper applied a mixed-method approach. A survey of academic and trade literature on a qualitative basis to establish identified ethical concerns in AI relevant to education was applied. A non-experimental questionnaire was then deployed to explore attitudes toward the identified concerns while recognizing additional concerns not defined by the initial literature review. Further literature review was then completed to support discussion.

Literature Review

An earlier literature review conducted by the author to explore the topic of bias in AI discovered that existing scholarly literature examining the ethics of AI in education is relatively small in volume. As such, the review conducted for this effort was broadened to include trade publications, conference papers, and literature from other fields. Inclusion of works from the fields of AI and machine learning was of significant benefit, and several works from the fields of law and medicine were useful. A significant amount of overlap occurs in the discussion of ethical concerns in AI, and a number of terms are used interchangeably, reducing the feasibility of measuring concerns by quantifying keywords. By expanding the focus and easing the rigor of the review, and applying some interpretation and synthesis, seventeen terms and common concerns were able to be identified in order to construct the questionnaire. See figure 3.

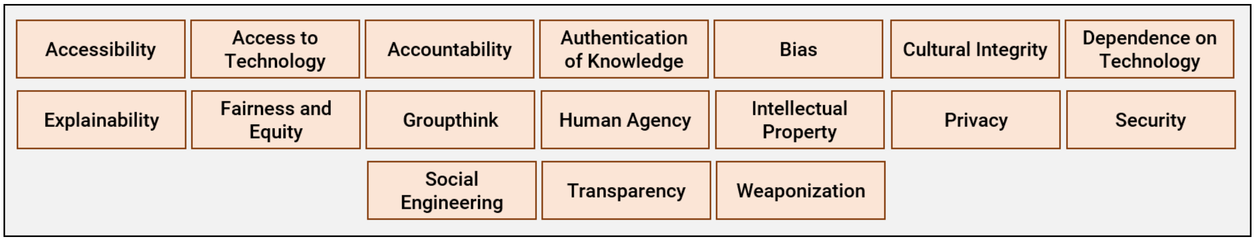

Figure 3. Seventeen initially defined ethical concerns for education in AI.

Questionnaire

Participants. The questionnaire was targeted at two groups of people, those with a background in education or a related field, and those with a background in AI or a related field. The questionnaire and script were posted to several professional communities on social media platforms. The questionnaire received 42 responses, of which 41 were valid. One data set was deleted as the participant provided false answers and explained that they had done so in order to provide feedback on the questionnaire itself. Of the 41 valid participants, 39 described themselves as having a background in education while 9 described themselves as having a background in AI. Only one participant described themselves as having a background in AI without also having a background in education, and one participant aligned with neither field. All participants did so on a voluntary basis, and no compensation was offered or rendered.

Materials. The questionnaire constructed for this survey included seventeen items to be rated along a four-point rating scale as well as an open-ended question. Each of the seventeen items provided a term or phrase representing the concerns identified during the initial literature review. Due to the significant overlap and interchangeability of some terms, each of these items was accompanied by a brief description to ensure consistency internal to this survey. The rating scale for these items included “Not of Concern,” “Of Minor Concern,” “Of Moderate Concern,” and “Of Significant Concern.” No option to state uncertainty or lack of opinion was offered to avoid participants defaulting to that option rather than applying judgment (Krosnick et al., 2002). The questionnaire also included three initial questions to allow users to align them with having an education background, an AI background, and to provide a job title. The background questions were designed for sorting purposes, not elimination purposes. The job title field and the open-ended question were optional responses, but the other questions used were all required in order to submit the form. The questionnaire was hosted and distributed on the internet. (See Appendix A)

Procedure. After opening the questionnaire link, participants were informed of the overall aim of the questionnaire and invited to participate further. Respondents were advised that all responses would be considered, but that AI and education backgrounds were sought, as indicated by the first two questions. After choosing whether to include a job title, participants proceeded onto the seventeen rating items, each covered by the same initial prompt asking them how concerning they believed the item was regarding the incorporation of AI in education. After these items, the final open-ended question requested any additional issues or concerns that the participant felt should be considered in a discussion of the use of AI by educational institutions. Again, this final question was optional. Upon pressing the submit button, participants were thanked for their answers.

Results. After the close of the questionnaire, all results were gathered, and a scoring scheme was applied to the rating scale questions, with scores of O, $.\overline{3}$, $.\overline{6}$, and 1 assigned to the four possible ratings. Based on this scoring, a response of 0 would indicate no concern expressed for the inclusion of AI in education, and 1 would express most significant concern. After application of this scoring, the mean average for the seventeen items fell between .51 and .85 when considering all responses. When averaged across all items, a score of .73 was found. Privacy, accountability, and bias scored the highest of the seventeen items, and dependence on technology, intellectual property, and groupthink scored the lowest (see appendix B). These results allow for emphasis in the subsequent discussion, though further research with larger and more diverse sampling is necessary to draw more meaningful comparisons.

For the purpose of discussion, these results were further broken down according to those with only an education background, and those with an AI background, with or without education backgrounds. The relatively small sample size and the low response from participants with backgrounds in AI limits comparison. One participant claimed both education and AI backgrounds and provided data which was notably different from all other respondents, providing ratings of “Not of Concern” on fourteen of the seventeen items. Because of the low number of participants with AI backgrounds, this outlier data would have an inordinate impact on mean averages for that group and was included in the total results, but otherwise isolated.

The education group, comprised of thirty-one participants, largely steered the overall rating, and their average score on each item was very close to the overall score averages. The AI group was comprised of eight participants and showed a few noteworthy differences. The average of all items for the education group was .72, while the AI group average across all items was .81. The AI group rated all items as more concerning than the education group with the exception of intellectual property and dependence on technology, which were rated lower by .07 and .05, respectively. Several of the other items were given similar marginally higher scores. Eight items were rated as being more concerning by .1 or above by the AI group, however. Most notably among these was groupthink. Groupthink was rated at .44 by the education group, marking it as of least concern among all items, while the AI group rated groupthink at .71, a difference of .27. Although groupthink was given more attention by the AI group, it is worth noting that it still placed in fifteenth place of the seventeen items overall. The ranking of items for both groups remains similar, though the low number of data sets in the AI group caused several tied placements. Other items given higher scores by the AI group included accessibility, weaponization, and transparency.

The responses also included seventeen replies to the open-ended question. Most notably, responses included two answers raising concerns about the impact on child development were received, identifying an important consideration specific for AI in education. Multiple other responses included comments aimed at budgetary concerns or profit receiving priority over the needs of learners. Concerns regarding automation displacing human labor and eventual self-awareness of AI were also voiced.

Discussion

The rapid growth of computing technologies over the past several decades has frequently left the examination of ethical use and regulation to play catch-up (Zimmerman, 2018). Analysis of the issues created or amplified by AI is occurring across multiple disciplines, and government and professional organizations have stepped forward to steer this conversation. The Association for the Advancement of Artificial Intelligence (AAAI), the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery (ACM), and the International

Organization for Standardization (ISO) have each advanced standards and principles to support ethical development and implementation of AI and data analytics (Stahl & Wright, 2018). With AI proliferating in all manner of industries and consumer goods, the disruption it may bring must be discussed and prepared for.

From the lack of scholarly discussion regarding AI in education, it appears that the development of technology appears to be outpacing thoughtful consideration of what is most effective or most ethical for the future of education. Artificial intelligence for education may be viewed as yet another piece of educational technology, but its potential for disruption must be accounted for. As AI becomes commonplace, discussions must occur in the academy and in classrooms at all levels in order to prepare for and examine the place of AI in education and in our lives. The following discussion of issues, ordered according to the results of the previously detailed questionnaire, aims to highlight their importance for educators, students, and educational institutions. Whittlestone, Nyrup, Alexandrova, and Cave (2019) suggest that any discussion of the ethical issues in AI should focus on the tensions between these issues and principles which emerge to address them, and this discussion will draw upon some of those tensions as well as the reinforcing connections between issues.

Privacy

Privacy is possibly the biggest single focal point of ethics in AI (Stahl & Wright, 2018). Ensuring that sensitive personal data is protected is a priority of existing legislation and standards, to ensure that privacy is included in AI by design but also backed by law (Stahl & Wright, 2018). With numerous disclosures of the extent of data access and breaches by technology companies, this issue is very much in the public eye and zeitgeist. Basic access to many services online requires granting access to potentially sensitive data. Perhaps because of this, there exists a perception that younger people are more willing to give up their privacy, and yet there is some evidence to suggest that young people care more about privacy than earlier generations (Blank, Bolsover, & Dubois, 2014).

Data is a valuable resource economically, however, and adequate data access is a functional requirement for developing and improving AI (Gilliard, 2018). One of the clearest tensions in AI ethics is that between privacy and efficacy. Machine learning allows for the processing of enormous amounts of data, and it is this access to data which allows for the development of AI models with relevant applications (Whittlestone et al., 2019; Zimmerman, 2018). Giving up some degree of data privacy comes with the promise of potential increases in efficacy and accuracy in AI (Zimmerman, 2018). This is especially clear in examples of healthcare diagnostics. Allowing sensitive diagnostic data to be used for training AI models for diagnosis and treatment comes with clear benefits, but privacy must first be waived.

Presently, consent is held up as a protective measure by many organizations that collect personal data, especially for healthcare (Dare, 2019). However, the efficacy of consent is doubtable. As Jones, Kauffman, and Edenberg (2018) point out, “whether intentional or not, the complex jargon and lack of transparency in many privacy policies obfuscate a user’s choices” (p. 68). Informed consent is an important tool, but the minimal hurdles faced by users of services, typically a checkbox to indicate having acknowledged a multiple-page terms and conditions document behind a separate link online, or a signature on a form at a doctor’s office, do not constitute true informed consent. Indeed, the overwhelming amount of things which many individuals consent to may contribute to a consenting by default because of “consent desensitization” (Schermer, Custers, and van der Hof, 2014, quoted in Jones et al., 2018). Rendering consent moot and automatic turns an important tool into an avenue for undermining human agency. If consenting to privacy is required for accessing education, coercion invalidates consent further. This is further exacerbated, along with concerns about surveillance, given the likelihood that AI systems used in public education might find their data subject to federal oversight and scrutiny.

Concerns about privacy may also be symptomatic of underlying concerns about bias, accountability, security, surveillance, and transparency as well, and these issues are intricately interwoven. An example given by Dare (2019) is that of sexual orientation, whereby privacy is not the true concern, but rather avoiding exposure to discrimination due to cultural values regarding sexual orientation; this privacy is not truly beneficial to the individual in this case and can interfere with the delivery of relevant healthcare services. Because data does have economic value, anxieties regarding the ownership and use of data become a matter of fairness and transparency as much as a matter of privacy, especially in education. In the case of Arizona State University collecting and utilizing data about student behavior and movement, it is somewhat understandable that privacy concerns might be expressed. However, by anonymizing the data and reporting predictive outcomes only to the students' advisers, are privacy concerns satisfied? (Gilliard, 2018). If the interconnected concerns are addressed and protections against unfair use of data are in place, perhaps privacy will become less of a focal point. Dare (2019) goes on to say, "We have come to think of that secrecy as normal and important, but it is not clear we are right" (p. 6). On the other hand, whereas AI has been used for policing, and even predictive policing, can learners feel secure if they feel surveilled? Especially for marginalized groups, data analysis could further embed the power structures and threatening surveillance that Foucault spoke of in discussing his conception of the panopticon (King, 2013).

Accountability

This discussion is considering accountability in terms of recognizing errors, failure, or harm. It is tempting and too easy to blame the AI itself, but it is worth noting that the AI is no more to blame than any other piece of software. Because people falsely believe that AI is intelligent and autonomous, this tendency is somewhat understandable, but it leads to "deresponsibilisation" of developers and users alike (de Saint Laurent, 2018, p. 742). As Allen, Wallach, and Smit (2006) point out, "an autonomous system that ignorantly causes harm might not be morally blameworthy, anymore than a toaster that catches fire can itself be blamed" (p. 13), but that lack of blame does nothing to correct errors, failure, or harm. Even in this explanation, there may be confusion and anxiety as interpretation of what autonomy means for machines varies. Autonomous machines can make decisions on their own but are presently not invested with the free will or self-interest that an intelligent animal demonstrates (Johnson & Verdicchio, 2017).

In the estimation of de Saint Laurent (2018), it is the responsibility of those who develop the AI and those who elect to deploy it who must be held responsible. This is all further complicated by the present inability of AI to explain itself or be interpreted. The risk here is that AI may produce harm for learners or educational institutions, but the ultimate reasoning or trigger for the harmful inaction or action may remain unknown, leaving the harmed party to suffer while no correction or recourse is available. International standards, such as those produced by the ACM and ISO, do include measures to improve accountability, but concerns about accountability cannot be fully addressed with self-policing by companies who have a financial stake in matters (Katyal, 2019). The tension between accountability and financial interests is present in most aspects of running an institution, and AI is not immune to these concerns. One of the primary concerns regarding accountability, in this paper and in the literature, is evidence of bias in AI applications.

Bias

For the purpose of this discussion, bias created in or expressed by algorithms is the focus. Bias is a natural feature of machine learning and, indeed, of human understanding. As stated by Professor Jan Willem de Graff (2019), "data is always a limited 'reflection' (measurement) of a limited part of reality. In essence, it's always biased" (p. 18). The predictive powers created in a machine learning model are sensitive to human biases and biases introduced by data sets and can express them in their outputs. Indeed, these biases may appear even when the creator of the model did not intend, or was not even aware of, the bias (Stahl & Wright, 2018). A simple example can be found in recognizing that Google Translate may assign gender along stereotypical roles when translating gender-neutral language (Miller, Katz, & Gans, 2018). This is not a qualitative decision made by Google Translate, but rather a quantitative measure. The data shows that language is most commonly used in this manner, so the model applies it accordingly.

The same basic dynamic explains how an AI application trained to help pre-screen candidates for an engineering position might favor male applicants in the case that it was trained predominantly with desirable examples that were male. In cases like this, or in policing and criminal justice, concerns about bias carry significant implications. Even when these models are trained in the absence of explicit reference to gender, ethnicity, or other legally protected classes, it may develop bias based on patterns and traits secondary to the absent label. These secondary attributes can be referred to as proxy attributes (Katyal, 2019; Kim, 2018). When a trait primarily associated with one group of people, such as extracurricular activities associated more with one gender than another, postal codes for areas with higher residency among certain socioeconomic groups or ethnicity, or risk factors for truancy based on historic data overrepresenting one group, then that trait may cause continued bias to appear. Algorithmic bias present in existing systems, such as advertisements that target men over women for high-paying job postings or crime-prediction software emphasizing certain minor offenses for further attention by police (Casacuberta & Guersenzvaig, 2019). When discussing the use of algorithmic decision-making systems in public services, one service provider suggested pointedly:

> My neighbour might be shooting up heroin and their six year old is out in the street, but they have private insurance so their records aren't part of this system. The computer tool is only capturing people who have to use public health so there's a bias to poorer people in the system. (Brown, Chouldechova, Putnam-Hornstein, Tobin, & Vaithianathan, 2019, p. 8)

Bias can be adjusted for in training AI, but first it must be recognized and accounted for. Since AI is unable to explain why it makes the choices it does presently, this vigilance and planning are crucial from the outset before harm is created. Miller and colleagues (2018) recommend several strategies for countering bias, including awareness training for parties involved in training models, inclusion, oversight, and mechanisms for airing grievances. Various methods can be used to correct bias with additional training and reinforcement of desirable model outcomes (Kose, 2018). However, since bias can come from human beings themselves, awareness training and inclusion may not be enough to avoid all bias. Algorithms designed to respond adaptively to humans, including AI, are composed of "opinions embedded in code" (Mann & O'Neil, 2016 quoted in Raub, 2018). More work must be done to understand implicit bias in human beings before we can more fully address how it works its way into AI (Katyal, 2019).

Security

Concerns about privacy, social engineering, and weaponization are further amplified when taken into consideration with concerns about security. A multitude of data breaches and security culture lapses have led to headline news in recent years and is both a potential tool to improve security, a tool to thwart security, and a vector for security lapses. While AI can be trained to augment cybersecurity practices and improve defensive capabilities, it can also be used in an offensive manner to augment efforts to bypass security (Darraj, Sample, & Justice, 2019). The development of tools to serve security efforts are in competition with the development of tools to thwart those efforts and vice versa. The importance of this dynamic is clear given the volume of data that AI is exposed to. Existing legal protections cover sensitive data in many countries, but additional proactive measures are needed.

Further concern can be directed at undermining or subverting AI itself. Where AI is used to serve a variety of purposes, including operational functions, AI itself may be targeted for disruption by actors seeking to damage or subvert an organization's efforts. In addition to crippling systems, the possibility exists to subvert systems by providing "adversarial examples" (Kose, 2018, p. 189). Adversarial data is data used to challenge the existing AI model and can be used for productive ends, but it may also be used to confuse or redirect the model by threat actors (Kose, 2018). If threat actors can gain significant access to models responsible for education, misinformation, and manipulation of social dynamics are possible consequences.

Fairness and Equity

The issue of fairness and equity shares a close relationship with bias, accessibility, and access to the necessary technology. A system that demonstrates significant bias or which cannot be accessed by many cannot be equitable, and so those discussion items contribute to this issue. Additionally, one of the reoccurring themes in the final questionnaire item was participants sharing concern about economic interests being emphasized over the needs of learners. Should AI education prove effective and financially advantageous, one concern expressed by experts is that disadvantaged school systems may become overly reliant on AI at the cost of reduced human teacher presence (Santry, 2018). While the potential for AI applications in education seems significant if this potential only benefits some learners and not others, the risk of "school apartheid" is real (Santry, 2018). In present capacity especially, the inclusion of sustained teacher-student and student-student interaction in approaches to education remains necessary for positive outcomes (Zhang, Zhang, Zou, & Huang, 2018). If AI is granted a role in any functions which involve the discipline, welfare, or rights of learners, ensuring fairness must be considered. In discussing the tensions they foresee in AI development and application, Whittlestone and colleagues (2019) suggest that the technical accuracy of AI systems may run counter to what is fair or equal.

Authentication of Knowledge

Authenticating the knowledge and predictions of AI becomes more important when AI is used for education since the further spread of inaccurate or outdated content could defy educational goals and further reinforce false information. Outside actors could produce adversarial data to contaminate AI models (Darraj et al., 2019). Considering bias and the potential for disruption of AI models, their accuracy can be called into question, but ensuring that correct and contemporary information is represented in AI is universally important.

When still-forming predictive models are overly depended upon, false beliefs can become action. Criticism of using facial recognition and analysis for predictive purposes, in the absence of explainability, shows the danger of this dependence (Pasquale, 2018). Factual data must be verified, and it can be. Subjective matters, like the grading of essays, are more difficult to judge and verify. With AI writing evaluation employed heavily in China in recent years, educators and AI specialists worldwide can take note of the success in supporting learners in writing well, but the failures of such systems to judge the logic and coherence of learner essays should also be noted (Lu, 2019).

Transparency

The word transparency is used in multiple ways in the literature, especially in relation to the important and interrelated issue of explainability. As a distinct concern, this paper is focused on transparency regarding institutions disclosing when and how AI is used. Without this sort of transparency, accountability cannot be found, and privacy is at risk. There is the basic consideration of whether or not learners need to be alerted to the fact that they are interacting with AI and not a person, as in the case of Professor Goel's Jill Watson secretly serving in the capacity of teaching assistant (Eicher, Polepeddi, & Goel, 2018). Beyond that, there are questions about student rights to their own data, how that data is shared, and the predictions made using that data (Blumenstyk, 2016). With data being described as having a value equivalent to oil, monetization and sharing of data by educational institutions may become further incentivized at the expense of privacy (Gilliard, 2018).

Transparency and fairness can come with increased tension regarding accountability and what is financially expedient, so stakeholders may need to insist on pushing this issue. Institutions may find it useful to increase transparency; however, as Blumenstyk (2016) describes, student opinion about data collection at the University of Maryland University College shifted when they were better informed about how data would be used. To return to the topic of consent, giving individuals the information they need to make a decision is what grants consent in what Jones and colleagues (2018) call “morally transformative power” (p. 65) and authority.

Social Engineering

Artificial intelligence has the potential to be used for the spread of disinformation, propaganda, and the shaping of social and cultural values. As with security, AI may play a role in both delivering and protecting from misinformation and attempts to manipulate. Darraj and colleagues (2019) cite misinformation campaigns as one of the key AI-related cybersecurity issues that must be monitored. Artificial intelligence systems have been employed to flood social media with content designed to disrupt, but AI systems can also be employed to stop and remove these threats (Strickland, 2018). One of the most noteworthy uses of AI for its potential misuse is in the creation of deepfakes, utilizing AI video tools to swap the face of one person for that of another (Güera & Delp, 2018). Whittlestone and colleagues (2019) suggest that one of the key tensions at play in developing AI further is between promoting highly individual and tailored experiences and promoting mutual solidarity and citizenship.

Education itself, of course, plays a role in the shaping of social and cultural values so the threat of disruption to those processes is more direct. Not all of the threat comes from outside actors. Bias and social norms may come into play. Introducing tools to aid in surveilling student behavior risks normalizing omnipresent surveillance and social control (Bali, 2017). While the shaping role of education is normalized and accepted, changes brought with AI could lead to unexpected results, and the absence of explainability may leave educators unable to detect or identify these results until they have already occurred.

Weaponization

Through discussing security, privacy, and social engineering, is the threat of weaponization of AI. Greater debates about the weaponization of AI are largely focused on the use of AI for combat purposes, but in the context of education, this paper considered weaponization as the potential for expert systems and data sets to be used against human beings. Offensive AI actors, as described in previous sections, are examples of such weaponization. Education may be of less concern than some other sectors, but whereas educational institutions have access to large amounts of sensitive data, they may become a target, and that data may become a vector for further compromise. Even systems designed for defensive purposes could be subverted for offensive uses (Darraj et al., 2019).

Accessibility

Through approaches like Universal Design for Learning, AI could support greater accessibility than current methods of classroom and digital education. As discussed, AI can support the same sorts of multimodal education endorsed by Universal Design for Learning, as described by Meyer, Rose, & Gordon (2014). Accessibility may be threatened by inaction or expediency in AI, as well as overreliance on specific modes of delivery, however. The potential for accessibility in AI is limited by the necessity for awareness and intentional inclusion of accessibility practices in designing platforms for education, and avoiding bias introduced by learners with disabilities falling outside of statistical norms or primary methods of biometric data collection.

Cultural Integrity

Cultural integrity, as discussed here, is both in line with and in tension with social engineering. Much of the discussion of bias and social engineering is relevant here, as we discuss the disruption of culture, especially for marginalized groups. Artificial intelligence may play a disruptive role in dominant culture centers, such as the United States and China, but it may also serve as a vector for furthering the cultural influence of those dominant cultures over the subcultures and external cultures they interact with. Zbrzebnicki (2017) suggests that, even in discussing and defining ethics for AI, AI may serve as a colonial agent for influence and spreading of norms by the countries producing AI. Much as the forces of mass media and capital exert neocolonial influence, AI may become another way in which culture is undermined. As with privacy, Foucauldian analysis of power dynamics, normalization, and carceral society can be applied to the influence of technology. The discussion of building cultural norms into AI is already ongoing (Hadfield-Menell, Andrus, & Hadfield, 2019; Malle, Bello, & Scheutz, 2019; Serramia, 2018). Discussion must also occur regarding whose cultural norms are represented. Whittlestone and colleagues (2019) suggest that disparities in benefits of AI among different groups of peoples is a tension to monitor.

Human Agency

Artificial intelligence may empower learners, but it also has the potential to remove choice and control from the hands of learners and educators by dictating the pace and content provided (Bali, 2076). Similar conversations are occurring in law and medicine wherein AI which can predict outcomes on par or surpassing the lawyers and doctors they are meant to support is a growing reality (Froomkin, Kerr, & Pineau, 2019; Nunez, 2017). This concern can be balanced with recognition that learners and educators are already bound to pace and content set by external parties in many cases. As with accessibility, this concern may be addressed or not by the intentional development of AI for education. Artificial intelligence may increase or decrease learner and stakeholder choices and control. Whittlestone and colleagues (2019) suggest that convenience must be weighed against empowerment as a key tension emerging from further development of AI.

Access to Technology

Artificial intelligence in education as it currently exists is largely accessed through online interfaces and depends on computing devices as their primary point of interaction. As such, access to devices and sufficient Internet access is a pre-requisite for the use of AI for education. In regions where socioeconomic conditions or lack of broadband infrastructure reduce access to devices and the Internet, there is a risk that learners and institutions will also miss out on any benefits derived from incorporating AI in education. In the United States, 24 million people live in areas without broadband access, largely in rural areas (Bauerly, McCord, Hulkower, & Pepin, 2019). Broadband access has been cited as a social determinant of health, employment, and education, each with negative consequences where access is absent (Bauerly et al., 2019). Studies have repeatedly demonstrated negative effects for learners in areas with no broadband or poor-quality access, especially for post-secondary students (Skinner, 2019).

Explainability

Despite its relatively low rating among these items, explainability is one of the better-discussed issues being discussed regarding AI. Explainability refers to the ability of AI models to justify why decisions are made. Explainability is also linked with the terms interpretability and transparency. Due to the nature of complex AI systems, especially those with multiple hidden layers of processing, users are unable to definitively understand errors that are made or double-check decision rationales in high-stakes situations (Bansal, 2018). Theorodorou, Wortham, and Bryson (2017), in discussing the real-time inspection of autonomous robots, layout four of the reasons why this capacity is important; these include ensuring a lack of deception, as a measure of reliability, as a mechanism to report unexpected behavior, and as a way to expose the decision-making process (Theorodorou et al., 2017, pp. 232–233).

In some applications, such as applied research where the reason for the answers given is as important for understanding as the answer itself, explainability may become a crucial stage in AI development. In high-stakes situations, like the use of AI in criminal justice, this lack of explainability creates serious concerns (Pasquale, 2018). When interviewed regarding concerns about the incorporation of AI in public services, participants shared concerns about bias and the ability to interpret decisions (Brown et al., 2019). In other cases, perhaps including some education uses, it may be enough that the system arrives at the correct answer. Dare (2019) argues that, in some cases, “what matters is not transparency, or ‘explainability’, but whether there is evidence of reliability [...] it is evidence of reliability — rather than transparency — that we should insist on in the case of automated decision-making systems” (p. 6).

Groupthink

In the literature, groupthink is defined as "nondeliberate suppression of critical thoughts as a result of internalization of the group's norms" (Cheshire, 2017, p. 7). Cheshire (2017) points out the potential for AI to heighten the effect of groupthink and create a new effect dubbed "loopthink" (p. 8). This can be illustrated in cases like that of the legal research AI, ROSS. If ROSS is inclined to generate similar answers to similar queries, the risk of certain case law being represented over other case law may arise; in turn, ROSS may be exposed to more reinforcement of this selection due to the overrepresentation it generated. Further, this can be understood in light of the previous discussion of deresponsilization, whereby users receiving data from an AI system may not question whether or not it is the only, or even best, relevant data. This is especially true in cases where existing selection bias or confirmation bias by learners causes them to seek out and interpret information in a way that reinforces their existing individual and group beliefs. These biases come from human beings, not the AI system, but they may create "a feedback loop that perpetuates those biases" (Katyal, 2019, p. 77).

Groupthink received one of the lowest scores in the questionnaire portion of this survey, but it is worth pointing out that the difference between the education group and the AI group was significant, with the latter scoring the item as being of more concern. Like explainability, the item may have more significant relevance to specific fields or activities, such as research. The lack of explainability may also amplify this issue; when systems are biased toward one answer and the resulting groupthink reinforces that answer further and, in the absence of explainability, leads to ignorance by the collective agreement on a shared answer (Cheshire, 2017). When this collective agreement is over the best audience for content, to the exclusion of some groups of people, this feedback loop creates the grounds for bias for or against labels that might be applied to human beings (Katyal, 2019). Opaque algorithms already shape what results we see when we enter a query, and the potential for individualized results based on past behavior and group alignment created by machine learning could lead us toward increasing groupthink (Ojala, 2018).

Intellectual Property

Intellectual property interacts with AI in the regard that AI depends on data to learn and may construct its own works considered intellectual property. The ownership of the intellectual property in both cases may be debated. For users, this concern may exist as a subset of concerns about transparency and privacy. As individuals, however, AI may come with some change regarding what can be reasonably expected in terms of control over one’s own data and creations. Levendowski (2018) goes so far as to recommend intentional relaxing of copyright law in order to grant AI broader access to data sources in order to support improved diversity in data and reduction of bias. Wachter and Mittelstadt (2018), on the other hand, argue that enhanced data protection measures must be put in place, especially for personal data. Intellectual property law typically seeks to balance the individual and societal benefits created by new properties, but AI may necessitate a significant reconsideration of the balance (Cubert & Bone, 2018). This issue was rated as being of relatively low concern by both the education and the AI group in this survey’s questionnaire but remains relevant.

Dependence on Technology

The lowest rated item of those included in the questionnaire employed for this survey was dependence on technology. This can be discussed both for the reliance of individuals and institutions on technology to meet their goals and needs, but also in terms of addiction to technology. Froomkin and colleagues (2019) discuss overreliance on machine learning processes, but they do so from the very specific perspective of the interaction between machine learning and the likely implications for medical malpractice law and practice. Overreliance on technology is a conversation that seems to occur at the advent of any significant development in technology, and one which seems to be forgotten by the advent of the next significant development in technology. Outside of specific concerns, as in medical diagnostics and its relationship with malpractice law, this concern may be accurately treated by the participant results in this survey. Dependence in terms of addiction to technology may be of more significant concern and deserving of more consideration but is not specific to AI.

Other Issues

The open-ended question at the conclusion of the questionnaire asked participants to identify any additional issues or concerns they might have. Several participants took advantage of this field to share such concerns, exposing several additional facets to the above-listed issues as well as two issues that may not be properly represented by those named issues alone. These new concerns were for the impact on childhood development, should AI continue to be integrated into activities for young students, and the general concern, mentioned above, that the availability of financially expedient AI tools and the possibility of profiting off of learner data could lead to financial considerations being prioritized over sound ethical and educational practices.

As with discussing dependence on technology, the concern related to childhood development being impacted is not specific to AI, but a concern which may be further aggravated by the proliferation of educational applications for younger students. The American Academy of Pediatrics previously recommended limiting children under the age of eight to less than two hours of digital media use per day (Globokar, 2018). In reviewing present research, Globokar (2018) points out that digital technology use has been argued as a negative factor in emotional development and social development. Further, Aric Sigman (2017) suggests that over-exposure to on-screen interactions may impact childhood neurodevelopment and increase the risk of addictive relationships with digital media.

Regarding financial considerations being prioritized over sound practices, this is, again, not explicitly an issue of AI, but rather a concern which is amplified by the potential of AI. This can also be viewed as a subset issue of fairness and equity but deserves further comment. As discussed above, Santry (2018) points out the concern held by some that AI will replace teachers out of convenience and financial necessity rather than due to efficacy. The potential harm created by institutions selling access to learner data may put learner needs at odds with institutional solvency. One significant cause of anxiety and discussion regarding AI is the potential for the displacement of human workers via automation (Wolla, Schug, & Wood, 2019). This issue is much larger than AI in education and was not a primary focus of this survey, but it does deserve consideration for the impacts it will bring to education. In the event of massive disruption to the role of human labor in society, the nature and role of education will undoubtedly shift and develop in a similar upheaval.

Recommendations

Given the rapid and continued development of new AI applications, education professionals at all levels should take note of existing and emerging uses in their institutions and take an active role in shaping policy. Educators should seek to establish themselves as stakeholders in the development and implementation of emerging educational technology and not depend on the limited agency given to consumers of these technologies. Transparent policy and practice standards should be developed and matched with regulatory protections for student data. Should implementation and efficacy continue to expand, strategies for ensuring access and equity must also be in place. If efficacy increases allow for uses which replace traditional lecture, deskwork, and homework activities, educators should look toward redefining their roles and learning interventions to focus more on establishing new best practices in the classroom, such as focusing on human interaction and facilitating social learning. Further research should be done in areas where existing literature is insufficient. Developmental and neurological impacts of interacting with technology, especially for young children, should be a continued area of exploration. If AI developers can develop systems that explain their decisions, solutions for concerns such as bias, groupthink, and authentication may become clearer. Broader and more sophisticated questionnaire approaches than employed in this survey may reveal any validity to the beliefs about AI exposed in this paper, as well as discover gaps in understanding and the need for increased education and cooperation between AI professionals and education professionals.

Conclusion

Given the growing diversity and availability of AI applications for education, education professionals must look at their use and existing practices critically. The incorporation of AI applications in educational institutions is likely to bring with it benefits, but also significant risks. Defining and examining these risks as applications are implemented, and new applications that emerge will be necessary on an ongoing basis. It may be that some of these concerns are unfounded and that current beliefs about education should be re-evaluated, but a sober and thoughtful examination can reveal important truths and obstacles. In the middle ground between sensational utopian and dystopian beliefs regarding AI, there is still a pressing need for this sober examination, and the impact of AI on society will not be that of just another educational technology.

References

Artificial intelligence (AI). (2019). In *Encyclopædia Britannica*.

Artificial intelligence. (2019). In *Funk & Wagnalls New World Encyclopedia*.

Neural network. (2016). In *Encyclopædia Britannica*.

Allen, C., Wallach W., & Smit, I. (2006). Why machine ethics?. *IEEE Intelligent Systems*, 21(4), 12–17. https://doi.org/10.1109/MIS.2006.83

Arruda, A. (2017). An ethical obligation to use artificial intelligence?. *American Journal of Trial Advocacy*, *40*(3), 443–458. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=124647905&site=ehost-live

Asaro, P. M. (2019). AI ethics in predictive policing. *IEEE Technology & Society Magazine*, *38*(2), 40–53. https://doi.org/10.1109/MTS.2019.2915154

Bali, M. (2017). Against the 3A's of edtech. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/blogs/profhacker/against-the-3as-of-edtech-ai-analytics-and-adaptive-technologies-in-education/64604

Bauerly, B.C., McCord, R.F., Hulkover, R., & Pepin, D. (2019). Broadband access as a public health issue: The role of law in expanding broadband access and connecting underserved communities for better health outcomes. *Journal of Law, Medicine & Ethics*, *47*, 39–42.

Bendici, R. (2018). Rise of the machines. *University Business*, *21*(10), 53-54.

Blank, G., Bolsover, G., & Dubois, E. (2014, August) *A new privacy paradox: Young people and privacy on social network sites*. Paper presented at Annual Meeting of the American

Sociological Association, San Francisco, California. Available at SSRN: https://ssrn.com/abstract=2479938 or http://dx.doi.org/10.2139/ssrn.2479938

Bloomfield, B., & Vurdubakis, T. (2008). IBM's chess players: On AI and its supplements. *Information Society*, *24*(2), 69–82. https://doi.org/10.1080/01972240701883922

Blumenstyk, G. (2018). Can artificial intelligence make teaching more personal?. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/Can-Artificial-Intelligence/243023

Bonderud, D. (2019). Artificial intelligence, authentic impact: How education AI is making the grade. *EdTech*. Retrieved from https://edtechmagazine.com/k12/article/2019/08/artificial-intelligence-authentic-impact-how-educational-ai-making-grade-perfcon

Brown, A., Chouldechova, A., Putnam-Hornstein, E., Tobin, A., & Vaithianathan, R. (2019, May). *Toward algorithmic accountability in public services*. Paper presented at CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland. https://doi.org/10.1145/3290605.3300271

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications: Profound change is coming, but roles for humans remain. *Science*, *358*(6370), 1530–1534. https://doi.org/10.1126/science.aap8062

Burke, L. (2019). Your interview with AI. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2019/11/04/ai-assessed-job-interviewing-grows-colleges-try-prepare-students

Casacuberta, D., & Guersenzvaig, A. (2019). Using Dreyfus' legacy to understand justice in algorithm-based processes. *AI & Society*, *34*(2), 313–319. https://doi.org/10.1007/s00146-018-0803-2

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'dood society': The US, EU, and UK approach. *Science & Engineering Ethics*, *24*(2), 505–528. https://doi.org/10.1007/s11948-017-9901-7

Crowe, D., Lapierre, M., & Kebritchi, M. (2017). Knowledge based artificial augmentation intelligence technology: Next step in academic instructional tools for distance learning. *TechTrends*, *61*(5), 494-506. doi:http://dx.doi.org.ezproxy.lib.umb.edu/10.1007/s11528-017-0210-4

Cubert, J. A., & Bone, R. G. A. (2018). The law of intellectual property created by artificial intelligence. In W. Barfield & U. Pagallo (Eds.). *Research handbook on the law of artificial intelligence*. Northampton, MA: Edward Elgar Publishing.

de Graff, W. (2019). Sorting machine. *Education Journal*, *378*, 18.

de Saint Laurent, C. (2018). In defence of machine learning: Debunking the myths of artificial intelligence. *Europe's Journal of Psychology*, *14*(4), 734–747. https://doi.org/10.5964/ejop.v14i4.1823

Dare, T. (2019). Ethics of artificial intelligence and health care. *New Zealand Medical Student Journal*, (28), 5–7. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=138364984&site=ehost-live

Darraj, E., Sample, C., & Justice, C. (2019). Artificial intelligence cybersecurity framework: Preparing for the here and now with AI. In T. Cruz, & P. Simoes (Eds.) *18th European Conference on Cyber Warfare and Security*, 131–142. Reading, UK: Academic Conferences and Publishing International Limited.

Desmarais, M., & Baker, R. S. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, *22* (1-2), 9-38.

Donahoe, E., & Metzger, M. M. (2019). Artificial intelligence and human rights. *Journal of Democracy*, *30*(2), 115–126. https://doi.org/10.1353/jod.2019.0029

du Boulay, B. (2019). Escape from the skinner box: The case for contemporary intelligent learning environments. *British Journal of Educational Technology*, *50*(6), 2902–2919. https://doi.org/10.1111/bjet.12860

Eicher, B., Polepeddi, L., & Goel, A. (2018). Jill Watson doesn't care if you're pregnant: Grounding AI ethics in empirical studies. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 88-94. New York, NY. DOI: https://doi.org/10.1145/3278721.3278760

Friedman, S. (2019a). AI research to focus on academic and social-emotional growth. THE Journal. Retrieved from https://thejournal.com/articles/2019/05/07/ai-research-to-focus-on-academic-and-social-emotional-growth.aspx

Friedman, S. (2019b). Online student monitoring platform prevents suicides. THE Journal. Retrieved from https://thejournal.com/articles/2019/09/16/online-student-monitoring-platform-prevents-suicides.aspx

Froomkin, A. M., Kerr, I., & Pineau, J. (2019). When AIs outperform doctors: Confronting the challenges of a tort-induced over-reliance on machine learning. *Arizona Law Review*, *61*(1), 33–99. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=135523331&site=ehost-live

Gilliard, C. (2018). How ed tech is exploiting students. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/How-Ed-Tech-Is-Exploiting/243020

Globokar, R. (2018). Impact of digital media on emotional, social and moral development of children. *Gulhane Medical Journal*, *60*(4), 545–560. https://doi-org.ezproxy.lib.umb.edu/10.31192/np.16.3.8

Gose, B. (2016). When the teaching assistant is a robot. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/When-the-Teaching-Assistant-Is/238114

Güera, D., & Delp, E.J. (2018, November) *Deepfake video detection using recurrent neural networks*. Paper presented at 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand. doi: 10.1109/AVSS.2018.8639163

Hadfield-Menell, D., McKane, A., & Hadfield, G.K. (2019, January) *Legible normativity for AI alignment: The value of silly rules*. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 417–422. Honolulu, Hawaii.

Heller, C. H. (2019). Near-term applications of artificial intelligence: Implementation opportunities from modern business practices. *Naval War College Review*, *72*(4), 79–105. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=139043132&site=ehost-live

Henderson, P., Sinha, K., Angelard-Gontier, N., Ke, N. R., Fried, G., Lowe, R., & Pineau, J. (2018). *Ethical challenges in data-driven dialogue systems*. In Proceedings of

AAAI/ACM Conference on AI Ethics and Society (AIES-18), New Orleans, Lousiana, USA.

Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., & Hlosta, M. (2019). A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective. *Educational Technology Research & Development*, *67*(5), 1273–1306. https://doi.org/10.1007/s11423-019-09685-0

Herder, E., Sosnovsky, S., & Dimitrova, V. (2017). Adaptive intelligent learning environments. *Technology Enhanced Learning*, 109–114. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-02600-8_10

Hoeschl, M.B., Bueno, T.C.D., & Hoeschl, H.C. (2017, November). *Fourth industrial revolution and the future of engineering: Could robots replace human jobs? how ethical recommendations can help engineers rule on artificial intelligence*. Paper in 7th World Engineering Education Forum, 21–26, Kuala Lumpur, doi: 10.1109/WEEF.2017.8466973

Johnson, D., & Verdicchio, M. (2017). Reframing AI discourse. *Minds & Machines*, *27*(4), 575–590. https://doi.org/10.1007/s11023-017-9417-6

Jones, M.L., Kaufman, E., & Edenberg, E. (2018). AI and the ethics of automating consent. *IEEE Security & Privacy*, *16*(3), 64–72. doi: 10.1109/MSP.2018.2701155

Kay, J. (2012) AI and Education: Grand Challenges. IEEE Intelligent Systems, *27*(5), 66–69. doi: 10.1109/MIS.2012.92

Kim, P. (2018) Big Data and artificial intelligence: New challenges for workplace equality. *University of Louisville Law Review*. Advance online publication. Available at SSRN: https://ssrn.com/abstract=3296521

Kim, Y., Soyata, T., & Behnagh, R.F. (2018). Towards emotionally aware ai smart classroom: current issues and directions for engineering and education. *IEEE Access*. *6*, 5308–5331, 2018. doi: 10.1109/ACCESS.2018.2791861

King, C. (2013). A Foucauldian analysis of NCLB: Student data as panoptic surveillance. (Unpublished doctoral dissertation). Georgia State University, Georgia, USA.

Kokku, R., Sundararajan, S., Dey, P., Sindhgatta, R., Nitta, S., & Sengupta, B. (2018). *Augmenting classrooms with AI for personalized education*. Paper presented at IEEE International Conference on Acoustics, Speech and Signal Processing, 6976–6980, Calgary, AB, Canada. doi: 10.1109/ICASSP.2018.8461812

Kompella, K. (2018). AI-enabled visual search. *EContent*, *41*(4), 34. Retrieved from http://search.ebscohost.com.ezproxy.lib.umb.edu/login.aspx?direct=true&db=a9h&AN=132405478&site=ehost-live

Köse, U. (2018). Are we safe enough in the future of artificial intelligence? A discussion on machine ethics and artificial intelligence safety. *BRAIN: Broad Research in Artificial Intelligence & Neuroscience*, *9*(4), 184–197. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=133436011&site=ehost-live

Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Michael Hanemann, W., Kopp, R. J., & Moody, W. R. (2002). The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice?. *Public Opinion Quarterly*, *66*(3), 371-403.

Krueger, N. (2017). Artificial intelligence has infiltrated our lives. Can it improve learning? *International Society for Technology in Education (ISTE).* Retrieved from https://www.iste.org/explore/Empowered-Learner/Artificial-intelligence-has-infiltrated-our-lives.-Can-it-improve-learning%3F

Lee, S. H., Maenner, M. J., & Heilig, C. M. (2019). A comparison of machine learning algorithms for the surveillance of autism spectrum disorder. *PLoS ONE*, *14*(9), 1–11. https://doi.org/10.1371/journal.pone.0222907

Lu, X. (2019). An empirical study on the artificial intelligence writing evaluation system in China CET. *Big Data*, *7*(2), 121–129. doi: 10.1089/big.2018.0151

Malle, B.F., Bello, P., & Scheutz, M. (2019). Requirements for an artificial agent with norm competence. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 21–27. ACM, New York, NY, USA. doi: https://doi.org/10.1145/3306618.3314252

McArthur, D., Lewis, M. & Bishary, M. (2005). The roles of artificial intelligence in education: Current progress and future prospects. *Journal of Educational Technology*, *1*(4), 42–80.

McCarthy, T., Rosenblum, L. P., Johnson, B. G., Dittel, J., & Kearns, D. M. (2016). An artificial intelligence tutor: A supplementary tool for teaching and practicing braille. *Journal of Visual Impairment & Blindness*, *110*(5), 309–322.

McMurtrie, B. (2018). How artificial intelligence is changing teaching. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/How-Artificial-Intelligence-Is/244231

Medianik, K. (2018). Artificially intelligent lawyers: updating the model rules of professional conduct in accordance with the new technological era. *Cardozo Law Review*, *39*(4), 1497–1531. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=129621291&site=ehost-live

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. Wakefield, MA: CAST Professional Publishing

Miller, F. A., Katz, J. H., & Gans, R. (2018). AI × I = AI2 the OD imperative to add inclusion to the algorithms of artificial intelligence. *OD Practitioner*, *50*(1), 6–12.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012*). Foundations of machine learning.* MIT Press: Cambridge, MA, USA.

Nunez, C. (2017). Artificial intelligence and legal ethics: Whether AI lawyers can make ethical decisions. *Tulane Journal of Technology & Intellectual Property*, *20*, 189–204. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=127131782&site=ehost-live

Pasquale, F. (2018). When machine learning is facially invalid. *Communications of the ACM*, *61*(9), 25–27.

Pierce, D., & Hathaway, A. (2018) The promise (and pitfalls) of AI for education: Artificial intelligence could have a profound impact on learning, but it also raises key questions. *THE Journal*, *45*(3), 20–23.

Raub, M. (2018). Bots, bias and big data: Artificial intelligence, algorithmic bias and disparate impact liability in hiring practices. *Arkansas Law Review*, *71*(2), 529–570. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=134307912&site=ehost-live

Rodriguez, A. A., Boticario, J. G., Finat, C., & Petrie, H. (2017). Setting accessibility preferences about learning objects within adaptive elearning systems: User experience and organizational aspects. *Expert Systems*, *34*(4), n/a-N.PAG. https://doi.org/10.1111/exsy.12187

Santry, C. (2018). Artificial intelligence 'risks school apartheid'. *Tes*. Retrieved from https://www.tes.com/news/artificial-intelligence-risks-school-apartheid

Schaffhauser, D. (2019a). ALEKS to add AI-based nudge alerts. *THE Journal*. Retrieved from https://thejournal.com/articles/2019/01/07/aleks-to-add-ai-based-nudge-alerts.aspx

Schaffhauser, D. (2019b). AI in education shows most promise for the repetitive and predictable. *THE Journal*. Retrieved from https://thejournal.com/articles/2019/02/28/ai-in-education-shows-most-promise-for-the-repetitive-and-predictable.aspx

Schalow, T. (2016, August). *Human learning culture in an age of advanced artificial intelligence*. In Proceedings of The Multidisciplinary Academic Conference, Prague, Czech Republic.

Schroeder, R. (2019) Emerging roles of AI in education. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/digital-learning/blogs/online-trending-now/emerging-roles-ai-education

Serramia, M., López-Sánchez, M., Rodríguez-Aguilar, J.A., Morales, J., Wooldridge, M., & Ansótegui, C. (2018, February). *Exploiting Moral Values to Choose the Right Norms*. In AAAI/ACM Conference on AI, Ethics, and Society, 264–270, New Orleans, LA, USA. doi: 10.1145/3278721.3278735

Sigman, A. (2017). Screen dependency disorders: a new challenge for child neurology. *Journal of the International Child Neurology Association*, *17*(119), 1–13. https://doi.org/10.17724/jicna.2017.119

Skinner, B. T. (2019). Making the connection: Broadband access and online course enrollment at public open admissions institutions. *Research in Higher Education*, *60*(7), 960–999. https://doi.org/10.1007/s11162-018-9539-6

Stahl, B.C., & Wright, D. (2018). Ethics and privacy in ai and big data: Implementing responsible research and innovation. *IEEE Security & Privacy*, *16*(3), 26–33. doi: 10.1109/MSP.2018.2701164

Strickland, E. (2018). AI-human partnerships tackle "fake news": Machine learning can get you only so far-then human judgment is required. IEEE Spectrum, *55*(9), 12–13. doi: 10.1109/MSPEC.2018.8449036

Theodorou, A., Wortham, R. H., & Bryson, J. J. (2017). Designing and implementing transparency for real time inspection of autonomous robots. *Connection Science*, *29*(3), 230–241. https://doi.org/10.1080/09540091.2017.1310182

Timms, M.J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education*, *26*(2), 701–712.

Vanitha, C.N., Jeevaa, V.N., & Shriman, S.P. (2019, August). *Image and face recognition using CV lens machine learning android application*. In International Conference on Advanced Computing & Communication Systems, 972–975, Chennai, India. doi: 10.1109/ICACCS.2019.8728435

Van Laere, E., and van Braak, J. (2017). Auditory support in linguistically diverse classrooms: factors related to bilingual text-to-speech use. *Journal of Computer Assisted Learning*, *33*, 79–95. doi: 10.1111/jcal.12171.

Wachter, S., & Mittelstadt, B. (2018) A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review*, *2019*(2), 494–620.

Whittlestone, J., Nyrup, R., Alexandrova, A. & Cave, S. (2019, January). The role and limits of principles in ai ethics: Towards a focus on tensions. Paper presented at AAAI/ACM Conference on AI, Ethics, and Society, Honolulu, HI, USA. https://doi.org/10.1145/3306618.3314289

Wijekumar, K., Meyer, B., & Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Educational Technology Research & Development*, *60*(6), 987–1013. https://doi.org/10.1007/s11423-012-9263-4

Wolla, S., Schug, M.C., & Wood, W.C. (2019). The economics of artificial intelligence and robotics. *Social Education*, *83*(2), 84–88.

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. *British Journal of Educational Technology*, *50*(6), 3119–3137. https://doi.org/10.1111/bjet.12758

Zgrzebnicki, P. (2017). Selected ethical issues in artificial intelligence, autonomous system development and large data set processing. *Studia Humana*, *6*(3), 24–33. https://doi.org/10.1515/sh-2017-0020

Zhang, J., Zhang, Y., Zou, Q., & Huang, S. (2018). What learning analytics tells us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. *Journal of Educational Technology & Society*, *21*(2), 245–258. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=128981071&site=ehost-live

Zimmerman, M. (2018). *Teaching AI: Exploring new frontiers for learning*. Portland, OR: International Society for Technology in Education.

Appendix A

Questionnaire: Ethical Considerations for AI in Education

Thank you for agreeing to take part in this academic questionnaire on the topic of ethical concerns for incorporation of artificial intelligence (AI) in education. The data collected by this questionnaire will be used by the author, Dana Remian, for the purpose of discussing the topic in a Master of Education capstone project at the University of Massachusetts Boston. This questionnaire is intended primarily for individuals with some background in education and/or artificial intelligence, but all responses will be considered. This questionnaire should take 5–10 minutes for you to complete. This questionnaire will remain available through 11/8/2019.

Do you have a professional or academic background in artificial intelligence or a related field?

Yes [ ] No [ ]

Do you have a professional or academic background in education or a related field?

Yes [ ] No [ ]

What is your job title? [ ]

Issues and Concerns

Recent developments in AI and computing have placed AI-enhanced applications in various industries and a growing number of consumer products. AI platforms and services aimed at enhancing educational outcomes and taking over administrative tasks are becoming more common and appearing in more classrooms and offices. The following questions seek to clarify both emerging and unidentified concerns about the inclusion of these platforms and services in education. When considering these questions, note that this questionnaire is not restricted to a specific age-range or type of institution, and includes the use of AI to support education as well as the administrative functions of educational institutions.

How concerning are each of the following items in regard to the incorporation of artificial intelligence in education from an ethical perspective? Please pick from the four options for each of these items.

[Not of Concern, Of Minor Concern, Of Moderate Concern, Of Significant Concern]

Accessibility (exclusion of users due to lack of necessary modes of communication, cognition, etc.)
Access to Technology/Infrastructure (Exclusion of users due to lack of necessary equipment or infrastructure [i.e.- broadband])
Accountability (recognition of errors or failures in the absence of human actors)

Authentication of Knowledge (ensuring knowledge is accurate and current, free of excess gatekeeping, etc.)
Bias (algorithmic bias produced by data sets used in machine learning)
Cultural Integrity (disruption of culture)
Dependence on Technology (creation of group and/or individual dependence)
Explainability/Interpretability (lack of ability of to understand AI process and decisions)
Fairness and Equity (differentiation and flexibility for needs in and out of the learning environments)
Groupthink (promotion of irrational conformity due to shared expectations or process)
Human Agency (loss of individual/group control)
Intellectual Property (individual rights to authored content)
Privacy (use of data for non-learning purposes or surveillance)
Security (unanticipated access to data or assets)
Social Engineering (potential for intentional spread of misinformation, propaganda, etc.)
Transparency (institutional disclosure about use of AI)
Weaponization (potential for data or expert systems to be used against human beings)
Accessibility (exclusion of users due to lack of necessary modes of communication, cognition, etc.)

Are there other ethical issues and concerns that you believe should be considered in a discussion of the use of artificial intelligence by educational institutions? Please list. [ ]

Appendix B

Issue	**Overall Score** (mean avg of 41 responses)	**Ed Group Score** (mean avg of 31 responses)	**AI Group Score** (mean avg of 8 responses)
Accessibility	.73	.72	.88
Access to Technology/Infrastructure	.70	.70	.79
Accountability	.85	.84	.88
Authentication of Knowledge	.80	.80	.92
Bias	.83	.81	.92
Cultural Integrity	.73	.76	.83
Dependence on Technology	.51	.53	.48
Explainability/Interpretability	.68	.66	.75
Fairness and Equity	.81	.83	.88
Groupthink	.60	.44	.71
Human Agency	.71	.70	.83
Intellectual Property	.53	.55	.48
Privacy	.85	.88	.96
Security	.81	.81	.92
Social Engineering	.76	.75	.88
Transparency	.75	.73	.88
Weaponization	.74	.72	.88
Total Average	.73	.72	.81