EasyChair Preprint

№ 8935

Artificial Intelligence (AI) in Education: Addressing Societal and Ethical Challenges in K-12 Settings

Selin Akgun and Christine Greenhow

October 3, 2022

International Society of the Learning Sciences

Artificial Intelligence (AI) in Education: Addressing Societal and Ethical Challenges in K-12 Settings

Abstract: Artificial intelligence (AI) incorporates the applications of algorithms, machine learning, and natural language processing. AI has several applications in education, such as automated assessment and facial recognition systems, personalized learning tools and microblogging systems. These AI applications have the potential to increase capacity within education by supporting the social and cognitive development of students. Despite these affordances, AI applications pose critical ethical and societal drawbacks which are rarely considered in K-12 education. Integration of these algorithms in education may amplify societies' existing systemic biases and discrimination, perpetuate privacy, autonomy, and surveillance concerns for students from marginalized and underserved groups, and amplify existing gender and racial bias. In this paper, we interrogate applications of AI in K-12 education, highlighting their ethical risks. We introduce instructional resources to help educators navigate the challenges of integrating AI and advance K-12 students' understanding of AI and ethics. The paper concludes with recommendations for research.

Keywords: Artificial intelligence (AI) in education; AI and ethics; teacher education

Introduction and purpose

Artificial intelligence (AI) and its applications have become an inseparable part of our lives as we perform Google searches, read articles and emails, or join online tutorials and courses (Iman et.al., 2020; Regan & Jesse, 2019). During the COVID-19 pandemic, AI systems became even more important to everyday functioning; it is now nearly impossible to live in a modern society without using the applications powered by AI (Iman et.al., 2020). AI leads the fields of science, engineering, and technology, but also is actively present in education through machine-learning systems and algorithms (Ali et. al., 2019). AI has a variety of algorithmic applications in education, such as through automated assessment systems to lessen teachers' workload and support them in evaluating what students know; facial recognition systems to bring insights about student behaviors; and personalized learning systems to cultivate students' learning experiences (Remian, 2019). Algorithm systems are also dominant in education through various social network sites and social media outlets, especially within mobile applications and microblogging systems. Therefore, AI applications have the potential to increase capacity within education by leveraging the cognitive and social development of students (Murphy, 2019).

Despite this promise, these applications also pose critical societal and ethical drawbacks in education. In essence, AI algorithms do not occur in a vacuum; rather, they shape and are shaped by society's ever-evolving historical and institutionalized systemic biases, which ultimately transform into algorithmic bias (O'Neil, 2016). The biggest risks of integrating these algorithms in education (particularly in K-12 contexts) are: (a) amplifying societies' existing systemic bias and discrimination, (b) perpetuating privacy, surveillance, and autonomy concerns for students from mostly marginalized and disadvantaged groups, and (c) amplifying exiting gender bias, racism, xenophobia, and other forms of inequity and injustices (McMurtrie, 2018). As Ruha Benjamin's (2019, p.15) points out that: "New technologies reflect and reproduce existing inequities, but they are promoted and perceived as more objective or progressive than the discriminatory systems of previous era."

Considering these pressing concerns, we argue that it is our responsibility (as learning scientists, educators, and teacher educators) to educate teachers and students to recognize, critique and challenge the ethical and societal implications of algorithm use in education. To create a future generation where an inclusive and diverse citizenry can participate in the development of the future of AI, we need to study and develop AI- and ethics-based curricula and professional development opportunities for teachers and K-12 students as they engage with AI and apply a critical perspective (Authors, 2021).

Toward this end, in this paper, we synthesize existing literature on how applications and ethical dimensions of AI takes place in K-12 education. We first unpack the definition, benefits, and drawbacks of AI applications. Then, we provide strategies for practitioners to reap the benefits while navigating the ethical challenges of AI use in K-12 settings. Toward this end, we introduce different instructional and pedagogical resources, such as MIT Media Lab's "*AI and Ethics*" curriculum for elementary schools (Payne, 2019) and Code.org's "AI for ocean" investigation for elementary and high schools. We conclude with suggested directions for learning scientists seeking to build on such resources in future design studies.

Applications and benefits of AI in K-12 settings

Considering its capacity and affordances in fields of science, engineering, and education, we define AI as a form of technology that constructs systems to think and act like humans with the capability of solving problems and achieving goals (Iman et. al., 2020; Remian, 2019). Powered by algorithms and machine learning systems, AI takes place in education in the forms of a) personalized learning systems, b) facial recognition systems and predictive analytics, c) automated assessment systems and d) social networking sites and chatbots.

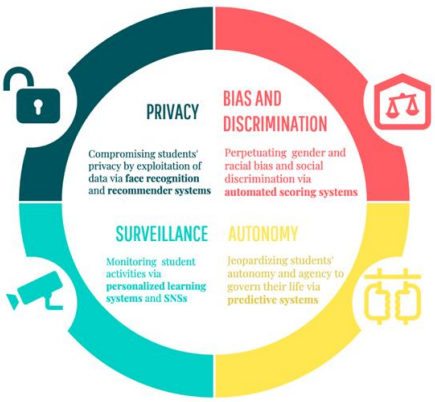
More specifically, ***personalized learning systems***, are one of the most common applications of AI in supporting students during their individual learning process. These systems provide students with access to various learning resources by considering each students' individual learning needs and based on specific subject area goals (McMurtrie, 2018). For example, rather than engaging in ideas and solving problems in physics on a worksheet or reading a textbook, students are able to use an interactive, multimedia space to help them progress toward learning objectives in their subject matter (McMurtrie, 2018). Microsoft's recent report (2018) from Singapore, the U.K., the U.S., and Canada shows that such personalized learning platforms promise to identify gaps in students' prior knowledge by offering different learning tools to promote students' growth. With the drastic shift to online education during the COVID-19 pandemic, personalized learning systems have become one of the promising distance learning options for students (Remian, 2019). Second, ***facial recognition systems*** are used to monitor and capture students' facial expressions. These systems aid teachers in recognizing and evaluating students' behaviors during the learning process. They also provide teachers with a space to intervene, which, in turn, can support teachers in developing student-centered classroom environments by increasing students' participation and engagement (Dishon, 2017).

Another prevalent tool powered by AI: ***automated assessment systems,*** can help meet the need for scoring students' exams, essays, and assignments, which are usually assessed by teachers. Assessment algorithms offer course management tools to decrease teachers' workload by increasing their productivity and capacity. For example, a scoring tool called "Gradescope" has been used by a number of universities to improve and streamline the assessment process (Blank et al., 2014). These systems have the potential to support students' learning process by giving them feedback and offering guidance to develop and revise their writing. Finally**,** ***social networking sites*** (SNSs) monitor students' well-being, provide space to deepen teacher-student relationships, and widen learning opportunities for students beyond the formal classroom setting (Asterhan & Rosenberg, 2015). Different scholars point out that the integration of social media can promote students' collaboration skills, active learning, and connections with their communities (Asterhan & Rosenberg, 2015; Authors, 2020; 2019).

Ethical and societal risks of integrating AI applications in education

In addition to their affordances, AI systems pose serious ethical risks and challenges. In essence, the risks become visible as AI systems are introduced and marketed as value-neutral and objective tools. In contrast, algorithms carry the values of their builders who hold positions of power (Remian, 2019). In the process of building algorithms, society's systemic and historical biases are represented within the codes, which ultimately transform into algorithmic bias. Even though the biased code is not integrated into the algorithmic model with an explicit intention, such representation can perpetuate existing racial and gender biases in different AI platforms (Benjamin, 2019). In the sections below, we focus on four major concerns in relation to privacy, autonomy, surveillance, and bias and discrimination (see Figure 1).

Figure 1. Ethical and Societal Challenges of AI applications in Education (Authors, 2021)



One main ethical issue raised in AI use in K-12 education relates to ***privacy*** concerns for teachers and students (Remian, 2019; Kang, 2021). As people expose an intensive amount of personal data (their language, location, biographical data, racial identity) in different online platforms, privacy violations can occur. Although legislation has been created to protect sensitive personal information, tech companies' violations (such as the privacy lawsuit aimed at Facebook (Kang, 2021)) with respect to security and data access, escalate concerns about privacy and human agency as independent and introspective individual thought is reduced (e.g., AI users may not realize the extent to which personal data is being shared) (McMurtrie, 2018). ***Surveillance*** is another notable ethical concern which relates to tracking preferences and actions of students and teachers. AI-based surveillance or tracking systems not only monitor activities but also determine the future actions of their users (Regan & Jesse, 2019). Patrolling and monitoring students' actions can be interpreted as part of a teacher's role and responsibility (e.g., to intervene in cases like exposure to sexual content or cyber-bullying). However, these systems may limit students' learning process by making them feel uncomfortable and unsafe. How can students feel secure and safe to take ownership for their ideas when they realize the AI systems are policing and surveilling their actions and thoughts? (Remian, 2019).

Connecting to these surveillance concerns, concerns about ***autonomy*** also stem from limiting and diminishing people's ability and freedom to act based on their own values and interests. For instance, predictive algorithms' predictions in relation to people's future actions based on their metadata raise serious questions about amplifying existing bias and prejudices of social stratification (Murphy, 2019). Finally, ***bias and discrimination*** concerns introduce another critical issue for AI ethics in K-12 education (Asterhan & Rosenberg, 2015). For instance, gender bias and gender-specific stereotypes become one of the visible forms of this problem. As just one example, when students in language learning courses use Google Translate, they see the Turkish equivalent of "S*he/he is a nurse*" translated into the feminine form, while the phrase of "*She/he is a doctor*" translated into the masculine form (Johnson, 2020). Racial bias against non-dominant groups is also present in voice-based assistants Murphy, 2019). Similarly, racial bias is present in facial recognition systems towards darker skinned people that has led to wrongful arrests (Murphy, 2019).

Teaching ethical and societal dimensions of AI in K-12 settings

The four major concerns above illustrate the need to introduce students and teachers to risks of AI use in K-12 education. To meet this need, nonprofits and research groups (e.g., IBM's Educator's AI Classroom Kit, Google's Teachable Machine, and Machine Learning for Kids) have created open-access instructional resources related to AI and ethics. Next we highlight two such resources: the MIT Media Lab's *AI and Ethics* curriculum and Code.org's *AI and Oceans* activity.

The MIT Media Lab provides an open-access AI and ethics curriculum for the middle school level. Through lesson plans and activities, the curriculum aims to support students' in learning the terminology of AI systems and their ethical implications (Ali et al., 2019). The curriculum includes learning objectives, such as introducing students to basic components of AI (supervised machine, datasets and algorithms) and helping them to figure out the process of creating bias in algorithms. For instance, through the "Algorithms as opinions" investigation, students consider algorithms as recipes. First, they write an algorithm to make their **"**best" jelly sandwich and peanut butter. In doing so, they raise the question of what it means to be **"**best." They see how their thoughts and definitions are reflected in their algorithms. Thus, students engage with the idea that algorithms can have various goals and motives. Following this investigation, students also build their own "Ethical matrix." They refer to their "best" jelly sandwich and peanut butter algorithms. They further discuss what counts as the "best" sandwich for themselves (most practical, healthy, delicious). Using the ethical matrix (chart), they identify different stakeholders (their parents, friends, teacher) who care about their jelly sandwich and peanut butter. In this way, stake-holders' opinions and values are also embedded into their algorithms. These investigations assist students and teachers in realizing a) the role and agency of different stakeholders in a society (or system), and b) how they are able to build their values and biases into the algorithms in the ethical matrix.

Another resource for practitioners is provided by Code.org, a nonprofit organization focused on increasing students' participation and involvement in computer science. Sponsored by tech companies like Microsoft and Google, Code.org aims to bring opportunities for students (grades 3-12) to learn about machine learning systems, algorithms, and algorithmic bias (Code.org, 2020). Through their online activity called "AI for oceans," students build their own machine learning models by exploring how to train, classify and model data**,** as well as how personal bias takes place in machine-learning systems. Students initially classify the objects as either "not fish" or "fish" to remove trash from the ocean and develop their data set by adding other sea creatures.

Creating such instructional and curricular resources are extremely crucial to support teachers and their students a) in negotiating the meaning and role of AI systems within the machine learning applications and b) in

figuring out the potential risks, limits, and ethical risks of AI applications in education to reap the benefits and minimize the costs of AI use.

Conclusion and implications

In this paper, we emphasize four potential ethical and societal drawbacks of using AI applications in education. To highlight these risks, we focused on several instructional resources and strategies for practitioners who are seeking to incorporate AI applications and their ethical issues in K-12 settings. These instructional strategies might help students and teachers to reap the benefits of AI while criticizing and disrupting the ethical challenges related to privacy, autonomy, surveillance, and bias.

Existing research on how best to support practitioners in teaching AI and ethics is still in its infancy; it is critical to conduct future research on designing professional learning opportunities for teachers as well as on the integration of ethics-focused curriculum where AI is adopted in schools by centering culturally responsive and relevant pedagogies. In other words, centering students' funds of knowledge (Bang et. al., 2012) and family and community background are important to addressing the concerns of surveillance, privacy, autonomy, and bias within the AI systems. More specifically, further research is needed that focuses on questions of a) how existing AI curriculum materials and activities on AI and ethics are developed and taught and b) in what ways do professional learning sessions support teachers in coming to understand and teach the risks and benefits of using AI in K-12 classrooms are needed. In such collaborative research projects, students might have a space to voice their own contextual and cultural experiences while trying to critique and disrupt existing power structures (Ladson-Billings, 1995).

We believe that this work is essential for learning scientists and teacher educators to understand and build on existing instructional materials and resources with respect to AI and ethics. It would also contribute to the work of the Learning Science community by increasing awareness and action on the critical use of AI and algorithms to create more inclusive, fair, and equitable AI education in K-12 settings.

References

Ali, S. A., Payne, B. H., Williams, R., Park, H. W., & Breazeal, C. (2019). Constructionism, ethics, and creativity: developing primary and middle school artificial intelligence education. Paper presented at International Workshop on Education in Artificial Intelligence (EDUAI). Palo Alto, CA, USA.

Asterhan, C.S.C., Rosenberg, H. (2015). The promise, reality and dilemmas of secondary school teacher–student interactions in Face- book: the teacher perspective. *Computer Education, **85***, 134–148.

Authors, 2021.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new jim code. Polity.

Code.org. (2020). AI for Oceans. https://code.org/oceans

Dishon, G. (2017). New data, old tensions: big data, personalized learning, and the challenges of progressive education. *Theory Res. Educ. **15***(3), 272–289.

Iman, M., Arabnia, H. R., & Branchinst, R. M. (2020). Pathways to artificial general intelligence: a brief overview of developments and ethical issues via artificial intelligence, machine learning, deep learning, and data science. Paper presented at ICAI. Las Vegas, Nevada, USA.

Johnson, M.: A scalable approach to reducing gender bias in Google translate. Accessed 26 March 2021

Kang, C. (2021) Mark Zuckerberg will be added to a Facebook privacy lawsuit. Accessed 19 November 2021.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J. **32***(3), 465–49.

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

Murphy, R. F. (2019). Artificial intelligence applications to support k–12 teachers and teaching: a review of promising applications, challenges, and risks. *Perspective.* 1–20. https://doi.org/10. 7249/PE315

O'Neil, C. (2016). Weapons of math destruction: how big data increases inequality and threatens democracy. Penguin Books.

Payne, B. H. (2019). An ethics of artificial intelligence curriculum for middle school students. MIT Media Lab. https://www.media.mit. edu/projects/ai-ethics-for-middle-school/overview/

Regan, P.M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking. *Ethics Inf. Technol. **21***, 167–179.

Remian, D. (2019). Augmenting education: ethical considerations for incorporating artificial intelligence in education (Unpublished master's thesis). University of Massachusetts, Boston.