

## **THE ETHICS OF ARTIFICIAL INTELLIGENCE IN EDUCATION: PRACTICES, CHALLENGES, AND DEBATES**

**Wayne Holmes, Kaška Porayska-Pomsta (Eds.)**

New York and London: Routledge, 2022, p. 312

### **Introduction**

The book discusses the field of Artificial Intelligence in Education (AIED). The authors believe that AIED is a diverse field that encompasses aspects of philosophy, learning and teaching, research, and engineering. They argue that AIED practitioners need to take a broader approach to define the purpose of the field and be more engaged with more general societal issues. The authors call for AIED systems to be designed with transparency, accountability, and user control, to ensure fairness and equity in education. They also suggest a shift away from the traditional model of AI design, which assumes a fixed objective, to a more collaborative model that allows for negotiation between humans and AI to set individual goals.

### **Structure of the book**

The book is organised in two separate sections, each with five chapters. Section I has chapters 1-5, while section II has chapters from 6- 10.

Chapter 1, "Learning to Learn Differently" by Jutta Treviranus, begins Part I. She notes that AI technology usually depends on big homogeneous datasets that underrepresent outliers and minorities. Despite the best intentions, this increases injustices and automates disparities, with disadvantaged pupils in the tails of the normal distribution. She also critically engages with AIED's function in "personalisation", which, as she points out, involves

bringing all students towards "perfected conformity," which she thinks can mask critical flaws in our education systems. She closes by advocating for a Wabi-sabi aesthetics in education, which, contrary to machine logic, promotes "the imperfect, transitory, and incomplete" and values humanity's vast diversity.

Alison Fox discusses how to incorporate ethics into AIED research in Chapter 2. Firstly, she introduces the CERD ethical evaluation framework based on the consequentialist, ecological, relational, and deontological ethical thought, emphasising context-specific ethical justification. These are then used to create a framework for critical considerations, including the ethics of data sources and their long-term curation, to which AIED researchers are beholden, avoiding harm and minimising risk (such as building on Chapter 1, the risk of normalisation through data cleansing); respect and duty of care; transparency and choice for genuine consent; and legal, moral, methodological, and ethical obligations. She closes with a plea for ethics by design approach in which researchers must be aware of both their research's positive and negative effects and be open to alternatives.

Ivana Bartoletti's Chapter 3, "AI in education: An opportunity riddled with challenges", examines AI's potential to help the school system improve equality and socioeconomic mobility. It asks if AI may be a "springboard" for opportunity or will it undermine the human agency and perpetuate economic and cultural disparities. Bartoletti then discusses six significant dangers for AIED implementation: scaling up weak pedagogical ideas; the negative impact of monitoring student emotions; the ethics of nudging; the erosion of human agency; privacy invasions; and teaching AI. She finishes with a plea for an ethics framework based on purpose,

impact, justice and fairness, and human values, reiterating that technology is never neutral and it reflects cultural, social, and political norms.

Lionel Bossi, Ana María Castillo, and Sandra Cortesi argue in Chapter 4, “Student-centred standards for the ethics of AI in education,” that the “student voice” should lead AIED development, design, deployment, and regulation. Based on UNESCO’s Beijing Consensus on Artificial Intelligence and Education (2019)<sup>1</sup>, they suggest that most young people are excluded from AI narratives, which often shape their futures. Women, girls, individuals with disabilities, and people from underdeveloped countries and rural areas are also marginalised. The authors discuss HabLatam, a groundbreaking Latin American research programme on youth and digital technology that shows how students’ voices may be heard.

Chapter 5, by Nathalie Smuha, concludes the Part I. She discusses AIED’s ethical challenges in terms of the seven Trustworthy AI requirements listed in the Ethics Guidelines of the European Commission’s High-Level Expert Group on AI: respect for human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental wellbeing, and accountability. Smuha adds that AIED is multifaceted, with numerous technologies gathering different sorts of data and having diverse aims and effects on instructors and students, so there is no one-size-fits-all strategy. The power and knowledge asymmetries in all educational relationships and the variety of stakeholders (students, parents, instructors, schools,

policymakers) whose needs must be considered complicate this. Before deploying any AI solution in education, we should understand better whether the identified need has to be solved and whether AI is the right tool.

Chapter 6, “Equity and Artificial Intelligence in education” by Kenneth Holstein and Shayan Doroudi, explains that AI-supported education is the future wave and can revolutionise how we learn. If implemented correctly, AI-supported education can play a pivotal role in closing the existing gaps in education and creating a more equitable learning landscape for all. However, it is also essential to consider the unintended consequences that AI-supported education can have, such as the potential for amplifying existing inequalities. To ensure that AI-supported education leads to greater equality in education, it is essential to recognise the various social and technical components and continually have honest, critical discussions about implementing this technology. By doing so, we can ensure that AI-supported education is utilised to its full potential, ultimately leading to a more equitable and just education system.

Chapter 7, “Algorithmic fairness in education” (René F. Kizilcec and Hansol Lee), discusses the increasing use of algorithmic systems and educational technologies. These systems use data and predictive models to provide support and insights to students, instructors, and administrators. However, using these technologies raises questions about their impact on stakeholders, including fairness, bias, and discrimination. The text focuses on the issue of algorithmic fairness in education and how it can be mitigated by considering three steps in the process: measurement, model learning, and action. The authors build on research from other domains to examine standard measures of algorithmic fairness

<sup>1</sup> UNESCO (2019). Beijing Consensus on Artificial Intelligence and Education. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>.

and explore how a better understanding of threats to algorithmic fairness can advance the responsible use of Artificial Intelligence in education. Another chapter in the book by Holstein and Doroudi (2021) examines the socio-technical system surrounding AI in education and its risks for amplifying educational inequalities.

Chapter 8 entitled “Beyond “fairness” Structural (in)justice lenses on AI for education” is written by Michael Madaio, Su Lin Blodgett, Elijah Mayfield and Ezekiel Dixon-Román. The authors argue that educational policies and practices, as instantiated through academic AI systems, reflect and further entrench societal visions and values, particularly in a settler colonialist, white supremacist, patriarchal, cis-heteronormative, and ableist society. The authors call for a radical, liberatory form of education that challenges and transforms these inequitable structures. The authors examine the issue of structural injustice in educational AI systems and argue for changing the epistemology of academic AI towards justice and equity. They do not offer solutions, but argue that the current crisis of the COVID-19 pandemic presents an opportunity to re-imagine education systems, including educational AI, and to reshape the priorities of academic research towards addressing structural injustice in educational AI.

In Chapter 9, entitled “The overlapping ethical imperatives of human teachers and their Artificially Intelligent assistants”, Benedict du Boula, based on the text presented, concluded that when considering the deployment of AIED systems, ethical dilemmas can arise that could be seen to involve teachers not doing their best, or exploiting their position. To help avoid this, four guidelines should be followed: (1) using an iterative co-design process that involves the immediate stakeholders, (2) providing training for users on

how best to use the system, (3) providing training for teachers on how to integrate the system into their work best, and (4) ensuring that all stakeholders are aware of the data that the system extracts and creates, where it is stored, and who has what rights over it. Following these guidelines should ensure that AIED systems are used responsibly and ethically.

Chapter 10, entitled “Integrating AI ethics across the computing curriculum”, written by Iris Howley, Darakhshan Mir, and Evan Peck, concludes that integrating computing technology into people’s daily lives requires re-examining how we teach about algorithms and AI. This includes teaching AI ethics as a critical component of the computing curriculum, as it reflects the importance of ethics in human lives. The authors provide an overview of existing approaches to teaching AI ethics and point to various educational resources, including Casey Fiesler’s “Tech Ethics Curricula”. The chapter also acknowledges that AI ethics teaching is rapidly advancing, and new resources are becoming available as instructors recognise the importance of ethical skills development. The authors recommend the reader to investigate new resources that have become available since the finalisation of the chapter.

AI is widely used in education through applications like personalised learning systems, automated assessment systems, and social media platforms. AI has the potential to improve education outcomes and upskill students, but it is essential to develop ethical frameworks to ensure that AI is used responsibly and without perpetuating existing inequalities. UNESCO has adopted the “Recommendation on the Ethics of AI” to ensure that AI aligns with human rights and values and includes considerations for data governance, privacy, and reducing biases. The recommendation

stresses the importance of developing critical thinking and AI ethics skills in learners and calls for multi-stakeholder collaboration in data governance and a “digital commons” approach to data sharing. If used ethically, AI can improve education and benefit society.

### **Editors**

Wayne Holmes is an expert in artificial intelligence and education, having served as a consultant for the IRCAI, UNESCO, and the Council of Europe, and has taught courses in learning sciences and innovation at University College London in the United Kingdom.

Kaska Porayska-Pomsta is a professor in the Department of Culture, Communication, and Media in the Faculty of Education at University College London in London, United Kingdom, where she specialises in the use of artificial intelligence in the classroom.

### **Concluding remarks**

The book discusses the importance of ethics in developing AI in education and the need for ongoing conversations about the topic. Editors and authors argue that the ethics of AI in education should be based on explicit values and ready to

address the subject’s political issues. The book provides guidance for AI researchers, learning scientists, educational technologists, and others on how to ensure that AI impacts learning positively. It offers questions, frameworks, guidelines, policies and regulations to help address the ethical concerns raised by AI in education. The book draws on the perspectives of experts both within and outside the AI in education scholarly community. The editors and authors hope that this book and the future works will inspire collective discussions about the ethical implications of AI in education and encourage a more deliberate and ethical approach to developing AI solutions. This approach may require a shift from rapid growth ambitions to a focus on slow and careful consideration.

Valentin Kuleto

Milena Ilić

Faculty of Contemporary Arts,

University Business Academy in Novi Sad, Serbia

Maja Ivović,

Savremena Primary and Secondary School,

Belgrade, Serbia