Shih, JL. et al. (Eds.) (2023). Proceedings of the 31st International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Ethical Challenges and Best Practices for Transparency in AIED: A Literature Review and Learner Centric Guidelines

Ram Das RAI*, Meera PAWAR & Ramkumar RAJENDRAN
Indian Institute of Technology Bombay, India

Abstract: Artificial intelligence in education (AIED) is the use of artificial intelligence techniques and tools to enhance and support learning and teaching processes. However, AIED also raises ethical challenges related to transparency, which is the ability to understand how AIED systems make decisions that influence educational outcomes. This paper reviews the literature on transparency in AIED, and presents a set of learner centric guidelines for ensuring transparency. The paper uses the Transparency Index Framework by Chaudhry et al. (2022) as a guiding tool, and discusses the best practices for providing information, allowing oversight, and respecting rights and choices in AIED. The paper also illustrates these practices with examples of ethically and transparently designed AIED systems. The paper concludes by highlighting the trade-offs and precautions involved in ensuring transparency in AIED, and suggests future research directions. This paper contributes to the ethical discourse on transparency in AIED by providing a comprehensive overview, identifying the best practices and challenges, and proposing the learner centric guidelines.

Keywords: artificial intelligence in education, ethics, transparency, accountability, safety, fairness, explainability

1. Introduction

Artificial intelligence in education (AIED) is a rapidly developing field that aims to create personalized learning experiences for each student, tailored to their individual needs and learning styles (Xu, 2020). AIED can also be used as a technological engine of educational innovation, promoting the reform and development of educational modes and teaching methods. Furthermore, AIED can be used as a technical tool to open the "learning black box," through which educational, psychological, and social knowledge can be accurately quantified and made explicit. However, the adoption of AIED has also led to increasing ethical risks and concerns regarding several aspects such as personal data and learner autonomy (Nguyen et al., 2023). Moreover, AIED poses ethical challenges such as the potential for AI to be used to discriminate against students from marginalized groups, violate students' privacy, and create a culture of surveillance in the classroom. Therefore, it is important to develop ethical guidelines for the use of AIED that ensure that this technology is used in a way that is beneficial to all.

Some of the ethical aspects of AIED that have been identified by researchers and practitioners include cultural integration, accountability, fairness, equity, affordability, security, privacy, and human dignity (Holmes et al., 2022). These aspects reflect the values and principles that should guide the design, development, and deployment of AIED systems. Among these aspects, transparency stands out as a key ethical principle that guides the use of artificial intelligence in education (AIED). Transparency means that AIED systems should provide sufficient information and explanation about how they make decisions and what data they use. Transparency is important for ensuring fairness, accountability, and trust in AIED systems. However, transparency is often challenged by the issue of "black boxes" in AI, which are software that do not reveal how they analyze input data to arrive at their outputs. This can create ethical dilemmas and risks in various domains, such as medical education, where AI

systems are intended to assist physicians with patient care (Katznelson & Gerke, 2021). Therefore, ethicists emphasize the moral requirement for transparency in AI and argue that algorithms should be accompanied by explainability (Farrow, 2023). Moreover, transparency is valued by educators and learners who use AIED systems. For example, Holmes et al. (2022) found that educators might prefer an AI system that sacrifices some accuracy in favor of providing greater transparency in decision-making processes. Yu and Yu (2023) also stated that "transparency stands out as the primary ethical principle guiding the advancement of AI implementation in the field of education" (p. 9).

Previous researchers have explored the ethical challenges of AIED (Adams et al., 2023, Holmes et al., 2022, Nguyen et al., 2023). However, most work has been broad and few studies have focused on specific aspects like transparency or learner autonomy. Chaudhry, Cukurova, and Luckin (2022) addressed transparency in AIED using an ML pipeline framework. They accounted for transparency at various stages of the AI development process from data collection to deployment. They shortlisted studies using popular frameworks for documentation, robustness and reproducibility for each ML pipeline. They also improved their framework based on interviews of education stakeholders. Their framework was robust and new, but mainly technical. Farrow (2023) emphasized using a socio-technical lens to look at ethics in AIED. This means looking beyond technical solutions and incorporating educational, legal and policy dimensions. There is also a need to consider trade-offs when ensuring transparency, such as sacrificing accuracy. This is a complex issue that needs discussion. Therefore, more research is needed in this area using a systematic and socio-technical approach to ethics in AIED.

The purpose of this paper is to review the literature on transparency in artificial intelligence in education (AIED), and to present a set of learner centric guidelines for ensuring transparency. Transparency is an ethical challenge in AIED, as it affects the ability to understand how AIED systems make decisions that influence educational outcomes. The paper uses the Transparency Index Framework by Chaudhry et al. (2022) as a guiding tool, and discusses the best practices for providing information, allowing oversight, and respecting rights and choices in AIED. The paper also illustrates these practices with examples of ethically and transparently designed AIED systems. The paper aims to contribute to the ethical discourse on transparency in AIED by providing a comprehensive overview of the existing literature, identifying the best practices and challenges, and proposing the learner centric guidelines. The paper is guided by the following research questions:

- What are the causes and consequences of the problem of lack of transparency in artificial intelligence in education (AIED)?
- How can effective and ethical strategies be developed and implemented to address the problem of lack of transparency in AIED and its associated challenges and trade-offs?

In this paper, the authors first conceptualized transparency based on the Transparency Index Framework proposed by Chaudhry, Cukurova, and Luckin (2022). This framework offers a comprehensive and multidimensional perspective on transparency in AI in education. They then derived a keyword string from this framework: (Artificial Intelligence OR AI) AND (Education) AND (Ethics) AND (Transparency OR Accountability OR Safety OR Fairness OR Explainability OR Autonomy). They conducted a systematic search using the PRISMA framework and shortlisted 15 papers that met their inclusion criteria (Moher, D., et al., 2015). In addition to these papers, the authors also referred to the document Recommendation on the Ethics of Artificial Intelligence by UNESCO (2021), which provides a global standard for ethical AI in the educational domain. The details of the PRISMA process are illustrated in Fig 1.

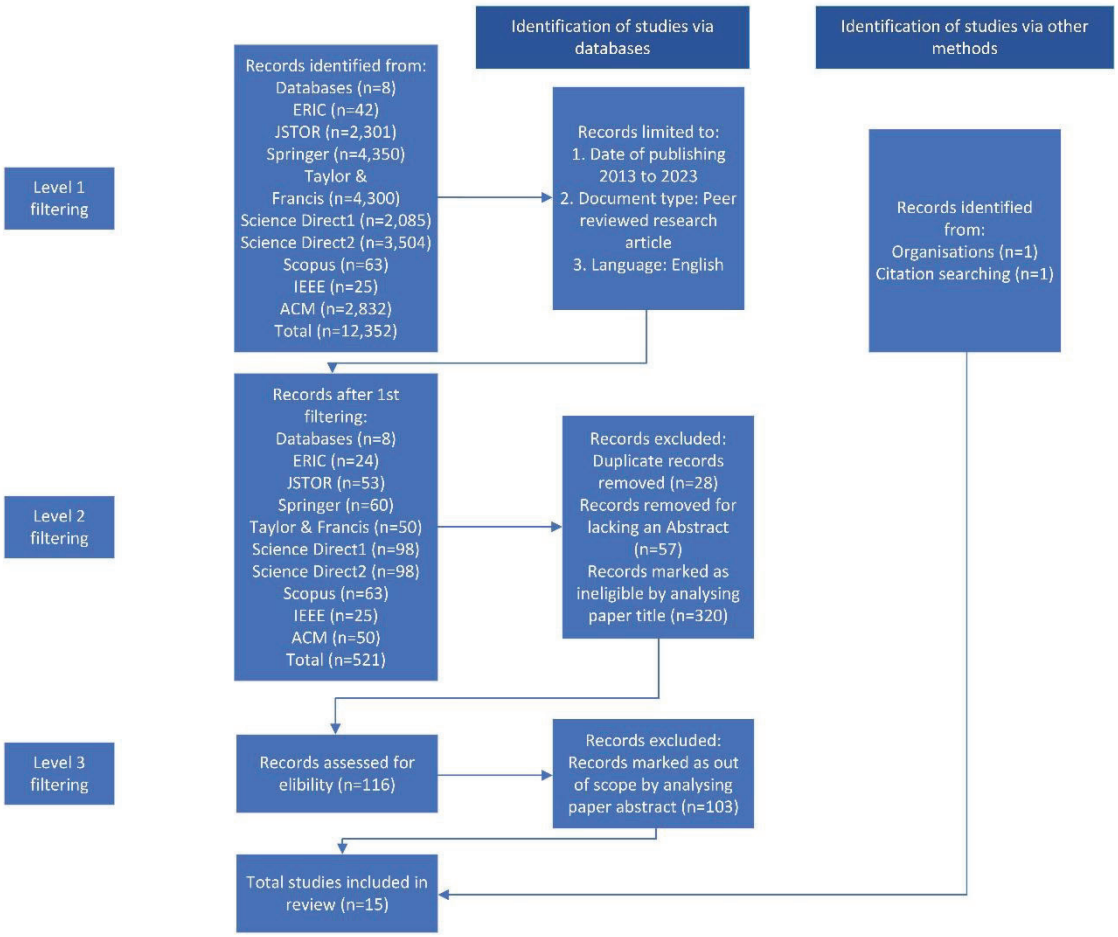

Fig. 1. PRISMA analysis details

This paper has three sections. Section 2 explains the objectives and method of the paper. Section 2.1 discusses the lack of transparency in AIED and its implications for decision-making, learning outcomes, and ethical values. Section 2.2 explores some strategies to address the ethical challenges of AIED, such as legal and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions. Section 3 discusses the trade-offs involved and the precautions that must be taken when implementing transparency in AIED. It also concludes the paper with a summary and some suggestions for future research.

2. Findings

This section presents the findings of the systematic review on ethical transparency in AIED. The section is organized into two subsections: the first one discusses the lack of transparency in AIED and its ethical implications, and the second one proposes strategies to address these ethical challenges. These findings are illustrated in Fig 2 which lists down the findings of this paper.

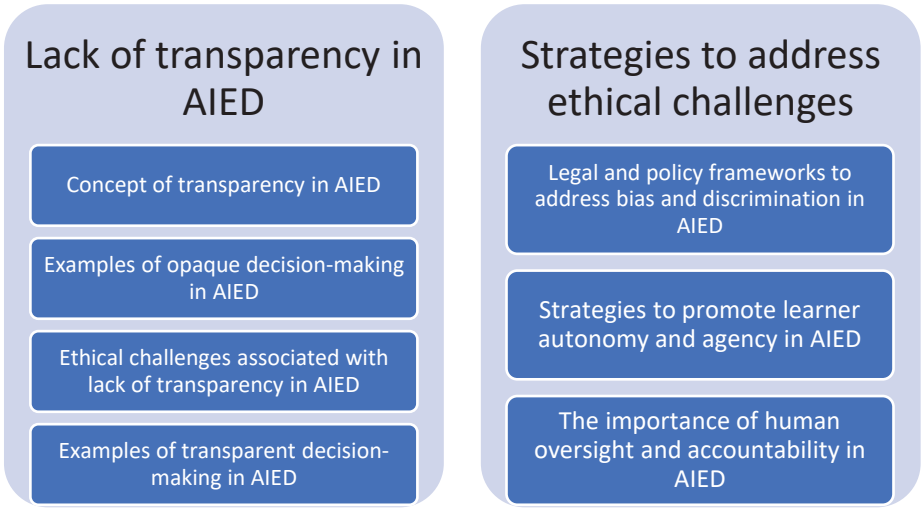

Fig. 2. A list of the findings on lack of transparency in AIED and strategies to address ethical challenges

2.1 Lack of transparency in AIED

2.1.1 Concept of transparency in AIED

Transparency is a key concept in artificial intelligence in education (AIED) that has implications for decision-making processes. Transparency can be defined in different ways depending on the perspective and the context of AI use in education or information transfer. However, a common theme is that transparency refers to the availability and accessibility of information about how AI systems make decisions and what data they use. This section presents some possible definitions of transparency from the literature. Turilli and Floridi (2009) defined transparency as the availability of information, the degree of availability, or the role of information in the decision-making process (as cited in Yu & Yu, 2023). Jobin et al. (2019) defined transparency as the requirement that when AI is used in education or information transfer, the specific parameters, source, responsibility, investment, and impact of AI should be disclosed to enhance accountability, modification, interpretation, and communication between AI users, educational researchers, and practitioners (as cited in Adams et al., 2023). Siegel et al. (2018) defined transparency as the degree to which the determinations or predictions of AI systems are revealed to relevant parties in ways that those parties prefer and can understand (as cited in Coghlan et al., 2021). These definitions illustrate the complexity and diversity of transparency in AIED. Transparency is also closely related to other concepts such as explainability, explicability, responsibility, and accountability in the context of AI and data (Nguyen et al., 2023). Yu and Yu (2023) emphasized that transparency has become a prominent concern in the use of AI for educational purposes. Transparency is important for ensuring fairness, trust, and communication in AIED systems.

2.1.2 Examples of opaque decision-making in AIED

Transparency in artificial intelligence in education (AIED) is important for understanding how AI systems are making decisions and what are the implications of those decisions for the users. However, transparency can be challenged by the existence of AI systems that are not transparent, meaning that they do not reveal how they make decisions or predictions, or what data or models they use. This is also known as the ‘black box’ problem, or the lack of transparency of the processes and workings that convert input to output, which is a major structural aspect of the ‘black box’ model of computation (Farrow, 2023). The ‘black box’ problem can create various ethical issues and risks for the AIED community and stakeholders, such as:

False trust and confidence: Users may not understand, question, or challenge the AI decisions or predictions that are accurate but not transparent (Holmes et al., 2022).

Negative impact on autonomy, agency, and self-regulation: AI systems may influence the learning process without providing a detailed account of how the algorithm works, or how it affects the learners' behavior, motivation, or outcomes (Farrow, 2023).

Reduction of engagement, motivation, and satisfaction: AI systems may deliver feedback to learners without explaining the criteria, evidence, or rationale behind it, or how it can be used to improve learning (Farrow, 2023).

Unfair or discriminatory outcomes: AI systems may use biased or incomplete data sets that favor certain groups of learners over others, or that do not account for the diversity and variability of learners' needs, preferences, and contexts (Holmes et al., 2022).

Violation of rights, interests, and well-being: AI systems may not disclose the purposes, methods, or outcomes of their data collection, analysis, or reporting, or that do not obtain informed consent or respect the privacy and ownership of the users' data (Sahlgren, 2023).

Confusion, inconsistency, or ambiguity: AI systems may not reveal the regulatory roles, responsibilities, or conditions of their design and use, or that do not address the ethical challenges or concerns raised by the AIED community or other stakeholders (Sahlgren, 2023).

2.1.3 Ethical challenges associated with lack of transparency in AIED

The lack of transparency in different aspects of artificial intelligence in education (AIED) systems can pose various ethical challenges for the AIED community and stakeholders. Chaudhry et al. (2022) classified the lack of transparency into three categories: data collection and analysis, decision making, and deployment of the AIED environments.

Data collection and analysis. The lack of transparency in data collection, analysis, and reporting may raise issues of informed consent, data ownership, data privacy, data quality, and data accountability. These issues can affect the rights, interests, and well-being of the individuals whose data are collected, processed, or used by the AIED systems (Holmes et al., 2022).

Decision making. The lack of transparency in the teaching decisions and the system's model of the learner may affect the trust, understanding, and agency of the learners and educators, and may limit their ability to challenge or improve the AIED systems. These issues can affect the learning outcomes, experiences, and opportunities of the learners and educators who interact with the AIED systems (Holmes et al., 2022).

Deployment of the AIED environments. The lack of transparency in the design and deployment of AIED systems may result in unintended or harmful consequences for the learners, educators, and wider society, such as bias, discrimination, manipulation, or exploitation. These issues can affect the quality, equity, and accountability of the AIED systems and their impact on various aspects of society (Holmes et al., 2022).

In addition to these three categories, there are also other ethical issues that can arise from the lack of transparency in AIED systems, such as:

Ethical guidelines and regulations. The lack of transparency in the ethical guidelines and regulations of AIED systems may create confusion, inconsistency, or ambiguity among the AIED researchers and developers, and may hinder their ethical responsibility and accountability. These issues can affect the validity, reliability, and credibility of the AIED research and development and their compliance with ethical standards and legal obligations (Holmes et al., 2022).

Benefits and costs of AIED innovation. The lack of transparency in the potential benefits and costs of AIED innovation may prevent the AIED community from balancing the ethical risks and opportunities of their work, and may undermine their social impact and value. These issues can affect the innovation potential, sustainability, and scalability of the AIED systems and their contribution to social good (Holmes et al., 2022).

2.1.4 Examples of transparent decision-making in AIED

Some examples of transparency in artificial intelligence in education (AIED) systems are as follows. One example is bots or intelligent tutoring systems (ITS) that provide clear and understandable feedback, guidance, and explanations to learners and educators about their learning goals, progress, and outcomes. These systems can help to enhance the trust, understanding, and agency of the users and support their learning processes and outcomes (Farrow, 2023). Another example is automated assessment tools that reveal the criteria, methods, and evidence used to evaluate learners' performance and provide constructive suggestions for improvement. These tools can help to ensure the validity, reliability, and fairness of the assessment and foster the learning and development of the users (Farrow, 2023). A third example is predictive analytics dashboards that display the data sources, assumptions, and limitations of the models used to generate predictions and recommendations for learners, educators, and institutions. These dashboards can help to inform the users about the rationale, accuracy, and uncertainty of the predictions and recommendations and enable them to make informed decisions and actions (Farrow, 2023). A fourth example is plagiarism checkers that show the original sources of the text that are matched with the learners' submissions and indicate the degree of similarity and originality. These checkers can help to promote the academic integrity, honesty, and creativity of the users and prevent the misuse or abuse of the AI systems (Farrow, 2023). A fifth example is fairness-promoting algorithms that monitor and mitigate the potential bias or discrimination in the data sets, models, or outcomes of AIED systems and ensure that they benefit all learners equally. These algorithms can help to protect the rights, interests, and well-being of the users and foster the quality, equity, and accountability of the AIED systems (Holmes et al., 2022).

2.2 Strategies to address ethical challenges

2.2.1 Legal and policy frameworks to address bias and discrimination in AIED

To address the issues of bias and discrimination in AIED and ensure transparency, several legal and policy frameworks have been proposed by researchers. One of them is a code of best practice that the AIED community can rely on in designing and deploying AIED technologies in diverse educational contexts, incorporating guidance addressing the issues of fairness, accountability, transparency, bias, autonomy, agency, and inclusion specifically at the intersection of AI and the learning sciences (Holmes et al., 2022). Another framework is an educational framework that embeds the principles of AI ethics in education, and also ethics issues such as responsibility, inclusion, fairness, security and explainability in conducting educational research, to raise awareness and foster critical thinking among educators and learners (Nguyen et al., 2023). Furthermore, a website that publishes and updates the ethical principles for AIED, inviting feedback and improvement suggestions from various stakeholders, such as researchers, educators, learners, policymakers, and the public, has been suggested as a way to facilitate communication and collaboration on AIED ethics (Nguyen et al., 2023). Additionally, an automatic method by text analysis that can evaluate the ethical implications of AIED research papers and systems, providing feedback and recommendations to the authors and reviewers, has been developed to enhance the quality and accountability of AIED research (Holmes et al., 2022). Finally, a regulatory framework that ensures that AIED systems are subject to rigorous testing and auditing before and after deployment, to detect and mitigate any potential harms or biases that might affect the users or the society, has been advocated as a necessary measure to ensure the trustworthiness and responsibility of AIED systems (Holmes et al., 2022).

2.2.2 Strategies to promote learner autonomy and agency in AIED

Researchers have suggested several strategies to promote learner autonomy and agency in AIED and ensure transparency. One of them is developing negotiation-based adaptive

learning systems that allow learners to choose the type and frequency of support, scaffolding of not only knowledge but also metacognition and self-regulation skills, and feedback on their learning progress and outcomes (Nguyen et al., 2023). Another strategy is providing learners with control over their data, such as the ability to access, edit, delete, or share their data with others, and informing them about how their data is collected, stored, processed, and used by the AIED system (Holmes et al., 2022). Moreover, adopting fairness-promoting algorithms that can detect and mitigate any potential biases or harms that might affect different sub-populations and demographics of learners, and ensuring that the AIED system does not favor or disadvantage any groups of learners based on their characteristics or backgrounds, is a crucial strategy to ensure equity and justice in AIED (Holmes et al., 2022). Furthermore, explaining the role and function of the AIED system to the learners in a simple and understandable way, such as using visualizations, examples, or analogies, and providing them with opportunities to ask questions, give feedback, or challenge the system's decisions, is a key strategy to enhance trust and engagement in AIED (Farrow, 2023). Finally, encouraging learners to develop AI literacy skills, such as understanding the basic concepts and principles of AI, its capabilities and limitations, its ethical and social implications, and how to collaborate effectively with AI systems, is a vital strategy to empower learners and prepare them for the future of education (Nguyen et al., 2023).

2.2.3 The importance of human oversight and accountability in AIED

Transparency is a fundamental principle for ethical AIED, and it requires human oversight and accountability from the various stakeholders involved in the design and use of AIED systems. One of the reasons why human oversight and accountability in AIED are important is that they ensure that AIED systems are developed and used with good intentions, integrity, and careful consideration of potentially harmful factors, such as bias, discrimination, or manipulation (Nguyen et al., 2023). Another reason is that they empower the human stakeholders, such as educators and learners, to have control over the design and use of AIED systems, and to make informed decisions about when and how to rely on AIED systems for their educational purposes (UNESCO, 2021). Moreover, human oversight and accountability in AIED allow the attribution of ethical and legal responsibility for any stage of the life cycle of AIED systems, as well as in cases of remedy related to AIED systems, to physical persons or to existing legal entities (UNESCO, 2021). Furthermore, human oversight and accountability in AIED enable the protection of personal information throughout the life cycle of the AIED system, by adopting a privacy by design approach that considers the potential risks and harms of data collection, processing, and use (UNESCO, 2021). Finally, human oversight and accountability in AIED facilitate the auditability, traceability, explainability, and communicability of the AIED systems, by providing clear and open information about the data, the system, and the business model of AIED, as well as the purposes, conditions, roles, responsibilities, and trade-offs involved in AIED (Nguyen et al., 2023).

3. Discussion and Conclusion

This paper has reviewed the literature on transparency in artificial intelligence in education (AIED), and presented a set of best practices for ensuring transparency in AIED systems. Transparency is an ethical challenge in AIED, as it affects the ability to understand how AIED systems make decisions that influence educational outcomes. The paper has discussed the best practices for providing information, allowing oversight, and respecting rights and choices in AIED, and illustrated them with examples of ethically and transparently designed AIED systems. The paper has also examined the role of transparency in fostering communication and collaboration among the different stakeholders involved in AIED, such as educators, ed-tech experts, and AI practitioners. The paper has contributed to the ethical discourse on transparency in AIED by providing a comprehensive overview of the existing literature, identifying the best practices and challenges, and proposing the learner centric guidelines. The paper has also acknowledged the trade-offs and precautions involved in ensuring

transparency in AIED, such as privacy and security issues, information and cognitive overload, data or algorithm misuse or abuse, ethical dilemmas or conflicts, and unintended or adverse outcomes. In the ensuing discussion, the paper offers the authors' perspective on the initially posed questions and provides key takeaways for the readers.

3.1 Causes and Consequences of Lack of Transparency in AIED

The origins of the deficiency in transparency in AIED are intricately rooted in a combination of factors. These factors include biased algorithms, concerns surrounding data privacy, and a general lack of transparency in AI-driven decision-making processes, all of which contribute to this problem. Consequently, the implications are extensive, impacting not only the impartiality and fairness of educational outcomes but also eroding confidence in AI-driven educational technologies. These challenges, while not insurmountable, demand purposeful and ethically grounded approaches.

3.2 Strategies to Address Ethical Challenges

The review outlined a range of strategies to tackle the ethical challenges associated with transparency in AIED. Legal and policy frameworks play a pivotal role in ensuring accountability, fairness, and transparency in the development and deployment of AIED systems. Approaches such as the formulation of a code of best practice, the integration of AI ethics principles into educational frameworks, and the establishment of websites dedicated to ethical guidelines hold significant promise.

Promoting learner autonomy and agency in AIED is imperative to empower students and ensure transparency. Granting learners control over their data, embracing fairness-promoting algorithms, and facilitating comprehension through explanations are measures that can foster trust and engagement in AI-enhanced educational environments. Additionally, equipping learners with AI literacy skills prepares them to confront the ethical challenges presented by an AI-driven world.

3.3 The Importance of Human Oversight and Accountability

Transparency in AIED cannot be achieved without the presence of human oversight and accountability. Active participation in the design and utilization of AIED systems by educators, learners, policymakers, and other stakeholders is paramount. Such involvement ensures that systems are developed and employed with ethical considerations, integrity, and due regard for potential adverse consequences. Moreover, it facilitates the allocation of ethical and legal responsibility and safeguards personal information via a privacy-by-design approach. Crucially, human oversight and accountability enhance the auditability, traceability, and explainability of AIED systems, thereby fostering transparency and trust.

3.4 Trade-offs involved when implementing ethical use of AIED to ensure transparency

Ethical use of AIED, particularly to ensure transparency, involves some trade-offs that need to be carefully considered and balanced by the researchers, developers, and users of AIED systems. One of the trade-offs is between accuracy and explainability of AI algorithms, as some AI models, such as deep neural networks, may achieve high performance but are difficult to interpret or justify their decisions (Nguyen et al., 2023). Another trade-off is between scalability and contextualization of AIED systems, as some AIED systems may be designed to reach a large and diverse audience but may not account for the specific needs, preferences, and cultures of different groups of learners or educators (Holmes et al., 2022). Moreover, a trade-off exists between innovation and regulation of AIED research and development, as some AIED research and development may aim to explore new possibilities and opportunities but may face ethical challenges or constraints from existing policies or guidelines (Holmes et

al., 2022). Furthermore, a trade-off occurs between privacy and personalization of AIED systems, as some AIED systems may require a large amount of data from the users to provide tailored and adaptive support but may also pose risks to the users' data protection and ownership (Nguyen et al., 2023). Finally, a trade-off arises between autonomy and oversight of AIED systems, as some AIED systems may enable the users to have more control and agency over their learning but may also require human intervention or supervision to ensure quality and accountability (Holmes et al., 2022).

3.5 Precautions that must be taken while ensuring transparency in AIED

Ensuring transparency in AIED requires some precautions that must be taken by the researchers, developers, and users of AIED systems. One of the precautions is to respect the privacy and confidentiality of the users' information and to obtain well-informed consent from the users before using their data for any purpose (Nguyen et al., 2023). Another precaution is to explain and justify the AI algorithms for specific educational purposes, but to avoid revealing sensitive or personal information about the users, such as their identity, performance, behavior, or preferences (Nguyen et al., 2023). Moreover, a precaution is to be transparent, traceable, explainable, and communicable in the process of establishing, conducting, monitoring, and controlling regulations of AIED, but to balance the interests and rights of different stakeholders, such as developers, educators, policymakers, and users (Nguyen et al., 2023). Furthermore, a precaution is to be transparent in the audibility of the regulation of AIED and to address acknowledgment and responsibility for each stakeholder's actions involved in the design and use of AIED, but to consider the potential trade-offs, risks, and harms that may arise from the use of AIED systems (Nguyen et al., 2023). Finally, a precaution is to accompany the transparency of AIED systems with the provision of guidance, feedback, and support to the users, especially the learners, to help them understand, trust, and challenge the AIED systems, and exercise their agency and rights over their own data and learning outcomes (Nguyen et al., 2023). Additionally, a precaution is to ensure transparency with respect to the needs of the user. An excessive amount of transparency can potentially overwhelm users, for example, when intricate details concerning the AI Model are disclosed to individuals lacking technical expertise (S. Amer-Yahia, personal communication, July 14, 2023). Striking the right balance between transparency and privacy is equally imperative. For instance, in the context of an AIED (Artificial Intelligence in Education) system, disclosure of the reasons behind a student's failure in a specific test should remain undisclosed unless expressly requested by the student, as this information must not be accessible to others (S. Amer-Yahia, personal communication, July 14, 2023).

3.6 Limitations and future research directions

This paper has some limitations that should be acknowledged and addressed in future research. First, it may not capture the full range and diversity of perspectives and approaches on transparency in AIED, as it may have excluded some relevant papers that used different terms or frameworks. Second, it does not consider other ethical principles that are equally important for AIED, such as privacy, beneficence, non-maleficence, justice, and human dignity. These principles are interrelated and may have implications for transparency in AIED. Third, it proposes some strategies to address the ethical challenges of AIED, such as legal and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions. However, these strategies are not exhaustive or definitive, and they may have their own limitations and challenges. To advance the research on transparency in AIED, we suggest the following future directions based on the limitations of this paper: (1) conduct a more comprehensive and systematic review of the literature on transparency in AIED, using a broader range of search terms and databases, and applying a more rigorous and transparent selection and analysis process; (2) explore the interplay and trade-offs between transparency and other ethical principles in AIED, such as privacy, beneficence, non-maleficence, justice, and human dignity; (3) evaluate the effectiveness and feasibility of the proposed strategies to address the ethical challenges of AIED, such as legal

and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions.

The paper hopes to stimulate more discussion and debate on the ethical challenge of transparency in AIED, as well as to inspire more innovation and collaboration among the AIED community. Transparency is not only a technical or legal issue, but also a moral and social one. Transparency is not only a means to an end, but also an end in itself. Transparency is not only a challenge for AIED, but also an opportunity for AIED. By embracing transparency as an ethical value and a design principle, AIED can become more trustworthy, responsible, and beneficial for education.

Acknowledgements

We would like to thank Dr. Shitanshu Mishra and the rest of the members from the UNESCO MGIEP International Working Group on Ethics of AIED for their help with this paper.

References

Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. Computers and Education: Artificial Intelligence, 4, 100131. https://doi.org/10.1016/j.caeai.2023.100131

Chaudhry, M. A., Cukurova, M., & Luckin, R. (2022). A Transparency Index Framework for AI in Education. In M. M. Rodrigo, N. Matsuda, A. I. Cristea, & V. Dimitrova (Eds.), Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium (pp. 195–198). Springer International Publishing. https://doi.org/10.1007/978-3-031-11647-6_33

Coghlan, S., Miller, T., & Paterson, J. (2021). Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies. Philosophy & Technology, 34(4), 1581–1606. https://doi.org/10.1007/s13347-021-00476-1

Farrow, R. (2023). The possibilities and limits of XAI in education: A socio-technical perspective. Learning, Media and Technology, 48(2), 266–279. https://doi.org/10.1080/17439884.2023.2185630

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Katznelson, G., & Gerke, S. (2021). The need for health AI ethics in medical school education. Advances in Health Sciences Education, 26(4), 1447–1458. https://doi.org/10.1007/s10459-021-10040-3

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews, 4(1), 1-9.

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. Education and Information Technologies, 28(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Sahlgren, O. (2023). The politics and reciprocal (re)configuration of accountability and fairness in data-driven education. Learning, Media and Technology, 48(1), 95–108. https://doi.org/10.1080/17439884.2021.1986065

UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence—UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Xu, L. (2021). The Dilemma and Countermeasures of AI in Educational Application. Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence, 289–294. https://doi.org/10.1145/3445815.3445863

Yu, L., & Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer. Frontiers in Psychology, 14. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1061778.