ACCOUNTABILITY AND RELIABILITY OF AI SUPPORTED SYSTEMS IN THE EDUCATIONAL SECTOR

Sabarish.S, Sastra Deemed University

Prasanthi @ Bharathi, Sastra Deemed University

ABSTRACT

With the growing integration of artificial intelligence (AI) in the educational sector, concerns about accountability and reliability have become paramount. This research paper delves into the multifaceted landscape of AI-supported systems in education, examining the intricate relationship between technological advancements and the accountability and reliability demands of educational stakeholders. The study navigates through a comprehensive literature review, establishing the groundwork for understanding AI's role in education and highlighting past challenges and successes. Grounded in a framework that combines principles of accountability and reliability, the research employs a meticulous methodology to investigate the current state of AI-supported systems in educational settings. Case studies offer practical insights into real-world applications, exploring both the triumphs and pitfalls of implementing AI in education. The paper identifies key challenges faced in ensuring accountability and reliability within these systems, considering legal, ethical, and technical dimensions. Drawing from the analysis, the study formulates recommendations for policymakers, educators, and technologists to enhance the accountability and reliability of AI-supported systems in education. By addressing these crucial aspects, the research aims to contribute to the ongoing discourse surrounding the responsible deployment of AI in education, fostering a more transparent, accountable, and reliable technological landscape for the benefit of learners and educators alike.

Keywords: Accountability, Reliability, AI- Supported system, Education.

1. Introduction

The educational landscape is in the midst of a transformative era marked by the pervasive integration of artificial intelligence (AI). The digitization of educational processes, coupled with the advent of intelligent algorithms, has given rise to a spectrum of AI-supported systems designed to enhance learning experiences, streamline administrative tasks, and personalize educational content. Yet some of the challenges and potentials remain unsolved or undiscovered which may be in part due to the largely unchanged structure of education[1] However, as these technologies become deeply ingrained in educational practices, questions surrounding accountability and reliability have gained prominence, necessitating an in-depth exploration of their implications. However, AI become deeply ingrained in educational practices, questions surrounding accountability and reliability have gained prominence, necessitating an in-depth exploration of their implications. The technological perspective cannot achieve a deep understand of the complex roles of AI in instructional and learning processes and its relationship with other educational elements[2] The foundational shift towards AI in education is not without its challenges. The traditional roles of educators, the dynamics of student-teacher interactions, and the very fabric of pedagogical methodologies are undergoing significant transformations. The rise of AI-driven solutions introduces a complex interplay of technological innovation, educational theory, and societal expectations. In this context, understanding how AI-supported systems impact accountability and reliability becomes imperative for educators, policymakers, and technologists alike. Still AI is a commodity in education, and its effect is explored primarily in a qualitative and general, rather than pragmatic and specific level[3]. The significance of this study lies in addressing the fundamental pillars of accountability and reliability in AI-supported education. Accountability ensures that stakeholders in the educational process are transparent, responsible, and fair in their use of AI systems. Reliability, on the other hand, focuses on the consistent and accurate performance of these systems over time, acknowledging the importance of trustworthy outcomes. Accountability of AI systems needs to be handled at both a micro and macro level and further account for the lawmakers, regulatory system, and social systems of morality in

---

[1] AI grand challenges for Education
http://refhub.elsevier.com/S2666-920X(23)00031-0/sref58

[2] Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artifcial intelligence in education. Computers and Education: Artifcial Intelligence, 1, 100001.
https://doi.org/10.1016/j.caeai.2020.100001

[3] http://refhub.elsevier.com/S2666-920X(23)00031-0/sref16

general[4]. Educational institutions continue to navigate the integration of AI, this research serves as a valuable resource for stakeholders seeking insights into fostering responsible and dependable AI practices in education. The study's findings and recommendations aim to contribute to the ongoing discourse surrounding the ethical and effective use of AI in shaping the future of education. Whatever happens AI, a critical technological element, has interdependent relations with other elements, which would cause changes of the whole system and formation of emerging characteristics[5]. Traditionally, face-to-face mentoring requires high effort in one-to-one situations because of its holistic conversational processes and therefore does not scale easily for numerous students. Still, an active relationship and frequent interaction between the two parties are crucial parts for successful mentoring[6]

2. AI-Supported Systems in the Educational Sector

Heading on to the discussion regarding AI-Supported Systems in the Educational Sector , it is important to seek to some examples they are

i. Intelligent Tutoring Systems (ITS):

Intelligent Tutoring Systems leverage AI to provide personalized guidance to students. These systems assess individual learning styles and adapt the instructional approach accordingly. ITS can cover various subjects, from mathematics to language arts, offering immediate feedback and additional support in areas where a student may be struggling. The way of instructor-student interaction is also influenced by intelligent tutor, which has potential to transform instructor role in the instructional processes and even replace instructor[7]

Example: Carnegie Learning's MATHia

MATHia is an AI-powered tutoring system designed to teach mathematics. It assesses individual student progress, adapts to their learning style, and provides real-time feedback and

[4] From Automation to Autonomous system, ugo pagallo http://refhub.elsevier.com/S2666-920X(23)00031-0/sref38

[5] . More than tools? Making sense of the ongoing digitization of higher education. International Journal of Educational Technology in Higher Education, 15, 22. https://doi.org/10.1186/s41239-018-0109-y

[6] Cornelius, V., Wood, L., and Lai, J. (2016). Implementation and Evaluation of a Formal Academic-Peer-Mentoring Program in Higher Education. Active Learn. Higher Educ. 17, 193–205. doi:10.1177/1469787416654796

[7] Vision, challenges, roles and research issues of artifcial intelligence in education. Computers and Education: Artifcial Intelligence, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001

guidance. The system continuously refines its approach based on the student's responses to ensure personalized learning experiences.

ii. Adaptive Learning Platforms:

Adaptive learning platforms utilize AI algorithms to tailor educational content based on a learner's performance and preferences. These platforms can adjust the difficulty of assignments, recommend supplementary materials, and provide a personalized learning path. By adapting to individual needs, adaptive learning platforms aim to enhance the efficiency and effectiveness of the learning process.

Example: DreamBox. It is an adaptive learning platform that focuses on mathematics education. It adapts to each student's proficiency level, providing targeted lessons and adapting difficulty levels based on performance. The platform incorporates AI algorithms to create personalized learning paths. In recent years, the development of smart classrooms has begun to integrate AI technology, and to automatically recognize and record most learning events and behaviours on the spot, so as to provide immediate feedback to teachers and support students' adaptive learning, but most research is still in the initial and feasible test development stage[8]

iii. Automated Grading Systems:

Automated grading systems employ AI to assess and grade student assignments. These systems can evaluate written essays, code submissions, or other assignments with a high degree of accuracy and efficiency. Automated grading not only saves time for educators but also provides students with prompt and constructive feedback, facilitating continuous improvement.

Example: Gradescope

Gradescope uses AI to streamline the grading process for educators. It can handle various types of assignments, from handwritten essays to coding assignments. The system employs machine learning to understand grading patterns, providing consistent and efficient grading while allowing for customization by instructors. The purpose of automated assessment tools is to replace human instructors in grading students' learning outcomes (eg. essays). Learner models that are enabled by algorithms can be used to automatically monitor and describe the online

[8] https://doi.org/10.1287/mnsc.41.8.1328

behaviours of students, and then create and suggest individualized learning paths for them. However, since most AI technologies are not created expressly for the educational sector, using them does not ensure that students will receive instruction of a high calibre or achieve desired learning outcomes.[9]

iv. Natural Language Processing (NLP) Applications:

NLP applications in education involve language-related tasks such as language learning, translation, and proficiency assessments. Language learning apps, for instance, may utilize NLP to enhance interactive language lessons, provide pronunciation feedback, and facilitate real-time language practice.

Example: Duolingo

Description: Duolingo, a language-learning platform, utilizes NLP to offer interactive language lessons. It employs AI to adapt lessons based on individual progress, assess pronunciation through voice recognition, and provide instant feedback. The NLP component enhances the platform's effectiveness in teaching languages.

v. Data Analytics for Educational Insights:

AI-driven data analytics tools help educators and administrators gain insights into student performance, behaviour, and learning trends. These tools can analyse vast amounts of data to identify patterns, predict student outcomes, and inform decision-making. Data analytics in education supports evidence-based practices and helps institutions optimize their educational strategies.

Example: Bright Bytes. It is an education analytics platform that utilizes AI-driven data analysis to provide insights into various aspects of education, including student performance, digital citizenship, and classroom technology usage. It helps educators make data-informed decisions to improve learning outcomes.

[9] Ouyang, F. & Jiao. P. (2021). Artifcial intelligence in education: The three paradigms. Computers and Education: Artifcial Intelligence, 2, 100020. https://doi.org/10.1016/J.CAEAI.2021.100020

3. Implementation Challenges:

a) Privacy Concerns: The privacy and security of student data are issues raised by the implementation of AI in education. To solve these difficulties, it is imperative to guarantee compliance with data protection rules, establish efficient safety measures, and maintain honest communication regarding data usage.

b) Lack of Standardization: Lack of standard frameworks for using AI in education might cause problems with connectivity and make it more difficult to integrate various technologies smoothly. In order to promote cooperation, data sharing, and the growth of a coherent educational AI ecosystem, standardization initiatives are essential. At present Connectivity across various educational AI systems is hampered by the lack of common frameworks.

c) Resistance to Change: Due to concerns about job displacement, a lack of experience with technology, or a perceived threat to conventional teaching techniques, educators, administrators, and students may be resistant to the integration of AI. Techniques like thorough training programs, outlining the advantages clearly, and include stakeholders in the decision-making process can all aid in reducing opposition. For example, by encouraging a culture of technology adoption, professional development initiatives, workshops, and cooperative decision-making procedures involving educators can assist in addressing opposition.

   When employed in a high-risk atmosphere, AI-based assessments encounter some resistance and present new technical challenges. The notion that assessments of people's abilities could be diverse but still accurate and fair challenges the beliefs of many parents, legislators, and students regarding equality, demonstrating that technological advancements in AI present society with just as many social and behavioural issues as they do technological ones.

d) Ethical Issues: Potential biases in AI algorithms, equitable access to educational resources, and the appropriate use of student data are some ethical issues. Enforcing algorithmic fairness, putting ethical standards into practice, and regularly assessing and monitoring the effects of artificial intelligence on a number of student categories are all necessary to ease these worries. An ethical evaluation survey is carried out by Kong et al.[10] with an emphasis on autonomy, beneficence, and fairness. The survey's findings indicate that students valued privacy and transparency in addition to these three elements. The work made it possible to incorporate AI ethics into projects rather than doing so after the fact.

e) Technical limitations: Reliability, scalability, and the requirement for constant upkeep and updates are examples of technical difficulties. Investments in technical infrastructure, frequent testing, and cooperation with technology providers to meet changing educational needs are necessary to ensure the durability of AI systems.

4. Accountability in AI-Supported Education

The most important prerequisite for reliable AI is responsibility. It is possible to interpret this as a restatement of all the previously stated requirements along with an assurance that they will be both declared and carried out. It means that throughout the AI system's whole life cycle, humans must accept accountability for the results and ramifications of the systems they create and use, and they must be prepared to face consequences when unfavourable effects materialize. The auditability of the AIED system is a crucial component in enabling accountability. When a system's data, algorithms, and design process can be examined and evaluated by outside auditors in addition to internal auditors, then it is considered auditable. Independent oversight of the auditing process, such as by public entities or authorized third-party auditors, is not an unnecessary luxury given the significance of education in society.

In descriptive terms, the context of artificial intelligence and higher education, accountability is defined as either the collection of mitigating or preventive measures that hold creators, owners, and users of artificially intelligent algorithms accountable. There may be multiple levels of accountability in higher education since strategic decision-making may be done by human committees or stakeholders or by algorithms. According to Bezuidenhout and Ratti

[10] Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. Educational Technology & Society, 26 (1), 16–30. https://doi.org/10.30191/ETS.202301_26(1).0002

(2021)[11], accountability and explicability are synonymous, meaning that any program that can be explained is deemed accountable. In the words of Shin et al.[12], accountability is a concept that aims to hold automated decision system providers accountable for the outcomes of their programmed decision-making. More precisely, Beerkens[13] approaches the concept of accountability from the standpoint of an educational audience, suggesting that the importance of indicator specificity over indicator simplicity is heavily dependent on the intended educational audience. According to Pagallo[14], it is unclear who should be held responsible—that is, the manufacturer or producer of the AI system or the data supplier. In a similar vein, systems based on algorithmic models are thought to lack accountability[15]. The design of AI programs may be partially to blame for the absence of accountability. AI is having implications for accountability in a discourse about changing authority. Nonetheless, opinions on whether humans, AI, or both should be held accountable differ. when teachers use AI in their work they come to shift or share accountability with AI programs. This necessitates a clear understanding of both the teacher and AI of one another to take responsibility as a united force or team. When there is a lack of understanding, either the teacher or AI may take the blame or reward when in fact it was not warranted

The results of these audits could then be made available to the public to improve the system's credibility as well as the processes that surround it, such as the choices made by humans regarding its use. The (effectiveness) of actions taken to reduce unjust bias, protect the technical stability and safety of the system, increase transparency in the decision-making processes, and guarantee sufficient human oversight can also be shown in these evaluation reports. The evaluation and reporting of any possible adverse effects of the AIED application, as well as the steps done to mitigate them, are closely related to this requirement. This kind of evaluation must be appropriate for the educational setting in which the AI system is being used, the users (teachers, students, and others), and the risks involved. Furthermore, in order to enable ongoing assessment, the assessment must be conducted on a regular basis if the AIED system

---

[11] Bezuidenhout, L., & Ratti, E. (2021). What does it mean to embed ethics in data science? An integrative approach based on the microethics and virtues. AI & Society, 36(3), 939–953. https://doi.org/10.1007/s00146-020-01112-w

[12] Shin, D., Rasul, A., & Fotiadis, A. (2022). Why am I seeing this? Deconstructing algorithm literacy through the lens of users. Internet Research, 32(4), 1214–1234. https://doi. org/10.1108/INTR-02-2021-0087

[13] Beerkens, M. (2022). An evolution of performance data in higher education governance: A path towards a 'big data' era? Quality in Higher Education, 28(1), 29–49. https:// doi.org/10.1080/13538322.2021.1951451

[14] http://refhub.elsevier.com/S2666-920X(23)00031-0/sref38

[15] Ungerer, L., & Slade, S. (2022). Ethical considerations of artificial intelligence in learning analytics in distance education contexts. https://doi.org/10.1007/978-981-19-0786-9_8

is actually deployed.

The significance of a protective framework for those who report on adverse impacts of AIED—be it an unfair bias, a safety concern, an unfounded claim of effectiveness, or another deficiency—in situations where such reporting could go against the interests of the school or of other stakeholders, or more generally raise difficulties, is an aspect that frequently goes unnoticed when ensuring accountability for AIED. In order to hold AI developers and implementers more accountable, whistleblowers have already been instrumental in bringing attention to some of the dubious applications of AI in other contexts[16]. It is critical that, in the context of education, parents, teachers, employees of AIED-developing companies, and other actors feel free to voice concerns about an application's possible negative effects without fear of any kind of negative fallout.

Lastly, it is critical that redress be easily sought whenever the creation or application of AIED results in an unfair harm to students, instructors, or any other stakeholder. Those in charge, such as the educational institution using AIED, should make it apparent that there are channels for recourse.

5. Reliability in AI-Supported Education

Putting in place robust human oversight mechanisms is one of the most important ways to address the use of AI that could (un)intentionally interfere with human agency or have other negative effects. The Guidelines list three specific types of oversight that can assist in achieving this, and they should be used in addition to one another based on the use case.

The first is a human-in-command (HIC) approach, in which the educational AI system's overall operations are monitored, along with any wider ethical, societal, legal, or economic ramifications. This includes making the initial decision about whether or not to use an AI system in a particular scenario and, if so, how the system should be used. To help provide insights into teachers' performance, for example, it is one thing for a school and its teachers to decide to use an AI-based value-added model. It is another matter entirely for a school to

[16] Crawford, Kate, Roel Dobbe, Theodora Dryer, Genevieve Fried, Ben Green, Elizabeth Kaziunas, Amba Kak, et al. 2019. AI Now 2019 Report. New York: AI Now Institute. https://ainowinstitute.org/AI_Now_2019_Report.pdf

determine that teachers could be fired based only on the results of such a model[17].

The human-on-the-loop (HOTL) approach is the second oversight strategy that is discussed in the Guidelines. This method guarantees that, once the AI system is in use, its use is closely monitored. It also allows for human intervention during the AI system's design phase.

The third strategy is known as "human-in-the-loop" (HITL), and it allows for human intervention into each and every AI system decision cycle. More ongoing human intervention may be required as a safety measure for AI applications that have the potential to have a major negative impact on the lives of teachers or students, even though the latter strategy would be too demanding for an AI system that only poses minimal or no tangible risks. In order to make sure that the system satisfies the requirements throughout its lifecycle, it is the responsibility of all parties involved—including the organizations that develop the system and the entities that choose to use it—to determine which oversight mechanism is necessary on an individual basis. Prior contemplation of the different rights and interests at risk during system implementation is also required for this assessment. In this context, it is particularly pertinent to remember the UN Convention on the Rights of the Child, which enumerates several crucial rights that must be taken into account whenever children are involved. Furthermore, ongoing projects like that map the precise effects of AI on children can also offer a useful framework (UNICEF 2020)[18].

AI systems should be protected against technical vulnerabilities by making them resilient to attacks on their software (comprising their data and model) or on their hardware. As with all software systems, AI is only as good as its technical infrastructure. Developers and deployers of AI should work to avoid this kind of situation, as it can still result in a variety of negative outcomes. In particular, where AI depends on students' personal data, appropriate security measures must be taken to guarantee that no data leaks can occur. The ability of the system to make accurate or correct decisions is another sign of technical robustness. Depending on the application, incorrect judgments, forecasts, or suggestions may have unfavourable effects on the parties concerned. In the context of education, examples include AI systems that incorrectly assess students or teachers, suggest inappropriate content based on erroneous evaluations of

[17] O'Neil, Cathy. 2017. Weapons of Math Destruction. Penguin Books Ltd

[18] https://www.unicef.org/globalinsight/media/1171/file/UNICEF-Global-Insight-policy-guidance-AI-children-draft-1.0-2020.pdf

their abilities, or give them inaccurate feedback.

In addition to presenting serious privacy risks, accuracy issues are particularly pertinent for AI systems that assert to be able to read cognitive or emotional states[19]. This kind of technology is being utilized more and more in other industries, such as analysing video interviews with potential employees[20]. It has also begun to be implemented in educational settings. Crucially, though, a number of researchers have highlighted the unscientific foundation of these technologies as well as their inherent shortcomings[21].

6. Conclusion

The concept of reliability encompasses key components such as consistency, accuracy, stability, and robustness, which collectively contribute to the system's dependability in diverse educational contexts. As AI continues to play a significant role in shaping educational experiences, it is imperative to address various factors that can impact reliability. Factors affecting reliability, including data quality, algorithm complexity, adaptability, user interaction, and system maintenance, highlight the multifaceted nature of maintaining consistent performance in dynamic educational environments. The quality of data used for training, the adaptability of algorithms to changing contexts, and the system's ability to engage users effectively are critical considerations that influence the overall reliability of AI in education. To uphold and enhance reliability, quality assurance measures are essential. Continuous monitoring, regular audits, user feedback mechanisms, adherence to best practices, and robust testing protocols collectively form a comprehensive approach to ensuring the ongoing reliability of AI systems. These measures not only identify potential issues but also contribute to the system's ability to adapt, improve, and align with evolving educational needs. As educational institutions increasingly rely on AI to support and enhance learning experiences, prioritizing reliability becomes a shared responsibility among developers, educators, administrators, and policymakers. Regular assessments, user involvement, and a

[19] Lieberman, Mark. 2018. 'Sentiment Analysis Allows Instructors to Shape Course Content around Students' Emotions'. Inside Higher Ed. 20 February 2018. https://www.insidehighered.com/digital-learning/article/2018/02/20/sentiment-analysis-allows-instructors-shape-course-content

[20] Harwell, Drew. 2019. 'A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job'. Washington Post, 6 November 2019. https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/

[21] Bjørnsten, Thomas Bøgevald, and Mette-Marie Zacher Sørensen. 2017. 'Uncertainties of Facial Emotion Recognition Technologies and the Automation of Emotional Labour'. Digital Creativity 28 (4): 297–307. https://doi.org/10.1080/14626268.2017.1383271

commitment to best practices contribute to building and sustaining reliable AI systems that can positively impact education. The pursuit of reliability in AI-supported education underscores the importance of fostering trust, transparency, and effectiveness in the integration of artificial intelligence within educational ecosystems. Ultimately, a reliable AI infrastructure in education lays the foundation for positive learning outcomes, equitable access to educational resources, and the advancement of educational practices in the digital era.