IGNATIAN
INTERNATIONAL JOURNAL FOR MULTIDISCIPLINARY RESEARCH
www.icceph.com
ISSN 2984-9942

EXPLORING DATA ETHICS AND RESPONSIBLE AI: A LITERATURE REVIEW IN STUDENT LEARNING CONTEXT

Jolly Hope T. Dancila, Sherylle Dawn D. Saramosing,
April John Tom Gomez, Jaylord Maligmat, Gretel T. Ricalde

Department of Information Technology, Misamis University- Oroquieta Unit, Sen. J. Ozamiz St., Pob. 1 Oroquieta City, Misamis Occidental, Philippines

https://doi.org/10.5281/zenodo.18401774

ABSTRACT

This literature review examines the ethical dimensions of Artificial Intelligence (AI) in student learning environments, focusing specifically on fairness, privacy, and accountability. As AI becomes increasingly integrated into education through intelligent tutoring systems, automated assessments, predictive analytics, and learning support tools, concerns regarding algorithmic bias, transparency, and data protection continue to grow. Drawing exclusively from existing literature, case studies, and documented reports from 2021 to 2025, this review synthesizes current findings on how AI can both enhance and challenge ethical practices in education. The analysis reveals that while AI tools have the potential to promote fairness through personalized learning and standardized evaluation, they also present risks related to cultural bias, privacy intrusion, and unequal access. Case studies consistently highlight issues such as over-reliance on AI systems, lack of transparency in decision-making, and disparities in technological resources across learners and institutions. These concerns underscore the importance of designing AI systems with strong ethical safeguards and implementing institutional policies that prioritize responsible use. The findings suggest that the ethical impact of AI in education depends on careful regulation, transparency, AI literacy, and adherence to robust data governance practices. This review contributes to ongoing academic and policy discussions on how educational stakeholders can ensure that AI tools remain trustworthy, equitable, and supportive of meaningful student learning outcomes.

Keywords: *Artificial Intelligence in Education, Data Ethics, Fairness, Privacy, Accountability, Student Learning, Ethical AI, Educational Technology*

INTRODUCTION

Artificial Intelligence (AI) has increasingly become a transformative component of modern education by reshaping teaching strategies, learning processes, and assessment practices through intelligent and data-driven technologies (Harry, 2023). Educational systems worldwide are integrating AI as a central component of digital learning environments to address evolving instructional and learner needs (Imamguluyev et al., 2024). Many educational institutions now adopt AI-powered tools such as intelligent tutoring systems, automated grading platforms, adaptive learning software, and generative AI applications to enhance instructional efficiency (Wang et al., 2024). These technologies support educators by automating routine academic tasks and enabling data-informed instructional decision-making (Mustafa et al., 2024).

AI-driven learning systems facilitate personalized education by adapting content delivery, pacing, and feedback based on individual learner profiles (Bettayeb et al., 2024). Such personalization is especially beneficial in diverse classrooms where students demonstrate varying academic abilities and learning preferences (Anierobi et al., 2025). Learning analytics and predictive modeling enable AI systems to identify learning gaps and recommend targeted interventions to improve student performance (Wang et al., 2024). The emergence of generative AI tools has further expanded opportunities for interactive learning by offering real-time explanations and formative feedback (Zhai et al., 2024).

Studies indicate that students perceive AI-supported learning environments as helpful in improving comprehension of complex academic concepts (Zhang et al., 2024). AI tools are also reported to reduce academic stress by providing immediate assistance outside traditional classroom hours (Mustafa et al., 2024). This level of continuous support is particularly valuable in online and blended learning settings where instructor access may be limited (Wang et al., 2024). As a result, AI technologies are increasingly viewed as academic support mechanisms that enhance learner confidence and autonomy (Zhai et al., 2024).

Despite these educational benefits, the rapid integration of AI into learning environments has raised significant ethical concerns related to student rights and institutional responsibility (Ramnani, 2024). Scholars warn that technological efficiency alone does not guarantee equitable or ethical educational outcomes (Williamson et al., 2023). Without appropriate safeguards, AI systems may unintentionally reinforce existing social and educational inequalities (Holmes, 2023). These risks are more pronounced in under-resourced educational contexts where access to technology and digital literacy remains uneven (Al-Zahrani, 2024).

Fairness is a critical ethical issue in AI-enhanced education because algorithmic systems may perpetuate bias through unrepresentative training datasets (Chinta et al., 2024). AI models trained on limited cultural or socioeconomic data may produce biased academic recommendations and evaluations (Tao et al., 2024). Research demonstrates

that biased AI outputs can negatively affect grading accuracy and student profiling processes (Johnston et al., 2024). Such outcomes disproportionately disadvantage marginalized learner groups and undermine educational equity (Holmes, 2023).

Students have expressed concern that unequal access to AI tools may result in unfair academic advantages among learners (Al-Zahrani, 2024). Differences in AI literacy levels among students further contribute to disparities in learning outcomes (Johnston et al., 2024). These concerns highlight the need for inclusive access and ethical oversight in AI deployment (Ramnani, 2024).

Student data privacy represents another major ethical challenge due to the extensive data collection required by AI systems (Huang, 2023). Educational AI platforms frequently process personal, behavioral, and academic data to generate insights and recommendations (Khan, 2024). The continuous data input required by generative AI tools raises concerns regarding surveillance and informed consent (Chan & Hu, 2023). Inadequate data governance and unclear privacy policies weaken student trust in AI-based learning technologies (Williamson et al., 2023).

Recent studies emphasize that poor transparency in data usage discourages ethical AI adoption in educational institutions (Nazaretsky et al., 2025). Students are more likely to resist AI systems when they are uncertain about how their data are stored and used (Bettayeb et al., 2024). Clear data protection policies are therefore essential to maintain ethical standards in AI-enhanced education (Khan, 2024).

Accountability and transparency further complicate the ethical use of AI in education due to the opaque nature of many algorithmic systems (Nazaretsky et al., 2025). Many AI tools operate as “black boxes” that provide limited explanations for their outputs or recommendations (Pitts & Motamedi, 2025). This lack of explainability reduces student and educator trust in AI-generated decisions (Zhang et al., 2025). Concerns regarding accountability become more severe when AI systems influence high-stakes academic outcomes (Pitts & Motamedi, 2025).

Earlier research on AI-based proctoring systems highlights ethical tensions between academic integrity and student privacy (Nigam et al., 2021). Automated monitoring technologies have raised concerns about excessive surveillance in learning environments (Khan, 2024). These ethical tensions underscore the need for balanced and transparent AI governance frameworks (Williamson et al., 2023).

Scholars argue that the effectiveness of AI in education depends not only on technological innovation but also on ethical governance and institutional responsibility (Mustafa et al., 2024). AI literacy among students and educators is essential for informed and ethical use of AI technologies (Ramnani, 2024). Institutions are encouraged to implement policies that prioritize fairness, transparency, and accountability in AI adoption (Al-Zahrani, 2024). Therefore, this literature review examines scholarly works from 2021 to 2025 to explore how principles of fairness, privacy, and accountability can guide the

ethical and responsible use of Artificial Intelligence in student learning environments (Wang et al., 2024).

Research Questions

1. How can AI tools be designed and implemented to mitigate bias and ensure fairness in diverse student learning environments?
2. What are the current measures in place to protect student data privacy within AI-powered educational platforms, and how effective are they?
3. What accountability frameworks should be established to ensure transparency and responsibility in AI-driven educational decision-making?

Research Objectives

General Objective

To examine how data ethics principles—fairness, privacy, and accountability—can be ensured in the use of Artificial Intelligence within student learning environments.

Specific Objectives

Connected to Fairness: To analyze how AI tools can be designed and implemented to avoid bias and promote fairness among diverse groups of students.

Connected to Privacy: To evaluate the effectiveness of current measures in protecting student data privacy when AI systems are applied in education.

Connected to Accountability: To identify mechanisms and policies that establish clear accountability for AI-driven decisions and outcomes in student learning environments.

Scope and Limitations of the Study

Scope

This study examines the ethical dimensions of Artificial Intelligence (AI) in student learning environments, focusing on fairness, privacy, and accountability through a comprehensive review of existing scholarly literature, published case studies, and institutional reports from 2021 to 2025. The research is limited to secondary sources only and does not involve primary data collection such as surveys, interviews, or school-based participation. By analyzing documented evidence from global and local contexts, the study explores how AI tools influence learning equity, manage personal data, and shape institutional responsibilities. This scope enables the research to draw meaningful insights

from established findings and real-world examples while avoiding direct involvement with students, teachers, or educational institutions.

Limitations

1. The study does not involve the technical design or programming of AI systems; rather, it examines their ethical implications in education.
2. It relies on secondary data sources (literature, case studies, reports) and does not include large-scale experimental implementation of AI tools.
3. The discussion is limited to the three ethical principles highlighted in the title—fairness, privacy, and accountability—and does not cover other dimensions such as transparency, accessibility, or sustainability in depth.

METHODOLOGY

This study employed a qualitative research design to explore issues of data ethics and responsible AI within the context of student learning. The analysis focused on published research studies by scholars in the field, which served as the primary sources of data. Information was gathered using a literature-based survey approach, guided by key questions, including how data ethics and responsible AI are defined, what ethical challenges are highlighted, and what recommendations are offered for their implementation in educational settings. A purposive sampling method was used to select studies that were most relevant to the research questions, ensuring the findings remained closely aligned with the study's objectives. The collected data were examined through thematic analysis, enabling the identification and organization of recurring themes, which were then synthesized to address the research questions. Supporting references from relevant studies were incorporated throughout to strengthen the analysis and situate the findings within the broader academic discourse, enhancing the reliability and credibility of the study.

RESULTS & DISCUSSION

1. Do you believe AI tools in education can help promote fairness among students?

   Survey responses show that many students believe AI tools can help promote fairness in education by offering standardized assessment, timely feedback, and personalized support. This aligns with Wang et al. (2024), who highlighted that AI technologies—including adaptive learning, intelligent assessment, and student profiling—have the potential to tailor instruction and reduce disparities in learning outcomes. However, they also noted challenges in adoption and design, suggesting that fairness depends heavily on how AI is implemented.

Similarly, Mustafa et al. (2024) emphasized that although AI tools support teaching and learning, their benefits are not yet equitably distributed across regions, learner groups, and educational levels. With most AIED studies concentrated in the U.S. and China, fairness gaps may persist on a global scale. Together, these findings suggest that while students perceive AI as promoting fairness, the broader literature warns that equitable design and access remain crucial.

2. Do you use AI tools more often for subjects you find difficult than for subjects you find easy?

The results indicate that students tend to use AI tools more frequently for subjects they find difficult. This is supported by Zhang et al. (2024), who found that academic stress and low self-efficacy increase AI reliance, especially when students expect AI to help them perform better. However, they also warn that overuse can lead to reduced creativity and weakened critical thinking.

Zhai et al. (2024) similarly noted that students often accept AI outputs uncritically when tasks feel challenging, which can adversely affect decision-making and analytical skills.

Anierobi et al. (2025) add that AI can boost confidence and engagement, especially in difficult subjects, but caution that excessive dependence may hinder critical thinking. Overall, the findings support the idea that difficulty level drives AI use, but balanced usage is necessary to preserve cognitive skills.

3. Have you personally noticed AI tools providing less helpful or accurate information for assignments about different cultures or groups?

Responses suggest that some students have encountered culturally biased or less accurate AI outputs. This aligns with Tao et al. (2024), who found that AI systems frequently generate content aligned with Western, English-speaking cultural norms, which can lead to inaccuracies when addressing culturally diverse topics. Although cultural prompting can reduce this bias, it does not eliminate it entirely. This supports the observation that AI tools may be less reliable when dealing with multicultural or culturally specific assignments.

4. Do you believe relying on AI for schoolwork might give some students an unfair advantage?

Survey data shows mixed student opinions on whether AI offers unfair advantages. Johnston et al. (2024) found that students generally

accept the use of basic tools such as grammar checkers but oppose relying on AI for full essay writing, which they consider unfair. Students with higher writing confidence were less supportive of AI use, highlighting uneven academic benefits.

Al-Zahrani (2024) adds that inequities in access to reliable AI tools, as well as issues of transparency and teacher readiness, may create unfair advantages for students with better resources. These findings align with respondents who believe AI can create inequalities if not implemented with clear guidelines.

5. Have you ever felt that an AI tool was asking for too much personal information to let you use it for school works?

Some respondents shared concerns about AI tools requesting personal information. Chan and Hu (2023) similarly reported that students appreciate the benefits of generative AI but remain concerned about data privacy and the extent of personal information required. This finding indicates that privacy concerns are a legitimate barrier to AI adoption among students.

6. Do you feel comfortable using AI tools even if you don't fully understand how they keep your data private?

Survey results indicate that students may still use AI tools despite uncertainty about data protection. Bettayeb et al. (2024) found that ChatGPT enhances engagement and provides helpful support, but students remain concerned about privacy and ethical safeguards. This suggests that comfort levels are mixed and depend on both perceived usefulness and trust in institutional policies.

7. When an AI tool gives you an answer, do you usually check it with another source to make sure it's correct?

Findings suggest that some students do not consistently verify AI outputs. Zhang et al. (2025) found that higher AI trust and AI literacy correlate with greater dependency, sometimes reducing the likelihood of cross-checking AI-generated information. This highlights a need for stronger emphasis on critical evaluation and information verification.

8. Do you trust AI results if the system does not explain how it made its decision?

Most students expressed hesitation to trust AI without clear explanations. Nazaretsky et al. (2025) demonstrated that lack of

transparency reduces trust and willingness to adopt AI tools. Pitts and Motamedi (2025) further explained that trust varies depending on how human-like the AI seems, but opaque decision-making consistently reduces confidence. These findings confirm that explainability is essential for building appropriate trust in AI.

9. Have you used an AI tool to complete a test or quiz when it was prohibited?

While a small number of students admitted the possibility, most respondents indicated they have not used AI in prohibited assessments. This aligns with Johnston et al. (2024), who found that although students are aware of AI tools, the majority oppose using them for tests or quizzes due to academic integrity concerns. This suggests that students generally understand and value ethical academic practices.

Conclusions & Recommendations

The findings of this study show that students recognize both the advantages and challenges of using AI tools in education. Many believe that AI can support fairness through consistent assessment, personalized learning, and timely feedback. Students also tend to use AI more often in subjects they find difficult, relying on it for clarification, confidence-building, and academic support. However, concerns remain regarding cultural bias, privacy, inequitable access, over-reliance, and the lack of transparency in AI decision-making. These issues highlight that while AI has the potential to elevate learning, its benefits depend on responsible, ethical, and well-regulated implementation.

To address these concerns and strengthen the effective use of AI in education, the study recommends that institutions promote AI literacy programs to guide students in verifying information, protecting their privacy, and managing reliance on AI. Schools are also encouraged to develop clear policies that ensure fair access and responsible use of AI tools. Providing students and educators with training on data privacy, ethical guidelines, and the use of explainable AI systems can help build trust and confidence. Additionally, ensuring equitable access to quality AI resources can prevent unfair academic advantages. Lastly, fostering balanced AI use—where AI supports learning without replacing essential cognitive skills—will help students benefit from technology while still strengthening their critical thinking, creativity, and decision-making abilities.

Compliance with Ethical Standards

This research complied with institutional ethical guidelines. Participation was voluntary, with respondents informed of the study's purpose, confidentiality measures, and their right to withdraw at any time. No identifying or sensitive personal data was collected. Data were used solely for academic purposes, and all sources were properly cited to maintain academic integrity.

Acknowledgements

The researchers extend their gratitude to the students of Misamis University–Oroquieta Unit for their participation and cooperation. Special thanks to the faculty members and academic advisors for providing guidance and support throughout the research process. Appreciation is also given to the authors and scholars whose works contributed to the foundation of this study.

REFERENCES

Al-Zahrani, A. M. (2024). Unveiling the shadows: Beyond the hype of AI in education. Heliyon, 10(9).

Anierobi, E. I., Amjad, A. I., & Agogbua, V. U. (2025). Artificial Intelligence Utilization: A Determinant of Academic Self-Efficacy, Engagement, and Satisfaction of Undergraduates. Environment and Social Psychology, 10(3), 3504.

Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024, July). Exploring the impact of ChatGPT: conversational AI in education. In Frontiers in Education (Vol. 9, p. 1379796). Frontiers Media SA.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1), 43.

Chinta, S. V., Wang, Z., Yin, Z., Hoang, N., Gonzalez, M., Quy, T. L., & Zhang, W. (2024). FairAIEd: Navigating fairness, bias, and ethics in educational AI applications. arXiv preprint arXiv:2407.18745. https://doi.org/10.48550/arXiv.2407.18745

Harry, A. (2023). Role of AI in education. Interdisciplinary Journal & Humanity (INJURIY), 2(3).

Holmes, W. (2023). The unintended consequences of artificial intelligence and education. Retrieved from: https://discovery.ucl.ac.uk/id/eprint/10179267/

Huang, L. (2023). Ethics of artificial intelligence in education: Student privacy and data protection. Science Insights Education Frontiers, 16(2), 2577-2587.https://doi.org/10.15354/sief.23.re202

Imamguluyev, R., Hasanova, P., Imanova, T., Mammadova, A., Hajizada, S., & Samadova, Z. (2024). AI-Powered Educational Tools: Transforming Learning in the Digital Era. International Research Journal of Modernization in Engineering Technology and Science, 6, 920–929. https://www.doi.org/10.56726/IRJMETS65040

Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. International Journal for Educational Integrity, 20(1), 2.

Khan, W. N. (2024). Ethical challenges of AI in education: Balancing innovation with data privacy. AI EDIFY Journal, 1(1), 1-13. . Retrieved from: https://researchcorridor.org/index.php/aiej/article/view/238

Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., ... & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. Smart Learning Environments, 11(1), 59.

Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2025). The critical role of trust in adopting AI-powered educational technology for learning: An instrument for measuring student perceptions. Computers and Education: Artificial Intelligence, 8, 100368.

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A systematic review on AI-based proctoring systems: Past, present and future. Education and Information Technologies, 26(5), 6421-6445. . Retrieved from: https://link.springer.com/article/10.1007/s10639-021-10597-x

Pitts, G., & Motamedi, S. (2025). Understanding human-ai trust in education. arXiv preprint arXiv:2506.09160.

Ramnani, S. (2024). Exploring Ethical Considerations of Artificial Intelligence in Educational Settings: An Examination of Bias, Privacy, and Accountability. International Journal of Novel Research and Development (IJNRD), 9(2), 2456-4184. . Retrieved from: https://www.researchgate.net/profile/Rami-Kamak/publication/388890763_Exploring_Ethical_Considerations_of_Artificial_Intelligence_in_Educational_Settings_An_Examination_of_Bias_Privacy_and_Accountability/links/67abba98207c0c20fa84b4ac/Exploring-Ethical-Considerations-of-Artificial-Intelligence-in-Educational-Settings-An-Examination-of-Bias-Privacy-and-Accountability.pdf

Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2024). Cultural bias and cultural alignment of large language models. PNAS nexus, 3(9), pgae346.

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. Expert Systems with Applications, 252, 124167.

Williamson, B., Eynon, R., Knox, J., & Davies, H. (2023). Critical perspectives on AI in education: Political economy, discrimination, commercialization, governance and ethics. In Handbook of artificial intelligence in education (pp. 553-570). Edward Elgar Publishing.https://doi.org/10.4337/9781800375413.00037

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International journal of educational technology in higher education, 16(1), 1-27.

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11(1), 28. Retrieved from: https://radensa.ru/wp-content/uploads/2024/05/Role_of_AI_in_Education.pdf

Zhang, D., Wijaya, T. T., Wang, Y., Su, M., Li, X., & Damayanti, N. W. (2025). Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers. Scientific Reports, 15(1), 14281.

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. International Journal of Educational Technology in Higher Education, 21(1), 34.

APA Citation:

Dancila, J. H. T., Saramosing, S. D. D., Gomez, A. J. T., Maligmat, J., & Ricalde, G. T. (2026). EXPLORING DATA ETHICS AND RESPONSIBLE AI: A LITERATURE REVIEW IN STUDENT LEARNING CONTEXT. Ignatian International Journal for Multidisciplinary Research, 4(1), 873–882. https://doi.org/10.5281/zenodo.18401774

Corresponding authors: *sherylle.saramosing@gmail.com, dancila.hope2001@gmail.com*