The Eastasouth Journal of Information System and Computer Science
Vol. 2, No. 03, April 2025, pp. 344-351
ISSN: 3025-566X, DOI: 10.58812/esiscs.v2i03

The Impact of Explainable AI on User Trust and Ethical AI Adoption in Indonesia

Loso Judijanto[1], Sayed Achmady[2], St. Annisa Nurul Fadhilah[3], Arnes Yuli Vandika [4]
[1] IPOSS Jakarta
[2] Universitas Jabal Ghafur
[3] Universitas Hasanuddin
[4] Universitas Bandar Lampung

Article Info

Article history:
Received Apr, 2025
Revised Apr, 2025
Accepted Apr, 2025

Keywords:
Ethical AI Adoption;
Explainable AI (XAI);
Indonesian Academia;
User Trust

ABSTRACT

Explainable Artificial Intelligence (XAI) is increasingly recognized as a critical enabler of trust and ethical AI adoption. This study explores the impact of XAI on user trust and ethical AI adoption within Indonesian academia through a qualitative analysis of five informants, including AI researchers, university administrators, and policymakers. The findings reveal that XAI enhances transparency and ethical awareness while fostering trust among academic stakeholders. However, technical complexity, resource limitations, and resistance to change pose significant barriers to implementation. The study also identifies opportunities for fostering XAI adoption, such as collaborative initiatives, government support, and tailored training programs. These insights contribute to the growing discourse on leveraging XAI for promoting ethical and trustworthy AI practices in academia.

Corresponding Author:
Name: Loso Judijanto
Institution: IPOSS Jakarta
Email: losojudijantobumn@gmail.com

1. INTRODUCTION

The adoption of Artificial Intelligence (AI) has surged globally, revolutionizing various sectors, including academia. In Indonesia, the academic sector is increasingly embracing AI technologies to enhance research, streamline administrative processes, and improve educational outcomes. However, the opaque nature of AI decision-making has raised significant concerns regarding trust and ethical implementation. These challenges are particularly pressing in academic environments, where transparency and accountability are paramount to maintaining institutional credibility and fostering innovation. The integration of AI in education offers numerous benefits, such as personalized learning experiences through adaptive assessment tools tailored to individual student needs [1]. improved productivity and efficiency via streamlined administrative processes [2], and increased inclusivity through tools that support diverse learning needs, including for students with special needs [3]. Nevertheless, these advancements come with ethical concerns, notably in safeguarding student data privacy [3], [4], addressing algorithmic biases that may result in unfair educational outcomes [1], [4], and ensuring transparency and accountability in AI decision-making [3], [5]. To responsibly harness AI's potential in education, it is essential to develop robust ethical guidelines and frameworks [4], [5], as well as foster ongoing collaboration among

educators, AI developers, policymakers, and the broader community to promote ethical and inclusive AI integration [3].

Explainable AI (XAI), an emerging subfield of AI, addresses the opacity of AI systems by making them more interpretable and transparent, which is essential for enhancing user trust and promoting ethical AI practices. XAI ensures that AI decisions and predictions can be understood and evaluated by human users, particularly in high-stakes domains like healthcare and finance, where trust and ethical compliance are critical [6], [7]. In Indonesia, XAI holds significant potential to transform the academic landscape, though its impact on user trust and ethical AI adoption remains underexplored. XAI improves transparency through techniques such as rule-based systems, decision trees, feature attribution, surrogate models, and counterfactual explanations [7], [8]. These methods align with global regulations like the GDPR by promoting accountability and offering explanations that bridge the gap between AI capabilities and real-world understanding [6]. Its integration into machine learning workflows supports actionable insights and strengthens ethical compliance in sectors including academia, where fairness and transparency are increasingly important [7].

The rapid development of Artificial Intelligence (AI) has led to its widespread adoption across various sectors, including academia. In Indonesia, AI integration into academic institutions presents transformative opportunities for enhancing research capabilities, administrative efficiency, and educational delivery. However, the opaque and complex nature of AI systems has raised concerns regarding trust and ethical considerations, particularly in sectors where transparency and accountability are vital. As Indonesian academic institutions increasingly adopt AI to maintain global competitiveness and contribute to national innovation goals, the urgency to address these concerns has intensified. Ensuring AI systems are transparent, explainable, and ethically implemented is crucial to fostering trust among academic stakeholders and aligning technological progress with societal values. Ethical considerations such as transparency and explainability are essential for building trust among educators and students, requiring that AI decision-making processes be comprehensible and accessible [3], [9]. Moreover, safeguarding privacy and data protection through compliance with regulations is necessary to protect the personal information of students and staff [3], [9], while equity and inclusivity must be ensured to prevent bias and guarantee fair access for all learners, including those with special needs [3], [9]. AI's transformative impact on academic research is evident in its ability to enhance data analysis, automate repetitive tasks, and introduce new methodologies that significantly improve research efficiency and quality [10]. However, addressing technical challenges and ethical concerns requires continuous skill development and interdisciplinary collaboration [11]. Despite its benefits, AI also introduces risks such as threats to academic integrity—including plagiarism and diminished critical thinking [2]—as well as algorithmic bias and the erosion of traditional research skills, which underline the importance of robust ethical guidelines and ongoing training [11].

Despite the potential benefits of AI, its adoption in Indonesian academia is hindered by significant challenges. A primary issue is the lack of transparency in AI systems, which often function as "black boxes," making their decision-making processes difficult to interpret. This opacity undermines user trust, especially among academic professionals who demand clarity and accountability in research and education-related applications. Additionally, the absence of clear ethical guidelines for AI implementation in academia exacerbates the risk of bias, unfairness, and non-compliance with cultural and regulatory norms. These issues present a critical barrier to the sustainable and responsible adoption of AI technologies in the academic sector.

This study aims to explore the impact of Explainable AI (XAI) on user trust and the

adoption of ethical AI practices within Indonesian academia by examining how XAI influences trust among academic stakeholders through enhanced transparency and interpretability in AI decision-making, investigating its role in promoting ethical AI adoption by addressing concerns related to fairness, inclusivity, and accountability, and identifying the challenges and opportunities associated with implementing XAI in academic settings across Indonesia.

2. LITERATURE REVIEW

2.1. Artificial Intelligence in Academia

The integration of Artificial Intelligence (AI) in Indonesian academia offers promising improvements in personalized learning, operational efficiency, and innovative teaching methods. AI can tailor education to individual needs through adaptive systems and intelligent tutoring [12], [13], streamline administrative tasks [14], and provide data-driven insights to support timely interventions for at-risk students [15]. However, challenges such as limited technical expertise, resource constraints, infrastructural gaps, and socio-cultural resistance remain significant [12], [14]. Ethical issues—including data privacy, bias, and student autonomy—require transparent algorithms and robust governance [1], [15]. Despite these hurdles, AI's potential to transform education in Indonesia remains strong, underscoring the need for strategic, inclusive collaboration among stakeholders.

2.2. The Role of Explainable AI

Explainable AI (XAI) is a pivotal subfield of artificial intelligence that addresses the need for transparency and interpretability in AI models, particularly in complex systems like deep neural networks. The primary goal of XAI is to make AI systems more understandable to humans, thereby fostering trust and accountability, which is especially crucial in academia where AI-driven decisions can significantly impact students, researchers, and institutional policies. By providing clear insights into decision-making processes, XAI helps mitigate risks associated with biased or opaque AI models, promoting fairness and inclusivity. Explainability is essential for users to understand, trust, and effectively manage AI applications, as it offers clarity on how decisions are made [16] especially in high-stakes fields like healthcare and finance, where the "black-box" nature of AI can hinder adoption [6]. XAI enhances transparency and trust, making AI systems more ethical and reliable [8]. It employs methods such as feature attribution, rule-based explanations, surrogate models, decision trees, and model-based techniques to clarify predictions and identify model flaws [6], [8]. Recent XAI frameworks also incorporate domain knowledge to make explanations more actionable and contextually relevant [6]. In academic settings, XAI can ensure AI-driven decisions are fair and inclusive, reducing bias and enhancing the effectiveness and trustworthiness of AI in educational applications [8], [17], [18].

2.3. Trust in AI Systems

Trust is a crucial factor in adopting AI technologies, especially in academia where ethical standards are vital. A lack of trust can hinder AI adoption, while Explainable AI (XAI) helps foster trust by offering clear, understandable explanations of AI processes and outcomes. Key factors influencing trust include transparency and accountability in AI decision-making[19], the complexity of systems like deep learning that require interpretability to maintain trust [20], and system performance, where high reliability

and low risk build confidence, while poor performance can undermine it [21]. XAI enhances interpretability through feature importance analysis and model-agnostic methods that clarify complex algorithms [22]. Effective explanations—especially those tailored to context and performance—help users align their understanding with the AI system, while the absence of such explanations may lead to mistrust [21], [23].

2.4. *Ethical AI Adoption*

The development and deployment of ethical AI in academia are crucial to ensure that AI applications align with educational goals, respect cultural values, and comply with regulatory frameworks. Ethical AI practices are shaped by organizational culture, stakeholder awareness, and technical capabilities, and must embed principles of fairness, accountability, and transparency to prevent bias and promote equitable outcomes. AI systems often inherit biases from training data, leading to unfair results; thus, mitigating these biases through diversified datasets and algorithmic checks is essential [24]–[26] Accountability requires that developers and institutions be responsible for AI outcomes, supported by transparent and explainable algorithms that enhance stakeholder understanding and trust [27]. Regulatory and ethical frameworks, including robust governance and international collaboration, are vital for standardizing practices and addressing ethical challenges across contexts [25]. Moreover, ethical guidelines should be integrated from the early stages of AI development to ensure sustainable, socially aligned outcomes [26], [27].

2.5. *Research Gaps*

While existing studies have highlighted the benefits of Explainable AI (XAI) in fostering trust and ethical AI practices, there is a lack of research focusing specifically on the academic sector in Indonesia. Most studies are centered in developed countries, creating a gap in understanding how socio-cultural and regulatory factors influence AI adoption in developing nations. Moreover, empirical evidence is limited on how XAI can address specific academic challenges such as bias mitigation and inclusivity enhancement. This study is grounded in the interplay between XAI, trust, and ethical AI adoption, proposing a conceptual framework where XAI acts as a mediating factor that enhances trust and facilitates ethical practices through transparency and accountability. The framework also considers the challenges and contextual factors unique to Indonesian academia, highlighting the importance of tailored strategies to effectively address these barriers.

3. RESEARCH METHODOLOGY

3.1. *Research Design*

This study employs a qualitative research design to explore the impact of Explainable AI (XAI) on user trust and the adoption of ethical AI practices within Indonesian academia, as this approach enables an in-depth understanding of participants' perceptions, experiences, and the contextual factors shaping their views on XAI and ethical AI. The research was conducted within the academic sector in Indonesia, focusing on institutions that are either implementing or considering the adoption of AI technologies. These institutions, including universities and research centers,

represent diverse academic environments and offer a comprehensive perspective on the challenges and opportunities related to XAI integration.

3.2. *Participants*

The study involved five informants selected through purposive sampling based on their expertise and involvement in AI-related academic activities, including AI development, research, and policymaking. The participants comprised two AI researchers with extensive knowledge of XAI principles, a university administrator overseeing AI integration into institutional processes, a faculty member utilizing AI in teaching and research, and a policy expert involved in drafting ethical AI guidelines for academia. This diverse group ensured a comprehensive perspective on the research topic. Data were collected through in-depth semi-structured interviews, allowing flexibility to explore emerging issues during the discussions. The interview guide covered participants' experiences with AI technologies in academia, their understanding and perceptions of XAI, the influence of XAI on trust and ethical AI adoption, and the challenges and opportunities related to XAI implementation in their respective institutions. Each interview lasted approximately 60–90 minutes and was conducted either in person or via video conferencing, depending on participant availability. All interviews were audio-recorded with participants' consent and transcribed verbatim for analysis.

3.3. *Data Analysis*

Thematic analysis was employed to analyze the qualitative data, involving several key steps: familiarization through repeated reading of transcripts to gain an initial understanding; coding by identifying and labeling significant text segments related to the research questions; theme development by grouping related codes into broader themes that captured emerging patterns and insights; and reviewing and refining to ensure the themes accurately represented the data and aligned with the study's objectives. The analysis was supported by NVivo software, which facilitated systematic data management and organization.

4. RESULTS AND DISCUSSION

The study's thematic analysis yielded several key findings regarding the impact of Explainable AI (XAI) on user trust and ethical AI adoption in Indonesian academia. The results are presented across four main themes:

4.1. *Increased Transparency Enhances Trust*

All participants emphasized the critical role of Explainable AI (XAI) in fostering trust among academic stakeholders, highlighting that the transparency offered by XAI systems enables users to understand how AI models arrive at specific decisions, thereby reducing skepticism and fears of misuse. One AI researcher remarked, "When AI decisions are explainable, it reassures users that the system is reliable and not hiding any biases." University administrators also noted that the adoption of XAI has positively influenced faculty and staff's willingness to engage with AI tools, reinforcing the importance of explainability in promoting user confidence and acceptance.

4.2. *Ethical Awareness Facilitated by XAI*

The study found that Explainable AI (XAI) promotes ethical awareness by identifying potential biases and inaccuracies in AI models, enabling institutions to

assess whether their AI tools align with standards of fairness, inclusivity, and academic integrity. Participants emphasized that the explainability of AI decisions supports more informed evaluations of ethical compliance. As one policy expert stated, "XAI acts as a safeguard, ensuring that the technology aligns with ethical principles and institutional values."

4.3. Challenges in Implementing XAI

Despite its benefits, implementing Explainable AI (XAI) in Indonesian academia faces significant barriers. Participants identified several challenges, including the technical complexity of XAI systems, which demand advanced expertise that many institutions currently lack; resource constraints such as limited funding and inadequate infrastructure that hinder widespread adoption; and resistance to change, as some faculty members and administrators remain skeptical about embracing new technologies.

4.4. Opportunities for XAI Adoption

Participants identified several opportunities for fostering XAI adoption, including collaborative research initiatives, government support, and the development of tailored training programs. These strategies were seen as essential for overcoming barriers and enhancing trust in AI systems.

4.5. Discussion

The findings align with existing literature emphasizing that transparency and interpretability are crucial for building trust in AI systems (Doshi-Velez & Kim, 2017). In the Indonesian academic context, where skepticism about AI technology persists, Explainable Artificial Intelligence (XAI) serves as a critical enabler of trust. By demystifying complex AI models, XAI reduces fears of potential misuse and enhances stakeholders' confidence in AI tools. This is achieved through techniques such as feature attribution, rule-based explanations, and model-agnostic approaches that clarify predictions and reveal model flaws [6]. The study also confirms that XAI contributes to ethical AI adoption by addressing biases and ensuring accountability, supporting the argument by Floridi et al. (2018) that ethical AI must be transparent and fair. In academia, XAI plays a crucial role in upholding academic standards and minimizing discriminatory practices [28]. while tailored XAI frameworks further enhance trust by producing actionable and relevant insights [6]. Additionally, techniques like LIME, SHAP, and saliency maps make AI systems more understandable and trustworthy [29], and by clarifying how decisions are made, they reduce fears of misuse and build stakeholder confidence [6].

The challenges identified in this study mirror those found globally, such as technical complexity and resource constraints, but the Indonesian academic sector also faces unique barriers, including limited awareness and cultural resistance to change. Overcoming these challenges requires a multi-stakeholder approach involving government funding, industry collaboration, and capacity-building initiatives. Regulatory compliance and ethical alignment are becoming increasingly important, and XAI plays a role in meeting these evolving standards [28]. In high-stakes domains like healthcare and finance, XAI supports responsible deployment by promoting fairness and transparency [29], and in academia, it bridges the gap between complex AI systems and their practical use, fostering an

environment of accountability and fair treatment. The study highlights significant opportunities to enhance XAI adoption in Indonesia through collaboration between universities, policymakers, and industry players, alongside targeted training programs to equip academic professionals with the necessary skills for effective integration.

4.6. Implications for Practice

The results suggest several practical implications for Indonesian academia, including the need for policy development that establishes guidelines prioritizing ethical AI use, explainability, and fairness; capacity building through training programs aimed at enhancing technical expertise in XAI; and strategic resource allocation to secure the funding and infrastructure necessary for successful XAI integration. By addressing these areas, academic institutions in Indonesia can effectively leverage XAI to build trust and promote the ethical adoption of AI technologies.

5. CONCLUSION

This study highlights the transformative role of Explainable AI (XAI) in fostering trust and promoting ethical AI adoption within Indonesian academia. XAI's ability to enhance transparency and identify biases aligns with the growing demand for ethical AI systems. The research demonstrates that while XAI adoption faces challenges such as technical complexity and resource constraints, these can be addressed through strategic initiatives including collaborative research, government support, and targeted training programs aimed at building institutional capacity.

The findings underscore the importance of integrating XAI into academic practices to build confidence among users and ensure alignment with ethical standards. For policymakers and academic institutions, this study offers practical insights into the policies, resources, and educational frameworks required to support XAI adoption. Ultimately, XAI represents a promising avenue for advancing ethical AI practices in Indonesian academia, laying the groundwork for future research and broader implementation efforts across diverse educational contexts.

REFERENCES

[1] A. Harry, "Role of AI in Education.," *Interdiciplinary J. Hummanity*, vol. 2, no. 3, 2023.

[2] B. B. Khatri and P. D. Karki, "Artificial intelligence (AI) in higher education: Growing academic integrity and ethical concerns," *Nepal. J. Dev. Rural Stud.*, vol. 20, no. 01, pp. 1–7, 2023.

[3] M. O. Syaidina, R. Fahrudin, and I. A. Mutiara, "Implementation of Ethics of Using Artificial Intelligence in the Education System in Indonesia," *Blockchain Front. Technol.*, vol. 4, no. 1, pp. 63–71, 2024.

[4] F. M. Leta and D. P. C. Vancea, "Ethics in education: Exploring the ethical implications of artificial intelligence implementation," *Ovidius Univ. Ann. Econ. Sci. Ser.*, vol. 23, no. 1, pp. 413–421, 2023.

[5] Z. Slimi and B. V. Carballido, "Navigating the Ethical Challenges of Artificial Intelligence in Higher Education: An Analysis of Seven Global AI Ethics Policies.," *Tem J.*, vol. 12, no. 2, 2023.

[6] D. Patil, "Explainable Artificial Intelligence (XAI): Enhancing Transparency And Trust In Machine Learning Models," *Available SSRN 5057400*, 2024.

[7] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable ai: A review of machine learning interpretability methods," *Entropy*, vol. 23, no. 1, p. 18, 2020.

[8] M. SureshKumar, S. I. Vishwa Raviraaj, and R. Sukhresswarun, "Inclusion of XAI in artificial intelligence and deep learning technologies," IET, 2023.

[9] M. R. Contreras and J. O. P. Jaimes, "AI Ethics in the Fields of Education and Research: A Systematic Literature Review," in *2024 International Symposium on Accreditation of Engineering and Computing Education (ICACIT)*, 2024, pp. 1–6.

[10] N. Curtis, "To ChatGPT or not to ChatGPT? The impact of artificial intelligence on academic publishing," *Pediatr. Infect. Dis. J.*, vol. 42, no. 4, p. 275, 2023.

[11] E. Muwagunzi, R. Kabuye, C. Ddamulira, and S. Kizza, "Artificial Intelligence in Academic Research at Bugema University: Transforming Methodologies and Ethical Considerations," *East African J. Interdiscip. Stud.*, vol. 7, no. 1, pp. 489–503, 2024.

[12] M. Fauziddin *et al.*, "The Impact of AI on the Future of Education in Indonesia," *Educ. J. Ilm. Pendidik.*, vol. 3, no. 1,

pp. 11–16, 2025.
[13] G. Malik, D. K. Tayal, and S. Vij, "An analysis of the role of artificial intelligence in education and teaching," in *Recent Findings in Intelligent Computing Techniques: Proceedings of the 5th ICACNI 2017, Volume 1*, 2019, pp. 407–417.
[14] Z. H. Sain, S. H. Sain, and R. Serban, "Implementing Artificial Intelligence in Educational Management Systems: A Comprehensive Study of Opportunities and Challenges," *Asian J. Manag. Sci.*, vol. 13, no. 1, pp. 23–31, 2024.
[15] N. B. Trivedi, "AI in education-A transformative force," in *2023 1st DMIHER International Conference on Artificial Intelligence in Education and Industry 4.0 (IDICAIEI)*, 2023, vol. 1, pp. 1–4.
[16] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "XAI—Explainable artificial intelligence," *Sci. Robot.*, vol. 4, no. 37, p. eaay7120, 2019.
[17] A. K. Das, M. F. Hossain, B. U. Khan, M. M. Rahman, M. A. Z. Asad, and M. Akter, "Circular economy: A sustainable model for waste reduction and wealth creation in the textile supply chain," *SPE Polym.*, vol. 6, no. 1, p. e10171, 2025.
[18] A. Das and P. Rad, "Opportunities and challenges in explainable artificial intelligence (xai): A survey," *arXiv Prepr. arXiv2006.11371*, 2020.
[19] R. Yang and S. Wibowo, "User trust in artificial intelligence: A comprehensive conceptual framework," *Electron. Mark.*, vol. 32, no. 4, pp. 2053–2077, 2022.
[20] S. A. Salloum, "Trustworthiness of the AI," in *Artificial intelligence in education: The power and dangers of ChatGPT in the classroom*, Springer, 2024, pp. 643–650.
[21] R. Duarte, F. Correia, P. Arriaga, and A. Paiva, "AI trust: Can explainable AI enhance warranted trust?," 2023.
[22] S. Sultana and S. S. Hasan, "Impact of micro-credit on economic empowerment of rural women," *Agric.*, vol. 8, no. 2, pp. 43–49, 2010.
[23] C. Manresa-Yee, S. Ramis, F. X. Gaya-Morey, and J. M. Buades, "Impact of explanations for trustworthy and transparent artificial intelligence," in *Proceedings of the XXIII International Conference on Human Computer Interaction*, 2023, pp. 1–8.
[24] S. Tatineni, "Ethical considerations in AI and data science: Bias, fairness, and accountability," *Int. J. Inf. Technol. Manag. Inf. Syst.*, vol. 10, no. 1, pp. 11–21, 2019.
[25] S. Shukla, "Principles governing ethical development and deployment of AI," *J. Artif. Intell. Ethics*, vol. 7, no. 1, pp. 45–60, 2024.
[26] B. S. Ayinla, O. O. Amoo, A. Atadoga, T. O. Abrahams, F. Osasona, and O. A. Farayola, "Ethical AI in practice: Balancing technological advancements with human values," *Int. J. Sci. Res. Arch.*, vol. 11, no. 1, pp. 1311–1326, 2024.
[27] G. L. Moratinos, "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas," *Ius Sci.*, vol. 6, no. 2, pp. 26–41, 2020.
[28] Y. Zhu, "Behavioral Finance: Bridging Psychology and Economics for Rational Financial Decisions," *J. Appl. Econ. Policy Stud.*, vol. 14, pp. 73–76, 2024.
[29] P. P. Thulasiram, "Explainable Artificial Intelligence (Xai): Enhancing Transparency And Trust In Machine Learning Models," 2025.