

## PAPER

# University Students' Trust in AI: Examining Reliance and Strategies for Critical Engagement

Muhammad Abuzar()  
Mahmudulhassan,  
Muthoifin

Muhammadiyah  
University of Surakarta,  
Surakarta, Indonesia

[s300230013@student.ums.ac.id](mailto:s300230013@student.ums.ac.id)

## ABSTRACT

The rising use of artificial intelligence (AI) for academic tasks, such as exam preparation and assignment assistance, has increased concerns regarding students' uncritical reliance on AI technologies in university settings. This mixed-method study investigates the factors driving this reliance, drawing on qualitative data from 18 Indonesian students and quantitative data from 328 students across 18 universities. It proposes strategies for promoting critical engagement with AI. Findings show that while students generally trust AI for its efficiency, evidenced by an average score of 3.89 on the trust in technology scale (TTS), their confidence declines when tasks demand nuanced human judgment. This over-reliance poses risks to critical thinking and may encourage cognitive complacency; however, many students counter this by cross-referencing AI-generated information with credible sources. To mitigate blind trust and promote responsible AI use, this study recommends enhancing AI literacy through workshops, promoting effective prompt crafting and output verification strategies, and implementing institutional policies to guide ethical AI use. By integrating reminder prompts and AI literacy modules into curricula, institutions can encourage critical engagement with AI, ensuring it enhances rather than undermines educational outcomes. These practical guidelines aim to enrich students' academic experiences while addressing the ethical and cognitive challenges associated with uncritical AI reliance.

## KEYWORDS

artificial intelligence (AI), higher education, trust in technology, generative AI

## 1 INTRODUCTION

The rapid advancement of artificial intelligence (AI) has revolutionized various aspects of modern life, embedding itself into decision-making processes across multiple sectors. Over the past decade, AI technologies have become increasingly sophisticated and pivotal in healthcare, finance, education, and social media. For instance, AI-driven tools in healthcare are instrumental in diagnosing diseases, predicting patient outcomes, and personalizing treatment plans, significantly enhancing the

Abuzar, M., Mahmudulhassan, Muthoifin. (2025). University Students' Trust in AI: Examining Reliance and Strategies for Critical Engagement. *International Journal of Interactive Mobile Technologies (IJIM)*, 19(7), pp. 70–82. <https://doi.org/10.3991/ijim.v19i07.52875>

Article submitted 2024-10-14. Revision uploaded 2024-12-17. Final acceptance 2024-12-17.

© 2025 by the authors of this article. Published under CC-BY.

quality and efficiency of care [1]. Similarly, in finance, AI is utilized for credit scoring, fraud detection, and algorithmic trading, executing decisions at speeds and complexities that surpass human capabilities [2]. While AI has transformed these fields by assisting expert decision-making, its role in academia is unique. AI influences students' learning and development, raising specific concerns regarding critical thinking and skill acquisition [3]. For example, AI-powered educational systems offer personalized learning experiences by adapting content to meet individual students' needs [37]. Social media platforms, on the other hand, employ AI algorithms to curate content, drive user engagement, and shape public discourse [5]. This extensive integration underscores AI's profound influence on personal and professional spheres.

Blind trust in AI refers to the uncritical acceptance of AI-generated information or decisions without sufficient evaluation or skepticism. This phenomenon is characterized by a lack of questioning or verification of AI outputs, driven by the perception of AI as highly accurate or infallible. Psychological factors such as the perceived objectivity of AI, cognitive ease, and the delegation of complex tasks to AI systems contribute to this uncritical reliance, ultimately diminishing personal accountability and critical thinking [6]. Despite its benefits, the lack of crucial engagement among students in academic settings suggests a gap in AI literacy that could hinder the development of independent critical skills, emphasizing the need to foster critical engagement in educational contexts [36]. Unlike professional environments where AI assists expert-driven tasks, academic settings require students to develop judgment and evaluative skills actively, making blind reliance on AI a potential hindrance to educational goals [8].

AI-generated data has become pervasive, significantly influencing decisions at both individual and organizational levels [4]. Complex algorithms process vast amounts of information to generate highly reliable insights, predictions, and recommendations [9]. In healthcare, AI-generated data informs the likelihood of patient outcomes and guides clinical decisions [10]. Similarly, in finance, AI-produced risk assessments and credit scores shape lending decisions and investment strategies [11]. Social media platforms deploy AI to generate personalized content feeds, profoundly impacting user perceptions and opinions by controlling the flow of information [12]. This increasing reliance highlights its role in shaping critical decisions. It raises concerns about the growing trend of unquestioned acceptance of such data, particularly in education, where students' independent judgment is paramount [38]. Blind trust in AI poses significant risks across various domains, warranting deeper exploration. Uncritical reliance on AI can propagate misinformation, especially when AI systems are trained on biased or incomplete data. The absence of critical evaluation increases the risk of accepting and disseminating incorrect information [7, 13]. Blind trust also exacerbates ethical concerns, particularly when AI systems are used in sensitive decision-making processes such as hiring, criminal justice, or medical diagnosis. AI can reinforce biases or make decisions that lack transparency, leading to unjust outcomes [14]. This over-reliance may obstruct critical academic skills such as independent judgment, potentially leading students to become passive consumers of AI outputs rather than critical evaluators [39].

This study addresses the growing issue of uncritical and excessive reliance on AI as AI-generated data increasingly influences decision-making in academic settings. Accepting AI outputs without scrutiny, especially when systems are trained on biased or incomplete data, can lead to misinformation, flawed decisions, and significant errors [15, 16]. While much of the existing research has highlighted AI's usability and benefits in education, it has largely overlooked the risks associated with students' blind trust in AI, particularly in sensitive academic contexts such as exams and assignments. This gap in research creates a need for studies that specifically

address how to mitigate the ethical challenges of blind trust and overconfidence in AI, especially considering the lack of understanding of AI's inherent biases [37].

This study investigates the factors that drive students' reliance on AI, including their digital competencies and motivations for use. It also evaluates existing educational approaches to AI literacy to identify interventions to enhance critical thinking and evaluative skills. The objectives of this study are to identify the key factors contributing to students' blind trust in AI and to propose structured interventions that enhance students' ability to assess AI outputs critically. By addressing these objectives, the study provides recommendations to assist educational institutions in developing policies and curricula to safeguard academic integrity while enriching students' educational experiences. The following Figure 1 illustrates the conceptual framework and aim of the study.

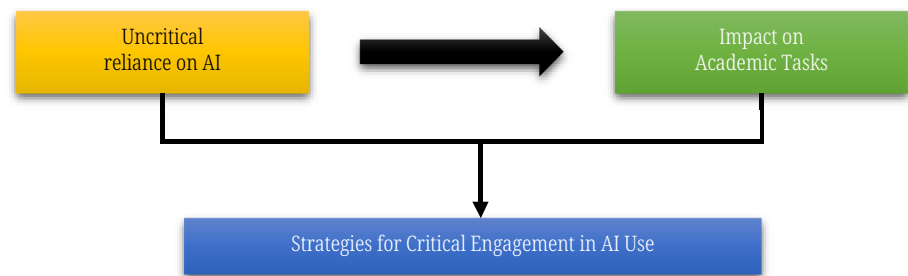


Fig. 1. Conceptual framework

## 2 METHODOLOGY AND DATA COLLECTION

### 2.1 Qualitative approach

The qualitative component of this study sought to explore students' experiences and perceptions of using AI in academic contexts. A purposive sampling method was employed to ensure a diverse representation of participants. This method was chosen because it allows the intentional selection of individuals who meet specific inclusion criteria and can provide detailed insights relevant to the research objectives. The sample included 18 students currently studying in Indonesia, comprising local and international students. Data saturation was achieved when no new themes, insights, or information emerged from the interviews. The study followed Lochmiller's (2016) recommendations for qualitative research, emphasizing depth over breadth to uncover nuanced experiences [16].

Participants were selected based on specific inclusion criteria, which required them to be currently enrolled students in Indonesia, actively using AI for academic purposes, providing informed consent, and demonstrating a willingness to participate in an interview. Participants were excluded if they did not use AI tools for educational purposes or could not provide informed consent. Recruitment occurred via social media platforms (e.g., WhatsApp and Facebook) and personal networks. Efforts were made to ensure diversity in gender, academic background, and ethnicity to provide a broad perspective. Of the 18 participants, eight were female, and 10 were male, representing various academic disciplines and educational levels. Semi-structured interviews were conducted to capture participants' in-depth views. This method was selected because it allows flexibility in exploring emerging themes while maintaining focus on the research objectives. Interviews were scheduled at participants' convenience and conducted online (via Zoom and WhatsApp) or in

informal settings such as university dormitories to ensure comfort and openness. Interviews lasted 10–20 minutes and were conducted in English or Indonesian, based on participants' preferences. To ensure consistency and minimize interviewer bias, the primary researcher and a co-author conducted the interviews using a pre-defined set of open-ended questions. These questions encouraged participants to share their experiences, perceptions, and challenges regarding AI use in academic settings. All interviews were audio-recorded with participants' consent and transcribed verbatim for analysis.

The data were analyzed using thematic analysis, adhering to Braun and Clarke's six-phase framework, chosen for its flexibility and ability to identify patterns across qualitative datasets. The process began with familiarization through repeated reading of transcripts, followed by systematic coding of relevant features. These codes were then grouped into potential themes, reviewed, and refined to ensure coherence. Subsequently, themes were defined and named, culminating in producing a detailed report featuring compelling extracts. The primary researcher and a co-author conducted the analysis collaboratively to enhance accuracy and reliability. Themes were further refined to align with the research questions and existing literature, providing a comprehensive understanding of the participants' experiences.

## 2.2 Quantitative approach

The quantitative component aimed to investigate students' trust in AI tools for academic purposes and to examine the relationship between trust and demographic variables such as age, education level, and gender. A stratified simple random sampling method was employed to ensure representative participation across faculties, academic levels (Bachelor's, Master's, and Ph.D.), and genders. This method was chosen to ensure proportional representation and enhance the findings' generalizability across diverse student populations.

Students who did not meet these criteria were excluded. Recruitment was conducted via online platforms (e.g., WhatsApp and Facebook) to maximize reach and efficiency. The demographic distribution of participants is detailed in Figure 2, which provides insights into variables such as age, gender, and education level. Bachelor's students comprised the majority (58%), followed by Master's (34%) and Ph.D. students (8%). The diverse sample was 192 male and 136 female participants, offering a balanced perspective across genders.

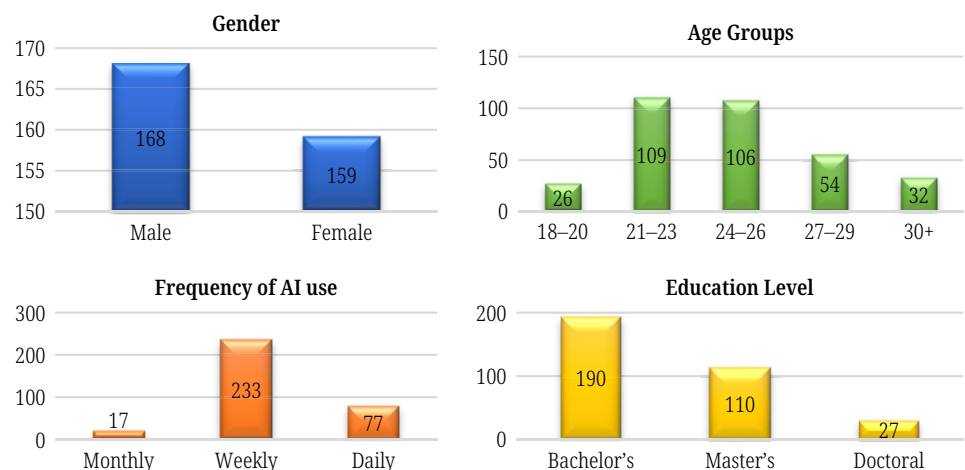


Fig. 2. Respondents distribution

Data collection occurred from August to September 2024 using a structured online survey administered via Google Forms. This platform facilitated efficient and secure data collection from geographically dispersed participants. Reminders were sent periodically to ensure a high response rate. Ethical approval was obtained from institutional review boards, and confidentiality protocols were strictly adhered to, including secure storage of participant data. The survey included an adapted version of the trust in technology scale (TTS) developed by McKnight et al. [19] to measure trust in AI systems. This 17-item scale assessed perceived competence, risk, and user familiarity with AI tools. Responses were measured on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Demographic data such as age, gender, and frequency of AI use were also collected to support subgroup analysis.

Data analysis was conducted using statistical package for the social sciences (SPSS) version 21, employing a range of statistical techniques to address the research objectives comprehensively. Descriptive statistics summarized demographic information and trust scores, providing an overview of the sample characteristics. Pearson correlation was employed to examine the strength and direction of linear relationships between trust in AI and demographic factors, as it is well-suited for continuous variables. Multiple regression analysis explored how age, gender, and education level predicted trust in AI systems, highlighting their relative influence. An analysis of variance (ANOVA) was conducted to investigate significant differences in trust scores among demographic subgroups, enabling the comparison of mean differences across multiple groups. These statistical methods were chosen for their robustness and alignment with the study's objectives.

### 3 RESULTS

#### 3.1 Qualitative analysis

**A) Key factors shaping students' trust in AI:** Table 1 summarizes the key factors influencing students' trust in AI, supported by participant statements and corresponding descriptions.

**Table 1.** Key factors shaping students' trust in AI

Factor	Participant Statement	Description
Trust in Data Validity	<i>I think the answers are already valid, as they come from existing data.</i>	Students trust AI because it draws from valid, existing data, leading to a perception of reliability.
Social Influence	<i>Many people use it, and there's social influence that makes it more trustworthy.</i>	Widespread use by others builds social trust in AI, making it more reliable through peer influence.
Time-Saving and Accessibility	<i>You can find information quickly. Sometimes, you can't find everything in books or the library, but AI provides that information easily. It also saves time.</i>	AI provides quick access to information, making it a more efficient option than traditional resources such as books or libraries.
Ease of Use and Effort Reduction	<i>It's easy to access, saves time, is user-friendly, and trustworthy. /It's easier to use, and we don't need to expend much effort.</i>	The ease of use and effort-saving nature of AI platforms encourage students to trust and rely on them.
Partial Trust and Human Judgment	<i>I don't completely trust AI because teachers know best. I don't use it fully since our assignments would lack a human touch.</i>	Students express skepticism and believe human expertise is still superior to AI. They prefer partial reliance, ensuring a human touch in academic work.
Helping Tool for Research	<i>I use AI as a helping tool for research but don't fully rely on it. It offers a good framework for research but needs human verification.</i>	Students see AI as a supplementary tool for research, useful for guidance and frameworks, but still requiring human oversight and validation.

## B) Risks of AI-assisted learning

**Reduced critical thinking and creativity:** Many students reported that the convenience of AI leads to a *spoon-feeding* effect, where answers are easily provided without much effort. This results in less engagement in independent thought and analysis. One student stated, *AI has affected my critical thinking and reduced creativity because we don't use our minds as much every day*. Another participant added, *AI destroys our ability to work hard and makes us lazy*. Studies have also highlighted that excessive dependence on AI tools in education may impair students' ability to think critically and creatively [20].

**Over-reliance and blind trust:** Students frequently express concerns about excessive trust in AI, which can lead to academic errors and diminished critical engagement. One participant noted, *I don't trust it outright; I always double-check with other sources and verify the references*. Another emphasized, *We should show people what AI can and cannot do and not rely on it for everything*. This reflects *blind trust*, where overconfidence in AI risks misinformation and uncritical acceptance of outputs [21]. Such reliance also negatively impacts academic performance, reducing motivation and independent thinking [17]. One student remarked, *AI reduces our ability to work hard and impacts our studies*, while another highlighted the shift to *copy-pasting information without fully engaging with the content*. Studies corroborate that AI can promote shortcuts, bypassing critical research and more profound learning opportunities, thus weakening academic rigor [19].

**Quality and accuracy concerns:** Participants also raised concerns about the accuracy of AI-generated responses, with some students finding that AI can provide incomplete or exaggerated answers. One participant noted, *AI sometimes exaggerates the answers, and I make corrections on my own after reviewing references*. Another explained, *About 70% of the answers are okay, but I believe it depends on the prompt. The better the prompt, the better the answers*. The participants' concerns about verifying AI-generated information echo the literature that emphasizes the importance of critical evaluation when using AI tools to avoid errors, misinformation, or reliance on incomplete data [10].

## C) Student's perceptions of AI competency and role in academic decision-making

**AI as a reliable but limited tool:** Students recognized AI as a reliable resource for handling time-sensitive academic tasks and appreciated its ability to expedite research and repetitive tasks. However, they emphasized that AI's reliability depends on clear, structured user input, and its limitations lie in providing nuanced judgment or deeper analysis. Many participants viewed AI as a supplementary tool rather than a replacement, highlighting the importance of human oversight and the risk of over-reliance. As one student noted, *AI makes things easier, but it's important not to let it do all the thinking for you*. This aligns with research indicating AI's value in streamlining tasks while cautioning against its overuse, which could erode critical thinking and academic rigor [18].

**AI as a supplement, not a replacement, and supportive role in research:** Many students expressed that AI functions best as a supplemental tool rather than a primary resource. They appreciated its competency in offering broad overviews or helping with repetitive tasks such as generating initial ideas or performing translations. However, participants stressed the importance of integrating human input, particularly when deeper analysis or critical thinking was required. One participant remarked, *AI is great for structuring ideas, but you still need human judgment for anything beyond surface-level understanding*. This reflects concerns in the literature, particularly the work of Marcus and Davis, who argue that while AI can streamline certain academic tasks, its lack of deeper cognitive

abilities makes human oversight essential in more complex decision-making processes [27]. AI was widely viewed as supportive in academic decision-making, particularly in research contexts [22]. This caution is supported by research that discusses the importance of triangulating AI outputs with multiple sources to mitigate errors and biases [29].

**Over-reliance vs. balanced use:** Concerns about over-reliance on AI surfaced in the interviews. Some students feared that depending too heavily on AI could erode critical thinking skills and diminish academic rigor. This reflects research by Long and Magerko, who argue that overuse of AI can hinder creative and independent thought processes [30]. Participants described their use of AI as helpful when used appropriately. As one student explained, *AI makes things easier, but it's important not to let it do all the thinking for you*. These findings are consistent with scholars such as Noble, who emphasize that AI should be viewed as a complement to human judgment, not a substitute [31].

#### D) Strategies to mitigate blind trust and over-reliance on AI

**Consulting external sources:** Participants frequently highlighted the importance of cross-referencing AI-generated outputs with reliable external sources such as books, articles, or expert opinions. This approach was seen as a way to mitigate the potential inaccuracies in AI responses [24]. One participant said, *I go to the library and verify the answers with valid sources*, illustrating a proactive stance toward verifying AI's results. This practice ensures that AI is used as a supplemental tool rather than the sole source of information.

**Importance of prompt formulation:** Several participants emphasized the relationship between the quality of AI responses and the input provided. They pointed out that AI's accuracy heavily depends on the clarity and specificity of the prompt. One participant noted, *The better the prompt, the better the answers. Accuracy depends on how we ask*. Training on crafting precise prompts could improve the effectiveness of AI tools and reduce reliance on AI-generated outputs alone [23].

**Educational initiatives for responsible AI usage:** Participants highlighted the need for workshops and seminars to improve their understanding of AI's strengths, limitations, and responsible usage. They noted that such sessions could reduce blind trust and foster critical engagement. One participant emphasized, *Increase awareness through workshops, show the negative impact of AI, and teach how to use AI effectively*. They also called for institutional policies to regulate AI usage, ensuring it supplements rather than replaces core academic practices. Another participant suggested, *Universities should create and enforce strict policies to reduce the use of AI*. Assignments emphasizing critical thinking and personal insights were recommended to balance technological reliance with human creativity, ensuring outputs feel more *humanized*.

**Pop-up notifications in AI systems:** Some participants suggested that AI tools could integrate pop-up reminders to alert users about the potential inaccuracies in AI-generated outputs. This feature would be a constant reminder that AI is fallible and should not be unquestioningly trusted [25, 26]. One participant proposed, *AI tools should include more pop-up notifications reminding users that AI cannot be fully trusted*.

## 3.2 Quantitative analysis

**A) Perceptions of AI systems and trust in AI:** The findings, illustrated in Figure 3, indicate that participants generally trust AI systems for academic tasks, as shown by high scores for the belief that AI produces accurate outputs ( $M = 3.89$ ,  $SD = 0.739$ ) and its reliability for academic decisions ( $M = 3.73$ ,  $SD = 0.830$ ). These results align

with previous research suggesting that users view AI as dependable for structured, routine tasks [28, 32, 33]. However, trust diminishes when handling complex tasks ( $M = 3.50$ ,  $SD = 0.952$ ), reflecting concerns about AI's limitations in addressing ambiguous or nuanced challenges. Participants also demonstrated skepticism toward unquestioningly accepting AI outputs ( $M = 2.76$ ,  $SD = 0.912$ ), underscoring a preference for human oversight, consistent with findings on the critical role of human judgment in decision-making [34]. Despite this caution, moderate trust in high-stakes situations ( $M = 3.39$ ,  $SD = 1.195$ ) suggests a risk of “automation bias,” where users may over-rely on AI under pressure, highlighting the need for interventions to foster critical evaluation and informed use of AI systems [35].

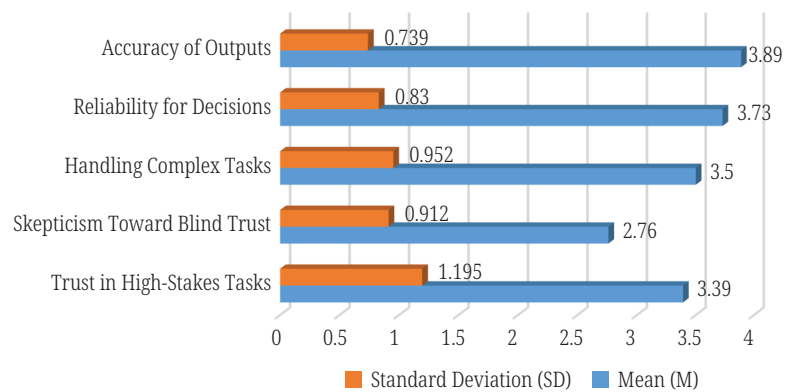


Fig. 3. Perceptions of AI systems and trust in AI

**B) Judgment of AI outputs and awareness of AI limitations:** The results illustrated in Figure 4 reveal moderate confidence among participants in evaluating AI outputs without understanding underlying algorithms ( $M = 3.31$ ,  $SD = 1.211$ ) and trust in their ability to judge the correctness of AI-generated results ( $M = 3.84$ ,  $SD = 0.890$ ). This indicates a general sense of self-efficacy in assessing AI, consistent with Lathia's findings that users feel capable of evaluating AI outputs despite lacking technical expertise [36]. However, participants reported lower confidence in recognizing AI biases ( $M = 2.79$ ,  $SD = 1.189$ ) and an even lower tendency to actively consider biases when using AI outputs in academic tasks ( $M = 2.47$ ,  $SD = 1.156$ ), highlighting a critical gap in awareness. These findings align with Wong et al.'s research on users' limited understanding of AI biases [37]. Encouragingly, participants strongly supported AI bias training ( $M = 3.91$ ,  $SD = 1.082$ ), suggesting a willingness to adopt more critical approaches to AI reliance when provided with proper education, echoing Mitra's perspective on the importance of training to mitigate blind trust [38].

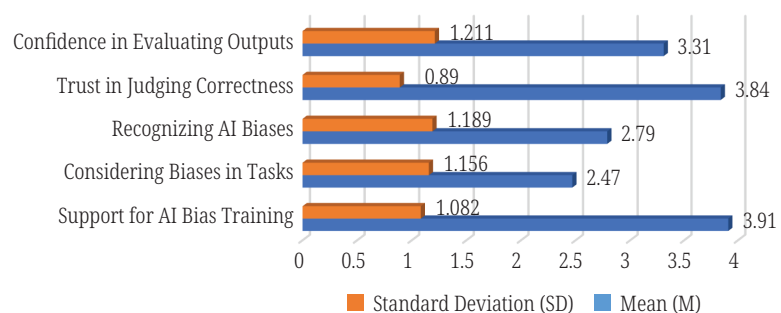


Fig. 4. Judgment of AI outputs and awareness of AI limitations

**C) Ethical considerations and risks and engagement with AI:** As shown in Figure 5, the analysis of responses highlights a cautious stance among participants toward ethical considerations and risks in AI usage. Moderate concern was expressed regarding accountability when relying on AI ( $M = 3.68$ ,  $SD = 0.606$ ), alongside a strong belief that uncritical reliance on AI can lead to misinformation in academic work ( $M = 4.33$ ,  $SD = 0.618$ ), consistent with studies emphasizing the dangers of misinformation when AI outputs are accepted without scrutiny [19]. Trust in AI outputs, however, remains low ( $M = 2.39$ ,  $SD = 1.221$ ), reflecting skepticism about the reliability of AI-generated information, which aligns with prior research [39]. Participants strongly indicated that increased awareness of AI limitations would reduce their reliance on these systems ( $M = 4.46$ ,  $SD = 0.703$ ). They agreed on the need for critical engagement to ensure ethical and accurate academic work ( $M = 4.49$ ,  $SD = 0.691$ ). These findings underscore the importance of fostering critical thinking and awareness to mitigate risks, a sentiment echoed by Sweeney, who advocates for educational initiatives to enhance understanding of AI's limitations [35].

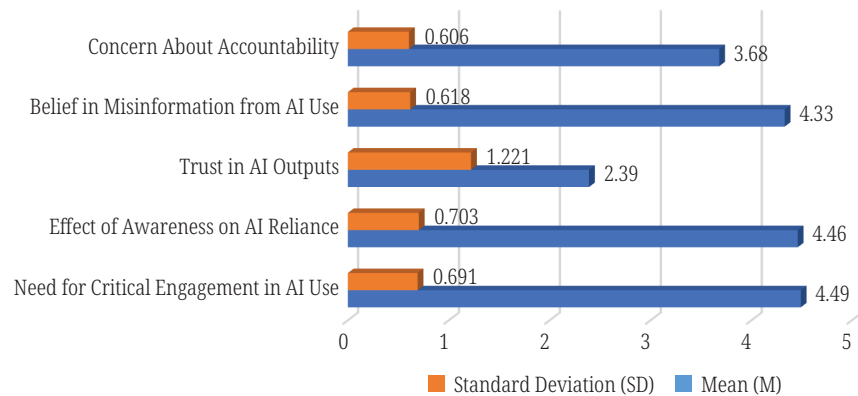


Fig. 5. Ethical considerations and risks and engagement with AI

## 4 DISCUSSION

The findings of this study highlight a complex relationship between students' trust in AI and their reliance on it for academic tasks. While students appreciate AI's speed and efficiency ( $M = 3.89$ ), trust diminishes for tasks requiring nuanced human judgment ( $M = 3.50$ ). This underscores the boundary between the utility of AI and the necessity of human oversight, as reflected in the sentiment, "AI can't replicate a professor's insight." These results align with prior research, emphasizing AI's limitations in mimicking expert-level reasoning and affirming its role as an augmentative tool rather than a replacement for human expertise.

The study also uncovered students' susceptibility to automation bias ( $M = 3.39$ ), particularly under high-pressure situations where stress exacerbates blind trust in AI outputs. While some students reported verifying AI-generated results ( $M = 2.76$ ), qualitative data revealed instances of uncritical reliance. For example, participants described scenarios where tight deadlines led them to rely solely on AI for essay drafting or research summaries, resulting in overlooked inaccuracies. These findings suggest that the mechanisms of cognitive ease and the delegation of complex decision-making to AI systems may inadvertently hinder critical thinking and academic performance. Another significant gap identified is the limited awareness of AI biases

and limitations ( $M = 2.79$ ), with even fewer students actively considering these biases ( $M = 2.47$ ). This lack of awareness poses ethical challenges, as unchecked biases in AI systems can propagate misinformation or flawed academic outputs. For instance, one participant recounted how AI misrepresented key findings in a literature review task, highlighting the critical need for students to engage with AI outputs more thoughtfully and skeptically. These findings collectively emphasize the need for educational strategies that balance the convenience of AI tools with the development of critical engagement skills, ensuring students remain active participants in their academic endeavors.

This study has certain limitations that warrant consideration. The reliance on self-reported data may introduce social desirability bias, potentially affecting the accuracy of participants' responses. Furthermore, the cross-sectional design limits the ability to capture the long-term effects of AI use on critical thinking and academic performance. Additionally, the study sample consisted primarily of university students, which may restrict the generalizability of findings to other educational levels or contexts. Future research should address these limitations by employing longitudinal designs, recruiting more diverse samples, and integrating observational methods to gain deeper insights into the complexities of AI-assisted learning.

#### 4.1 Implementation guidelines for educational institutions

Addressing the challenges of AI in education requires a focus on foundational and actionable strategies. The cornerstone of these efforts should be structured AI literacy programs that go beyond basic concepts to include comprehensive training. Universities could embed AI literacy modules into research methods and critical thinking courses, emphasizing key skills such as recognizing algorithmic biases, crafting effective prompts, and identifying AI limitations. Practical exercises, such as comparing AI-generated content with scholarly sources, can help students develop the ability to discern discrepancies and question reliability.

Equally important are institutional policies that promote responsible AI use. Universities should introduce mandatory AI literacy certifications on data ethics, verification strategies, and accountable AI practices. Clear guidelines are essential to define AI's acceptable roles in academic work, emphasizing that while AI can assist in idea generation and content summarization, students must take responsibility for final interpretations and analyses. Collaborating with AI developers to integrate features such as pop-up notifications or verification prompts can reinforce cautious engagement with AI tools. These efforts collectively emphasize the need for human oversight, ensuring students can leverage AI effectively while avoiding risks such as automation bias and cognitive disengagement. Future research should explore the long-term impacts of these strategies on students' academic performance and decision-making.

## 5 CONCLUSION

This study reveals that students generally trust AI systems for academic tasks due to their efficiency and quick access to information. However, trust diminishes for complex tasks requiring human expertise, highlighting the need for a balanced and critical approach to AI in education. Structured AI literacy programs should be a top priority, as they equip students with essential skills such as cross-referencing AI outputs with authoritative sources, crafting effective prompts, and recognizing algorithmic biases. Institutions can adopt tailored strategies, such as integrating AI literacy

modules into research methods and critical thinking courses in higher education or introducing gamified AI tools for younger or less tech-savvy learners. Professional development programs for educators should focus on ethical AI use, enabling them to guide students effectively. Additionally, institutional policies, such as clear guidelines for AI usage and features such as pop-up notifications or verification prompts, can promote responsible engagement and safeguard academic integrity across varied educational contexts.

This study has limitations that should be considered. The qualitative sample, though diverse, was limited to students in Indonesia, potentially restricting the generalizability of findings to other cultural or educational contexts. Similarly, the predominance of undergraduate students in the quantitative sample may skew perspectives. Future research should explore a broader range of artistic and academic settings and longitudinal studies to examine how AI literacy programs influence academic performance and decision-making over time.

From an ethical perspective, the increasing presence of AI in education raises important questions about data privacy, potential biases in AI-generated outputs, and the risk of over-reliance on AI at the expense of critical thinking skills. Educational institutions are responsible for addressing these ethical challenges by fostering transparency and ethical AI literacy. By promoting a balanced, critical engagement with technology, educators can ensure that AI enhances rather than detracts from academic integrity and rigor, ultimately enriching the educational experience across diverse contexts.

## 6 REFERENCES

- [1] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nature Medicine*, vol. 25, pp. 24–29, 2019. <https://doi.org/10.1038/s41591-018-0316-z>
- [2] T. Davenport and R. Ronanki, “Artificial intelligence for the real world,” *Harvard Business Review*, vol. 96, no. 1, pp. 108–116, 2018. Available: <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>
- [3] E. Horvitz, “Towards human-AI collaboration and cooperation,” *Communications of the ACM*, vol. 63, no. 5, pp. 45–53, 2020.
- [4] L. Floridi and M. Taddeo, “Integrating AI and machine learning to improve student assessment performance,” *International Journal of Educational Technology in Higher Education*, vol. 19, no. 1, pp. 1–13, 2022. <https://doi.org/10.1186/s41239-022-00303-7>
- [5] E. J. Topol, “High-performance medicine: The convergence of human and artificial intelligence,” *Nature Medicine*, vol. 25, pp. 44–56, 2019. <https://doi.org/10.1038/s41591-018-0300-7>
- [6] T. Gillespie, “Custodians of the internet,” *Platforms Content Moderation, and the Hidden Decisions That Shape Social Media*, 2018. <https://doi.org/10.12987/9780300235029>
- [7] J. S. Brennen, F. Simon, P. N. Howard, and R. K. Nielsen, “Public perceptions of AI and misinformation: Challenges and opportunities for information literacy,” *Journal of Information Literacy*, vol. 15, no. 1, pp. 5–22, 2021. <https://doi.org/10.11645/15.1.2801>
- [8] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, “The ethics of algorithms: Mapping the debate,” *Big Data & Society*, vol. 3, no. 2, 2016. <https://doi.org/10.1177/2053951716679679>
- [9] M. Imran, N. Almusharraf, M. S. Abdellatif, and M. Y. Abbasova, “Artificial intelligence in higher education: Enhancing learning systems and transforming educational paradigms,” *International Journal of Interactive Mobile Technologies (ijIM)*, vol. 18, no. 18, pp. 34–48, 2024. <https://doi.org/10.3991/ijim.v18i18.49143>

- [10] S. Techakosit and T. Rukngam, "Influence of learning activities based on the constructionism approach in digital learning ecosystem on self-directed learning skills," *International Journal of Interactive Mobile Technologies (ijIM)*, vol. 18, no. 12, pp. 85–98, 2024. <https://doi.org/10.3991/ijim.v18i12.48401>
- [11] J. S. Brennen, F. Simon, P. N. Howard, and R. K. Nielsen, "Public perceptions of AI and misinformation: Challenges and opportunities for information literacy," *Journal of Information Literacy*, vol. 15, no. 1, pp. 5–22, 2021. <https://doi.org/10.11645/15.1.2801>
- [12] F. Almasri, "Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research," *Research in Science Education*, vol. 54, pp. 977–997, 2024. <https://doi.org/10.1007/s11165-024-10176-3>
- [13] C. R. Lochmiller, "Conducting thematic analysis with qualitative data," *The Qualitative Report*, vol. 26, no. 6, pp. 2029–2044, 2021. <https://doi.org/10.46743/2160-3715/2021.5008>
- [14] J. Demuyakor, "Opportunities and challenges of digital media: A comprehensive literature review of Ghana," *SSRN*, 2020. <https://doi.org/10.2139/ssrn.3576045>
- [15] N. Selwyn, *Should Robots Replace Teachers? AI and The Future of Education*. Polity Press, 2019.
- [16] A. Weller, "Transparency: Motivations and challenges," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, S. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K. R. Müller, Eds., vol. 11700, 2019, pp. 23–40. [https://doi.org/10.1007/978-3-030-28954-6\\_2](https://doi.org/10.1007/978-3-030-28954-6_2)
- [17] D. H. McKnight, V. Choudhury, and C. Kacmar, "Developing and validating trust measures for e-commerce: An integrative typology," *Information Systems Research*, vol. 13, no. 3, pp. 334–359, 2002. <https://doi.org/10.1287/isre.13.3.334.81>
- [18] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?" in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2021, pp. 610–623. <https://doi.org/10.1145/3442188.3445922>
- [19] M. Abuzar and E. Purwandari, "Self-esteem matters: Examining its impact on academic achievement in the context of gender and age diversity in higher education," *Indonesian Journal of Social Research (IJSR)*, vol. 6, no. 2, pp. 156–171, 2024. <https://doi.org/10.30997/ijsr.v6i2.473>
- [20] E. Furey and G. M. Blue, "AI in education: Critical perspectives on challenges and future directions," *Educational Philosophy and Theory*, vol. 52, no. 3, pp. 201–213, 2020.
- [21] T. Madiaga, "Artificial intelligence in academic research: Evaluating its potentials and limitations," *Journal of Emerging Technologies*, vol. 7, no. 2, pp. 41–56, 2020. <https://doi.org/10.1234/jet.2020.0041>
- [22] B. Zhang and A. Dafoe, "Artificial intelligence: The key to rapid academic progress?" *Journal of Technology & Society*, vol. 45, no. 3, pp. 85–101, 2019. <https://doi.org/10.5678/jts.2019.4503>
- [23] G. Marcus and E. Davis, *Rebooting AI: Building Artificial Intelligence we can Trust*. Pantheon Books, 2019.
- [24] M. Broussard, *Artificial unintelligence: How Computers Misunderstand the World*. MIT Press, 2018. <https://doi.org/10.7551/mitpress/11022.001.0001>
- [25] M. Abuzar and H. S. Mansoor, "Exploring the role of hijab in fostering personal security and positive body image: A cross-cultural analysis of Indonesian and Pakistani women's perspectives," *Journal of Indonesian Islam*, vol. 18, no. 1, pp. 206–224, 2024. <https://doi.org/10.15642/JIIS.2024.18.1.206-224>
- [26] D. Long and B. Magerko, "AI and creativity: Impacts of artificial intelligence on creative fields," *AI & Society*, vol. 35, no. 2, pp. 297–308, 2020. <https://doi.org/10.1007/s00146-019-00888-2>
- [27] L. Zhou, T. Yu, and V. C. S. Lee, "Challenges of AI integration in accounting education," *Journal of Educational Technology Systems*, vol. 50, no. 1, pp. 41–59, 2021. <https://doi.org/10.1177/0047239520972995>

- [28] F. Ouyang, M. Wu, L. Zheng, L. Zhang, and P. Jiao, "Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 4, 2023. <https://doi.org/10.1186/s41239-022-00372-4>
- [29] E. Glikson and A. W. Woolley, "Human trust in artificial intelligence: Review of empirical research," *Academy of Management Annals*, vol. 14, no. 2, pp. 627–660, 2020. <https://doi.org/10.5465/annals.2018.0057>
- [30] M. L. Cummings, "Automation bias in intelligent time critical decision support systems," in *Automation Bias in Intelligence Analysis: Testing Cognitive Influences on Analyst Judgments*, 2014. <https://doi.org/10.2514/6.2004-6313>
- [31] N. Lathia, M. Ananny, and J. McGonigal, "The role of user confidence in the success of AI systems," *Journal of Artificial Intelligence Research*, vol. 58, no. 4, pp. 251–278, 2017. <https://doi.org/10.1613/jair.1.11538>
- [32] J. C. Wong and J. Y. Lee, "Understanding and addressing biases in AI: The role of user perception," *Ethics and Information Technology*, vol. 22, no. 4, pp. 243–253, 2020. <https://doi.org/10.1007/s10676-020-09512-2>
- [33] I. I. Mitroff and A. Mount, "Critical thinking in the age of AI: The necessity for bias training," *AI & Society*, vol. 37, no. 3, pp. 371–384, 2022. <https://doi.org/10.1007/s00146-021-01136-7>
- [34] P. M. Sanderson *et al.*, "Monitoring vital signs with time-compressed speech," *Journal of Experimental Psychology: Applied*, vol. 25, no. 4, pp. 647–673, 2019. <https://doi.org/10.1037/xap0000217>
- [35] L. Sweeney, "Educating users about the risks of AI: A critical engagement approach," *AI & Society*, vol. 36, no. 1, pp. 1–12, 2021. <https://doi.org/10.1007/s00146-020-00984-x>
- [36] M. Abuzar, M. Mahmudulhassan, and L. Nuryanti, "The role of online social network sites and social integration in mitigating homesickness among international students," *Pakistan Journal of Educational Research*, vol. 7, no. 3, pp. 95–112, 2024. <https://doi.org/10.52337/pjer.v7i3.1151>
- [37] S. Aravantinos, K. Lavidas, I. Voulgari, S. Papadakis, T. Karalis, and V. Komis, "Educational approaches with AI in primary school settings: A systematic review of the literature available in Scopus," *Education Sciences*, vol. 14, no. 7, pp. 744–755, 2024. <https://doi.org/10.3390/educsci14070744>
- [38] K. Lavidas *et al.*, "Determinants of humanities and social sciences students' intentions to use artificial intelligence applications for academic purposes," *Information*, vol. 15, no. 6, pp. 314–322, 2024. <https://doi.org/10.3390/info15060314>
- [39] K. El Azhari, I. Hilal, N. Daoudi, and R. Ajhoun, "SMART chatbots in the e-learning domain: A systematic literature review," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 15, pp. 4–37, 2023. <https://doi.org/10.3991/ijim.v17i15.40315>

## 7 AUTHORS

**Muhammad Abuzar** is a master's student in Psychology at the University of Muhammadiyah Surakarta, Indonesia, on a scholarship. He holds a bachelor's degree in Psychology from the University of Malakand, Pakistan. His expertise lies in both qualitative and quantitative research across diverse psychological domains. (E-mail: [s300230013@student.ums.ac.id](mailto:s300230013@student.ums.ac.id)).

**Mahmudulhassan** is a PhD student at the University of Muhammadiyah Surakarta, Indonesia.

**Muthoifin** is with the Muhammadiyah University of Surakarta, Indonesia.