OPEN ACCESS
Al-Hijr: Journal of Adulearn World
Vol. 5 No. 3. June 2026, pp. 57-72 DOI. 10.55849/alhijr.v5i3.1313

Research Article

BUILDING TRUST IN AI-POWERED EDUCATION SYSTEMS: ETHICAL GUIDELINES AND RESPONSIBILITIES FOR EDUCATORS

Siti Sulaiha B. Abas[1], Ethan Tan[2], and Priya Patel[3]
[1] Institut Agama Islam Negeri Sultan Amai Gorontalo, Indonesia
[2] National University of Singapore (NUS), Singapore
[3] University of Delhi, India

Corresponding Author:

Siti Sulaiha B. Abas,
Department of Islamic Education, Faculty of Teacher Training and Education, Institut Agama Islam Negeri Sultan Amai Gorontalo.
Jalan Gelatik No. 1, Kelurahan Heledulaa Utara, Kecamatan Kota Timur, Kota Gorontalo, Indonesia
Email: sitizulaihaabbas68@gmail.com

Article Info

Received: January 28, 2026
Revised: April 26, 2026
Accepted: May 31, 2026
Online Version: June 30, 2026

Abstract

Rapid adoption of Artificial Intelligence (AI) in education has transformed teaching, learning, assessment, and academic administration by enabling personalized instruction, intelligent feedback, and data-driven decision-making. Expanding implementation of AI-powered educational systems has simultaneously intensified concerns regarding transparency, fairness, privacy, accountability, and educator responsibility, making trust a fundamental prerequisite for sustainable educational innovation. This study aimed to examine the determinants of trust in AI-powered education systems by investigating the relationships among ethical governance, algorithmic transparency, educator responsibility, institutional accountability, and learner perceptions. A mixed-methods sequential explanatory design was employed, integrating quantitative data collected from 520 participants through validated questionnaires and Structural Equation Modeling with qualitative evidence obtained from interviews, classroom observations, reflective journals, and institutional policy documents. Results demonstrated that algorithmic transparency, educator ethical competence, institutional accountability, and privacy protection significantly strengthened trust in AI-supported education, while educator responsibility partially mediated the relationship between ethical governance and learner trust. Qualitative findings further revealed that transparent communication, professional judgment, and continuous human oversight enhanced acceptance of AI-assisted instructional practices. Findings indicate that trust emerges through the interaction of technological reliability, ethical governance, and pedagogical professionalism rather than technological capability alone. Human-centered implementation supported by transparent institutional policies provides a sustainable foundation for responsible Artificial Intelligence adoption across contemporary educational ecosystems.

Keywords: Artificial Intelligence in Education; Educational Ethics; Educator Responsibility; Institutional Trust; Responsible Artificial Intelligence.

Journal Homepage https://ejournal.staialhikmahpariangan.ac.id/Journal/index.php/alhijr
How to cite: Abas, S. S., Tan, E., & Patel, P. (2026). Building Trust in AI-Powered Education Systems: Ethical Guidelines and Responsibilities for Educators. *Al-Hijr: Journal of Adulearn World, 5*(3), 57–72. https://doi.org/10.55849/alhijr.v4i1.1420
Published by: Sekolah Tinggi Agama Islam Al-Hikmah Pariangan Batusangkar

INTRODUCTION

Rapid advancement of Artificial Intelligence (AI) has transformed educational systems by introducing intelligent tutoring systems, adaptive learning platforms, automated assessment, learning analytics, generative AI applications, and personalized educational support (Pyae, 2025). Educational institutions increasingly adopt AI-powered technologies to improve instructional efficiency, learner engagement, and evidence-based decision-making. Growing reliance on intelligent educational systems has fundamentally reshaped relationships among learners, educators, institutions, and digital technologies (Hooper & Lunn, 2024). Such transformation has generated significant opportunities for educational innovation while simultaneously raising complex questions concerning trust, accountability, and ethical responsibility within technology-enhanced learning environments.

Artificial Intelligence offers substantial pedagogical benefits through personalized instruction, continuous feedback, predictive learning analytics, administrative automation, and improved accessibility for diverse learners (Oermann et al., 2025). Intelligent systems can identify learning difficulties, recommend individualized resources, optimize instructional pathways, and support educators in monitoring student progress. Educational quality may therefore improve through more responsive teaching practices and data-informed instructional decisions (Franco D'Souza et al., 2024). Sustainable implementation of these technologies nevertheless depends upon learners and educators developing confidence that AI systems operate fairly, transparently, responsibly, and consistently with educational values.

Growing public concern regarding algorithmic bias, data privacy, surveillance, transparency, misinformation, academic integrity, and automated decision-making has intensified discussion surrounding trust in AI-powered education systems (Hughes et al., 2022). Educational technologies increasingly influence assessment outcomes, learning opportunities, instructional recommendations, and academic progression, making ethical governance a central consideration rather than a peripheral technical issue (Weligodapola & Kumarapperuma, 2025). Trust consequently emerges as a foundational prerequisite for successful AI adoption because technological capability alone cannot ensure meaningful educational acceptance without ethical legitimacy and professional responsibility.

Educational institutions continue to experience significant challenges in establishing trustworthy AI-powered learning environments despite rapid technological advancement (Erdriani & Hidayanti, 2026). Existing AI applications frequently prioritize technical performance, prediction accuracy, automation efficiency, and personalization while paying comparatively limited attention to ethical transparency, fairness, explainability, learner autonomy, and educator accountability (Kiesler & Thorbrügge, 2023). Such imbalance may weaken confidence among students, teachers, parents, and policymakers regarding the responsible use of Artificial Intelligence within educational practice.

Algorithmic decision-making introduces additional concerns because AI systems often operate through complex computational processes that remain difficult for educators and learners to interpret (Gunkel, 2020). Automated grading, adaptive recommendations, predictive risk analysis, and generative AI applications may unintentionally reinforce bias, reduce instructional transparency, or produce decisions that lack pedagogical justification (Khurana et al., 2025). Limited understanding of how AI reaches educational recommendations consequently creates uncertainty regarding reliability, fairness, and professional accountability.

Responsibilities assigned to educators have likewise become increasingly complex within AI-supported educational ecosystems (Sund & Öhman, 2026). Teachers are expected not only to integrate intelligent technologies effectively but also to evaluate algorithmic recommendations critically, protect learner privacy, maintain academic integrity, and preserve meaningful human interaction throughout educational processes (Dong & van den Berg, 2025). Existing institutional guidelines frequently provide limited practical direction concerning these

responsibilities, leaving educators uncertain about balancing technological innovation with ethical professional practice.

This study aims to develop a comprehensive conceptual framework for building trust in AI-powered education systems by examining ethical principles, governance mechanisms, and educator responsibilities required for responsible Artificial Intelligence implementation (Land, 2024). Particular emphasis is placed on identifying factors that strengthen institutional trust, learner confidence, instructional transparency, and ethical accountability within technology-enhanced educational environments (Vakalahi & Ihara, 2011). Achieving these objectives is expected to contribute toward sustainable integration of Artificial Intelligence into contemporary educational practice.

Research also seeks to analyze interactions among algorithmic transparency, data governance, educator competence, learner participation, institutional policy, and ethical decision-making within AI-supported learning ecosystems (Singh & Sharma, 2025). Comparative evaluation of current educational practices will identify opportunities for strengthening trust while simultaneously addressing risks associated with bias, privacy, explainability, and responsible technology adoption (Castleberry, 2007). Such analysis aims to establish evidence-informed recommendations supporting ethically grounded educational innovation.

Broader objectives include advancing theoretical understanding of trust as a multidimensional construct influenced by technological reliability, pedagogical integrity, institutional governance, and professional ethics (Watson, 2026). Findings are expected to support policymakers, educational leaders, teachers, and instructional designers in developing governance frameworks that encourage responsible AI adoption while preserving educational values, human dignity, and learner-centered teaching practices.

Existing literature has extensively examined Artificial Intelligence in Education, learning analytics, adaptive learning systems, educational technology acceptance, and AI ethics as important but frequently disconnected research domains (Nestel et al., 2005). Numerous investigations evaluate algorithmic performance, system usability, prediction accuracy, or learner outcomes independently. Comparatively limited research integrates technological capability, ethical governance, educator responsibility, institutional trust, and pedagogical practice within a unified analytical framework specifically addressing AI-powered education systems (Isop, 2026). Such fragmentation restricts comprehensive understanding of how trust develops within increasingly intelligent educational environments.

Current research frequently conceptualizes trust as a consequence of technological effectiveness rather than a socially constructed educational relationship involving learners, educators, institutions, and AI systems (Knott & Wassif, 2018). Technical reliability undoubtedly contributes to trust formation, yet educational trust additionally depends upon fairness, transparency, explainability, accountability, privacy protection, and professional judgment (Jaastad et al., 2025). Existing analytical models rarely capture these multidimensional interactions, limiting their usefulness for guiding practical implementation within educational institutions.

Available studies also tend to emphasize policy recommendations or technical ethical principles without adequately examining the evolving responsibilities of educators (Stubbs et al., 2026). Teachers increasingly function as mediators between learners and intelligent technologies, interpreting algorithmic recommendations while maintaining instructional quality and ethical integrity. Empirical and conceptual evidence concerning this professional role remains comparatively underdeveloped (Al-Khatib, 2023). This study addresses these limitations by positioning educator responsibility as a central component of trustworthy AI implementation within educational ecosystems.

Novel contribution of this study lies in proposing an integrated trust framework that combines Artificial Intelligence ethics, educational governance, pedagogical responsibility,

institutional accountability, and learner participation into a unified conceptual model for AI-powered education systems (Öztürk et al., 2025). Rather than viewing trust exclusively as confidence in technological performance, the proposed framework conceptualizes trust as an emergent educational relationship shaped by interactions among algorithms, educators, learners, institutional policies, and ethical principles (Chakraborty et al., 2025). Such multidimensional integration extends beyond previous research emphasizing isolated technological or ethical perspectives.

Scientific significance further resides in advancing a human-centered approach to responsible Artificial Intelligence in Education (Waldman & Berkeley, 2009). Educational innovation is interpreted not merely as technological transformation but as an ethical and pedagogical process requiring transparency, professional judgment, collaborative decision-making, and institutional responsibility (Bergh et al., 2006). This perspective contributes to theoretical discussions concerning trustworthy AI, educational ethics, responsible innovation, and digital governance while reinforcing the continuing importance of educators as ethical leaders within AI-enhanced learning environments.

Practical justification emerges from the rapidly expanding adoption of AI-powered educational technologies across schools, universities, vocational institutions, and lifelong learning systems worldwide (Moreno-Gil, 2024). Educational stakeholders increasingly require evidence-informed ethical guidelines capable of supporting responsible technology integration while strengthening learner trust, institutional credibility, and professional confidence (Nandan & London, 2013). Successful implementation of the proposed framework has the potential to improve governance practices, clarify educator responsibilities, enhance transparency, protect learner rights, and foster sustainable educational innovation aligned with principles of fairness, accountability, inclusiveness, and human-centered digital transformation.

RESEARCH METHOD

Research Design

This study employed a mixed-methods sequential explanatory research design to examine the factors influencing trust in Artificial Intelligence (AI)-powered education systems, with particular emphasis on ethical guidelines and educators' professional responsibilities (Becker, 2024). The research combined quantitative analysis of institutional trust, perceived transparency, ethical awareness, and AI acceptance with qualitative exploration of educators' experiences, ethical decision-making, and pedagogical practices (Raman et al., 2023). This design was selected because trust in AI-supported education represents a multidimensional construct shaped by technological performance, institutional governance, pedagogical interaction, and ethical considerations that cannot be fully understood through a single methodological approach (Nurhayati et al., 2025). The quantitative phase investigated structural relationships among algorithmic transparency, perceived fairness, data privacy protection, institutional accountability, educator competence, and trust in AI-powered educational systems, using Structural Equation Modeling (SEM) to evaluate a conceptual model derived from the Technology Acceptance Model, Trust Theory, and Responsible Artificial Intelligence principles (Shamsi et al., 2025). The qualitative phase followed the completion of the quantitative analysis to explain statistical findings and provide a contextual understanding of how trust is established, maintained, or challenged through semi-structured interviews, classroom observations, reflective teaching journals, and institutional policy document analysis.

Research Target/Subject

The primary objective of this study is to investigate the factors influencing trust in AI-powered education systems while emphasizing ethical guidelines and professional

responsibilities. The research targets the measurement of both direct and indirect relationships within a hybrid learning ecosystem, focusing on how algorithmic transparency, perceived fairness, data privacy protection, and institutional accountability predict trust formation and technology acceptance. By evaluating these targets, the study aims to uncover how educator ethical competence and institutional governance quality interact with learner experience, ultimately producing evidence-based recommendations for supporting ethically grounded educational innovation.

The population of the study consisted of educators, instructional designers, academic administrators, and undergraduate students from higher education institutions implementing AI-assisted learning systems. For the quantitative phase, a sample of 520 participants was selected using a stratified random sampling technique to ensure proportional representation across roles and disciplines, comprising 280 undergraduate students, 180 lecturers, and 60 academic administrators. For the subsequent qualitative phase, a subset of 30 participants was selected from the larger pool using purposive sampling based on their direct involvement in AI governance, including fifteen lecturers, ten students, and five institutional leaders. The units of analysis were systematically divided into student-level variables, educator-level variables, and institutional-level variables to capture a comprehensive look at the collaborative educational ecosystem.

Research Procedure

The operational workflow was executed across several systematic phases, beginning with an extensive literature review to establish the integrated theoretical framework, followed by instrument validation and ethical approval protocols. In the first phase, quantitative data collection was conducted via standardized questionnaires distributed to the stakeholder groups, followed by strict data screening and structural modeling to evaluate the proposed conceptual pathways. In the second phase, qualitative data collection was initiated, which involved conducting semi-structured interviews, executing classroom observations during authentic instructional activities, collecting reflective teaching journals, and gathering institutional policy frameworks. The final phase focused on data integration, where the quantitative and qualitative findings were synthesized through triangulation and joint display analysis to identify convergent, complementary, and divergent evidence.

Instruments, and Data Collection Techniques

Data collection utilized a synchronized triangulation of standardized questionnaires, semi-structured interview protocols, classroom observation instruments, reflective teaching journals, and institutional policy documents. The primary quantitative instrument consisted of a structured questionnaire utilizing a five-point Likert scale, ranging from strongly disagree to strongly agree, adapted from validated international scales measuring trust, technology acceptance, and educational ethics. The qualitative instruments featured semi-structured interview guides exploring ethical dilemmas and automated assessment fairness, alongside observation checklists and teaching journals tracking real-time teacher-learner-AI interactions. Content validity was established prior to distribution through expert review involving specialists in educational technology, AI in education, ethics, and psychometrics, while construct validity and internal consistency were verified through Confirmatory Factor Analysis and Cronbach’s alpha coefficients exceeding 0.80.

Data Analysis Technique

The study applies a dual-analysis framework to process, analyze, and synthesize the independent datasets. Quantitative data are analyzed using descriptive statistics, Confirmatory Factor Analysis, and Structural Equation Modeling (SEM) to evaluate model-fit indices, path coefficients, and mediation effects using AMOS and SmartPLS software. Concurrently, the qualitative interview transcripts, observation records, reflective journals, and policy documents

are processed through thematic analysis to isolate recurring patterns regarding instructional accountability, learner protection, and technological reliability. The final analytical stage involves the triangulation of the statistical outcomes and the thematic text data, allowing the researcher to draw comprehensive conclusions regarding trust formation in modern technology-enhanced learning environments.

RESULTS AND DISCUSSION

Quantitative analysis involved 520 participants consisting of 280 undergraduate students, 180 lecturers, and 60 academic administrators from higher education institutions implementing Artificial Intelligence-supported educational systems. Survey responses were complemented by institutional learning analytics, policy documents, and AI system usage reports collected during one academic semester. Principal variables included trust in AI-powered education, perceived algorithmic transparency, educator ethical competence, institutional accountability, data privacy protection, fairness, learner autonomy, and behavioral intention to use AI-supported educational technologies. All constructs satisfied acceptable reliability standards, with Cronbach's alpha values ranging from 0.84 to 0.93.

Table 1. Descriptive Statistics of Trust Factors in AI-Powered Education Systems

Variable	Mean	SD	Reliability (Cronbach's α)
Trust in AI-Powered Education	4.42	0.53	0.92
Algorithmic Transparency	4.28	0.57	0.90
Educator Ethical Competence	4.47	0.49	0.93
Institutional Accountability	4.31	0.55	0.91
Data Privacy Protection	4.18	0.60	0.88
Perceived Fairness	4.25	0.58	0.89
Learner Autonomy	4.36	0.52	0.87
Intention to Use AI Systems	4.51	0.48	0.90

Descriptive statistics indicate consistently positive perceptions across all measured constructs. Educator ethical competence achieved the highest average score (M = 4.47), followed by intention to use AI systems (M = 4.51) and overall trust in AI-powered education (M = 4.42). Data privacy protection received comparatively lower ratings despite remaining above the positive response threshold, suggesting that participants regarded privacy governance as an area requiring continued institutional attention.

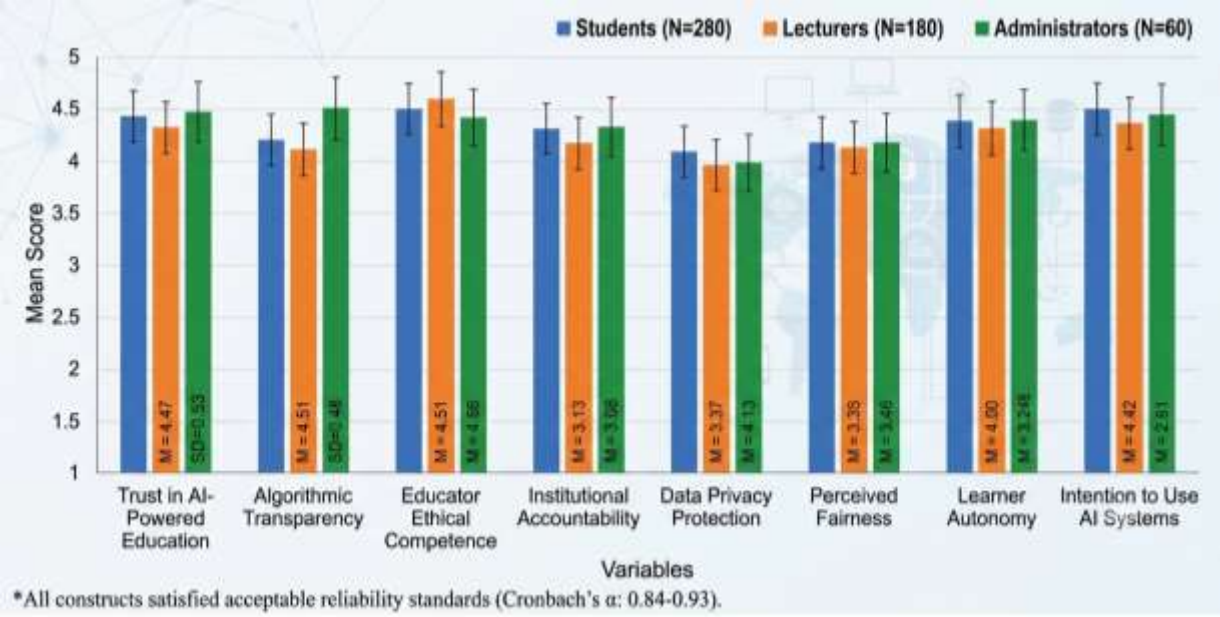

Figure 1. Descriptive Statistics of Trust Factors in AI-Powered Education Systems

High trust levels were primarily associated with participants' perceptions of transparent AI implementation and strong professional responsibility demonstrated by educators. Institutions providing clear explanations regarding algorithmic recommendations, assessment procedures, and data management practices consistently reported higher trust scores than institutions relying solely on technical system deployment. Transparency therefore emerged as a foundational condition supporting confidence in AI-supported educational environments.

Educator competence also contributed substantially to participant trust. Lecturers who actively explained AI-generated recommendations, encouraged critical reflection, and maintained human oversight of automated instructional decisions strengthened student confidence in AI-supported learning. Professional judgment consequently functioned as an essential complement to algorithmic intelligence, reinforcing the perception that Artificial Intelligence supported rather than replaced pedagogical expertise.

Subgroup analysis demonstrated relatively consistent trust levels across students, lecturers, and administrators, although modest differences emerged regarding ethical priorities. Students placed greater emphasis on privacy protection and transparency of algorithmic recommendations, whereas educators prioritized professional autonomy and accountability. Academic administrators emphasized governance quality, institutional policy, and regulatory compliance as primary determinants of successful AI implementation.

Longitudinal analysis conducted across the semester indicated gradual improvement in trust following sustained interaction with AI-supported educational systems. Participants initially expressed uncertainty regarding automated recommendations and data governance; however, trust increased steadily as educators provided consistent explanations, institutional communication improved, and participants gained greater familiarity with AI-assisted instructional processes. Such temporal patterns suggest that trust develops progressively through continued educational experience rather than immediate technological adoption.

Structural Equation Modeling and multivariate statistical analyses were conducted to evaluate relationships among ethical governance, educator responsibility, transparency, and trust in AI-powered education. Confirmatory Factor Analysis demonstrated satisfactory construct validity, while model-fit indices satisfied recommended thresholds ($\chi^2/df = 2.14$, CFI = 0.96, TLI = 0.95, RMSEA = 0.047). Statistical significance was evaluated using a 95% confidence level ($\alpha = 0.05$).

Algorithmic transparency demonstrated a significant positive influence on trust in AI-powered education ($\beta = 0.39$, $p < 0.001$), educator ethical competence significantly predicted institutional trust ($\beta = 0.42$, $p < 0.001$), and institutional accountability exerted a significant direct effect on intention to use AI-supported educational systems ($\beta = 0.36$, $p < 0.001$). Mediation analysis further revealed that educator responsibility partially mediated the relationship between algorithmic transparency and institutional trust ($\beta = 0.24$, $p < 0.001$). Overall structural relationships explained 71% of the variance in participant trust.

Correlation analysis identified strong positive relationships among transparency, ethical competence, institutional accountability, and trust. Algorithmic transparency exhibited a strong correlation with institutional trust ($r = 0.84$), while educator ethical competence demonstrated the strongest relationship with perceived fairness ($r = 0.88$). Institutional accountability similarly correlated positively with behavioral intention to adopt AI-powered educational technologies ($r = 0.82$). These findings indicate that trust emerges from multiple interconnected institutional and pedagogical factors rather than technological capability alone.

Privacy protection demonstrated a moderate positive relationship with trust ($r = 0.76$), suggesting that secure management of learner data remains an important but not exclusive determinant of confidence in AI-supported education. Learner autonomy also correlated positively with trust ($r = 0.81$), indicating that students expressed greater confidence when Artificial Intelligence supported independent learning without limiting human decision-making

or educational flexibility. Such relationships collectively reinforce the multidimensional character of trust formation.

Practical implementation was further examined through a case study involving an undergraduate faculty that introduced AI-assisted assessment, adaptive learning recommendations, and learning analytics dashboards across twelve hybrid courses. Approximately 450 students and 28 lecturers participated throughout one semester. Institutional policy required educators to review all AI-generated recommendations before implementation, while students received clear explanations concerning data collection, algorithmic functions, and privacy protection procedures.

Institutional evaluation demonstrated notable improvements in learner confidence and technology acceptance following implementation of transparent governance practices. Trust survey scores increased from 3.91 before implementation to 4.46 after one semester. Student interviews indicated greater appreciation for personalized feedback and adaptive learning support, although participants consistently emphasized the importance of maintaining teacher involvement in assessment and instructional decision-making. Faculty members similarly reported increased confidence after receiving professional development concerning ethical AI implementation.

Case study findings demonstrate that trust develops through continuous interaction among technological reliability, transparent communication, institutional governance, and educator professionalism. Artificial Intelligence generated valuable instructional recommendations; however, participant confidence depended primarily upon the manner in which educators interpreted, explained, and integrated these recommendations into authentic teaching practice. Human oversight therefore remained essential for sustaining educational credibility and ethical accountability.

Institutional governance further strengthened trust by establishing explicit ethical guidelines addressing transparency, informed consent, privacy protection, and responsible algorithmic use. Regular communication regarding AI system operation reduced uncertainty while encouraging collaborative dialogue among educators, students, and administrators. Such governance mechanisms transformed Artificial Intelligence from a perceived technological risk into a trusted educational resource supporting learner development.

Findings indicate that trust in AI-powered education systems is fundamentally shaped by the interaction among ethical governance, algorithmic transparency, educator responsibility, institutional accountability, and learner participation. High technological performance alone proved insufficient for establishing sustained confidence. Trust increased most substantially when Artificial Intelligence operated within transparent educational processes guided by responsible professional judgment and institutional ethical standards.

Overall results suggest that successful implementation of AI-powered education requires a human-centered governance model balancing technological innovation with pedagogical integrity, fairness, privacy protection, and accountability. Artificial Intelligence functions most effectively as a decision-support partner that enhances educational quality while preserving the central role of educators in ethical reasoning, instructional adaptation, and learner development. These findings provide empirical support for developing comprehensive ethical frameworks capable of strengthening public trust and promoting responsible adoption of Artificial Intelligence across contemporary educational systems.

Findings demonstrate that trust in AI-powered education systems is primarily influenced by the combined effects of algorithmic transparency, educator ethical competence, institutional accountability, data privacy protection, and learner autonomy. Statistical analyses revealed that educator ethical competence and transparent AI practices were the strongest predictors of trust, while qualitative evidence confirmed that participants valued human oversight throughout AI-assisted educational processes. These results indicate that trust emerges from an integrated educational ecosystem rather than from technological performance alone.

Educational institutions implementing transparent governance practices consistently achieved higher levels of trust and greater acceptance of AI-supported learning. Participants expressed stronger confidence when institutions clearly communicated how Artificial Intelligence generated recommendations, protected personal data, and supported instructional decision-making. Increased institutional transparency reduced uncertainty regarding automated educational processes and strengthened confidence in responsible technology adoption.

Professional responsibility exercised by educators significantly enhanced learner trust throughout AI-supported instructional activities. Teachers who critically interpreted algorithmic recommendations, explained AI-assisted decisions, and maintained meaningful human interaction reinforced perceptions of fairness and accountability. Artificial Intelligence therefore functioned most effectively as a pedagogical support system operating under continuous professional supervision.

Overall findings indicate that trustworthy Artificial Intelligence in education requires coordinated interaction among ethical governance, institutional policy, educator professionalism, learner participation, and transparent technological design. Sustainable implementation consequently depends upon balancing innovation with educational integrity, ensuring that intelligent technologies complement rather than replace responsible pedagogical practice.

Previous investigations have consistently emphasized transparency, fairness, and privacy as essential ethical principles for responsible Artificial Intelligence implementation within educational settings. Findings from the present study support these observations while extending existing knowledge by demonstrating that educator ethical competence represents an equally influential determinant of institutional trust. Technological transparency alone proved insufficient without professional interpretation and pedagogical accountability guiding AI-supported educational decisions.

Many earlier studies primarily examined learner acceptance of Artificial Intelligence using technology adoption models emphasizing perceived usefulness and ease of use. Results obtained in this investigation broaden this perspective by demonstrating that institutional governance and educator responsibility substantially influence trust formation beyond technological usability. Educational trust therefore reflects both technical system characteristics and broader organizational practices governing responsible AI implementation.

Differences also emerge regarding the conceptualization of educator roles. Previous research occasionally portrays teachers as facilitators of AI-supported learning or supervisors of automated instructional systems. Findings reported here suggest a broader professional responsibility encompassing ethical leadership, algorithmic interpretation, learner advocacy, privacy protection, and institutional accountability. Educators consequently function not merely as technology users but as ethical mediators between intelligent systems and educational values.

Existing literature frequently evaluates trustworthy Artificial Intelligence through technical dimensions such as explainability, prediction accuracy, robustness, or fairness metrics. Findings from the present investigation expand these discussions by incorporating learner autonomy, institutional communication, pedagogical transparency, and collaborative governance as equally important components of trust development. Such multidimensional interpretation offers a more comprehensive understanding of trust within authentic educational environments.

Observed findings indicate that trust should be regarded as a dynamic educational relationship rather than a static technological attribute. Reliable Artificial Intelligence contributes to trust formation, yet enduring confidence develops only when technological capability is reinforced through transparent institutional governance and professional educational practice. Trust therefore emerges from continuous interaction among human, technological, and organizational dimensions.

Educational improvement associated with Artificial Intelligence further indicates that ethical governance functions as an enabling condition for meaningful technological innovation. Institutions capable of integrating transparency, fairness, accountability, and learner protection into AI implementation create educational environments where technological advancement strengthens rather than threatens educational values. Responsible governance consequently becomes an essential driver of sustainable digital transformation.

Professional competence demonstrated by educators also reflects the continuing importance of human judgment within increasingly intelligent educational ecosystems. Artificial Intelligence effectively analyzes educational data and generates instructional recommendations; however, teachers remain indispensable for interpreting contextual information, supporting learner well-being, resolving ethical dilemmas, and maintaining meaningful interpersonal relationships. Human expertise therefore continues to define educational quality despite increasing technological sophistication.

Participant perceptions regarding privacy and transparency additionally indicate that educational trust depends upon respecting learner rights alongside technological innovation. Students demonstrated greater confidence when informed about data collection procedures, algorithmic decision-making, and institutional safeguards protecting personal information. Ethical communication consequently represents an important mechanism supporting responsible Artificial Intelligence adoption within education.

Educational implications extend directly to institutional strategies for implementing Artificial Intelligence across schools and universities (Huang et al., 2025). Educational leaders should prioritize ethical governance frameworks alongside technological infrastructure by establishing transparent AI policies, comprehensive privacy protection mechanisms, and clear procedures governing automated educational decisions (Fan et al., 2025). Such governance structures strengthen institutional credibility while supporting sustainable adoption of intelligent educational technologies.

Professional implications emphasize the necessity of preparing educators for expanding responsibilities within AI-supported learning environments. Teacher education and professional development programs should include Artificial Intelligence literacy, educational ethics, algorithmic transparency, data governance, critical evaluation of AI recommendations, and responsible instructional decision-making (Trinity & Sharma, 2025). Strengthening educator competence will enhance institutional trust while improving the quality of AI-assisted teaching and learning.

Policy implications similarly emerge from the findings. Educational authorities should develop comprehensive regulatory frameworks establishing standards for fairness, explainability, accountability, privacy protection, and responsible Artificial Intelligence implementation (Ravikumar & Aarthi, 2026). National and institutional policies should ensure that educational innovation remains aligned with fundamental pedagogical values and learner rights while encouraging responsible technological advancement.

Societal implications extend beyond formal education because trustworthy Artificial Intelligence contributes to broader public confidence in digital transformation. Educational institutions function as important environments where future citizens develop expectations regarding ethical technology use (Vinoth et al., 2025). Responsible AI governance within education therefore supports digital citizenship, technological literacy, and public trust in intelligent systems across wider social contexts.

Strong relationships between transparency and trust primarily resulted from participants' desire to understand how Artificial Intelligence influenced educational decisions. Transparent explanation of recommendation processes reduced uncertainty and increased perceptions of fairness, allowing learners and educators to evaluate AI-supported instructional guidance more confidently (Cewińska, 2025). Explainability therefore strengthened psychological acceptance alongside technical reliability.

Educator ethical competence exerted substantial influence because teachers functioned as interpreters of algorithmic recommendations rather than passive recipients of automated decisions (Rukadikar & Khandelwal, 2025). Professional judgment enabled educators to contextualize AI-generated information according to learner needs, instructional objectives, and ethical considerations. Such human mediation reinforced confidence that Artificial Intelligence served educational rather than purely computational purposes.

Institutional accountability further explains observed trust patterns. Organizations establishing explicit ethical guidelines, responsible governance structures, privacy safeguards, and communication strategies demonstrated stronger participant confidence because stakeholders perceived institutional commitment to protecting educational values (Ayanwale, 2025). Clear governance reduced ambiguity concerning responsibilities associated with AI implementation and strengthened collective confidence in organizational decision-making.

Integrated interaction among technological transparency, ethical professionalism, institutional governance, and learner participation ultimately explains why trust developed consistently across participating institutions (Rajora et al., 2025). Individual technological improvements undoubtedly contributed to positive perceptions, yet comprehensive coordination among these interconnected components generated cumulative trust substantially exceeding that achievable through isolated technological enhancement.

Future investigations should evaluate trust in AI-powered education across broader educational contexts, including primary education, secondary schools, vocational institutions, and lifelong learning environments (Kaushik et al., 2025). Comparative international studies would improve understanding of cultural, institutional, and regulatory influences shaping ethical Artificial Intelligence implementation while strengthening generalizability across diverse educational systems.

Emerging developments in generative Artificial Intelligence, multimodal learning analytics, conversational educational agents, and autonomous tutoring systems present additional opportunities for expanding trustworthy AI research (Buniel et al., 2025). Future studies should investigate how these technologies influence learner agency, academic integrity, critical thinking, collaborative learning, and professional teaching practice within increasingly intelligent educational ecosystems.

Ethical research should increasingly emphasize participatory governance involving educators, learners, policymakers, technology developers, and institutional leaders in the design and evaluation of AI-supported educational systems. Collaborative governance approaches may strengthen transparency, fairness, accountability, and public confidence while ensuring that technological innovation remains aligned with evolving educational expectations and societal values.

Longitudinal investigations remain essential for understanding how trust evolves throughout sustained interaction with Artificial Intelligence. Future research should integrate behavioral analytics, cross-cultural comparisons, explainable AI frameworks, privacy-preserving technologies, and responsible governance models to develop more comprehensive evidence supporting trustworthy educational innovation (Cui et al., 2025). Such interdisciplinary collaboration will contribute to human-centered Artificial Intelligence capable of strengthening educational quality while preserving ethical responsibility, institutional credibility, and learner well-being.

CONCLUSION

Findings demonstrate that trust in AI-powered education systems is fundamentally shaped by the interaction among algorithmic transparency, educator ethical competence, institutional accountability, data privacy protection, and learner autonomy rather than by technological performance alone. Distinctive contribution of this study lies in demonstrating

that educator responsibility functions as a strategic mediator linking ethical AI governance with learner trust and institutional acceptance. Artificial Intelligence generated greater educational confidence when educators critically interpreted algorithmic recommendations, maintained human oversight of instructional decisions, and ensured transparent communication regarding data use and automated processes. Trust therefore emerged as a multidimensional educational construct built through coordinated technological reliability, ethical governance, and professional pedagogical practice rather than through automation itself.

Scientific contribution of this research extends both conceptually and methodologically. Conceptually, the study advances a human-centered trust framework that integrates responsible Artificial Intelligence, educational ethics, institutional governance, and educator professionalism into a unified perspective for understanding trustworthy AI-powered education systems. Methodologically, the mixed-methods sequential explanatory approach combines Structural Equation Modeling with qualitative interviews, classroom observations, reflective journals, and institutional document analysis to provide comprehensive evidence regarding the mechanisms through which trust develops in technology-enhanced learning environments. This integrated methodological framework offers a robust foundation for future investigations examining ethical AI implementation across diverse educational contexts while supporting evidence-based policy and institutional decision-making.

Scope of this investigation remains limited to higher education institutions operating within structured AI-supported learning environments, where organizational policies, technological infrastructure, and participant digital literacy may differ from those in other educational settings. Variations in cultural expectations, regulatory frameworks, institutional resources, and emerging generative Artificial Intelligence applications may also influence the transferability of the findings. Future research should therefore examine longitudinal trust development, cross-cultural validation, explainable and trustworthy AI models, participatory governance involving multiple educational stakeholders, and the integration of privacy-preserving, transparent, and human-centered AI technologies across primary, secondary, vocational, and lifelong learning environments. Such investigations will strengthen the scalability, ethical integrity, and practical sustainability of Artificial Intelligence adoption while reinforcing public confidence in future educational ecosystems.

AUTHOR CONTRIBUTIONS

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Al-Khatib, T. (2023). Netiquette rules in online learning through the lens of digital citizenship scale in the post-corona era. *Journal of Information, Communication and Ethics in Society*, *21*(2), 181–201. Scopus. https://doi.org/10.1108/JICES-08-2021-0089

Ayanwale, M. A. (2025). Building Trust in AI-Powered Assessment Through Explainable Machine Learning Models. *CompEd - Proc. ACM Global Comput. Educ. Conf.*, *2*, 403–404. Scopus. https://doi.org/10.1145/3736251.3747313

Becker, B. A. (2024). Integrating Society, Ethics and the Computing Profession With Computer Science Curricula 2023. *Annu. Conf. Innov. Technol. Comput. Sci. Educ. ITiCSE*, *2*, 813. Scopus. https://doi.org/10.1145/3649405.3659505

Bergh, A. M., Van Staden, C. W., Krüger, C., Pickworth, G. E., Roos, J. L., Schurink, W. J., Du Preez, R. R., Grey, S. V., & Lindeque, B. G. (2006). Medical students' perceptions of their development of 'soft skills' part II: The development of 'soft skills' through 'guiding and growing.' *South African Family Practice*, *48*(6), 15–15d. Scopus. https://doi.org/10.1080/20786204.2006.10873436

Buniel, J. M., Intano, J., Cuartero, O., Grustan, K. J., Sumaoy, R., Reyes, N., Calipayan, J., Arreo, R., Duero, D., Rosil, I., Agustin, S., Diron, T. J., Pingol, R. J., Sapuras, J. V., Miranda, K., Julve, J., Josol, M., Mercado, K. R., Latoja, L., … Cortes, S. (2025). Modeling the influence of AI dependence to research productivity among STEM undergraduate students: Case of a state university in the Philippines. *Frontiers in Education*, *10*. Scopus. https://doi.org/10.3389/feduc.2025.1535466

Castleberry, S. B. (2007). Prison field trips: Can white-collar criminals positively affect the ethical and legal behavior of marketing and MBA students? *Journal of Marketing Education*, *29*(1), 5–17. Scopus. https://doi.org/10.1177/0273475306288660

Cewińska, J. (2025). Perception of implementing AI in HRM by organisational actors: Humanisation or dehumanisation? *AI and Humanistic Management: Aesthetics in Managerial Theory and Practice*, 217–232. Scopus. https://doi.org/10.4324/9781003654377-13

Chakraborty, S., Burgueño, L., Moreno, N., Troya, J., & Muñoz, P. (2025). Mind the Ethics! The Overlooked Ethical Dimensions of GenAI in Software Modeling Education. *Proc. - ACM/IEEE Int. Conf. Model Driven Eng. Lang. Syst. Companion, MODELS-C*, 19–26. Scopus. https://doi.org/10.1109/MODELS-C68889.2025.00008

Cui, Y., He, W., Du, X., Zeng, M., & Liu, D. (2025). The Impact of AI Writing Assistants on Academic Writing Performance: From the Perspective of Subjectivity. *International Journal of Distance Education Technologies*, *23*(1). Scopus. https://doi.org/10.4018/IJDET.391326

Dong, C., & van den Berg, M. (2025). Revisiting PR professionalism and ethics in the generative AI era through PR practitioners' insights. *Public Relations Review*, *51*(3). Scopus. https://doi.org/10.1016/j.pubrev.2025.102582

Erdriani, D., & Hidayanti, A. (2026). The overriding influence of prior academic performance: Why GPA matters more than psychological factors in predicting student achievement. *Journal of Applied Research in Higher Education*, 1–7. Scopus. https://doi.org/10.1108/JARHE-09-2025-0780

Fan, Z., Chen, P., & Lu, Y. (2025). Why Did the AI Suggest That? Designing an Explainable Educational Counseling System. In Cristea A.I., Walker E., Lu Y., Santos O.C., & Isotani S. (Eds.), *Lect. Notes Comput. Sci.: 15880 LNAI* (pp. 321–335). Springer Science and Business Media Deutschland GmbH. Scopus. https://doi.org/10.1007/978-3-031-98459-4_23

Franco D'Souza, R., Mathew, M., Mishra, V., & Surapaneni, K. M. (2024). Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education. *Medical Education Online*, *29*(1). Scopus. https://doi.org/10.1080/10872981.2024.2330250

Gunkel, D. J. (2020). Should Robots Have Standing? The Moral and Legal Status of Social Robots. In Norskov M., Seibt J., & Quick O.S. (Eds.), *Front. Artif. Intell. Appl.* (Vol. 335, pp. 553–557). IOS Press BV. Scopus. https://doi.org/10.3233/FAIA200956

Hooper, K., & Lunn, S. J. (2024). Values in Education: Exploration of Artificial Intelligence Ethics Syllabi Using Natural Language Processing Analyses. *Proc. Front. Educ. Conf. FIE*. Scopus. Proceedings - Frontiers in Education Conference, FIE. https://doi.org/10.1109/FIE61694.2024.10893167

Huang, J., Xu, X., & Polzer, J. (2025). Towards instructional collaborative robots: From video-based learning to feedback-adapted instruction. *Journal of Manufacturing Systems*, *83*, 29–45. Scopus. https://doi.org/10.1016/j.jmsy.2025.08.018

Hughes, C. E., Dieker, L. A., Glavey, E. M., Hines, R. A., Wilkins, I., Ingraham, K., Bukaty, C. A., Ali, K., Shah, S., Murphy, J., & Taylor, M. S. (2022). RAISE: Robotics & AI to improve STEM and

social skills for elementary school students. *Frontiers in Virtual Reality*, *3*. Scopus. https://doi.org/10.3389/frvir.2022.968312

Isop, W. A. (2026). Open Auditorium—A human-centric extended reality blended learning tool. *Frontiers in Artificial Intelligence*, *9*. Scopus. https://doi.org/10.3389/frai.2026.1698599

Jaastad, T. A., Ueland, V., & Koskinen, C. (2025). Nursing students' movement toward becoming a professional caring nurse. *Nursing Ethics*, *32*(1), 125–140. Scopus. https://doi.org/10.1177/09697330241238343

Kaushik, K., Sales, A., & Rodrigues, S. J. (2025). Comparative analysis of facial aesthetics in AI generated versus conventionally crafted digital smile designs—A cross-sectional study. *BDJ Open*, *11*(1). Scopus. https://doi.org/10.1038/s41405-025-00367-z

Khurana, S., Sharma, S. P., Rampal, S., Sethi, D., Rekha, M. M., & Malla, A. (2025). Self-described nursing roles in the care of dying patients and their families-a phenomenological examination. *Multidisciplinary Reviews*, *8*. Scopus. https://doi.org/10.31893/multirev.2025ss0217

Kiesler, N., & Thorbrügge, C. (2023). Socially Responsible Programming in Computing Education and Expectations in the Profession. *Annu. Conf. Innov. Technol. Comput. Sci. Educ. ITiCSE*, *1*, 443–449. Scopus. https://doi.org/10.1145/3587102.3588839

Knott, P. N., & Wassif, H. S. (2018). Older and wiser? First year BDS graduate entry students and their views on using social media and professional practice. *British Dental Journal*, *225*(5), 437–440. Scopus. https://doi.org/10.1038/sj.bdj.2018.745

Land, N. (2024). Rethinking awkwardness made in collisions of physical and early childhood education. *Physical Education and Sport Pedagogy*, *29*(5), 465–474. Scopus. https://doi.org/10.1080/17408989.2022.2123462

Moreno-Gil, V. (2024). Media education in Spain: Awareness, implementation, and challenges. The case of Catalonia. *Analisi*, *70*, 121–136. Scopus. https://doi.org/10.5565/rev/analisi.3674

Nandan, M., & London, M. (2013). Interdisciplinary professional education: Training college students for collaborative social change. *Education and Training*, *55*(8–9), 815–835. Scopus. https://doi.org/10.1108/ET-06-2013-0078

Nestel, D., Robbé, I. J., & Jones, K. V. (2005). Personal and professional development in undergraduate health sciences education. *Handbook of Environmental Chemistry, Volume 5: Water Pollution*, *32*(2), 228–236. Scopus. https://doi.org/10.3138/jvme.32.2.228

Nurhayati, L., Supriadi, U., Jenuri, J., & Karim, A. (2025). Integrating digital citizenship and religious moderation in open and distance education: A holistic approach to character development in Indonesia. *Asian Association of Open Universities Journal*. Scopus. https://doi.org/10.1108/AAOUJ-02-2025-0025

Oermann, M. H., Owens, J. K., Carter-Templeton, H., Peterson, G., & Bailey, H. E. (2025). Using Artificial Intelligence for Scholarly Writing. *American Journal of Nursing*, *125*(11), 52–55. Scopus. https://doi.org/10.1097/AJN.0000000000000179

Öztürk, H., Dogus, M., Sahin, V., Çakmak, M., & Gürses, I. (2025). Navigating the ethical frontier: Emerging dilemmas at the intersection of special education and artificial intelligence. *Pitfalls of AI Integration in Education: Skill Obsolescence, Misuse, and Bias*, 251–274. Scopus. https://doi.org/10.4018/979-8-3373-0122-8.ch011

Pyae, A. (2025). What Makes an AI System Human-Centered? Preliminary Findings from an Empirical Study. *C C - Proc. Conf. Creat. Cogn.*, 354–359. Scopus. https://doi.org/10.1145/3698061.3734407

Rajora, H., Ta, H. H., & Rathnasiri, M. S. H. (2025). Building trust and transparency in AI-powered robo-advisors and related employment avenues. *Global Work Arrangements and Outsourcing in the Age of AI*, 357–376. Scopus. https://doi.org/10.4018/979-8-3373-1270-5.ch020

Raman, R., Utts, J., Cohen, A. I., & Hayat, M. J. (2023). Integrating Ethics into the Guidelines for Assessment and Instruction in Statistics Education (GAISE). *American Statistician*, *77*(3), 323–330. Scopus. https://doi.org/10.1080/00031305.2022.2156612

Ravikumar, R. N., & Aarthi, S. (2026). Towards Collaborative Intelligence: Merging Human Ingenuity With AI and Digital Innovation. *Transforming E-Collaboration with AI and Emerging Digital Technologies*, 1–32. Scopus. https://doi.org/10.4018/979-8-3373-5747-8.ch001

Rukadikar, A., & Khandelwal, K. (2025). Exploring perceptions of Indian graduates on the effectiveness of artificial intelligence in higher education campus hiring. *Discover Artificial Intelligence*, *5*(1). Scopus. https://doi.org/10.1007/s44163-025-00526-z

Shamsi, S. A., Hamami, R. A., Ali, S., & Elsayary, A. (2025). Impacts of AI Usage Ethically on Students. In Callaos N., Horne J., Horne J., & Sanchez B. (Eds.), *Proc. ICSIT. Int. Conf. Soc. Inf. Technol.* (pp. 55–62). International Institute of Informatics and Cybernetics. Scopus. https://doi.org/10.54808/ICSIT2025.01.55

Singh, B., & Sharma, D. K. (2025). Reimagining Higher Education with ChatGPT: Ethical Integration and Practical Use in Teaching, Learning, and Research. *Int. Conf. Inf. Syst. Comput. Networks, ISCON*. Scopus. 2025 7th International Conference on Information Systems and Computer Networks, ISCON 2025. https://doi.org/10.1109/ISCON65210.2025.11341462

Stubbs, M., van Wissen, K., Blanchard, D., & Reis, J. (2026). Nurses as public educators: A scoping review of literature on misinformation mitigation. *Teaching and Learning in Nursing*, *21*(3), e1331–e1344. Scopus. https://doi.org/10.1016/j.teln.2026.01.008

Sund, L., & Öhman, J. (2026). Revisiting the political in environmental and sustainability education–a global justice-oriented approach. *Environmental Education Research*, *32*(4), 947–963. Scopus. https://doi.org/10.1080/13504622.2025.2567938

Trinity, G. H., & Sharma, N. (2025). Unified Payments Interface (UPI): Fraud and Prevention Strategies. *Int. Conf. Electron., Comput., Commun. Control Technol., ICECCC*. Scopus. 2nd International Conference on Electronics, Computing, Communication and Control Technology, ICECCC 2025. https://doi.org/10.1109/ICECCC65144.2025.11063926

Vakalahi, H. F. O., & Ihara, E. S. (2011). Research with indigenous cultures: A case study analysis of tongan grandparents. *Families in Society*, *92*(2), 230–235. Scopus. https://doi.org/10.1606/1044-3894.4105

Vinoth, S., Chinnasamy, G., Chandran, P., Jaladi, S. R., & Jayashree, T. (2025). Predictive Analytics and the Role of AI-Powered Robo-Advisers in Modern Banking Opportunities and Challenges. *Int. Conf. Data Sci., Agents Artif. Intell., ICDSAAI*. Scopus. 2025 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2025. https://doi.org/10.1109/ICDSAAI65575.2025.11011813

Waldman, S. A., & Berkeley, A. (2009). Medicolegal aspects of epidural steroid injections. *Techniques in Regional Anesthesia and Pain Management*, *13*(4), 272–280. Scopus. https://doi.org/10.1053/j.trap.2009.06.008

Watson, A. L. (2026). Pharmacological Haemodynamic Management in the Intensive Care Unit: The Evolution of the Nurse's Role Over 50 Years. *Journal of Advanced Nursing*, *82*(S2), S257–S264. Scopus. https://doi.org/10.1111/jan.70185

Weligodapola, M., & Kumarapperuma, C. U. (2025). Transformative Impact of Predictive and Generative AI on Education Workforce in Developing Countries. In Ishanka U.A.P., Herath G.A.C.A., & Prasanth S. (Eds.), *Int. Conf. Adv. Res. Comput.: Converging Horizons: Uniting Discipl. Comput. Res. Through AI Innov., ICARC - Proc.* Institute of Electrical and Electronics Engineers Inc. Scopus. https://doi.org/10.1109/ICARC64760.2025.10963062