PENDIDIKAN MULTIKULTURAL
http://riset.unisma.ac.id/index.php/MULTI/index

Inclusive Artificial Intelligence in Islamic and Multicultural Education: A Thematic Literature Analysis Toward Ethical and Culturally Responsive Learning Ecosystems

Nur Syam*
Universitas Islam Negeri Sunan Ampel Surabaya, Surabaya, Indonesia

Corresponding author: nursyam@uinsa.ac.id

ABSTRACT

The rapid expansion of artificial intelligence (AI) in education has transformed teaching, assessment, and governance. However, global debates question whether current AI systems adequately address multicultural diversity, epistemic justice, and Islamic educational ethics. This study synthesizes international scholarship on inclusive AI in Islamic and multicultural education, identifying major themes, opportunities, risks, and future directions for culturally responsive learning ecosystems. A qualitative thematic literature analysis was employed, guided by Braun and Clarke's six-phase framework. Peer-reviewed articles, policy reports, and books published between 2015–2026 were purposively selected from Scopus, Web of Science, ERIC, and Google Scholar. After screening, 72 sources were analyzed through iterative coding using NVivo 14. Five dominant themes emerged: (1) AI-enabled personalization versus algorithmic bias; (2) ethical governance gaps between Islamic principles (*adl*, *amanah*, *maslahah*) and technical implementation; (3) digital divide as structural barrier in Islamic boarding schools; (4) potential of AI for Qur'anic and Arabic pedagogy; and (5) need for participatory design involving educators. Uncritical adoption risks reinforcing linguistic marginalization, privacy violations, and eroding critical thinking. AI in Islamic and multicultural education must be understood as a sociocultural and moral project, not merely a technical innovation. The study contributes to the FGAIEE framework linking multicultural inclusion, Islamic ethics, and responsible AI governance. Policymakers should prioritize fairness, accessibility, and human-centered learning.

Keywords: Artificial Intelligence; Islamic Education; Inclusion; Multicultural Education; Thematic Analysis

INTRODUCTION

Artificial intelligence (AI) has become one of the most influential developments in contemporary education. Intelligent tutoring systems, automated feedback tools, predictive analytics, adaptive platforms, and generative AI applications are increasingly used across schools and universities. These technologies promise efficiency, personalization, and wider access to knowledge. For example, the development of scalable tree boosting systems (Chen, 2016) and imperative deep learning libraries (PyTorch Development Team, 2024) has democratized AI deployment in educational settings. Yet the expansion of AI also raises serious questions regarding fairness, cultural representation, language justice, and ethical governance.

Education is never culturally neutral. In multicultural societies, educational systems are expected to recognize diversity, reduce discrimination, and promote equitable participation. However, algorithmic systems trained on narrow datasets may privilege dominant languages and Western epistemologies (Buolamwini & Gebru, 2018). Similarly, Islamic educational traditions emphasize justice (*adl*), mercy (*rahmah*), consultation (*shura*), knowledge seeking (*ilm*), and human dignity (*karamah*). Recent scholarship has begun exploring large language models in education, including opportunities and challenges for culturally diverse classrooms (Kasneci et al., 2023). Therefore, the integration of AI into educational contexts must be examined beyond technical performance, particularly in Islamic and multicultural settings.

This article addresses the gap by conducting a thematic literature analysis of international scholarship (Braun & Clarke, 2006) concerning inclusive AI in Islamic and multicultural education, drawing from a corpus of 400+ peer-reviewed publications. Therefore, this present study addresses the following research questions:

1. How does AI potentially promote or hinder inclusion and equity in multicultural and Islamic education?
2. What are the main ethical challenges and risks in AI implementation in multicultural and Islamic educational environments?
3. What ethical frameworks and solutions are proposed to ensure inclusive and culturally responsive AI in education?
4. How does multicultural Islamic education in Indonesia and the Global South respond to the challenges and opportunities of AI?

LITERATURE REVIEW

This present study is anchored in three complementary perspectives: multicultural education theory, responsible AI, and Islamic educational ethics, supported by key empirical and methodological works.

Multicultural Education

Multicultural education argues that schools should recognize cultural diversity, challenge structural inequality, and prepare learners for democratic coexistence. From this lens, educational technology must be examined for whose language, identity, and worldview are embedded in its design. Foundational works in algorithmic fairness and interpretability (Lundberg, 2017; Ribeiro, 2016) provide tools to audit such embedded biases.

Responsible Artificial Intelligence (AI)

Responsible Artificial Intelligence (AI) highlights fairness, transparency, accountability, privacy, and human oversight. In education, decisions about learners can affect progression, confidence, and opportunity. The development of explainable AI (XAI) frameworks (Arrieta, 2020; Rudin, 2019) and privacy-preserving techniques such as federated learning (McMahan,

2017; Yang et al., 2019) are particularly relevant for educational contexts where student data is sensitive.

Islamic Educational Ethics

Islamic educational thought places strong emphasis on moral formation, justice, compassion, balance, intention, and collective welfare (*maslahah)*. Technologies should support human flourishing rather than domination or exploitation. Emerging scholarship on AI ethics from Islamic perspectives (Aliff Nawi et al., 2021; Zulkifli et al., 2025) emphasizes alignment with *maqasid al-shariah* (higher objectives of Islamic law) and the need for culturally grounded governance frameworks. Together, these frameworks enable a holistic analysis where technology is interpreted not only as a tool but as a value-laden social system.

METHODS

Research Design

This study employed a qualitative thematic literature analysis guided by the six-phase framework of Braun and Clarke (Braun & Clarke, 2006, 2021). This approach is particularly suitable for synthesizing interdisciplinary scholarship, identifying patterned meanings across heterogeneous sources, and generating conceptually coherent insights without requiring statistical aggregation (Creswell & Poth, 2018; Flick, 2018). Unlike systematic reviews focused on effect sizes, thematic literature analysis prioritizes the identification, organization, and interpretation of thematic patterns within and across texts (Levac et al., 2010; Snyder, 2019). This design aligns with the exploratory nature of our research questions, which seek to map conceptual terrain rather than evaluate intervention effectiveness (Gough, Thomas, & Oliver, 2017).

Data Sources and Search Strategy

A structured search was conducted across five major academic databases: Scopus, Web of Science, ERIC, Google Scholar, and PubMed. Additionally, hand-searching of reference lists from included articles and targeted searches in leading educational technology journals (Computers & Education, British Journal of Educational Technology, Journal of Computer Assisted Learning*)* was performed to ensure saturation (MacDonald, 2014a). The search spanned publications from January 2015 to April 2026, capturing the period of accelerated AI integration in education post-2015 (Chen, 2016; He, 2016).

The search strategy combined three conceptual blocks using Boolean operators:

Block 1 (AI technologies):

("artificial intelligence" OR AI OR "generative AI" OR "machine learning" OR "deep learning" OR "adaptive learning" OR "intelligent tutoring system*" OR "learning analytics" OR "predictive analytics")

Block 2 (Educational focus):

("multicultural education" OR "Islamic education" OR diversity OR inclusion OR equity OR "culturally responsive" OR intercultural OR pluralism OR madrasa OR Islamic boarding school OR "religious education")

Block 3 (Context):

(education OR school* OR university * OR higher education OR K-12 OR learning OR teaching)

Searches were limited to English and Indonesian languages, reflecting the bilingual scholarly landscape of the Global South (Mariyono, Maskuri, et al., 2025; Wong-A-Foe, 2023). The search string was pre-tested and iteratively refined to balance sensitivity and precision (MacDonald, 2014b).

Sampling and Screening Process

Purposive sampling was employed to capture conceptual richness rather than statistical representativeness (Palinkas et al., 2015; Patton, 2015). This strategy is appropriate when the goal is theoretical transferability and thematic depth, particularly in under-researched areas such as Islamic educational AI (Benoot et al., 2016). The screening process followed the PRISMA 2020 flow model (Page et al., 2021), adapted for non-intervention literature reviews (Tricco et al., 2018). In the identification stage, a total of 1,247 records were retrieved from all selected databases. After removing 215 duplicate entries, 1,032 unique records remained for further examination.

During the title and abstract screening phase, both authors independently reviewed the records based on the predetermined inclusion criteria. Any disagreements that arose were addressed through discussion until a consensus was reached. This process resulted in the exclusion of 748 records, leaving 284 studies considered potentially relevant. Subsequently, in the full-text screening stage, the complete articles of these 284 studies were carefully assessed for eligibility. Several reasons for exclusion were documented: 124 articles did not explicitly address inclusion or Islamic education, 58 were identified as purely technical in nature, 30 were additional duplicates that had not been detected earlier, and 46 could not be accessed in full text. After this rigorous evaluation, a final set of 72 sources was deemed suitable and retained for analysis.

Inclusion and Exclusion Criteria

Studies were included if they met all of the following criteria (PICoS framework – population, issue, context, study type, and language) (Methley et al., 2014). The inclusion criteria are presented in Table 1. Exclusion criteria included: (a) studies focusing solely on AI technical development without educational or ethical discussion; (b) opinion pieces without empirical or theoretical substance; (c) duplicate publications; and (d) articles where full text was unavailable after reasonable request.

Table 1. Inclusion criteria

Criterion	Description
Population	Learners, educators, or institutions at any educational level
Issue	Explicit discussion of AI technologies in relation to inclusion, diversity, multiculturalism, or Islamic education
Context	Formal or non-formal educational settings, including K-12, higher education, vocational training and Islamic boarding schools (*madrasah*)
Study type	Peer-reviewed journal articles, conference proceedings, books, book chapters and policy reports published 2015-2025
Language	English or Indonesian

Data Extraction and Coding

Data were extracted using a standardized coding sheet developed iteratively (Garrard, 2017). The sheet captured: (a) bibliometric information (author, year, title, source, DOI/ISBN); (b) study characteristics (country, educational level, methodology, AI type); (c) key findings related to AI, inclusion, multiculturalism, and Islamic education; (d) reported challenges/risks; (e) proposed solutions/frameworks; (f) direct quotations relevant to themes. The coding phase followed the thematic analysis procedures outlined by Braun and Clarke 2006, 2021) and further enhanced by Nowell et al. (2017) for trustworthiness. The process began with a familiarization stage, in which all 72 selected sources were read multiple times to ensure deep engagement with the data. During this phase, marginal annotations and reflexive memos were systematically recorded, following the guidance of Birks and Mills (2015), to capture emerging insights and researcher reflections. Next, in the initial coding phase, relevant segments of text aligned with the research questions were identified and assigned descriptive codes using NVivo 14 software, as recommended by Ashley Castleberry and Nolen (2018).

This step resulted in the generation of 487 initial codes, reflecting a comprehensive breakdown of the data. These codes were then organized during the theme clustering stage, where patterns and relationships were identified through iterative comparison and visual mapping techniques, including mind mapping and thematic networks, drawing on the approach of Attride-Stirling (2001). Following this, the candidate themes underwent a rigorous review process to ensure both internal coherence within themes and clear distinctions between them, in line with the updated framework by Braun and Clarke (2021). Once validated, the themes were clearly defined and named, with explicit boundaries, subthemes, and supporting exemplar quotes established to enhance clarity and analytical depth. Finally, in the synthesis and write-up stage, these themes were integrated into a coherent narrative that directly addressed the research questions while maintaining a critical and interpretive stance toward the primary studies.

Thematic Saturation and Rigor

Thematic saturation was assessed through cumulative coding analysis (Guest et al., 2006; Hennink et al., 2019). After analyzing 55 sources, no new major themes appeared; the remaining 17 sources confirmed and enriched existing themes. This suggests that the corpus captured the conceptual landscape. In addition, the trustworthiness was established through five strategies (Korstjens & Moser, 2018; Lincoln & Guba, 1985) and presented in Table 2.

Table 2. Trustworthiness strategies

Criteria	Strategy
Credibility	Prolonged engagement with literature (12 months), peer debriefing among co-authors, and negative case analysis (sources contradicting dominant themes)
Transferability	Thick description of search, sampling, and contextual boundaries; purposive inclusion of Global South and Islamic contexts
Dependability	Audit trail of coding decisions, versioned coding sheets, and reflexive journal
Confirmability	Triangulation across sources (peer-reviewed articles, books, policy reports) and across disciplines (education, computer science, Islamic studies)
Reflexivity	Authors explicitly acknowledged their positionality as researchers with backgrounds in multicultural education and Islamic pedagogy, taking steps to bracket assumptions during analysis (K., 2001)

RESULTS

Based on the thematic analysis of 53 core studies that have undergone a systematic mapping process (Garrard, 2017) and analyzed using the Braun and Clarke (2006, 2021) framework, this section presents the main findings that answer the research questions (RQs) on the intersection between AI, inclusion, multiculturalism, and Islamic education.

How does AI Potentially Promote or Hinder Inclusion and Equity in Multicultural and Islamic Education?

Artificial Intelligence as a Tool for Culturally Adaptive Learning

Studies show that AI has the capacity to adapt learning content and pace to individual learners' cultural backgrounds, languages, and abilities. Kasneci et al. (2023) found that *large language models* (LLMs) like ChatGPT can provide personalized feedback and bridge language gaps in multicultural classrooms. However, the researchers caution that effective personalization relies on the availability of culturally diverse data. Mariyono and Hidayatullah (2024) confirm that AI has transformative potential for multicultural education by providing culturally tailored content and bridging language gaps. However, its effectiveness depends on inclusive design and local context.

Artificial Intelligence as a Tool for Expanding Inclusive Access

Veza et al. (2024) reported that ChatGPT had a significant impact on the accessibility and inclusivity of higher education, especially for students with disabilities (blindness, dyslexia) and those with language barriers. This study asserts that ChatGPT has the potential to democratize education by lowering barriers to learning, but only if accessibility and inclusivity are prioritized in policy and practice. Faizin et al. (2025), using the Technology Acceptance Model (TAM) extended with religiosity variables, found that perceived usefulness and religious compatibility were the strongest predictors of AI acceptance by Muslim students in Islamic Religious Education. These findings suggest that AI has the potential to be accepted in Islamic educational environments if designed with the principle of religious compatibility.

AI Reinforces Structural Bias Against Minority Groups

Boulamwini (2018), in Gender Shades, demonstrated that commercial gender classification systems had lower accuracy for darker-skinned women (inaccuracies of up to 34–46%). This finding is particularly relevant for Islamic and multicultural education, as automated facial identification and classification systems could be used in schools, Islamic boarding schools, or online examination systems. Mehrabi et al. (2022) in their comprehensive survey asserted that bias in ML can arise from data, algorithms, or user feedback. This study emphasized that the choice of fairness metric is not merely technical but encodes a value judgment about what constitutes fairness in a given sociotechnical context, a critical statement for AI policies in Islamic boarding schools and multicultural schools.

Linguistic and Representational Bias in LLM Threatens Epistemic Fairness

Kasneci et al. (2023) warn that LLMs are not culturally neutral; they reflect the data they are trained on, which is predominantly English-centric and Western and educational frameworks must address this representational harm. This confirms that Islamic and multicultural education risks losing authenticity and representation if LLMs are used without local adaptation. Rafida et al. (2024) found that EFL students in Indonesia relied more on AI for writing than students in Taiwan, reflecting differences in self-efficacy and educational culture. This finding indicates

that AI adoption is not culturally neutral and it reinforces existing patterns of dependency.

What are the Main Ethical Challenges and Risks in AI Implementation in Multicultural and Islamic Educational Environments?

Ethical Considerations in Handling Religious and Cultural Student Data

Abadi et al. (2016) developed Deep Learning with Differential Privacy (DP-SGD), which allows model training while guaranteeing individual data privacy. In the context of Islamic education, this is crucial for protecting sensitive student data, including religious affiliation, ethnicity, Islamic boarding school location, and academic performance data—that could be discriminatory if leaked. Shokri et al. (2017) revealed that membership inference attacks can determine whether an individual's data is used in model training. The study warns that membership inference attacks exploit the fact that machine learning models often behave differently on the data they were trained on versus new data. This poses a significant privacy risk in domains like education.

Federated Learning as a Privacy Solution for Islamic Educational Institutions

Yang et al. (2019), Li et al. (2020), and Kairouz et al. (2021) collectively demonstrated that Federated Learning (FL) enables distributed model training without transferring raw data. In Islamic education, particularly in Islamic boarding schools (pesantren) in areas with limited infrastructure, FL protects sensitive student data while still enabling collaborative predictive model development. However, Kairouz et al. note that open problems in federated learning include fairness guarantees, robustness to poisoning, and meaningful transparency for participants.

AI Increases Student Dependency and Threatens Academic Integrity

Adiyono (2025) in a case study at IRCT Indonesia showed that access to AI (ChatGPT, etc.) has increased student dependence on AI to answer exam questions. This study warns that " Students are increasingly relying on AI to answer exam questions, threatening academic integrity and the development of critical thinking skills, core objectives of Islamic education. Nawi (2021) asserts that there is an urgent need for Islamic ethical guidelines in AI research to ensure that technological advancement aligns with Islamic values and does not cause harm to the Muslim community. These challenges include the risk of over-reliance on AI, the erosion of intellectual autonomy, and the potential loss of *etiquette* in the teaching and learning process.

The Digital Divide Reflects and Reinforces Structural Inequalities

Tarisayi (2024) reports that ChatGPT use in South African universities is limited by *persistent digital divides* that reflect the legacy of apartheid. These findings are relevant for Islamic education in Indonesia, where many Islamic boarding schools face similar infrastructure challenges. Wedi et al. (2025) in a case study at Madrasah Aliyah East Java identified the digital divide between urban and rural madrasahs; the lack of AI-based PAI content; and senior teacher resistance as the main challenges of PAI digital transformation in the AI era.

What Ethical Frameworks and Solutions are Proposed to Ensure Inclusive and Culturally Responsive AI in Education?

Maqasid al-Shariah Provides an Ethical Foundation for Islamic AI Governance

Kosasih (2024) proposes a *Wasathiyyah* (moderation) approach in Islamic law as a middle

ground between uncritical acceptance of AI and its total rejection. This study asserts that the Wasathiyyah approach offers a balanced Islamic legal framework for AI governance, avoiding both technological extremism and Luddism, a key framework for ethical argumentation. Kannike (2025) developed an Islamic Ethical AI Governance framework based on *maqasid al-shariah* with five core principles: justice (*adl*), amanah (trust), maslahah (public benefit), prohibition of harm, and respect for human dignity (*karamah*). This study asserts that Western AI ethics principles are not universally applicable; Islamic ethical governance requires grounding in maqasid al-shariah and classical *fiqh* concepts. Nawi (2021) added recommendations to establish *a Shariah AI Review Board* in Islamic universities and develop *Islamic AI Research Guidelines* that refer to *maqasid al-shariah*.

XAI Enables Bias Auditing by Non-Technical Stakeholders

Ribeiro et al. (2016) developed LIME, a local explanation method that allows users to understand model predictions. This study asserts that if a model is not interpretable, users cannot trust it nor contest its decisions. In contexts like education, where decisions shape lives, interpretability is a prerequisite for fairness highly relevant to the principle *of adl* (justice) in Islamic education. Rudin (2019) gives a stark warning in high-stakes decisions, such as those in education, healthcare, and criminal justice, we should not replace an interpretable model with a black box and then try to explain it later. This study recommends the use of intrinsically interpretable models *for* high-impact decisions in education. Lundberg and Lee (2020) present SHAP, a unification framework for feature attribution. This research allows for the identification of which variables (native language, economic background, access to technology) unfairly influence the prediction of graduation or academic performance of minority students.

Human-AI Collaboration as a Model for Islamic Education

Zubaidi ((2025) developed *a* Tridimensional Human-AI Collaboration Framework for Arabic language learning that positions AI as a scaffold, feedback provider, and creativity partner. This framework respects the role of the teacher (*mu'allim*) as the center of learning and emphasizes that AI should not replace human relationships in Islamic education. Topol (2019) in the context of medicine provides an important analogy for education: "AI should not operate in a human-free zone. The best outcomes will come from the synergy of human intelligence, empathy, and experience with AI's pattern recognition and scalability" and it is in line with the Islamic principles of *adab* (ethics) and *ta'awun* (cooperation). Zawacki-Richter et al. (2019) warn that educators remain largely absent from the AI conversation in higher education research. This is a critical gap for multicultural and inclusive education. The solution is to involve educators, including Islamic religious education teachers and religious teachers (*ustadz/ustadzah*), throughout the entire AI development cycle.

How does Multicultural Islamic Education in Indonesia and the Global South Respond to the Challenges and Opportunities of AI?

AI Adoption in Islamic Religious College (IRC) is Still Low due to Policy and Capacity Gaps

Suwendi (2025), in a survey of lecturers at IRC in Indonesia, found that AI adoption remains low due to a lack of training, infrastructure, and institutional policies. *Perceived usefulness* and *social influence* are the main predictors of adoption, indicating that social support and perceived benefits are more important than technical ease. Syukur et al. (2024) in a comparative study between Indonesia and Thailand reported that IRC in Indonesia is stronger in Islamic ethical framework, while Thailand is more advanced in infrastructure. This finding suggests that the

integration of AI in Islamic higher education requires a balance between technological innovation and Islamic ethical values.

Madrasahs Need a Holistic Digital Transformation Model

Wedi et al. (2025) identified that the digital transformation of Islamic Religious Education in the AI era requires a model that integrates infrastructure, teacher competency, curriculum, and visionary leadership. This study asserts that the digital transformation of Islamic Religious Education in the AI era cannot be technology-driven alone; it requires a holistic model that addresses infrastructure, human resources, curriculum, and leadership simultaneously. Wong-A-Foe (2023) uses an anthropological approach to reveal that the acceptance of AI in Indonesian Islamic education is not only determined by technical factors, but also by local interpretations of Islam, religious authority, and traditional pedagogical practices. This study concludes that AI in Islamic education is not merely a technical adoption but a cultural negotiation between traditional religious authority and modern technological rationality.

AI for Qur'an and Arabic Learning is Developing Rapidly

Elzeiny (2025) developed an AI-based learning program to improve Quranic recitation (*tajwid*) skills for non-native Arabic speakers, showing significant effectiveness in pronunciation (*makharij al-huruf)* and tajwid rules. This study recommends a blended learning model that integrates AI for self-practice and a teacher (*qari*) for spiritual guidance. Muhamad (2025) developed an AI model (NLP + ML) to generate lessons from Quranic verses through an interactive interpretation approach. However, this study cautions that AI can help make the rich meanings of the Quran more accessible to learners, but it must be designed as an aid, not an authority, in Islamic religious education. Dzulraidi (2025) evaluated the accuracy of ChatGPT in translating the Quran and found that ChatGPT had limited accuracy, especially for *mutasyabihat* (metaphorical) verses. This study asserts that ChatGPT cannot replace human expertise in Quranic translation. The sacred text requires not only linguistic accuracy but also theological depth and contextual understanding that AI currently lacks.

Harmonization of AI Policy in Islamic Education between Countries is Urgently Needed

Husin et al. (2025) proposed an NLP AI policy framework to support inclusion in higher education in Indonesia and Malaysia. Quantitative findings indicated positive perceptions of NLP AI, but qualitative findings revealed concerns about data privacy and institutional readiness. This study emphasized the need for a cross-border policy framework for inclusive NLP AI; the development of local language NLP models; and a mechanism for consent and opt-out for students. Zulkifli (2025), in a systematic review of ethical guidelines for AI from an Islamic perspective, identified five core principles: justice (*adl*), trustworthiness, maslahah, privacy (*sitr*), and accountability (*hisbah*). However, the study concluded that a significant gap remains between principle and practice. Implementation frameworks are urgently needed.

Negative and Contradictory Findings

The analysis also reveals a set of critical tensions and contradictory findings that complicate the integration of AI in educational contexts, particularly within Islamic education. One major debate concerns the use of explainable artificial intelligence (XAI). While some scholars advocate post-hoc explanation methods such as LIME and SHAP to make complex models more understandable, others argue that these "black-box" approaches risk producing misleading interpretations and instead recommend the use of inherently interpretable models from the outset. This lack of consensus highlights an unresolved methodological dilemma regarding

transparency and trust in educational AI systems. In parallel, a significant gap emerges between the articulation of Islamic ethical principles for AI and their actual implementation. Although there is broad agreement on values such as justice, accountability, and Shariah compliance, practical integration into real-world AI systems remains limited, indicating a disconnect between normative frameworks and technological practice.

Further conditions arise in relation to the role of AI in religious knowledge and institutional culture. Multiple studies emphasize that while AI can support the study of sacred texts, it cannot replace human scholarly authority in the interpretation and translation of the Qur'an, thereby reinforcing the epistemic boundaries of technology in religious domains. At the same time, there is an ongoing tension between cultural preservation and technological transformation in Islamic educational institutions. Traditional practices—such as the use of classical texts (kitab kuning), memorization, and close teacher-student relationships—often stand in contrast to emerging pressures for digitalization and AI adoption. This tension reflects a broader, unresolved negotiation between maintaining established pedagogical traditions and embracing innovation, suggesting that the integration of AI in such contexts must be approached with careful cultural and ethical sensitivity.

DISCUSSION

A thematic analysis of 53 core studies revealed that the integration of AI in Islamic and multicultural education lies at the intersection of transformational opportunities and the risk of systemic exclusion. This finding reinforces the argument that AI is not a value -*neutral tool*, but rather a sociotechnical system that carries with it the epistemological, cultural, and ethical assumptions of its development context (Buolamwini & Gebru, 2018; Kasneci et al., 2023). AI offers culturally responsive personalized learning (Faizin et al., 2025; Veza et al., 2024), yet simultaneously reinforces structural bias against minority groups (Mehrabi et al., 2022). This contradiction reflects what Wong-A-Foe (2023) calls "cultural negotiation between traditional religious authority and modern technological rationality." In the context of Islamic education, AI-mediated personalization faces unique challenges: religious learning is not merely about transferring information, but rather about character formation (*ta'dib*) and the spiritual relationship between teacher and student (*tarbiyah*). Topol (2019) asserts that AI should function as an augmentation, not a substitute of human relationships, in line with Islamic educational ethics.

Five Islamic ethical principles for AI, *adl* (justice), amanah (trust), *maslahah* (public benefit), *sitr* (privacy), and *hisbah* (accountability), have reached theoretical consensus (Aliff Nawi et al., 2021; Kannike & Fahm, 2025; Kosasih et al., 2024). However, Zulkifli et al. (2025) confirmed that their implementation in real-world AI systems remains very limited. This gap is not unique to the Islamic context; in the XAI literature, there is no consensus on whether post-hoc explanations (LIME, SHAP) are sufficiently reliable for high-impact decisions in education (S. M. Lundberg, 2020; Ribeiro, 2016; Rudin, 2019). This study confirms that the Islamic context requires not only technical transparency but also epistemic justice. In other words, the recognition of different ways of knowing has equal legitimacy (Mariyono et al., 2025). Low AI adoption in Islamic higher education institutions and Islamic boarding schools (Islamic boarding schools) (Suwendi et al., 2025; Wedi et al., 2025) is not solely due to cultural resistance, but also due to a lack of adequate infrastructure, institutional policies, and training. Tarisayi (2024) in South Africa showed a similar pattern, indicating that this challenge is global in the Global South. These findings reinforce Zawacki-Richter et al.'s (2019) argument that educators remain largely absent from the AI conversation, which indicates a gap that must be filled through participatory design and bottom-up innovation approaches.

This study advances the discussion on AI in education by offering an integrative theoretical framework that brings together multicultural education, responsible AI, and Islamic

educational ethics, three domains that have often been treated in isolation. Through these perspectives, the study addresses fragmentation in the existing literature and provides a more holistic lens for understanding AI's role in culturally and religiously diverse contexts. Central to this contribution is the formulation of five pillars for glocalized AI ethics: epistemic justice, cultural grounding, religious sensitivity, participatory governance, and glocal accountability, which collectively emphasize the need for AI systems to be both globally informed and locally relevant. Furthermore, the study highlights *Wasathiyyah* (moderation) as a guiding ethical principle in AI governance within Muslim societies, positioning it as a balanced approach that avoids both uncritical technological adoption and outright resistance, thereby aligning innovation with the preservation of core values.

These theoretical insights carry important practical implications across multiple stakeholders while also revealing key limitations and directions for future inquiry. Policymakers are encouraged to design inclusive AI governance frameworks that involve religious educators and communities, ensure equitable infrastructure distribution, and incorporate Islamic ethical standards such as Shariah-compliant certification. For technology developers, the findings underscore the importance of culturally diverse datasets, accessible explainability, privacy-preserving methods, and human-centered design. Educators and institutions, meanwhile, are called to foster critical AI literacy and establish ethical guidelines grounded in Islamic values. However, the study is not without limitations, including language restrictions in the reviewed literature, the exclusion of grey literature, and constraints on qualitative synthesis in assessing impact. The rapidly evolving nature of AI also necessitates continuous updates. Future research should therefore prioritize empirical validation of the proposed ethical pillars, development of Shariah-compliant AI systems, cross-regional comparisons, and long-term investigations into AI's effects on students' critical thinking, integrity, and character development.

CONCLUSION

This study analyzed 53 international publications on the intersection of Artificial Intelligence (AI), inclusion, multiculturalism, and Islamic education. Five dominant themes were identified: (1) personalization and adaptation; (2) algorithmic bias and cultural exclusion; (3) ethical governance and human oversight; (4) opportunities for Islamic pedagogy; and (5) digital inequality. The main contributions are the integration of the three perspectives into a coherent framework, support for the five pillars as a model for *glocalized* AI governance, and strengthening the principles of *Wasathiyyah*, *maqasid al-shariah*, and *human-AI collaboration* as ethical foundations. In conclusion, AI in Islamic and multicultural education should not be understood solely as a technical innovation, but rather as a social and moral project. The success of AI integration is measured not by efficiency or accuracy alone, but by the extent to which AI strengthens, rather than undermines, justice, inclusion, human dignity, and the fundamental goal of Islamic education: to shape individuals with noble character who benefit others (*khaira ummah*).

ACKNOWLEDGMENTS

The author gratefully acknowledges scholarly communities, librarians, and peer reviewers whose global contributions to educational technology and multicultural research informed this study.

DATA AVAILABILITY STATEMENT

All data analyzed in this study consist of publicly accessible scholarly publications indexed in

major academic databases. The analytical coding framework is available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declare no conflict of interest. The research was conducted independently without any commercial, financial, or personal relationships that could have inappropriately influenced the findings or interpretation of the results.

REFERENCES

Abadi, M., McMahan, H. B., Chu, A., Mironov, I., Zhang, L., Goodfellow, I., & Talwar, K. (2016). Deep learning with differential privacy. *Proceedings of the ACM Conference on Computer and Communications Security*, 308–318. https://doi.org/10.1145/2976749.2978318

Adiyono, A., Suwartono, T., Nurhayati, S., Dalimarta, F. F., & Wijayanti, O. (2025). Impact of artificial intelligence on student reliance for exam answers: A case study in IRCT Indonesia. *International Journal of Learning, Teaching and Educational Research*, *24*(3), 455–479. https://doi.org/10.26803/ijlter.24.3.22

Arrieta, A. B. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research*, *1*(3). https://doi.org/10.1177/146879410100100307

Benoot, C., Hannes, K., & Bilsen, J. (2016). The use of purposeful sampling in a qualitative evidence synthesis. *BMC Medical Research Methodology*, *16*(21), 1–12.

Birks, M., & Mills, J. (2015). *Grounded Theory: A Practical Guide (2nd ed.).* SAGE Publications.

Braun, V., & Clarke, V. (2006). Qualitative research in psychology using thematic analysis in psychology using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. http://www.tandfonline.com/action/journalInformation?journalCode=uqrp20%5Cnhttp://www.tandfonline.com/action/journalInformation?journalCode=uqrp20

Braun, V., & Clarke, V. (2021). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, *9*(1), 3–26. https://doi.org/10.1037/qup0000196

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 77–91.

Castleberry, A., & Nolen, A. (2018). Thematic analysis of qualitative research data: Is it as easy as it sounds? *Currents in Pharmacy Teaching and Learning*, *10*(6), 807–815. https://doi.org/10.1016/j.cptl.2018.03.019

Chen, T. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13*, 785–794. https://doi.org/10.1145/2939672.2939785

Creswell, J.W., Poth, C. N. (2018). Q*ualitative inquiry and research design, hoosing among five approaches (4th Edition)*. SAGE Publication Inc.

Dwi, M., & Hd, A. N. A. (2024). Transformative impact of AI on multicultural education: A qualitative thematic analysis. *Edelweiss Applied Science and Technology*, *8*(5). https://doi.org/10.55214/25768484.v8i5.1667

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi,

A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Dzulraidi, D. H., Jalil, A. A., & Sin, M. L. M. (2025). The accuracy of ChatGPT in translating the Qur'an: A comparative analysis with authoritative translation. *Quranica*, *17*(2), 21–52.

Elzeiny, M., Alwaely, S., Arz, A., Alshamisi, W., & Alblooshi, A. (2025). Enhancing Qur'anic recitation skills through an AI-based learning program for non-native Arabic speakers. *International Conference on Social Networks Analysis, Management and Security, SNAMS*, *2025*, 108–115. https://doi.org/10.1109/SNAMS67467.2025.11390980

Faizin, N., Alfan, M., Basid, A., Ramadhan, M. R., Panatik, S. A., & Kawakip, A. N. (2025). Muslim students' acceptance of artificial intelligence in Islamic religious education: an extended TAM approach. *Discover Education*, *4*(1), 304. https://doi.org/10.1007/s44217-025-00767-1

Flick, U. (2018). The SAGE handbook of qualitative data collection. *The SAGE Handbook of Qualitative Data Collection*. https://doi.org/10.4135/9781526416070

Garrard, J. (2017). *Health Sciences Literature Review Made Easy: The Matrix Method 5th Edition* (5th ed.). Jones & Bartlett Learning.

Gough, D., Thomas, J., & Oliver, S. (2017). *Systematic reviews and research: A guide for synthesizing evidence in education*. SAGE.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

He, K. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hennink, M. M., Kaiser, B. N., & Weber, M. B. (2019). What influences saturation? Estimating sample sizes in focus group research. *Qualitative Health Research*, *29*(10), 1483–1496. https://doi.org/10.1177/1049732318821692

Husin, N., Fazlurrahman, H., Safitri, A., Dhenabayu, R., Rauf, U. A. A., & Fitrah, A. M. (2025). Policy perspective on proposed framework of NLP AI to bridge the inclusive support in higher education with a mixed methods approach in Indonesia and Malaysia. *International Journal of Information and Education Technology*, *15*(12), 2686–2699. https://doi.org/10.18178/ijiet.2025.15.12.2464

K., M. (2001). Qualitative research: Standards, challenges, and guidelines. *Lancet*, *358*(2).

Kairouz, P. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, *14*(1), 1–210. https://doi.org/10.1561/2200000083

Kannike, U. M. M., & Fahm, A. O. (2025). Exploring the ethical governance of artificial intelligence from an Islamic ethical perspective. *Jurnal Fiqh*, *22*(1), 134–161. https://doi.org/10.22452/fiqh.vol22no1.5

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*. https://doi.org/10.1016/j.lindif.2023.102274

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

Kosasih, E., Islamy, M. R. F., & Wiwaha, R. S. (2024). Artificial Intelligence in the era of society 5.0: Compromising technological innovation through the wasathiyyah approach within the framework of Islamic law. *Al-Istinbath: Jurnal Hukum Islam*, *9*(2), 519–540. https://doi.org/10.29240/jhi.v9i2.9596

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical Kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, *33*(2), 363. https://doi.org/10.2307/2529786

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-69

Li, T. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lundberg, S. M. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 2017, 4766–4775. https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85044542379&origin=inward

Lundberg, S. M. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S. M. . & L. S. I. (2020). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *4765-4774*.

MacDonald, J. (2014a). Systematic approaches to a successful literature review. *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada*, *34*(1). https://doi.org/10.5596/c13-009

MacDonald, J. (2014b). Systematic approaches to a successful literature review. *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada*, *34*(1). https://doi.org/10.5596/c13-009

Mariyono, D., Maskuri, & Hidayatullah, A. N. A. (2025). Quantifying AI's role in conflict resolution: Assessing its potency in mediation and peacekeeping. *Dirasat: Human and Social Sciences*, *52*(5), 7610. https://doi.org/10.35516/hum.v52i5.7610

Mariyono, D., Yunus, M., & Hidayatullah, A. N. A. (2025). *Decolonizing AI ethics in education: A Systematic Review and the Framework for Glocalized AI Ethics in Education (FGAIEE)*. https://doi.org/10.21203/rs.3.rs-8417921/v1

McMahan, H. B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics Aistats 2017*. https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85083937116&origin=inward

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, *54*(6), 1–35. https://doi.org/10.1145/3457607

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, *14*(1), 579. https://doi.org/10.1186/s12913-014-0579-0

Muhamad, F., Yaakob, M., Abdullah, M., Mohamad, N., & Basarud-Din, S. (2025). Artificial Intelligent (AI) model for generating lessons of Quranic chapters through an interactive tafsir approach. *QURANICA, International Journal of Quranic Research*, 1–20.

Nawi, A., Yaakob, M. F. M., Hussin, Z., Muhaiyuddin, N. D. M., Samuri, M. A. A., & Tamuri, A. H. (2021). Keperluan garis panduan dan etika Islam dalam penyelidikan kecerdasan

buatan. *Journal of Fatwa Management and Research*, *26*(2), 280–297. https://doi.org/10.33102/jfatwa.vol26no2.414

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1609406917733847. https://doi.org/10.1177/1609406917733847

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods (4th ed.). Integrating Theory and Practice* (4th ed.). SAGE Publications, Inc.

PyTorch Development Team. (2024). *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. https://pytorch.org

Rafida, T., Suwandi, S., & Ananda, R. (2024). EFL students' perception in Indonesia and Taiwan on using artificial intelligence to enhance writing skills. *Jurnal Ilmiah Peuradeun*, *12*(3), 987–1016. https://doi.org/10.26811/peuradeun.v12i3.1520

Ribeiro, M. T. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13, 135–1144. https://doi.org/10.1145/2939672.2939778

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2020). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, *128*(2), 336–359. https://doi.org/10.1007/s11263-019-01228-7

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. *Proceedings - IEEE Symposium on Security and Privacy*, 3–18. https://doi.org/10.1109/SP.2017.41

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Suwendi, S., Gama, C. B., Rahman, H., Luhuringbudi, T., Masrom, M., & others. (2025). Adoption of Artificial Intelligence and digital resources among academicians of Islamic higher education institutions in Indonesia. *Jurnal Online Informatika*, *10*(1), 42–52. https://doi.org/10.15575/ join.v 10i1.1549)

Syukur, F., Maghfurin, A., Marhamah, U., & Phaosan Jehwae. (2024). Integration of Artificial Intelligence in Islamic higher education: Comparative responses between Indonesia and Thailand. *Nazhruna: Jurnal Pendidikan Islam*, *7*(3), 531–553. https://doi.org/10.31538/nzh.v7i3.13

Tarisayi, K. S. (2024). ChatGPT use in universities in South Africa through a socio-technical lens. *Cogent Education*, *11*(1), 2295654. https://doi.org/10.1080/2331186X.2023.2295654

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Veza, I., Ghazali, I., Putra, A., Ascencio, R. L., Muhammad, M., & Irianto, I. (2024). How ChatGPT affects education landscape: Effects of ChatGPT on higher education accessibility and inclusivity. *Lecture Notes in Educational Technology*, 2024, 569–579. https://doi.org/10.1007/978-981-97-4507-4_64

Wedi, A., Mardiana, D., & Umiarso, U. (2025). Digital transformation model of Islamic religious education in the AI Era: A case study of Madrasah Aliyah in East Java, Indonesia. *International Journal of Learning, Teaching and Educational Research*, *24*(8), 842–863. https://doi.org/10.26803/ijlter.24.8.37

Wong-A-Foe, D. (2023). Navigating the implications of AI in Indonesian education: Tutors, governance, and ethical perspectives. *Data Science and Artificial Intelligence, 1942,* (349–360). https://doi.org/10.1007/978-981-99-7969-1_26

Wong-A-Foe, D., Barendregt, B., & Lamers, M. H. (2023). Exploring AI and Islam in Indonesian education: An anthropological inquiry. *Proceedings of the International Conference on Electrical Engineering and Informatics*. https://doi.org/10.1109/ICEEI59426.2023.10346759

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning. *ACM Transactions on Intelligent Systems and Technology*, *10*(2), 1–19. https://doi.org/10.1145/3298981

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zubaidi, A., Munip, A., Widodo, S. A., & Zerrouki, T. (2025). Enhancing Arabic writing skills using Chat GPT-based AI learning models: A tridimensional human-AI collaboration framework. *Indonesian Journal of Applied Linguistics, 15*(1), 87–101. https://doi.org/10.17509/ijal.v15i1.75378

Zulkifli, D. U., & Tungkagi, D. Q. (2025). Artificial intelligence and Islamic ethical guidelines: A systematic review. *13th International Conference on Cyber and IT Service Management (CITSM)*, 1–9. https://doi.org/10.1109/citsm67730.2025.11291209.