Algorithmic Discernment: The Systematic Integration of Artificial Intelligence in Religious Education

Mulima Owen

Kwame Nkrumah University, Kabwe, Zambia

owen.mulima@nkrumah.edu.zm

https://orcid.org/0000-0002-2105-4038

ABSTRACT

This study examines the tension between the algorithmic efficiency of Artificial Intelligence (AI) and the fundamentally formative character of Religious Education (RE). While AI offers tangible administrative and pedagogical advantages, its integration into sacred instructional spaces risks undermining theological integrity and human relationality. Drawing on a qualitative library research methodology and thematic analysis, we evaluate institutional case studies and scholarly frameworks to identify the systemic impact of AI on spiritual instruction. Key findings underscore the necessity of human-in-the-loop systems to mitigate cognitive biases, particularly confirmation bias and algorithmic partiality, while leveraging AI for personalised learning and virtual religious exploration. The analysis establishes that AI functions most responsibly as an epistemological instrument rather than an authoritative replacement for the human educator. We conclude by proposing a human-centric hybrid pedagogical model. This framework prioritises ethical governance and contextual sensitivity, ensuring that digital transformation serves the transcendental goals of character formation and spiritual intelligence within diverse educational contexts, with specific attention to the Global South.

Keywords: algorithmic discernment; religious education; artificial intelligence; cognitive bias; digital theology; Global South pedagogy

1. INTRODUCTION

The incorporation of Artificial Intelligence (AI) into formal education has accelerated at a pace that few policymakers, curriculum designers, or theologians anticipated. From automated grading systems to intelligent tutoring platforms, AI has moved from experimental use to structural embedding across multiple educational sectors within a decade (Chen et al., 2020). Within this general expansion, however, Religious Education (RE) occupies a distinctly contested position. RE is not primarily concerned with the efficient transfer of factual content; it is a discipline oriented toward the formation of moral character, spiritual intelligence, and communal identity (Papakostas, 2025). The question of whether algorithmic tools can serve these goals without distorting them is not merely technical. It is fundamentally theological.

The central research question guiding this inquiry is: how can generative AI be integrated into Religious Education without compromising the theological integrity and relational formation that define the discipline? Three subsidiary questions follow. First, what patterns of cognitive bias does AI introduce into faith-based instructional environments? Second, how have specific institutions attempted to reconcile digital tools with the demands of spiritual formation? Third, what governance and pedagogical frameworks can mitigate the risks of uncritical AI adoption in RE settings?

To situate these questions, we define three foundational constructs. Generative AI refers to deep-learning systems trained on large datasets to produce novel outputs, including text, image, and audio, in ways that approximate human composition (Zhang et al., 2025). Pedagogical competence describes an educator's demonstrated capacity to facilitate differentiated

instruction, formative assessment, and reflective professional practice (Harefa, 2025). Digital theology denotes the systematic intellectual discipline that reflects on the intersection of digital culture, emerging technologies, and religious ontology, examining how faith, praxis, and community are reconfigured in digitally mediated environments (Papakostas, 2025).

The conceptual puzzle at the heart of this study lies in the convergence of algorithmic efficiency and the irreducibly human character of religious formation. Digital transformation in education is not optional: rising demand for distance learning, pressure on institutional resources, and shifting learner expectations have made some degree of technological adoption inevitable (Ahmad et al., 2021). Yet AI systems, as probabilistic pattern-matching engines, lack what we might term historical consciousness, the cultivated sensitivity to context, tradition, and communal memory that allows a skilled religious educator to interpret a learner's silence, redirect a theologically inaccurate claim, or situate a doctrinal question within its living tradition (Alimukhamedova et al., 2026). Machines can approximate the syntax of religious discourse; they cannot access its semantics or its sincerity.

This problem is not abstract. At St. Paul University Philippines (SPUP), researchers found that AI tools successfully supported administrative functions and expanded interfaith programming, but consistently required human oversight to preserve theological coherence (Lappay, 2025). At Agios Nikolaos Christian Junior High School in Indonesia, adaptive learning platforms improved accessibility for learners with diverse needs but demanded continuous pedagogical calibration by trained educators (Harefa, 2025). These cases illustrate a broader structural pattern: AI works in RE when it is governed; it becomes a liability when left to operate autonomously.

The significance of this inquiry extends beyond institutional case management. In the Global South, where the digital divide intersects with under-resourced educational systems and rich

oral-theological traditions, the stakes of uncritical AI adoption are particularly high (Choizhalanova, 2021). Algorithmic systems trained predominantly on Western, Protestant, and English-language datasets risk marginalising the theological particularities of Catholic, Orthodox, African traditional, and Islamic instructional frameworks (Papakostas, 2025). This is not a failure of intent; it is a structural consequence of how large language models are trained and deployed.

The study proceeds in four stages. The literature review maps the historical trajectory of educational technology and identifies the research gap in cognitive bias theory as applied to faith-based instruction. The methodology section details the qualitative library research approach and the analytical procedures governing source selection and thematic synthesis. The results section presents findings organised by empirical theme. The discussion interprets those findings through the lens of heuristics and bias theory and proposes a human-centric hybrid pedagogical model suited to diverse theological contexts.

2. LITERATURE REVIEW

2.1 The Historical Trajectory of Educational Technology

The integration of technology into formal instruction has a history that predates the digital era. B.F. Skinner's programmed instruction model (1954) established the earliest systematic framework for technology-mediated learning, premised on behaviourist principles of stimulus and reinforcement. By the 1960s, the PLATO system at the University of Illinois had introduced rudimentary online interaction and early forms of computer-assisted instruction. The 1970s and 1980s saw the emergence of rule-based Intelligent Tutoring Systems, most notably SCHOLAR (Carbonell, 1970) and the Andes Physics Tutor (Schulze et al., 2000), which simulated aspects of human tutoring through structured dialogue and adaptive feedback.

Across this trajectory, the defining assumption remained consistent: technology is instrumentally neutral, and its educational value depends on how it is designed and deployed. General education sectors adopted this logic productively, using quantifiable performance metrics to evaluate effectiveness. Religious Education resisted this instrumental logic because its core objectives are not quantifiable. Character formation, spiritual discernment, moral reasoning, and communal solidarity cannot be measured by completion rates or test scores (Papakostas, 2025). This divergence defines the conceptual tension that AI now intensifies.

Table 1. *Heuristic Divergence: Quantifiable Metrics vs. Transcendental Priorities*

General AI Education Priorities	**Religious Education Priorities**
Performance metrics: quantifiable test scores and completion rates	Character formation: moral maturation and internalisation of values
Efficiency: automation of grading and administrative processing	Spiritual intelligence: cultivation of wisdom, discernment, and relational depth
Information transfer: scalable delivery of content and micro-concepts	Moral discernment: ethical reasoning rooted in tradition, community, and lived experience

Note. Adapted from Papakostas (2025).

2.2 Pedagogical Frameworks for Theological Integration

Scholars have proposed several frameworks for thinking about how AI might be introduced into RE without collapsing its formative character. Ninian Smart's phenomenological framework offers one starting point: it positions RE as an exercise in cultivating empathetic understanding of religious phenomena across traditions (Papakostas, 2025). Within this framework, AI-mediated virtual reality has clear potential; immersive simulations of sacred sites such as Jerusalem, Mecca, or Varanasi can bring experiential proximity to traditions that learners would never otherwise encounter. The risk, however, is what phenomenologists call

the reduction of the lived world: the tendency for mediated experience to simplify complex ritual and reduce it to aestheticised spectacle.

Robert Jackson's interpretive or dialogical framework treats RE as an exercise in intercultural understanding, where learners explore religious traditions by engaging with their internal diversity and debating their claims (Papakostas, 2025). AI chatbots trained on diverse theological texts can function as dialogue partners in this framework, exposing learners to interfaith perspectives and prompting comparative analysis. The limitation is significant, however: these chatbots generate statistically probable responses, not theologically accountable ones. They do not hold convictions; they model the linguistic patterns of those who do.

Thomas Groome's shared praxis model offers the most relational of the three frameworks. It positions RE as a process of bringing faith traditions into conversation with lived experience through cycles of action and reflection (Papakostas, 2025). AI can contribute to this process by generating personalised reflection prompts, aggregating learner responses, and identifying patterns in moral reasoning across a cohort. Yet the shared praxis model ultimately depends on the educator's pastoral relationship with learners, a dimension that no algorithm can replicate.

2.3 Cognitive Bias and Algorithmic Partiality

The application of cognitive bias theory to AI-mediated learning environments represents an emerging and under-examined area of scholarship. Tversky and Kahneman's foundational work on heuristics and biases established that human judgement systematically deviates from normative models of rationality in predictable ways (Kahneman, 2011). This framework is now being applied to AI systems, which exhibit analogous patterns: confirmation bias, where algorithms prioritise content that reinforces pre-existing orientations; availability bias, where

frequently accessed sources dominate retrieval; and representativeness bias, where training data skews outputs toward dominant cultural patterns (Zhang et al., 2025).

In the context of RE, these biases carry particular weight. Confirmation bias in an AI tutor can reinforce theological exclusivism by consistently presenting content aligned with a learner's stated tradition, while systematically suppressing minority or dissenting perspectives. Algorithmic partiality, rooted in the demographic composition of training datasets, can functionally marginalise theological traditions that are under-represented in digital archives, including those of the Global South, indigenous communities, and oral theological traditions (Papakostas, 2025). The result is a technically functional but theologically partial instructional system.

2.4 Research Gap

Despite substantial scholarship on AI in general education (Ahmad et al., 2021; Chen et al., 2020; Harry & Sayudin, 2023) and growing attention to digital theology (Papakostas, 2025; Wilson, 2025), there is a notable absence of interdisciplinary studies that apply systematic cognitive bias theory specifically to faith-based instruction in the Global South. Existing studies either examine AI in education broadly, without attending to the theological specificity of RE, or discuss digital theology in abstraction, without empirical grounding in institutional practice. This study addresses that gap by combining bias theory with case-based institutional analysis and attending explicitly to the pedagogical implications for under-resourced theological contexts.

3. METHODOLOGY

3.1 Research Design

This study employs a qualitative library research design. Library research is appropriate when the research objective is conceptual synthesis rather than primary data collection, and when the

phenomenon under investigation is best understood through critical engagement with existing scholarly, institutional, and policy documentation (Hamzah, 2020). The nature of this inquiry, which concerns the structural and theoretical dimensions of AI integration in RE, makes empirical field research premature: institutional practice is still nascent, methodological frameworks are still forming, and the conceptual vocabulary needed to interpret empirical data is itself under construction. Library research, in this context, is not a methodological concession; it is the appropriate design choice.

3.2 Research Context

The study draws on documentation and scholarly accounts from four principal institutional contexts: St. Paul University Philippines (SPUP), which has piloted AI integration in its interfaith education programmes; Agios Nikolaos Christian Junior High School in Gunungsitoli, Indonesia, where adaptive learning strategies have been implemented in a multi-denominational setting; the Oxford Centre for Digital Theology based at Durham University, which has produced sustained theoretical work on the intersection of digital culture and religious formation; and Harvard Divinity School, whose interdisciplinary research on religion and technology provides comparative grounding (Lappay, 2025; Harefa, 2025; Papakostas, 2025). These contexts represent a range of theological traditions, geographic regions, and institutional scales, supporting analytical breadth.

3.3 Source Selection and Sampling Strategy

Sources were selected through purposive sampling, guided by three criteria. First, relevance: sources were included only where they addressed AI in education, RE, cognitive bias, or digital theology in ways directly pertinent to the study's research questions. Second, recency: with the exception of foundational theoretical texts, sources were required to have been published within the past five years. Third, credibility: sources were drawn from peer-reviewed journals,

institutional reports, and scholarly monographs, with preference given to works indexed in recognised academic databases. Grey literature was included where it offered institutional case documentation not available in peer-reviewed form.

3.4 Data Collection

Data collection proceeded through systematic database searches across Google Scholar, JSTOR, ERIC, and institutional repositories. Search terms included combinations of the following: artificial intelligence and religious education; algorithmic bias and pedagogy; digital theology; AI and spiritual formation; cognitive bias and learning environments; and inclusive education and AI. Searches were conducted iteratively, with each phase of conceptual development informing subsequent search refinements. A total of thirty-one sources met the inclusion criteria and are cited in this article.

3.5 Data Analysis

Analysis followed the Miles and Huberman interactive model, which involves three concurrent processes: data reduction, data display, and conclusion drawing (Mundofi, 2025). In the data reduction phase, collected sources were read closely and annotated for their primary arguments, methodological approaches, and empirical claims. Recurring themes were identified inductively. In the data display phase, themes were organised into an analytical matrix structured around three organising categories: pedagogical ecosystems, epistemic conformity, and adaptive learning. In the conclusion-drawing phase, patterns across categories were interpreted theoretically, with findings synthesised into the human-centric hybrid model proposed in the discussion.

Methodological transparency was maintained through content checkpoint techniques, whereby interpretive claims were verified against the primary source before inclusion (Sugiyono, 2008). Validity was assessed against the Mixed Methods Appraisal Tool (MMAT) criteria adapted for

qualitative library research, ensuring that the synthesis met standards of representativeness, coherence, and analytical rigour.

3.6 Ethical Considerations

This study involves no human participants and therefore did not require institutional ethics approval. All sources cited are publicly available and are used in accordance with standard scholarly fair use principles. The researcher declares no financial or professional conflicts of interest. The study does not misrepresent the findings of cited works, and where interpretive claims extend beyond the explicit conclusions of cited sources, this is noted explicitly in the text.

4. RESULTS

4.1 AI Applications in Religious Education: Spaces, Support, and Services

Findings from SPUP and Agios Nikolaos indicate that AI integration in RE can be meaningfully organised into three functional domains: spaces, support, and services (Lappay, 2025). Each domain addresses a distinct operational challenge, and each carries specific pedagogical implications.

In the domain of spaces, AI-mediated virtual pilgrimages offer learners immersive simulations of sacred sites, including Jerusalem, the Kaaba, and Buddhist temple complexes, through virtual reality environments. Lappay (2025) reports that at SPUP, these simulations expanded interfaith literacy among students with no direct exposure to traditions outside their own. Students engaged in facilitated post-simulation dialogue demonstrated greater capacity for perspective-taking than those who received equivalent content through text-based instruction. The simulations were most effective when accompanied by structured theological reflection led by a trained educator, rather than left as standalone digital experiences.

In the domain of support, adaptive learning platforms have been deployed to address learner diversity in RE classrooms. At Agios Nikolaos, these platforms provided dynamic content adjustment, including simplified linguistic explanations for complex theological concepts, extended processing time for learners with identified learning differences, and audio-based content delivery for students with low print literacy (Harefa, 2025). The educator's report from that institution is instructive:

> Adaptive learning is not solely a methodological approach but a pedagogical ecosystem where teacher expertise, ethical sensitivity, and reflective capacity converge to foster academic mastery and character formation. (Harefa, 2025, p. 91)

This finding supports the interpretation that AI tools enhance accessibility without independently generating the pedagogical conditions required for spiritual formation. The platform reduced friction; the teacher provided formation.

In the domain of services, natural language processing tools have been applied to facilitate and monitor interfaith dialogue. Yin et al. (2020) describe sentiment analysis applications that assess the affective tone of online religious discussions, flagging language patterns associated with disrespect, polarisation, or doctrinal dismissiveness. At SPUP, similar tools were used to monitor student interactions in online discussion boards, allowing educators to intervene before dialogue deteriorated into theological confrontation. Lappay (2025) notes that this application successfully reduced instances of hostile exchange without suppressing substantive theological disagreement, a distinction of considerable importance in RE.

4.2 The Emergence of Pedagogical Ecosystems

Across all three institutional contexts, the research identifies a consistent pattern: AI integration is most pedagogically effective when it operates within what Harefa (2025) terms a pedagogical ecosystem. This construct describes a structured environment in which AI tools, educator expertise, learner agency, and institutional governance interact dynamically rather than

operating independently. In the cases examined, AI tools that functioned autonomously, without consistent educator oversight and curricular integration, produced marginal or negative pedagogical outcomes. Those embedded within a structured ecosystem of professional reflection and ethical governance produced measurable gains in learner engagement and interfaith competency.

A second pattern concerns the redistribution of educator time. In both the SPUP and Agios Nikolaos contexts, the delegation of administrative and diagnostic tasks to AI platforms freed educators to invest greater attention in pastoral functions: one-to-one theological dialogue, moral coaching, and community-building activities. This redistribution was not automatic; it required deliberate institutional planning and ongoing professional development for educators unfamiliar with AI tools (Harefa, 2025; Lappay, 2025).

4.3 The Epistemological Shift: From Knowledge Transfer to Spiritual Inquiry

A third finding concerns the epistemological function of AI in RE. Across the case studies, AI tools facilitated a shift in instructional emphasis, from the transmission of doctrinal content to the facilitation of reflective theological inquiry. Mundofi (2025) documents this shift in the context of Islamic Religious Education, where AI-assisted platforms enabled students to engage with classical texts through guided questioning and comparative analysis, rather than passive recitation. The platform did not interpret the texts; it generated questions that prompted students to interpret them. The distinction matters: AI operated as an epistemological instrument that structured inquiry, not as a theological authority that resolved it.

5. DISCUSSION

5.1 Heuristics, Bias, and the Normative Problem

The findings described above cannot be interpreted without attending to the systematic biases that AI systems introduce into instructional environments. Following the analytical tradition

established by Tversky and Kahneman (1974), we identify three heuristic failures of particular relevance to AI-mediated RE.

The first is confirmation bias. AI recommendation systems are designed, structurally, to maximise engagement by prioritising content aligned with a user's prior behaviour and stated preferences. In an RE context, this design logic creates what Zhang et al. (2025) call epistemic conformity: a condition in which learners are systematically exposed to theological content that reinforces their existing tradition while algorithmically marginalising perspectives that might challenge or enrich it. The bias is not visible to learners; it operates through the architecture of the system, shaping what counts as relevant or authoritative without announcing its criteria. The normative corrective, in Tversky's terms, is to make the selection criteria transparent and to subject them to pedagogical scrutiny.

The second is algorithmic partiality. Because large language models are trained predominantly on English-language, Western, and Protestant-oriented datasets, their representations of theological diversity are structurally skewed (Papakostas, 2025). Minority traditions, including African Indigenous Religions, Coptic Christianity, Sufi Islam, and oral theological traditions, are under-represented in training data and, consequently, in AI outputs. This is not a trivial distortion; in RE, misrepresentation of a tradition is a pedagogical harm. Students in Zambia, Ghana, or Papua New Guinea who encounter their own theological traditions refracted through an algorithmically Western lens receive a systematically inaccurate education.

The third is the Turing Trap, a concept Brynjolfsson (2022) uses to describe the tendency to attribute human-like understanding to systems that are, in fact, producing statistically probable outputs. In RE, this heuristic operates as follows: because AI systems produce theologically fluent text, users infer that they possess theological understanding. This inference is false, and its consequences are significant. An AI chatbot that generates a coherent commentary on the

Book of Job is not exercising exegetical judgement; it is recombining patterns from texts in which such judgements were made by humans. The appearance of understanding is a product of fluency, not cognition. The normative solution is not to abandon AI but to teach learners to distinguish between fluency and understanding, a distinction that becomes a curriculum objective in itself.

5.2 Theoretical Implications

These findings carry implications for several established theoretical frameworks. For Smart's phenomenological framework, they suggest that AI-mediated simulations require active phenomenological scaffolding by educators to prevent the reduction of complex ritual into consumable spectacle (Papakostas, 2025). Experiencing a virtual pilgrimage is not equivalent to cultivating empathetic understanding of the tradition that makes pilgrimage meaningful; the latter requires interpretive guidance that AI cannot provide unassisted.

For Jackson's dialogical framework, the findings suggest that AI dialogue partners can introduce diverse perspectives but cannot model the epistemic humility and genuine openness that authentic interfaith dialogue requires. Balasubramanian (2023) and Camilleri (2024) both note that AI decision-making systems tend toward confident output even under conditions of uncertainty, a disposition that is antithetical to the spirit of genuine religious inquiry.

For Groome's shared praxis model, the implications are perhaps most generative. AI can support the action-reflection cycle by aggregating learner responses, identifying patterns in moral reasoning, and generating personalised prompts. But the praxis element, the living tradition that gives reflection its normative orientation, must come from human sources (Shofiyyah et al., 2024). AI can structure the conversation; it cannot anchor it in lived faith.

5.3 Pedagogical and Policy Implications

At the institutional level, the findings support a human-centric hybrid model in which AI serves explicitly as a cognitive assistant rather than an instructional authority. In this model, three governance principles are foundational. The first is ethical governance: institutions integrating AI into RE should establish curriculum review committees with explicit mandates to assess algorithmic bias, evaluate dataset representativeness, and enforce transparency in AI decision-making. The Rome Call for AI Ethics (2020), endorsed by a broad coalition of religious and technological institutions, provides one available framework for this governance work (Camilleri, 2024).

The second principle is contextual sensitivity. AI tools must be adapted to reflect local theological traditions, linguistic diversity, and cultural norms before deployment in RE settings. This is not a technical afterthought; it is a theological requirement. Deploying a system trained on Western Christian datasets in a Zambian Islamic RE classroom without contextual calibration is an act of pedagogical negligence (Almeida et al., 2025).

The third principle is pedagogical augmentation. AI tools should be deployed specifically to reclaim educator time for pastoral functions, not to replace them. Institutions should resist the economic logic that treats AI as a cost-reduction mechanism and instead treat it as a professional development resource that extends the educator's formative capacity (Kalip et al., 2025).

For the Global South specifically, the digital divide constitutes a structural equity problem that governance frameworks must address directly. Where AI infrastructure is unavailable or unreliable, the gap between technologically enriched and impoverished institutions functions as a form of educational stratification (Choizhalanova, 2021). Policy frameworks for AI in RE must include resource equity provisions, not only pedagogical guidelines.

6. CONCLUSION

This study set out to examine a specific and consequential question: how can AI be integrated into Religious Education without undermining the theological integrity and relational formation that define the discipline? The findings support three principal conclusions.

First, AI integration in RE is pedagogically productive when it operates within a structured ecosystem governed by trained educators, clear institutional frameworks, and explicit attention to cognitive bias. The case studies from SPUP, Agios Nikolaos, and the broader literature reviewed here all point in the same direction: AI tools that are ungoverned or left to operate autonomously produce pedagogical outcomes that are, at best, thin and, at worst, theologically distorting.

Second, the cognitive biases embedded in AI systems, particularly confirmation bias, algorithmic partiality, and the Turing Trap, are not incidental malfunctions. They are structural features of how current AI systems are built and trained. Addressing them requires not only technical mitigation but pedagogical curriculum work: students in RE need to learn to read AI outputs critically, to identify what such outputs systematically omit, and to understand the difference between theological fluency and theological understanding.

Third, the human-centric hybrid model proposed here is not a compromise between innovation and tradition. It is the appropriate institutional response to what AI actually is and what RE actually requires. AI is a powerful epistemological instrument. RE is a formative practice. The task is to use the former in service of the latter, with clear governance, contextual sensitivity, and unwavering commitment to the relational core of theological education.

6.1 Scholarly Contribution

This study makes three contributions to the existing literature. It applies systematic cognitive bias theory specifically to the domain of faith-based instruction, a connection that the existing literature addresses only partially. It grounds theoretical analysis in institutional case evidence

from geographically and theologically diverse contexts. And it proposes a governance framework adapted explicitly for RE in the Global South, attending to the structural equity dimensions of AI adoption that general educational technology scholarship tends to bracket.

6.2 Limitations

Several limitations must be acknowledged. The study relies on existing published documentation, and the institutional cases examined represent a limited geographic and denominational range. Data from SPUP and Agios Nikolaos, while valuable, reflects Christian institutional contexts, and may not generalise to Islamic, Jewish, Hindu, or indigenous RE settings without modification. The rapidly evolving nature of AI technology means that specific platform capabilities described in cited works may already have changed. And the library research design, while appropriate for the study's conceptual objectives, cannot capture the classroom dynamics, relational textures, and pedagogical micro-decisions that shape the lived experience of AI-integrated RE.

6.3 Recommendations

We offer four recommendations for practice and future research. Institutions considering AI integration in RE should conduct theological audit processes before deployment, assessing the representativeness of training datasets and the doctrinal coherence of AI outputs within their specific tradition. RE educator preparation programmes should include structured modules on AI literacy and cognitive bias, equipping teachers to serve as critical mediators rather than passive conduits of algorithmic content. Future empirical research should prioritise mixed-methods field studies in Global South RE contexts, generating primary data on how learners and educators in under-resourced settings experience and navigate AI integration. Finally, theological institutions and AI governance bodies should develop co-designed frameworks for

RE-specific AI certification, ensuring that tools deployed in faith-based classrooms meet minimum standards of theological adequacy and contextual sensitivity.

DECLARATIONS

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

Declaration of Generative AI and AI-Assisted Technologies: During the preparation of this work, the author used Open Grammarly for grammar checks and style suggestions. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of this study.

REFERENCES

Ahmad, S. F., Rahmat, M. K., Mubarik, M. S., Alam, M. M., & Hyder, S. I. (2021). Artificial Intelligence and its role in education. *Sustainability, 13*(22), 12902. https://doi.org/10.3390/su132212902

Al-Attas, S. M. N. (1984). *The concept of education in Islam: A framework for an Islamic philosophy of education.* Mizan.

Alimukhamedova, D., Altaeva, L., Danshova, A., & Dosmurzaeva, A. O. (2026). Ethical and pedagogical aspects of artificial intelligence application in inclusion: The balance between technology and human interaction. *Scientists.uz, 2*(1), 88-92.

Almeida, R., Santos, D., & Yu, C. (2025). Ethical AI integration in multicultural education settings. *Journal of Educational Technology Ethics, 17*(2), 45-63.

Arifin, Z. (2018). *Management of Islamic education curriculum development: Theory and practice.* Almuqsith Pustaka.

Aslan & Wahyudin. (2020). *Curriculum in the challenge of change.* Bookies Indonesia.

Balasubramanian, S. (2023). Ethical considerations in AI-assisted decision-making for end-of-life care in healthcare. *Power System Technology, 47*(1), 167-182.

Brynjolfsson, E. (2022). The Turing Trap: The promise and peril of human-like artificial intelligence. *Daedalus, 151*(2), 272-287. https://doi.org/10.1162/daed_a_01915

Camilleri, M. A. (2024). Artificial Intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems, 41*(2), e13406.

Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems, 11*(4), 190-202.

Cartolovni, A., Tomicic, A., & Lazic Mosler, E. (2022). Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. *International Journal of Medical Informatics, 161*, 104738.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in education: A review. *IEEE Access, 8*, 75264-75278.

Choizhalanova, A. T. (2021). New opportunities and barriers to social adaptation of persons with disabilities in the context of digital transformation. *Economic Sciences, 4*(3), 190-195.

Hamal, O., El Faddouli, N. E., Harouni, M. H. A., & Lu, J. (2022). Artificial Intelligence in education. *Sustainability, 14*(5), 2862.

Hamzah, A. (2020). *Library research methods.* Literasi Nusantara.

Harefa, A. (2025). Adaptive learning strategies for students with diverse backgrounds: Implementation of pedagogical competence at Agios Nikolaos Christian Private Junior High School Gunungsitoli. *International Transformative Education and Humanities Journal, 1*(2), 88-95.

Harry, A., & Sayudin, S. (2023). Role of AI in education. *Interdisciplinary Journal and Humanity (INJURITY), 2*(3), 260-268.

Huang, L. (2023). Ethics of Artificial Intelligence in education: Student privacy and data protection. *Science Insights Education Frontiers, 16*(2), 2577-2587.

Jose, D. (2024). Data privacy and security concerns in AI-integrated educational platforms. *Recent Trends in Management and Commerce, 5*(1), 87-91.

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux.

Kalip, K., Bakti, Y. D., Ebariani, E., & Dewi, C. (2025). Education in the digital age: Creating inclusive and adaptive learning environments. *Advances in Social Science, Education and Humanities Research, 932*, 536-544.

Kulaklıoglu, D. (2024). Ethical AI in autonomous systems and decision-making. *Human Computer Interaction, 8*(1), 87.

Lappay, A. P. M. (2025). Integration of Artificial Intelligence in SPUP's interfaith initiatives. *IOSR Journal of Humanities and Social Science, 30*(6), 30-37. https://doi.org/10.9790/0837-3006023037

Meurisch, C., & Muhlhauser, M. (2021). Data protection in AI services. *ACM Computing Surveys, 54*(4), 1-38.

Mundofi, A. A. (2025). Integration of deep learning approach in transforming Islamic religious education learning in schools: A pedagogical and technological study. *Journal of Asian Primary Education (JoAPE), 2*(1), 79-90. https://doi.org/10.59966/joape.v2i1.1787

Nasr, S. H. (1993). *An introduction to Islamic cosmological doctrines.* State University of New York Press.

Papakostas, C. (2025). Artificial Intelligence in religious education: Ethical, pedagogical, and theological perspectives. *Religions, 16*(563). https://doi.org/10.3390/rel16050563

Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2022). An Artificial Intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review, 33*(1), 100925.

Schulze, K. G., Shelby, R. N., Treacy, D. J., Wintersgill, M., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In *Lecture Notes in Computer Science* (Vol. 1839, pp. 133-142). Springer.

Shofiyyah, N. A., Lesmana, O., & Tohari, H. (2024). Metamorphosis of Islamic religious education learning method: Classic approach converted by Artificial Intelligence (AI). *Jurnal Pendidikan: Riset dan Konseptual, 8*(2), 265-275.

Sugiyono. (2008). *Qualitative, quantitative, and R&D research methods.* Alfabeta.

Tarigan, Y. P., Simangunsong, B., & Sihombing, W. F. (2023). Children's theology: Learning media for Sunday school children using Book Creator in Indonesia. *The International Journal of Education, Theology, and Humanities, 3*(1), 34-44.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Wilson, C. (2025, May 21). *Artificial Intelligence in religious education.* St Edmund's College.

Yin, H., Zhao, Q., & Zhang, L. (2020). AI and its role in moderating interfaith dialogues: A sentiment analysis approach. *Journal of Intercultural Communication, 45*(1), 78-94.

Zhang, J., Song, W., & Liu, Y. (2025). Cognitive bias in generative AI influences religious education. *Scientific Reports, 15*, 15720. https://doi.org/10.1038/s41598-025-99121-6