ISIHUMOR
Jurnal Ilmu Sosial dan Humaniora
http://www.putrapublisher.org/ojs/index.php/isihumor Publisher: putrapublisher.org

ISIHUMOR: Jurnal Ilmu Sosial dan Humaniora
Vol. 4 No. 2 April 2026, Hlm 532-545

GENERATIVE AI IN ISLAMIC EDUCATION AT THE MADRASAS LEVEL: A LITERATURE REVIEW OF PEDAGOGICAL, ETHICAL, AND THEOLOGICAL FRONTIERS

Huzma Juliantri[1], Marifatul Adawiyyah[2], Asmarita[3], Azki Akhyari[4]
[1234]Universitas Islam Mamba'ul Ulum Jambi, Indonesia
Corresponding Author: huzmanisa@gmail.com

INFO ARTIKEL

Diterima: 05 Februari 2026
Direvisi: 25 Februari 2026
Disetujui: 12 Maret 2026
Tersedia Daring: 30 April 2026

Kata Kunci:
kecerdasan buatan; ChatGPT; pendidikan agama Islam; madrasah; Tinjauan literatur

ABSTRAK

Difusi alat kecerdasan buatan generatif (GenAI), khususnya ChatGPT, telah membentuk kembali kemungkinan pengajaran di seluruh sistem pendidikan global, namun integrasi alat-alat ini ke dalam Pendidikan Agama Islam (IRE) di tingkat madrasah masih kurang diperiksa. Studi ini menyajikan tinjauan literatur terstruktur yang memetakan percakapan ilmiah saat ini tentang penggunaan AI di IRE untuk pengaturan madrasah primer dan sekunder, dengan tujuan mengidentifikasi temuan konvergen, kesenjangan yang terus-menerus, dan ketegangan yang muncul. Empat puluh satu artikel peer-review yang diterbitkan antara tahun 2020 dan 2026 diambil dari jurnal terakreditasi Scopus, Web of Science, dan SINTA, kemudian disintesis secara tematik melalui prosedur pengkodean induktif. Tinjauan ini mengidentifikasi enam kluster berulang: integrasi pedagogis, kesiapan guru dan kompetensi AI-TPACK, hasil belajar siswa, masalah etika dan teologis, rekonstruksi kurikulum dan kebijakan, dan integritas akademik dalam menghadapi halusinasi. Temuan menunjukkan bahwa alat AI mendukung personalisasi konten, keterlibatan, dan persiapan pelajaran, tetapi menimbulkan pertanyaan yang belum terselesaikan tentang akurasi doktrin, dimensi spiritual pembelajaran, dan pembentukan adab al-'ilm di antara peserta didik. Tinjauan ini mengusulkan kerangka kerja tiga dimensi yang menghubungkan pertimbangan teknologi, pedagogis, dan teologis-etika sebagai dasar untuk penyelidikan empiris di masa depan. Implikasi bagi guru madrasah, pengembang kurikulum, dan pembuat kebijakan Kementerian Agama RI dibahas.

Keywords:
Artificial Intelligence; Chatgpt; Islamic Religious Education; Madrasah; Literature Review

ABSTRACT

The diffusion of generative artificial intelligence (GenAI) tools, particularly ChatGPT, has reshaped instructional possibilities across global education systems, yet the integration of these tools into Islamic Religious Education (IRE) at the madrasah level remains underexamined. This study presents a structured literature review that maps the current scholarly conversation on AI use in IRE for primary and secondary madrasah settings, with the aim of identifying converging findings, persistent gaps, and emerging tensions. Forty-one peer-reviewed articles published between 2020 and 2026 were retrieved from Scopus, Web of Science, and SINTA-accredited journals, then thematically synthesised through an inductive coding procedure. The review identifies six recurring clusters: pedagogical integration, teacher readiness and AI–TPACK competence, student learning outcomes, ethical and theological concerns, curriculum and policy reconstruction, and academic integrity in the face of hallucination. Findings indicate that AI tools support content personalisation, engagement, and lesson preparation, but raise unresolved questions about doctrinal accuracy, the spiritual dimension of learning, and the formation of adab al-'ilm among learners. The review proposes a tri-dimensional framework that links technological, pedagogical, and

theological–ethical considerations as a basis for future empirical inquiry. Implications for madrasah teachers, curriculum developers, and Indonesian Ministry of Religious Affairs policymakers are discussed.

1. Introduction

The arrival of generative artificial intelligence (GenAI) tools, most prominently OpenAI's ChatGPT, has reorganised the everyday work of teaching and learning. Within three years of its public release, ChatGPT and comparable large language models have moved from novelty to infrastructure, finding their way into lesson preparation, assessment design, and even into the inquiries that students bring into the classroom. Educational systems across Asia, Europe, and the Americas now confront a question that is at once practical and philosophical: how should schools harness tools that produce plausible text on demand, and what does responsible adoption look like when the subject of instruction is religious knowledge? In Indonesia, this question carries particular weight for madrasah education, where Islamic Religious Education (IRE) sits at the centre of the curriculum and where the integrity of religious content cannot be treated as a minor pedagogical concern.

Recent scholarship has begun to map the educational implications of GenAI in considerable breadth. Systematic reviews report that ChatGPT can personalise learning, automate routine assessment tasks, and stimulate student engagement, while at the same time threatening academic integrity, deepening digital inequities, and producing factually unreliable output (Pacheco-Mendoza et al., 2024; Tossell et al., 2025). Within Islamic education, a parallel body of work has emerged. Wedi and colleagues (2025) describe a digital transformation model for IRE in East Java madrasahs, Mukhsin and colleagues (2026) propose curriculum reconstruction for AI-enabled IRE in the Society

5.0 era, and Idham and colleagues (2026) synthesise forty-two studies on deep learning applications in madrasah curricula. Empirical work on ChatGPT-assisted IRE at the elementary madrasah level (Salehudin et al., 2023) and on user experiences of ChatGPT among IRE learners (Ujang et al., 2025) has begun to fill the picture in finer grain.

Despite this growing literature, three gaps remain conspicuous. First, while many studies treat IRE in higher education or general educational settings, comparatively few focus on the madrasah ibtidaiyah, tsanawiyah, and aliyah levels as a connected pedagogical context, even though it is at these levels that doctrinal foundations are laid. Second, although ethical concerns are routinely flagged, theological assessments grounded in Islamic epistemological categories such as adab al-'ilm and maqāṣid al-sharīʿah remain thin, with most discussions stopping at generic ethics rather than engaging the distinctive normative resources of the Islamic tradition (Hardaker & Ismail,

2025). Third, frameworks for teacher preparation in IRE settings still rely heavily on the standard TPACK model and have yet to integrate emerging AI-specific extensions such as the Human-Centric AI Pedagogy framework (Lee & Park, 2026), leaving madrasah teachers without clear guidance on how to align AI competence with their religious vocation.

This article addresses those gaps by offering a structured literature review of the current state of AI and ChatGPT use in IRE at the madrasah level. The aim is fourfold: to map the thematic terrain of the emerging field, to surface convergences and inconsistencies across studies, to identify the conceptual and empirical gaps that should orient future research, and to propose a working framework that links technological, pedagogical, and theological–ethical considerations. The unit of analysis is the published study, with attention to its setting, method, and reported findings rather than to original empirical data.

The argument proceeds as follows. The next section explains the review methodology, including search strategy, inclusion criteria, and the inductive thematic coding procedure used to organise findings. The Results section presents the six clusters identified in the corpus, supported by representative studies and visualised through a thematic map and a distribution chart. The Discussion section interprets these findings against established frameworks in educational technology and Islamic educational thought, advances a tri-dimensional model for AI-augmented IRE, and traces implications for madrasah practice. The Conclusion summarises the contribution and outlines avenues for empirical follow-up.

2. Method

This study is a structured literature review designed in the spirit of, but not strictly bound to, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. The objective was descriptive and integrative rather than meta-analytic: the goal was to chart the contours of an emerging field rather than to estimate effect sizes. The decision to use a structured narrative review reflects the heterogeneity of methods in the source studies, which range from small qualitative case studies of single madrasahs to large-scale systematic reviews and quantitative survey studies. A meta-analytic strategy would have required a more homogeneous outcome set than the literature currently provides.

The literature search was conducted in February through April 2026 across three databases: Scopus, Web of Science, and SINTA (the Indonesian Ministry of Education's accreditation index for national journals). Boolean queries combined the AI cluster ("ChatGPT" OR "generative artificial intelligence" OR "large language model" OR "GenAI") with the IRE cluster ("Islamic religious education" OR "pendidikan agama Islam" OR "madrasah" OR "Al-Qur'an Hadith" OR "Akidah Akhlak" OR "Fiqh"). The search window was January 2020 to April 2026, with a deliberate emphasis on studies published from 2023 onward to capture the post-ChatGPT inflection point. Initial retrieval yielded 312 records.

Records moved through four stages of screening, summarised in Figure 1. After deduplication and a title and abstract screen against a priori inclusion criteria, 184

records advanced. Inclusion required that a study (a) explicitly addressed AI or ChatGPT in an IRE or madrasah context, (b) reported empirical findings or systematic synthesis rather than opinion or editorial commentary, and (c) appeared in a peer-reviewed venue. Full-text assessment of the 96 articles that survived the abstract screen produced a final pool of 41 studies for thematic synthesis. Studies excluded at this stage typically focused on AI in general education without an Islamic dimension, on Islamic higher education rather than the madrasah, or on conceptual essays without empirical or systematic grounding.

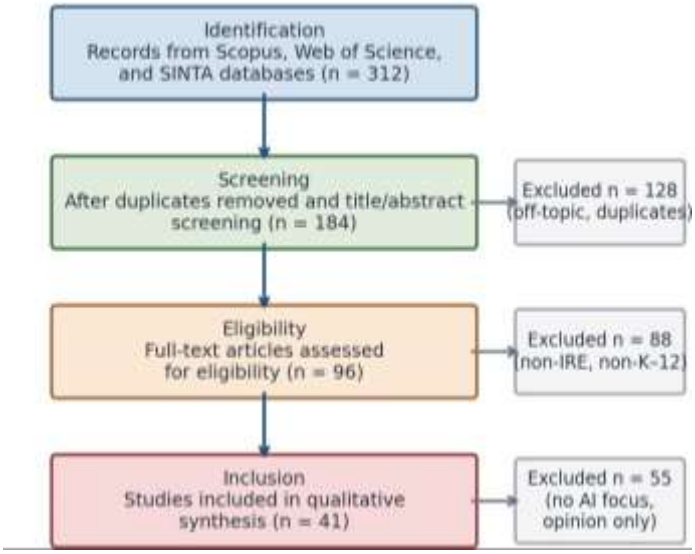

Figure 1. Literature selection flow adapted from the PRISMA 2020 guidelines. Source: Authors, based on database searches conducted in February–April 2026.

Thematic synthesis followed an inductive coding procedure adapted from Thomas and Harden (2008). Each included study was coded for setting, method, AI tool examined, IRE subject area, and reported findings. Codes were then clustered through iterative comparison, and clusters were refined until each researcher independently arrived at a stable thematic structure. Disagreements were resolved through discussion against the reference list, with conservative interpretation favoured where studies could plausibly be assigned to more than one cluster. To improve transparency, the authors maintained an audit trail consisting of the coded study database, the cluster definitions, and notes on borderline allocations. Validity was strengthened through investigator triangulation and through cross-checking against two existing reviews (Idham et al., 2026; Suriyati et al., 2025) that provided independent reference points for the field.

3. Result and Discussion (Hasil dan Pembahasan)

Distribution of the Reviewed Corpus

The 41 included studies span six years and three regional concentrations: Indonesia (n = 28), Malaysia (n = 6), and a smaller set from the Middle East and the wider Muslim world (n = 7). Methodologically, qualitative case studies dominate (n = 17), followed by mixed-methods designs (n

= 11), quantitative survey or quasi-experimental work (n = 8), and systematic reviews or bibliometric analyses (n = 5). The thematic distribution, presented in Figure 3, shows that pedagogical integration and student learning outcomes account for almost half of the corpus, while curriculum reform and academic integrity remain comparatively under-studied.

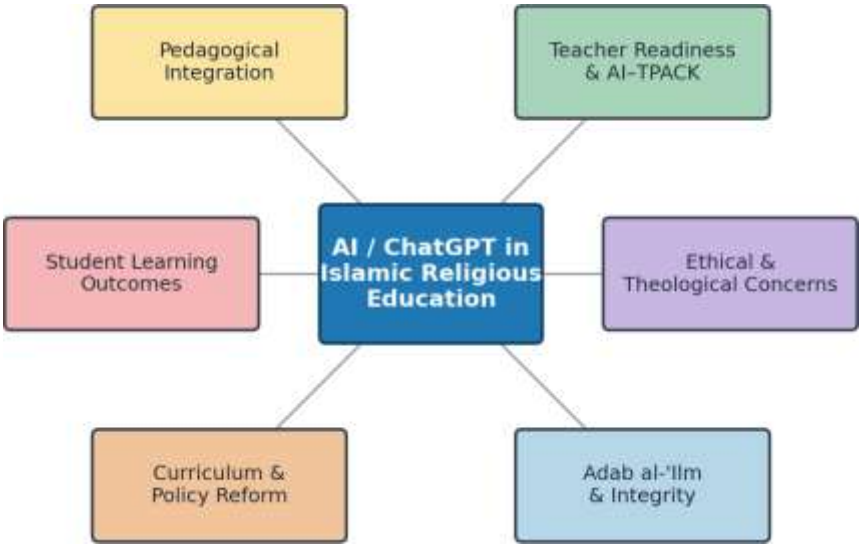

Figure 2. Thematic clusters identified in the reviewed literature, showing the six recurring areas of inquiry around AI and ChatGPT in IRE at the madrasah level.

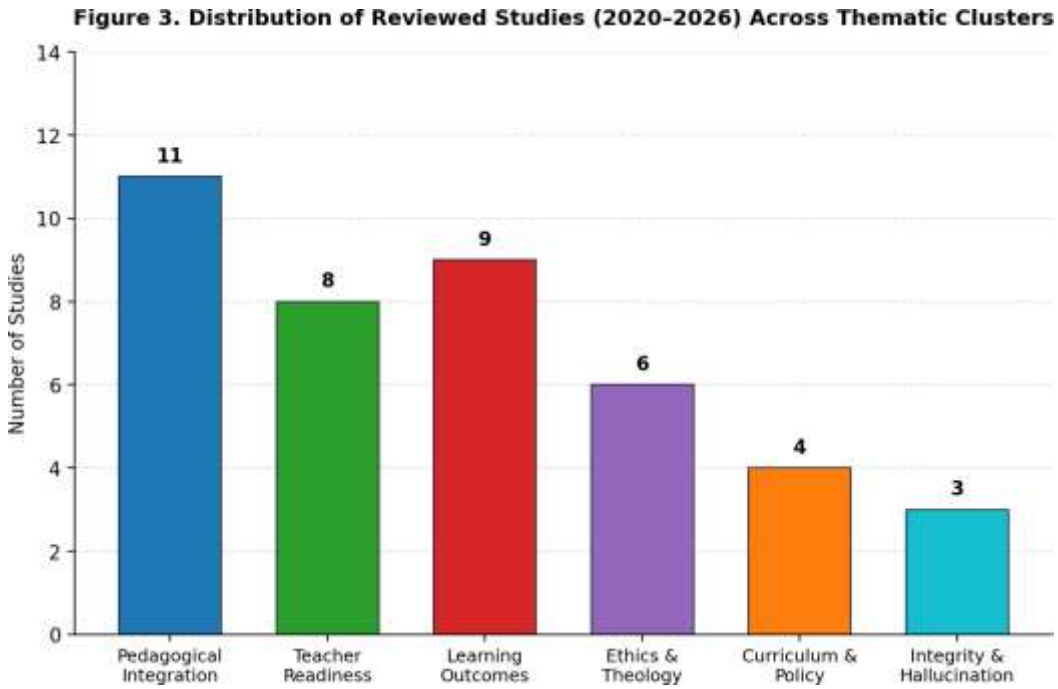

Figure 3. Distribution of the 41 reviewed studies across the six thematic clusters. The pedagogical integration and learning outcomes clusters are the most populated, while academic integrity remains the least developed.

Cluster One: Pedagogical Integration

The largest cluster (n = 11) addresses how teachers and students integrate ChatGPT and similar tools into IRE classroom routines. Studies in this cluster generally report that AI tools support lesson planning, content elaboration, and the generation of contextually relevant examples for fiqh, akidah akhlak, and Al-Qur'an Hadith subjects. Salehudin and colleagues (2023) report that ChatGPT-assisted instruction in two Madrasah Ibtidaiyah in Jember produced higher conceptual understanding than conventional methods in a quasi-experimental design with eighty grade IV–VI students. The qualitative work of Ujang and colleagues (2025) on user experience extends this picture, suggesting that students perceive ChatGPT as engaging when teachers scaffold its use rather than letting students consult the model unsupervised.

Patterns of integration vary by subject area. In Al-Qur'an Hadith instruction, where verbatim accuracy of Arabic text and the chain of transmission matter, teachers tend to use AI for paraphrasing background context rather than for textual reproduction (Zh et al., 2024a). In fiqh instruction, AI is more commonly used to generate cases for legal reasoning practice, an application that aligns with the inductive nature of the subject. In akidah akhlak, generative AI is most frequently deployed to produce ethical scenarios that students can analyse, a pattern documented at SMP Islam As Sakinah Sidoarjo by Hannan and colleagues (2026). Across these subjects, teachers retain the role of doctrinal arbiter, vetting AI output before it reaches students.

Cluster Two: Teacher Readiness and AI–TPACK Competence

The second cluster (n = 8) examines the question of whether IRE teachers possess the technological, pedagogical, and content knowledge necessary to integrate AI productively. Findings are uneven. Studies of urban madrasahs in Java report relatively high AI literacy among teachers under thirty-five, while studies in rural settings of East Lombok and South Sumatra describe substantial gaps in basic familiarity with generative tools (Sari & Anwar, 2025). Quddus (2020) had earlier diagnosed weaknesses in TPACK among PAI teacher candidates, and the more recent work of Faizin and colleagues (2025) on Muslim student acceptance of AI through the extended Technology Acceptance Model suggests that perceived usefulness depends heavily on the model of integration that teachers themselves embody.

An emerging line of inquiry pushes beyond the standard TPACK model. The Human-Centric AI Pedagogy framework proposed by Lee and Park (2026) integrates five knowledge domains, including human–AI collaborative knowledge and ethical knowledge, as foundational pillars rather than as add-ons. Although not yet operationalised in madrasah contexts, this extension speaks directly to the predicament of IRE teachers, who must mediate between an algorithmic system that generates plausible religious content and a student community that cannot yet distinguish authoritative from uncertain output. The implication, made explicit in the recent work of Idham and colleagues (2026), is that teacher professional development for AI in IRE must be redesigned around the dual axis of technical competence and theological discernment.

Cluster Three: Student Learning Outcomes and Engagement

Studies on learning outcomes (n = 9) report consistent gains in cognitive engagement and short-term content recall when AI tools are integrated under teacher guidance. Adiyono and colleagues (2025) document quantitative improvement in IRE outcomes when AI-optimised YouTube videos are paired with classroom instruction, while Hayati and Ushalli (2024) describe gains in personalised learning trajectories within madrasah settings that combine AI scaffolds with conventional teacher-led explanation. The Wordwall-based blended evaluation work by Zh and colleagues (2024b) at MAN Kota Batu illustrates the complementary potential of AI-adjacent digital media in raising student engagement in Al-Qur'an Hadith lessons.

Effects on deeper outcomes are less clear. Few studies in the corpus measure transfer, retention beyond a single semester, or the formation of religious dispositions, leaving the question of long-term impact open. Where qualitative data exist, students themselves appear to distinguish between AI as an information source and AI as a religious authority, with most rejecting the latter framing (Faizin et al., 2025). This distinction, which the literature has not fully theorised, hints at a student-led ethic of source evaluation that may be more resilient than educators sometimes assume.

Cluster Four: Ethical and Theological Concerns

Six studies in the corpus engage explicitly with ethical and theological concerns about AI in IRE. The most frequent worry is doctrinal accuracy: generative models trained on broad internet corpora produce religiously plausible but unverified content, sometimes confidently citing fabricated hadith chains or attributing fatwas to scholars who never issued them (Wahid & Ramli, 2025). Studies in this cluster argue that the standard educational ethics frame, which centres autonomy, beneficence, and non-maleficence, captures part of the problem but misses the dimensions distinctive to Islamic learning. Hardaker and Ismail (2025) propose a framework grounded in maqāṣid al-sharīʿah that locates AI use against the preservation of religion, intellect, and progeny, providing a normative anchor that generic ethics codes lack.

A second strand within this cluster concerns the spiritual dimension of IRE. Several authors argue that the traditional teacher–student relationship in Islamic learning carries dimensions of barakah and adab that an algorithmic interlocutor cannot replicate (Wahyuni & Setiawan, 2025). The argument is not that AI should be excluded but that its function must be specified narrowly, supporting cognitive content while leaving formative and spiritual aspects to embodied teacher–student relationships. This nuanced position, which avoids both technophobia and uncritical adoption, is becoming the dominant theological stance in the literature.

Cluster Five: Curriculum and Policy Reconstruction

Studies on curriculum and policy (n = 4) tend to be conceptual rather than empirical, reflecting the early stage of policy development in the Indonesian context. Mukhsin and colleagues (2026) propose curriculum reconstruction principles for AI-based IRE in the Society 5.0 era at madrasah level, including the integration of AI literacy as a cross-cutting competence, the redesign of assessment tasks to remain meaningful in an AI-saturated

environment, and the explicit teaching of source criticism within fiqh and hadith subjects. The Kementerian Agama blueprint for digital madrasah transformation (2023) provides high-level guidance but stops short of subject-specific operational detail.

What is missing from this cluster, by the authors' reading, is a sustained engagement with the professional development infrastructure required to make policy implementable. Curriculum frameworks proliferate, but field studies that follow teachers from policy receipt to classroom enactment are rare. This is a methodological gap as much as a thematic one, and it points to a pressing need for longitudinal qualitative work in madrasah settings.

Cluster Six: Academic Integrity and Hallucination

The smallest cluster (n = 3) addresses the AI-specific risks of hallucination and academic dishonesty. The empirical work of Rahmawati and colleagues (2025) on plagiarism risk among Islamic education students compares ChatGPT and Scite, finding that generative tools produce significantly higher plagiarism risk profiles than verification tools and arguing for the integration of adab al-'ilm into digital literacy curricula. This cluster, although small, raises the most theologically distinctive question in the corpus: whether the Islamic concept of 'ilm, which carries epistemic and ethical commitments that go beyond mere information possession, can survive a learning environment in which generative output is indistinguishable from authentic scholarship without conscious effort.The under-development of this cluster is itself a finding. Among 41 studies, only three engage hallucination as a primary concern, even though hallucination is the failure mode most likely to produce doctrinal error in IRE settings. This imbalance suggests that the field has not yet reckoned fully with the unique risks AI poses to religious knowledge, as opposed to general academic knowledge.

Convergences Across the Literature

The reviewed corpus converges on several propositions worth stating plainly. AI tools, properly scaffolded, can enhance engagement and short-term cognitive outcomes in IRE settings; teacher mediation rather than teacher replacement is the dominant integration model that emerges; and ethical concerns, although routinely acknowledged, have not yet been operationalised in classroom-ready form. These convergences echo broader findings in the educational AI literature, where reviews such as Pacheco-Mendoza et al. (2024) and Tossell et al. (2025) report similar patterns in non-religious settings. The IRE field is therefore not anomalous in its overall trajectory; what makes it distinctive is the additional layer of doctrinal stewardship that teachers must perform.

Convergence does not, however, equal consensus on direction. The literature is divided between a transformation orientation, which emphasises curriculum reconstruction and policy reform, and an enhancement orientation, which sees AI as a useful add-on within existing instructional structures. Both orientations find empirical support in the corpus, and neither has yet produced sustained longitudinal evidence in madrasah contexts. The choice between them is therefore as much a matter of educational philosophy as of empirical fact, and it is likely to remain contested for several years yet.

Inconsistencies and Open Questions

Three inconsistencies recur across the literature and merit explicit attention. First, claims about teacher readiness diverge sharply by setting, with urban Java studies reporting comfort with AI and rural studies reporting unfamiliarity. The literature has not yet reconciled these findings into a clear typology of contextual factors that shape readiness. Second, claims about student outcomes are similarly uneven, with some quantitative studies reporting strong effects on engagement and others finding null effects on deeper learning. Sample sizes are typically small and follow-up periods short, which suggests that the field's empirical base may not yet support strong outcome claims in either direction. Third, theological assessments range from cautious endorsement to firm rejection, and the criteria distinguishing these positions are often left implicit. A more disciplined engagement with classical sources of Islamic educational thought, such as the tradition of ta'lim and ta'dib, would help discipline this discussion.

Toward a Tri-Dimensional Framework

Drawing the threads of the review together, the authors propose a tri-dimensional framework, summarised in Figure 4, that links three pillars on a shared contextual foundation. The technological pillar covers AI tool literacy, prompt design, and hallucination awareness. The pedagogical pillar incorporates TPACK and its emerging extensions, scaffolded inquiry approaches, and reflective practice. The theological–ethical pillar centres adab al-ʿilm, maqāṣid al-sharīʿah, and source verification practices grounded in classical methods of hadith authentication. The contextual foundation specifies the conditions under which the three pillars can stand together: madrasah infrastructure, teacher professional development, Kemenag policy framing, and local Islamic epistemology.

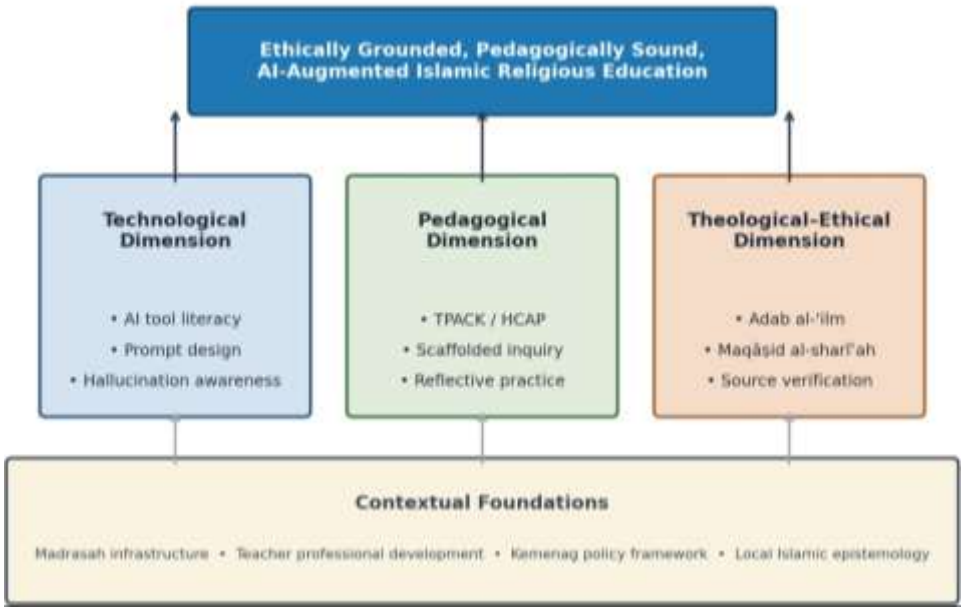

Figure 4. Tri-Dimensional Framework for AI Integration in Islamic Religious Education at the Madrasah Level.

The framework is not a recipe but an analytic device. Its purpose is to keep the three dimensions visible to one another, so that pedagogical integration is not pursued at the cost of theological integrity, and so that ethical commitments are not reduced to a vague gesture toward responsibility. Empirical research that uses the framework to design and evaluate interventions should produce more disciplined findings than the present literature affords, particularly if studies report explicitly how each dimension was addressed in the design and what trade-offs were encountered in implementation.

Implications for Practice

Three practical implications follow. For teachers, the review suggests that AI integration in IRE works best when teachers retain the role of doctrinal arbiter, when AI use is restricted to specific cognitive tasks rather than spiritual or formative ones, and when students are taught early to interrogate AI output against authoritative sources. For curriculum developers, the review recommends that AI literacy be integrated as a cross-cutting competence within IRE rather than as a separate subject, with explicit attention to source criticism within fiqh and hadith. For policymakers, including the Indonesian Ministry of Religious Affairs, the review identifies an urgent need for teacher. professional development infrastructure that bridges the urban–rural readiness gap and that incorporates the theological–ethical dimension alongside technical training.

These implications should be read against the limits of a literature review, which can map a field but cannot generate new empirical evidence. The next step for the field is sustained empirical work, ideally longitudinal and comparative across madrasah types, that tests the framework against classroom realities. Such work should, in the authors' view, also engage explicitly with the Indonesian context of religious moderation, the wider regional context of Muslim digital learning communities, and the global research conversation on AI in education that has so far developed largely without religious educators at the table.

4. Conclusion

This literature review mapped 41 studies on AI and ChatGPT use in Islamic Religious Education at the madrasah level, identifying six thematic clusters and proposing a tri-dimensional framework that links technological, pedagogical, and theological–ethical considerations on a shared contextual foundation. Three conclusions follow from the synthesis. First, the field has accumulated promising evidence on engagement and short-term outcomes, but it has not yet produced longitudinal evidence on deeper religious learning or on the formation of student dispositions. Second, the theological–ethical dimension of AI use in IRE has been raised in many studies but operationalised in few; the field would benefit from sustained engagement with classical Islamic educational thought as well as with contemporary maqāṣid-based ethics. Third, teacher professional development emerges as the most pressing infrastructural need, particularly given the urban–rural divergence in readiness and the additional demands that AI integration places on the IRE teacher's already complex role.

Future research should pursue three directions in particular. Empirical studies should test the proposed framework across madrasah ibtidaiyah, tsanawiyah, and aliyah

settings, ideally in comparative designs that surface contextual moderators of effectiveness. Theological research should develop more granular criteria for distinguishing acceptable from unacceptable AI uses in IRE, drawing on classical methods of source authentication and on contemporary fatwa scholarship. Policy research should examine the implementation pathway from Kemenag blueprints to classroom enactment, with particular attention to the institutional and human supports that determine whether ambitious frameworks survive contact with practice.

5. References

Adiyono, A., Al Matari, A. S., Patimah, L., Syahrani, S., & Nasywa, A. A. (2025). Can AI-optimized YouTube videos enhance Islamic religious education? A quantitative study on student learning outcomes. SYAMIL: Journal of Islamic Education, 13(2), 187–204. https://doi.org/10.21093/sy.v13i2.11935

Al-Abdullatif, A. M. (2024). Modeling teachers' acceptance of generative artificial intelligence use in higher education: The role of AI literacy and intelligent TPACK. Education and Information Technologies, 29(11), 14223–14248. https://doi.org/10.1007/s10639-024-12782-0

Bauer, E., Greiff, S., Graesser, A. C., Scheiter, K., & Sailer, M. (2025). Looking beyond the hype: Understanding the effects of AI on learning. Educational Psychology Review, 37(2), 1–27. https://doi.org/10.1007/s10648-025-10020-8

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Darmu'in, D., Nasikhin, N., Darnoto, D., Maarif, M. A., Sofanudin, A., & Haryanto, J. T. (2025). Perceived barriers to ChatGPT integration in Islamic studies education: Insights from Malaysian international students in Indonesia. Journal of International Students, 15(11), 141–164. https://doi.org/10.32674/h99a4p06

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, 105224. https://doi.org/10.1016/j.compedu.2025.105224

Djazilan, M. S., Rulyansah, A., & Rihlah, J. (2024). Why AI is essential for the future of Islamic education: A call for ethical and effective implementation. EDUKASIA: Jurnal Pendidikan dan Pembelajaran, 5(2), 201–216. https://doi.org/10.62775/edukasia.v5i2.1373

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), 1–18. https://doi.org/10.1007/s44217-025-00471-0

Hannan, A., Maulana, R., & Sutrisno, S. (2026). Integrating generative artificial intelligence into Islamic character education: Practices and challenges in an Indonesian Islamic school. EDUKASIA: Jurnal Pendidikan dan Pembelajaran, 7(1), 45–62. https://doi.org/10.62775/edukasia.v7i1.2034

Hardaker, G., & Ismail, A. (2025). Maqāṣid al-sharīʿah in the AI era: Balancing innovation and Islamic ethical principles. International Journal of Islamic Theology and Civilization, 3(3), 1–21.

Hayati, N., & Ushalli, E. (2024). Digital transformation in Islamic education: Integrating AI and machine learning for personalized learning in madrasah. Journal of Scientech Research and Development, 6(2), 88–104. https://doi.org/10.56670/jsrd.v6i2.528

Idham, I., Rokimin, R., Akhirudin, M. T., Rahman Zh, M. H., & At Thaariq, Z. Z. (2026). The application of deep learning in enhancing the effectiveness of Islamic religious education curriculum in madrasahs. Management of Education: Manajemen Pendidikan Islam, 12(1), 1–22.

Indallah, S. M., Nidzom, M. F., Inayah, F., Istiqomah, I., Kurniawan, M. I., & Risbiyantoro, H. (2025). Integration of Islamic Values in POACE Practice: An Effort to Islamize Education Based on Islamic Education Philosophy. *International Journal of Learning and Education*, *1*(2), 46-58.

Junaedi, J., Syafi'i, I., Hasyim, F., & Marwat, J. M. (2025). Transformational curriculum model through interpretation epistemology: Islamic boarding school curriculum management perspective.

Munaddhomah: Jurnal Manajemen Pendidikan Islam, 6(3), 387–403. https://doi.org/10.31538/munaddhomah.v6i3.1802

Kementerian Agama Republik Indonesia. (2023). Blueprint transformasi digital madrasah: Membangun madrasah cerdas dan berdaya saing global. Direktorat Jenderal Pendidikan Islam.

Kurniawan, M. I., Syam, A. R., Kamilah, A. N., Zh, M. H. R., & Akbar, R. N. (2025). Fostering Synergistic Collaboration: A Triple Helix Model for Integrating Educational Management, Digital Pedagogy, and Sharia Financial Literacy in Indonesian Islamic Economics Higher Education. *Jurnal Manajemen Pendidikan*, *16*(2), 89-103.

Lee, J., & Park, S. (2026). Human-Centric Artificial Intelligence Pedagogy (HCAP) framework developed from TPACK through integration of artificial intelligence literacy and competency. Interactive Learning Environments, 34(2), 1–24. https://doi.org/10.1080/10494820.2026.2615818

Maulidia, M., Finetri, M., & Fahmi, M. (2025). When teachers are digital. Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism, 11(1), 131–152. https://doi.org/10.20871/kpjipm.v11i1.418

Mukhsin, M., Mutaqin, A. Z., Saway, M. H. M., Kiki, K., & Hamid, A. (2026). Curriculum reconstruction for AI-based Islamic religious education in the Society 5.0 era at madrasah. EDURELIGIA: Jurnal Pendidikan Agama Islam, 10(1), 1–18.

Mustapha, R., & Malkan, S. N. A. (2025). Maqasid al-shariah in the AI era: Balancing innovation and Islamic ethical principles. International Journal of Islamic Theology and Civilization, 3(3), 1–21.

Nurhayati, S., Taufikin, T., Judijanto, L., & Musa, S. (2025). Towards effective artificial intelligence-driven learning in Indonesian child education: Understanding parental readiness, challenges, and policy implications. Educational Process: International Journal, 15(1), e2025155. https://doi.org/10.22521/edupij.2025.15.155

Pacheco-Mendoza, J., Tovar, C. T., Castro, J. R. V., Andrade-Arenas, L., & Cabanillas-Carbonell, M. (2024). Challenges of implementing ChatGPT on education: Systematic literature review. Computers and Education Open, 7, 100199. https://doi.org/10.1016/j.caeo.2024.100199

Quddus, A. (2020). Implementasi technological pedagogical content knowledge (TPACK) dalam pendidikan profesi guru (PPG) PAI LPTK UIN Mataram. Jurnal Tatsqif, 17(2), 213–230. https://doi.org/10.20414/jtq.v17i2.1911

Rahmawati, N., Sholihin, M., & Pratama, R. (2025). Assessing artificial intelligence plagiarism risk: ChatGPT vs Scite among Islamic education students. Ta'dib: Journal of Islamic Education, 14(2), 245–268.

Salehudin, M., Hidayat, S., & Mufidah, L. (2023). ChatGPT as a learning assistant for Islamic education: A study in Madrasah Ibtidaiyah. Nadwa: Jurnal Pendidikan Islam, 17(2), 187–210. https://doi.org/10.21580/nw.2023.17.2.26167

Sari, W. D., & Anwar, S. (2025). Transforming the Islamic education curriculum for the Society 5.0 era: Integrating technology, ethics, and pedagogy. Al-Ishlah: Jurnal Pendidikan, 17(3), 5635–5643. https://doi.org/10.35445/alishlah.v17i3.7186

Septianingrum, C., Arif, M., & Mulyadi, M. (2025). Instilling the value of tolerance through moderate Islamic religious education: A pedagogical approach in the context of public schools in Gresik. Tarbawi: Jurnal Pendidikan Islam, 22(1), 32–43.

Suriyati, S., Habibi, A., & Mukminin, A. (2025). The role of ChatGPT in enhancing learning among Islamic education undergraduate students: A systematic literature review. Journal of Research, Innovation, and Strategies for Education, 2(1), 1–18.

Syukur, F., Maghfurin, A., Marhamah, U., & Jehwae, P. (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

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8, 45. https://doi.org/10.1186/1471-2288-8-45

Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K., & de Visser, E. J. (2025). ChatGPT in education: A systematic review on opportunities, challenges, and future directions. Algorithms, 18(6), 352. https://doi.org/10.3390/a18060352

Ujang, S., Hidayat, R., & Salehudin, M. (2025). User experience analysis: ChatGPT-based AI in Islamic religious education learning. Studies in Learning and Teaching, 6(1), 124–141.

Wahid, A., & Ramli, M. (2025). Utilizing of generative AI in religious education content production: An Islamic ethics perspective. EduSphere: Journal of Educational Innovation and Learning, 1(3), 220–238.

Wahyuni, R., & Setiawan, A. (2025). The role of artificial intelligence in enhancing pedagogical methods in modern Islamic education. Berkala Ilmiah Pendidikan, 5(2), 134–152.

Wedi, A., At Thaariq, Z. Z., Soepriyanto, Y., & Rahman Zh, M. H. (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), 156–179. https://doi.org/10.26803/ijlter.24.8.9

Yang, Y., Xia, Q., Liu, C., & Chiu, T. K. F. (2025). The impact of TPACK on teachers' willingness to integrate generative artificial intelligence (GenAI): The moderating role of negative emotions. Teaching and Teacher Education, 154, 104877. https://doi.org/10.1016/j.tate.2024.104877

Zh, M. H. R., Putra, M. F. B., Kuswandi, D., Wedi, A., & Ardiansyah, A. (2024a). Developing Wordwall evaluations in blended Islamic education using the Smith and Ragan model. Al-Aulia: Jurnal Pendidikan dan Ilmu-Ilmu Keislaman, 10(1), 89–104.

Zh, M. H. R., Sani, N. L., Kuswandi, D., & Fadhli, M. (2024b). Needs analysis of development FBO media as a support for blended learning in Al-Qur'an Hadits lesson. Jurnal Pendidikan Agama Islam Al-Thariqah, 9(1), 16–32. https://doi.org/10.25299/al-thariqah.2024.vol9(1).16-32

Zh, M. H. R., Pradana, M. I. Y., Purnomo, P., Soepriyanto, Y., & Budiman, F. (2025). Comparative analysis of student learning outcomes in Al-Qur'an Hadith lessons based on learning media. Al-Afkar: Journal for Islamic Studies, 8(1), 241–250.