

ARTIFICIAL INTELLIGENCE IN ISLAMIC EDUCATION: OPPORTUNITIES, CHALLENGES, AND AN ETHICAL FRAMEWORK GROUNDED IN MAQASID AL- SHARI'AH

Muhammad Syafiq Abdul Aziz^{1*}

^{1}Postgraduate Student, Department of Fiqh-Usul and Applied Sciences, Academy of Islamic Studies,
Universiti Malaya, 50603 Kuala Lumpur, Malaysia. Email: md.syafiq.abdulaziz@gmail.com.
(Corresponding Author)*

ABSTRACT

Artificial intelligence has become part of the everyday machinery of teaching, and its reach now extends into the institutions and disciplines of Islamic learning. Scholarship on this encounter is growing quickly. It remains fragmented, however, and despite an almost universal concern with ethics, there is still limited research that integrates the technological applications of AI, its ethical challenges, and Islamic epistemology within a unified *maqasid*-based framework. This article sets out to address that gap. It examines how AI is currently used in Islamic education, weighs the difficulties this raises, and develops an ethical and governance framework to guide responsible adoption. The study employs a critical narrative literature review and thematic synthesis of peer-reviewed scholarship and the principal AI-governance instruments published between 2018 and 2026. AI is found to support five

domains, ranging from adaptive learning and conversational tutoring to assessment, language processing, and administration, while raising four recurring concerns: academic integrity and cognitive over-reliance, data privacy and algorithmic bias, the reconfiguration of teacher authority, and the doctrinal reliability of machine-generated content. Read against the UNESCO, OECD, European Union, and Malaysian frameworks, these largely procedural regimes are shown to leave the substantive ends of AI undetermined. The article proposes a three-layered framework integrating tawhidic epistemology, the *maqasid al-shari'ah*, and the principles of *maslahah*, *amanah*, *karamah insaniyyah*, and *adab*. It may serve as a guide for policymakers, educational institutions, and religious scholars seeking to govern AI responsibly.

Keywords: artificial intelligence; Islamic education; maqasid al-shari'ah; AI governance; generative AI

INTRODUCTION

Artificial intelligence has moved, within a comparatively short period, from a specialised field of computer science into the everyday infrastructure of teaching and learning. In its broadest sense the term denotes computational systems that perform tasks ordinarily thought to require human intelligence, including reasoning, perception, language understanding, and learning from data (Russell & Norvig, 2021). Systems that generate text, respond to open-ended questions, translate between languages, and tailor material to the individual learner are now widely accessible, and their uptake in education has outpaced that of most earlier technologies (Crompton & Burke, 2023; Kasneci et al., 2023). Quantitative evidence indicates that this enthusiasm is not unfounded: a meta-analysis of adaptive learning systems reported a medium-to-large effect on cognitive outcomes relative to non-adaptive instruction (Wang et al., 2024), and broader reviews associate AI tools with gains in engagement, feedback, and the scaffolding of motivation (du Plooy et al.,

2024; Luckin et al., 2016). At the same time, a substantial literature documents accompanying hazards, among them threats to academic integrity, the reliability of generated content, and the risk that habitual reliance may erode independent thought (Zhai et al., 2024; Farrokhnia et al., 2024; Gerlich, 2025).

These developments have not arrived in a settled field. An influential systematic review of artificial intelligence in higher education observed that the literature was dominated by computer scientists rather than educators, that it engaged only weakly with established pedagogical theory, and that it reflected too little critically upon the ethical and educational risks of the technologies it described (Zawacki-Richter et al., 2019). Complementary critical scholarship has situated these shortcomings within the longer history of the field, tracing a recurrent gap between technological promise and educational reality and urging closer attention to the social and institutional conditions of its adoption (Williamson & Eynon, 2020). The arrival of generative models has sharpened rather than resolved these deficiencies, extending the reach of AI while multiplying the questions of authority, integrity, and formation that its educational use provokes (Selwyn, 2019; Holmes et al., 2022).

Islamic education, understood here as both the teaching of the Islamic disciplines and the formation of the learner within an Islamic worldview, has been drawn into this transformation. Learners increasingly consult conversational systems for help with Qur'anic memorisation, Arabic grammar, and the meaning of texts, while institutions experiment with adaptive platforms, automated assessment, and administrative automation (Amilusholihah & Ramadhan, 2025; An et al., 2025; Rapi et al., 2024). The encounter is consequential because Islamic education is not epistemically neutral. Its disciplines rest on a conception of knowledge in which revelation and reason operate as complementary sources, transmitted through a chain of qualified authority (*sanad*) and directed towards the cultivation of character (*adab*) and sincerity of intention (*niyyah*) (al-Attas, 1995; Bakar, 1998). A technology capable of generating authoritative-seeming religious content therefore raises questions that an evaluation framed solely in terms of efficiency cannot resolve, and that recent work on the reliability and abstention behaviour of language models on religious questions has begun to document empirically (Atif et al., 2025).

These questions assume particular urgency in Malaysia, where national policy integrates technology and Islamic education rather than treating them as separate domains. The country adopted its National Artificial Intelligence Roadmap 2021 to 2025, issued National Guidelines on AI Governance and Ethics in 2024, established a National AI Office in late 2024, and has incorporated technical training, including AI, into national *tahfiz* and Islamic educational pathways (Ministry of Science, Technology and Innovation [MOSTI], 2021, 2024; National AI Office, 2024; Hng, 2025). This orientation accords with the wider international consensus that education systems must prepare both teachers and learners to engage with AI competently and ethically (Miao & Cukurova, 2024; Miao & Shiohira, 2024). For Malaysian institutions of Islamic learning, the pertinent question is therefore not whether to engage with AI but how to do so in a manner that is at once pedagogically effective and faithful to Islamic principles.

Notwithstanding a rapidly expanding body of publications, the field exhibits several persistent deficiencies. First, it remains dominated by conceptual and small-scale descriptive studies, with comparatively few large empirical investigations and limited representation of Malaysian higher education (Faizin et al., 2025; Hady et al., 2025). Second, although ethical concern is nearly universal in this literature, its articulation is frequently confined to general statements, and only a portion connects these concerns to established Islamic legal and ethical scholarship such as the *maqasid al-shari'ah* (Mohadi & Tarshany, 2023; Rosmaini, 2025; Elmahjub, 2023). Third, the literature has not been consolidated into a single account linking the practical applications of AI, the recurrent challenges, and a principled Islamic response oriented towards Muslim-majority settings. A fourth deficiency follows from the first three: although a global architecture of AI governance has rapidly taken shape, the relationship between these largely secular instruments and the *maqasid al-shari'ah* remains unexamined, and the *maqasid*-based scholarship that does exist has remained predominantly evaluative, stopping short of the operational guidance that institutions require. The consequence is that institutions adopt AI without a coherent, values-based framework for distinguishing beneficial from harmful uses.

Addressing these deficiencies, the present study pursues three objectives: to map the principal applications of AI in Islamic education as reported

in recent scholarship; to evaluate critically the challenges and risks that accompany them; and to develop an Islamic epistemological and ethical framework capable of guiding responsible adoption. Three questions follow. First, how is AI currently applied across the domains of Islamic education? Second, what challenges does its adoption present, and how do these intersect with the distinctive aims of Islamic learning? Third, what framework, grounded in the Islamic intellectual tradition, can govern its responsible integration? The study's originality lies in three connected moves: it consolidates a fragmented field into a single critically argued synthesis; it benchmarks the *maqasid al-shari'ah* against the principal instruments of global and national AI governance to specify their precise complementarity; and it operationalises the *maqasid* as both a design instrument and a recurring governance cycle, thereby converting diffuse ethical concern into applicable guidance for Islamic institutions in Malaysia and comparable settings. The article proceeds through a review of the literature, a statement of methodology, a combined presentation and discussion of the findings that incorporates the Islamic epistemological and ethical analysis, and a conclusion.

LITERATURE REVIEW

Artificial Intelligence in Education

Artificial intelligence in education denotes computational systems capable of adaptation, prediction, natural language understanding, and decision support applied to teaching and learning. It is useful to distinguish the principal forms relevant here, since the term spans technologies with differing implications. Rule-based systems encode predetermined logic and characterised early intelligent tutoring; machine-learning systems infer patterns from data and underpin adaptive learning and analytics; natural language processing enables the analysis and generation of text and speech; and generative AI, the most recent category, produces novel output and underlies the conversational agents now in widespread use (Crompton & Burke, 2023; Russell & Norvig, 2021). These forms differ in transparency, data requirements, and propensity to generate unverified content, a distinction to which the ethical analysis below remains attentive. The aspiration to deliver individualised, responsive instruction at scale is long-standing, and proponents have argued that AI could approximate the benefits of one-

to-one tutoring for far larger numbers of learners than human tutoring alone can reach (Luckin et al., 2016; Holmes et al., 2019).

The effectiveness of these technologies is increasingly well documented. Research on intelligent tutoring systems, the most thoroughly evaluated category, is instructive. A meta-analysis of such systems in higher education found a moderate positive effect on learning relative to most conventional methods, with intelligent tutors outperforming traditional classroom instruction, reading, and computer-assisted instruction, while remaining somewhat less effective than expert one-to-one human tutoring (Steenbergen-Hu & Cooper, 2014). A parallel meta-analysis reported consistent positive effects across educational levels and subject domains (Ma et al., 2014), and an earlier synthesis established that human tutoring still set the benchmark that automated systems approached but did not surpass (VanLehn, 2011). This pattern, in which automation augments but does not displace the human tutor, recurs throughout the present review and acquires particular significance in the religious domain.

A scoping review of adaptive learning systems found consistent positive effects on performance and engagement, attributed to self-paced progression, real-time feedback, and flexibility (du Plooy et al., 2024), while a meta-analysis reported a medium-to-large effect on cognitive outcomes for adaptive instruction relative to conventional teaching (Wang et al., 2024). The arrival of generative models intensified both enthusiasm and concern, prompting international responses. Building on earlier guidance that set out for policy-makers the opportunities, risks, and conditions for the equitable adoption of AI in education (Miao et al., 2021), the United Nations Educational, Scientific and Cultural Organization issued global guidance anchored in a humanistic vision that foregrounds human agency, inclusion, and data protection, proposing measures including the protection of data privacy and an age threshold for the independent use of generative systems (Miao & Holmes, 2023; Sabzalieva & Valentini, 2023). It subsequently published competency frameworks for teachers and for students that recast the classroom as a teacher–AI–learner relationship and specify the knowledge, skills, and values each party requires (Miao & Cukurova, 2024; Miao & Shiohira, 2024). Parallel scholarship has consolidated a recognised set of AI-ethics principles encompassing fairness, transparency, accountability, and the

avoidance of harm, alongside sustained attention to algorithmic bias (Baker & Hawn, 2022; Jobin et al., 2019; Floridi et al., 2018).

The Cognitive and Ethical Debate over Generative AI

A growing evidentiary base interrogates the cognitive consequences of generative tools, and its findings are mixed rather than uniformly favourable. Reviews report that such systems can improve accessibility and provide cognitive support, yet caution that their misuse may impede independent learning and foster cognitive offloading, with one systematic review linking over-reliance on AI dialogue systems to diminished cognitive engagement (Zhai et al., 2024). Recent empirical work has sharpened this concern. A mixed-methods study of 666 participants reported a strong negative correlation between frequent AI use and critical-thinking performance, mediated by cognitive offloading and most pronounced among younger users, with higher education appearing to buffer the effect (Gerlich, 2025). An experimental study using electroencephalography found that participants who wrote essays with the assistance of a large language model exhibited weaker neural connectivity and poorer recall of their own writing than those who wrote unaided, a pattern its authors described as the accumulation of ‘cognitive debt’ (Kosmyrna et al., 2025). A large survey of knowledge workers likewise found that greater confidence in AI was associated with less self-reported critical thinking, whereas greater self-confidence was associated with more (Lee et al., 2025). These findings do not establish that AI inevitably diminishes cognition, but they indicate that the manner of use is decisive.

Concerns regarding academic integrity feature prominently. Plagiarism is among the most frequently cited risks of generative AI in higher education (Farrokhnia et al., 2024), and detailed treatments have examined both the difficulty of detecting machine-assisted work and the institutional responses, including policy, assessment redesign, and the cultivation of ethical use, that universities have begun to adopt (Cotton et al., 2024). A more fundamental concern attaches to the nature of the systems themselves. Large language models generate text by modelling statistical regularities rather than meaning, a property that helps explain their fluency and their propensity to produce confident but unfounded output, sometimes characterised as a tendency to reproduce form without understanding (Bender et al., 2021; Kasneci et al., 2023). A

recurrent and important qualification across this literature is that the value of these tools is contingent on pedagogical mediation, with educator guidance proving decisive in directing learners towards verification and self-regulation rather than passive dependence (Saqr et al., 2024; Yu, 2024). This contingency frames the more specific question pursued here: how AI interacts with a domain whose aims extend beyond information transfer to the formation of the person.

The Epistemology and Aims of Islamic Education

Islamic education is distinguished by its epistemological foundations. Classical scholarship classified knowledge into the transmitted sciences (*'ulum naqliyyah*), grounded in revelation, and the rational sciences (*'ulum 'aqliyyah*), grounded in reason and observation, regarding the two as complementary expressions of a single, divinely ordered reality (Bakar, 1998). The unifying principle of this scheme is *tawhid*, the affirmation of the oneness of God, which integrates inquiry into the natural world with inquiry into revelation rather than partitioning them into rival domains (Bakar, 1991; Nasr, 1993). On this understanding, education is not merely the transmission of information but the formation of the learner, encompassing intellectual discernment, ethical disposition, and right intention. Al-Attas articulates this orientation through the concept of *ta'dib*, the instilling of *adab* understood as the just ordering of the self in accordance with knowledge of reality (al-Attas, 1995; Wan Daud, 1998).

The twentieth-century programme for the Islamization of contemporary knowledge gave this epistemology renewed institutional expression. Al-Faruqi (1982) advanced a systematic work plan for reconstructing the modern disciplines in light of the Islamic worldview, emphasising mastery of both the modern discipline and the Islamic legacy and their critical synthesis. Al-Attas located the deeper task in the prior dewesternization of knowledge and the restoration of an Islamic metaphysical framework, contending that the malaise of the community is fundamentally a crisis of knowledge requiring the purification of imported concepts before any constructive synthesis (Wan Daud, 1998). The two formulations differ in emphasis, al-Faruqi's procedural programme contrasting with al-Attas's metaphysical priority, yet both rest on *tawhid* and on the conviction that God is the origin of knowledge. This heritage supplies the lens through which any technology introduced

into Islamic education, AI included, must be assessed: not by efficiency alone but by its compatibility with the integrative and formative aims of Islamic learning.

Science and Technology in the Islamic Intellectual Tradition

The encounter between Islamic education and AI has antecedents in the tradition's longer engagement with the instruments of inquiry. The classical period was marked by a sustained cultivation of the rational sciences alongside the transmitted ones, exemplified by al-Khwarizmi in mathematics, Ibn al-Haytham in optics and experimental method, and al-Biruni in astronomy, all of whom pursued empirical and mathematical inquiry within, rather than against, a theistic worldview (Bakar, 1991; Nasr, 1993). The instruments of reason were not regarded as autonomous from revelation but as means whose legitimacy derived from their service to truth and to the welfare of the community. If the tradition welcomed the astrolabe and the algorithm as *'aqli* instruments disciplined by a tawhidic worldview, AI may be approached in the same spirit, as an instrument whose admissibility depends upon its subordination to the ends of revealed guidance rather than upon any intrinsic incompatibility with the tradition (al-Faruqi, 1982). The discontinuity AI introduces lies not in its rationality but in its capacity to generate authoritative-seeming content autonomously, a feature that raises the question of authority in a form the classical instruments did not, and that the framework developed later in this article is designed to address.

Emerging Scholarship on AI in Islamic Education

Research at this intersection has expanded markedly since 2018. A systematic review applying the PRISMA framework to eighteen of an initial 379 articles published between 2022 and 2024, retrieved from Scopus, Google Scholar, and ERIC, concluded that AI improves students' understanding, motivation, critical thinking, creativity, and efficiency, while identifying ethical issues, privacy concerns, and the changing role of the teacher as the principal challenges, and recommending that stakeholders develop guidelines aligning AI use with Islamic principles (Amilusholihah & Ramadhan, 2025). A further PRISMA-based review of forty-two articles published between 2020 and 2025 identified five thematic areas, namely adaptive and personalised

learning, automated content analysis, predictive analytics, intelligent assessment, and multimedia instruction, while cautioning that such applications must preserve the authenticity of religious knowledge (Idham et al., 2026). A SWOT-structured analysis reached a complementary conclusion, locating the strengths of AI in accessibility and personalisation and its threats in dependence, doctrinal error, and the dilution of the teacher's role (Abubakari et al., 2024).

Within Qur'anic studies, a systematic review published in a *Universiti Malaya* journal distinguished three strands of application: semantic analysis of Qur'anic text, recitation and memorisation support through voice recognition and intelligent feedback, and broader pedagogical and translation uses, while flagging concerns regarding authenticity, contextual appropriateness, and cultural sensitivity (Mohd Noor et al., 2025). A thematic review devoted specifically to the variant readings (*qira'at*) cautioned that the oral and transmitted character of this science poses distinctive difficulties for automated processing, since the integrity of recitation depends upon an unbroken human chain that no system can supply (Abdelgelil et al., 2025). A bibliometric analysis of the integration of deep learning and AI in Qur'anic education and research, drawing on the Scopus database, documented rapid growth in output and widening international collaboration, situating the field as one in formation rather than maturity (An et al., 2025). Comparable work has examined the application of natural language processing to Arabic and to Qur'anic studies, surveying both the technical challenges and the opportunities (Al-Shidi et al., 2025; Bashir et al., 2023). At the level of underlying technique, researchers have developed Arabic language models capable of machine reading comprehension over the text of the Qur'an, demonstrating both the feasibility and the difficulty of the task (Malhas & Elsayed, 2022).

Empirical contributions are beginning to complement these syntheses. In a study of 224 Muslim university students across twelve institutions, analysed through partial least squares structural equation modelling and framed by an extended technology acceptance model, perceived ease of use significantly shaped perceived usefulness, which in turn influenced attitudes and behavioural intention to adopt AI for Islamic religious learning, although actual use remained limited (Faizin et al., 2025; Davis, 1989). A quasi-experimental study in Indonesian Islamic higher

education reported that AI-driven tools improved measures of Qur'anic literacy relative to conventional instruction, while its authors cautioned that the gains required careful pedagogical framing (Ridwan, 2025). Exploratory work has examined the use of AI applications to support Qur'anic memorisation, including among students outside the Islamic-studies stream (Ab Alim et al., 2025), and other studies have considered the contribution of AI to inclusive provision for learners with additional needs (Anandal et al., 2024). Qualitative case studies of Islamic schools report that AI enhances Qur'anic learning, ethical discussion, and academic proficiency through adaptive feedback and automated assessment, while emphasising that it must complement rather than replace teacher-led instruction if the spiritual and moral aspects of Islamic education are to be preserved (Djazilan et al., 2024). Studies of the Indonesian higher-education sector have identified both the opportunities AI presents and the infrastructural and competence constraints that condition its adoption (Rapi et al., 2024). Critical contributions, finally, interrogate the implications of AI for the teacher's epistemic authority and for the chain of transmission of religious knowledge, and document the competencies that religious educators must now acquire (Tobing et al., 2025; Arif et al., 2025).

Global AI Governance Frameworks and Islamic Ethics

As AI has diffused across sectors, a layered architecture of governance has taken shape, extending from non-binding global standards through binding regional law to national strategy. Reading these instruments against the *maqasid al-shari'ah* clarifies both their convergence and a shared substantive limitation that an Islamic ethics is well positioned to address. Four instruments are particularly salient.

The UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted by acclamation by all member states in 2021, is the first global standard-setting instrument on the subject. It anchors AI governance in human rights and human dignity and foregrounds human oversight, transparency, fairness, accountability, privacy, and environmental sustainability, translating these values into concrete policy action areas spanning data governance, education, and the protection of vulnerable groups (UNESCO, 2022). Its force, though, lies more in orientation than in specification. Dignity is made the anchor of the whole edifice, yet the term itself is left largely undefined, and the Recommendation's

universalist idiom quietly assumes a shared content for it that different moral and religious traditions may in practice fill very differently. The OECD AI Principles, first adopted in 2019 and updated in 2024 to address generative and general-purpose systems, constitute the first intergovernmental standard on AI; they set out five values-based principles, namely inclusive growth and well-being, human-centred values and fairness, transparency and explainability, robustness and safety, and accountability, alongside recommendations for policymakers (OECD, 2024). Here the orientation is more technocratic than moral. Trustworthiness and economic vitality are kept to the fore, and the principles say little about the ends a human life should serve; goods that a religious tradition would call transcendent or formative fall outside their frame.

The European Union's Artificial Intelligence Act, which entered into force in 2024, is the first comprehensive and legally binding framework for AI. It adopts a risk-based architecture that graduates obligations according to the risk a system poses to health, safety, and fundamental rights: practices deemed an unacceptable risk are prohibited outright, while high-risk systems must satisfy requirements for risk management, data governance, transparency, human oversight, and post-market monitoring (European Union, 2024). Its great strength is that it binds where the UNESCO and OECD texts only commend. Even so, its gaze is fixed on risk to rights and to safety, and the spiritual and formative dimensions that weigh most heavily in religious education lie wholly beyond its reach. Malaysia's own governance combines strategy with ethics guidance. The National Artificial Intelligence Roadmap 2021–2025 sets out the policy architecture for an AI-driven economy (Ministry of Science, Technology and Innovation, 2021), and in 2024 the responsible ministry issued the National Guidelines on AI Governance and Ethics, which advance seven principles, namely fairness, reliability and safety, privacy and security, inclusiveness, transparency, accountability, and the pursuit of human benefit, explicitly modelled on the UNESCO, OECD, and EU instruments and addressed to users, policymakers, and developers (Ministry of Science, Technology and Innovation, 2024). Institutional coordination was strengthened by the establishment of a National AI Office in 2024 (National AI Office, 2024).

Malaysia thus brings unusual policy ambition to the question, even if, as the next section argues, its institutional readiness remains uneven.

Across these instruments a convergent core is evident, encompassing transparency, fairness, accountability, privacy, safety, and human oversight, a convergence that comparative analyses of the global AI-ethics landscape have documented even as they note substantial divergence in interpretation and implementation (Jobin et al., 2019; Floridi et al., 2018). Yet this convergence reveals a feature that is also a limitation. These frameworks are predominantly procedural and rights-based: they specify how AI should be constrained, requiring that it be transparent, fair, accountable, and subject to human oversight, without specifying the substantive human ends that AI should serve. They affirm that AI must respect human dignity and promote well-being, but they leave the content of dignity and flourishing largely undetermined, to be supplied by each society's own moral tradition (Elmahjub, 2023). It is precisely here that the *maqasid al-shari'ah* contributes. Where the international instruments establish a procedural floor, the *maqasid* furnish a substantive account of the goods to be preserved, namely religion, intellect, life, lineage, and property, and thereby specify, for a community oriented towards the knowledge of God, what the well-being that AI should serve actually consists in. Table 1 sets out this comparison. The relationship is complementary rather than competitive: a *maqasid*-based ethics does not displace the procedural safeguards of the international frameworks but supplies the substantive normative content that they presuppose yet do not themselves provide.

Table 1. Principal AI-governance instruments compared with a *maqasid*-based ethics.

Framework	Status and scope	Core orientation	Representative principles	Relation to a <i>maqasid</i> -based ethics
UNESCO Recommendation (UNESCO, 2022)	Non-binding global standard; adopted by all member states (2021)	Human rights and human dignity; human oversight	Transparency, fairness, accountability, privacy, oversight, sustainability	Shares the commitment to dignity but leaves its content open; <i>maqasid</i> specify the goods to be preserved

ARTIFICIAL INTELLIGENCE IN ISLAMIC EDUCATION: OPPORTUNITIES,
CHALLENGES, AND AN ETHICAL FRAMEWORK GROUNDED IN MAQASID AL-
SHARI'AH

Framework	Status and scope	Core orientation	Representative principles	Relation to a maqasid-based ethics
OECD AI Principles (OECD, 2024)	Non-binding intergovernmental standard; 2019, updated 2024	Trustworthy, human-centred AI	Inclusive growth; human-centred values and fairness; transparency; robustness; accountability	Procedural and economic emphasis; offers no account of transcendent or formative ends
EU AI Act (European Union, 2024)	Legally binding regulation; in force 2024	Risk-based protection of fundamental rights and safety	Risk tiers; human oversight; data governance; transparency	Rights-based and secular; silent on the doctrinal and spiritual dimensions central to religious education
Malaysia: Roadmap and AIGE (MOSTI, 2021, 2024)	National strategy plus voluntary ethics guidelines	Trusted and responsible AI; “AI for Malaysia, AI for All”	Fairness; safety; privacy; inclusiveness; transparency; accountability; human benefit	Adopts international principles but does not yet articulate a distinctively Islamic grounding despite the national context
Maqasid al-Shari'ah	Normative-ethical tradition of Islamic jurisprudence	Preservation of the higher objectives of the law	Hifz al-din, al-‘aql, al-nafs, al-nasl, al-mal; maslahah, amanah, karamah, adab	Supplies the substantive ends that procedural regimes leave undetermined; requires operationalisation (this study)

AI Governance and Islamic Education in the Malaysian Context

Malaysia is a particularly instructive setting because national policy treats technology and Islamic education as integrated rather than separate domains. The National AI Roadmap and the subsequent ethics guidelines position Malaysian institutions to adopt AI within an explicit national framework (Ministry of Science, Technology and Innovation, 2021, 2024), while the National AI Office, established in 2024 under the relevant ministry, provides centralised coordination of strategy, governance, and capacity-building and is tasked with developing further guidance, including a national code of AI ethics (National AI Office, 2024). This architecture gives Malaysian institutions of Islamic learning an unusually developed policy environment within which to situate their own adoption of AI.

Within this framework, national initiatives have sought to equip students in Qur’anic memorisation streams with technical competencies, including AI, so that graduates retain their religious formation while contributing across professional fields, and have promoted digital tools

within *tahfiz* education (Hng, 2025). More broadly, Islamic educational institutions have undergone a digitalisation of administration, assessment, and instruction, a transformation documented in case studies that nonetheless emphasise the indispensability of teacher mediation if the spiritual and moral aims of Islamic education are to be preserved (Djazilan et al., 2024). The opportunities are therefore accompanied by a recurrent caution that technological provision is a necessary but not a sufficient condition for benefit.

The principal constraint is human rather than technical. The documented gap between technological provision and teacher readiness indicates that policy ambition has outpaced institutional capacity (Rapi et al., 2024; Djazilan et al., 2024), and recent scholarship has begun to specify the digital and AI competencies that religious educators must acquire (Arif et al., 2025). The cultivation of AI literacy, understood as the capacity to understand, evaluate, and use AI critically (Long & Magerko, 2020), is therefore a precondition for responsible adoption, a need that international competency frameworks for teachers and students have made concrete (Miao & Cukurova, 2024; Miao & Shiohira, 2024). For institutions of Islamic learning, AI literacy carries an additional dimension: it must include the capacity to appraise machine-generated religious content against authoritative scholarship, so that critical verification becomes a habitual professional competence rather than an occasional safeguard. This requirement connects the Malaysian policy context directly to the epistemological analysis developed below.

Synthesis and Research Gap

Several observations follow from this review. First, there is broad and now meta-analytically supported agreement that AI confers tangible pedagogical benefits, particularly through personalisation and feedback (Wang et al., 2024; du Plooy et al., 2024; Steenbergen-Hu & Cooper, 2014). Second, an equally consistent literature documents risks, among which academic integrity, cognitive over-reliance, privacy, bias, and the reconfiguration of the teacher's role recur most frequently (Zhai et al., 2024; Gerlich, 2025; Tobing et al., 2025). Third, while ethical concern pervades the Islamic education literature, its articulation is frequently general, and few studies connect it systematically to the *maqasid al-shari'ah* or to the tawhidic epistemology that defines the domain (Mohadi & Tarshany, 2023; Elmahjub, 2023). Fourth, the emerging architecture of

AI governance, though increasingly influential, has not been brought into critical relation with the *maqasid*, and the *maqasid*-based scholarship that exists has remained evaluative rather than operational. These are the gaps the present study addresses. It is worth noting that the contradiction within the evidence, robust benefits set against documented cognitive and ethical risks, is precisely what a principled framework is required to adjudicate, since neither uncritical adoption nor blanket prohibition follows from the literature as it stands.

Conceptual Framework

The study is guided by a conceptual framework integrating three layers of the Islamic intellectual tradition. At its foundation lies the tawhidic conception of knowledge, which holds the transmitted (*naqli*) and rational (*aqli*) sciences to be complementary aspects of a single, divinely ordered reality, and which positions any rational instrument, AI included, as legitimately employable insofar as it serves and is disciplined by the ends of revealed guidance (Bakar, 1998; al-Faruqi, 1982). Upon this foundation rests the evaluative layer of the *maqasid al-shari'ah*, whose five objectives, the preservation of religion, intellect, life, lineage, and property, supply concrete criteria against which the effects of AI may be assessed (Mohadi & Tarshany, 2023; Auda, 2008; Kamali, 2008). The third layer comprises the regulating principles of *maslahah*, *amanah*, *karamah insaniyyah*, and *adab*, which govern the disposition and conduct with which AI is adopted in practice. This structure, depicted in Figure 2 and elaborated below, provides the lens through which the findings are interpreted and constitutes the study's principal theoretical contribution.

METHODOLOGY

This study is conducted as a critical narrative literature review, employing qualitative thematic synthesis and critical interpretive analysis and complemented by a comparative reading of the principal AI-governance instruments. The interdisciplinary character of the subject, which spans AI, Islamic education, and ethics, calls for an interpretive and critical approach rather than the purely systematic aggregation of evidence that a PRISMA-type protocol is designed to deliver. Rather than aggregating quantitative outcomes, the review interprets and synthesises a body of scholarship that spans several disciplines, and it culminates in the

construction of a normative framework. The sections that follow set out the research design, the literature search strategy, the basis for selecting sources, the analytical procedure, the appraisal of source quality, and the ethical considerations governing the study.

Research Design

A narrative literature review was selected as the most appropriate design for the questions this study addresses. Narrative reviews are well suited to topics that are interdisciplinary, conceptually complex, and still in formation, where the aim is to interpret and integrate diverse strands of scholarship rather than to measure the pooled magnitude of an effect. The subject of this study, the integration of artificial intelligence into Islamic education, draws simultaneously upon the empirical literature on educational technology, the philosophy of Islamic education and epistemology, the ethics and governance of AI, and the jurisprudential tradition of the *maqasid al-shari'ah*. These bodies of knowledge differ in method, vocabulary, and evidentiary standards, and they cannot be reconciled through statistical aggregation; they require interpretive synthesis. The review therefore combines thematic analysis, through which recurring patterns across the literature are identified and organised, with critical and normative analysis, through which those patterns are evaluated against the commitments of the Islamic intellectual tradition. The purpose throughout is interpretive synthesis and the construction of a coherent conceptual framework, not the quantitative measurement or generalisation of findings. This choice carries a recognised cost in systematicity, which the procedures described below are designed to mitigate by making the search, selection, and analytical steps explicit and reproducible in their logic even where they are not exhaustive in their coverage.

A systematic protocol of the PRISMA type was considered but, in the end, set aside. Such protocols are built to pool comparable empirical studies around a tightly bounded question. The present inquiry is of a different kind. It must hold together literatures that differ in method and in vocabulary, and it must move from describing what is known towards arguing what ought to follow. A strict systematic search would also have excluded the foundational and jurisprudential sources on which much of the analysis depends. What a narrative review gives up in exhaustiveness it recovers in interpretive reach. Its principal limitation is real, namely the

reliance on scholarly judgement in selecting and weighing sources, and the procedures described below are meant to keep that judgement transparent and disciplined rather than to pretend it away.

Literature Search Strategy

A structured search of the scholarly literature was undertaken across three databases (Scopus, Web of Science, and Google Scholar), chosen for their disciplinary breadth and their coverage of both the educational-technology and the Islamic-studies literatures. The search concentrated on the period from 2018 to 2026 in order to capture the expansion of scholarship that followed the wider availability of generative AI, while foundational works of Islamic scholarship lying outside this window, including those of al-Faruqi, al-Attas, and Bakar, were consulted where they were conceptually necessary to the analysis. The principal AI-governance instruments were retrieved from their issuing bodies. Search terms combined the vocabulary of the technology with that of the educational and religious context, pairing expressions such as artificial intelligence, machine learning, and generative AI with expressions such as Islamic education, Islamic religious education, Qur'an, madrasah, and *tahfiẓ*. Boolean operators were used to combine these expressions, and the searches were repeated with minor variations to reduce the risk that idiosyncratic indexing would exclude relevant work. The reference lists of pertinent reviews were consulted to locate further relevant scholarship, an approach, sometimes termed backward snowballing, that is consistent with the exploratory and interpretive character of a narrative review.

Literature Selection

Sources were selected through scholarly judgement rather than through a mechanical screening protocol, in keeping with the narrative-review design. Four considerations guided the selection. The first was relevance to the research objectives: whether a source illuminated the applications of AI in Islamic education, the challenges these present, or the ethical and epistemological questions they raise. The second was scholarly quality, with preference given to peer-reviewed and authoritative work. The third was contribution to the field, that is, whether a source advanced understanding rather than merely restating established positions. The fourth was significance to the specific intersection of AI

and Islamic education, which ensured that the review remained focused on its subject rather than diffusing into the much larger general literature on educational technology. Works that touched upon AI only incidentally, or that addressed unrelated domains, were not drawn upon, although a deliberately selected body of general educational, AI-ethics, and AI-governance scholarship was retained where it supplied the empirical, conceptual, or regulatory grounding that the more specialised literature presupposes. Table 2 summarises the considerations that informed the selection of literature.

Table 2. Considerations guiding the selection of literature for the narrative review.

Consideration	Basis for selection
Relevance to objectives	The source illuminates the applications, challenges, or ethical and epistemological dimensions of AI in Islamic education.
Scholarly quality	Preference for peer-reviewed journal articles, scholarly reviews, and authoritative policy or foundational works.
Contribution to the field	The source advances understanding rather than restating established positions.
Significance to the topic	The source bears directly on the intersection of AI and Islamic education, or supplies grounding general, ethical, or governance scholarship that the specialised literature presupposes.
Currency	Published 2018–2026, with foundational Islamic scholarship consulted where conceptually necessary.

Analytical Procedure

The selected literature was analysed through thematic synthesis, undertaken in a sequence of stages adapted to the interpretive aims of the study. The first stage was close and repeated reading, through which the author became familiar with the content, arguments, and evidentiary basis of each source. The second stage was coding, in which significant statements, findings, and claims were labelled according to their conceptual content, attending both to what the literature reported about AI in Islamic education and to the concerns it raised. The third stage was categorisation, in which related codes were grouped into broader analytical categories. The fourth stage was comparison, in which categories were examined across sources to identify convergences, divergences, and tensions, including the central tension between the documented benefits of AI and its documented risks. The final stage was

synthesis, in which the categories were consolidated into the themes that organise the findings: the applications of AI, the challenges of its adoption, and the requirements of an Islamic ethical governance, and through which the three-layered *maqasid*-based framework was constructed. Coding proceeded inductively from the literature while remaining attentive to the research questions, so that the themes reflected the scholarship itself rather than being imposed upon it. The overall analytical process is depicted in Figure 1.

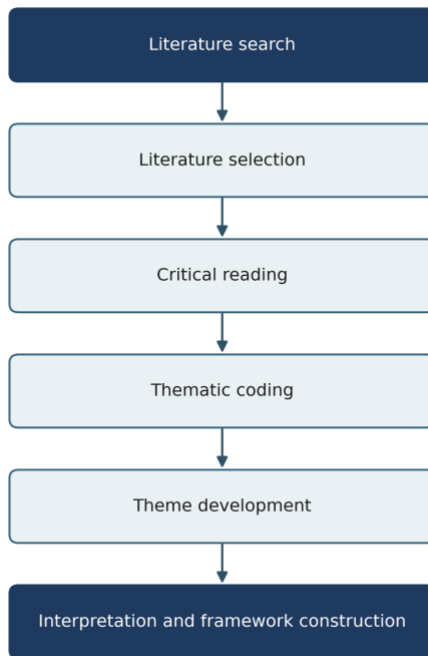


Figure 1. Analytical process of the narrative literature review and thematic synthesis.

Quality and Credibility of Sources

Because the credibility of a narrative review depends upon the standing of the sources on which it draws, priority was accorded to scholarship of

established authority. Peer-reviewed journal articles formed the principal evidentiary base, supplemented by systematic reviews and meta-analyses, which were drawn upon for the more robust quantitative claims, and by authoritative policy and governance documents, which informed the discussion of regulation. Foundational works of Islamic scholarship were relied upon for the epistemological and jurisprudential dimensions of the analysis. Where claims rested on smaller or more preliminary studies, this is signalled in the discussion, and greater interpretive weight is accorded to findings supported by stronger or more numerous sources, such as the meta-analytic evidence on adaptive learning and intelligent tutoring (Wang et al., 2024; Steenbergen-Hu & Cooper, 2014). This graded treatment of evidence runs throughout the findings and discussion, allowing the strength of the support for each claim to be made apparent to the reader.

Ethical Considerations

As a review of published scholarship, this study involved no human participants, and ethical approval was therefore not required. The ethical obligations pertinent to a study of this kind concern the integrity of scholarship rather than the protection of subjects. For that reason, sources were read critically and represented accurately, their arguments rendered in the author's own formulations so as to preserve scholarly integrity and avoid the displacement of original wording, and their claims attributed faithfully. Where the literature was divided, the division is reported rather than concealed, so that the synthesis reflects the state of the field as it is rather than as it might be wished to be. These commitments are themselves an instance of the principle of *amanah*, the faithful discharge of a trust, that the framework developed below identifies as central to the responsible conduct of inquiry in an age of generative tools.

RESULTS AND DISCUSSION

This section presents and interprets the findings of the thematic synthesis. It first reports the application domains of AI in Islamic education (4.1), the challenges and risks that accompany them (4.2), and the empirical evidence available in the literature (4.3). It then interprets these findings (4.4) and develops the Islamic epistemological and ethical framework through which they are best understood, including its

operationalisation as a design instrument and a governance process (4.5). The final subsections set out the study's theoretical contribution (4.6), its policy and practical implications (4.7), and its limitations and the resulting research agenda (4.8).

Applications of AI in Islamic Education

The reviewed literature points to five application domains that recur across the field, summarised in Table 3 and elaborated below.

Application domain	Description	Reported benefits (selected sources)
Personalised and adaptive learning	Adjustment of pace, sequence, and difficulty to the individual learner.	Moderate-to-large cognitive gains; improved understanding (Wang et al., 2024; Amilusholihah & Ramadhan, 2025).
Intelligent tutoring and conversational agents	On-demand support for Qur'an, recitation, tajwid, and Arabic; real-time feedback.	Engagement, motivation, accessibility (Mohd Noor et al., 2025; Steenbergen-Hu & Cooper, 2014).
Automated assessment and feedback	Rapid, formative feedback and reduced marking load.	Faster feedback loops; teacher time redirected (Idham et al., 2026).
Language and manuscript processing	NLP applied to Arabic and to interpretation of religious texts.	Wider access to sources; translation and comprehension support (Al-Shidi et al., 2025; Malhas & Elsayed, 2022).
Inclusive and accessible provision	Adaptation of materials and feedback for learners with additional needs.	Wider participation; individualised support (Anandal et al., 2024).
Institutional administration	Automation of scheduling, records, and routine tasks.	Operational efficiency (Djazilan et al., 2024).

Table 3. Application domains of AI in Islamic education and reported benefits.

Personalised and Adaptive Learning

Adaptive systems represent the most empirically substantiated application. In the general literature, a scoping review found consistent positive effects on performance and engagement, attributed to self-paced progression and real-time feedback (du Plooy et al., 2024), and a meta-analysis reported a medium-to-large effect on cognitive outcomes relative to non-adaptive instruction (Wang et al., 2024). The evidence on intelligent tutoring points in the same direction, with automated systems improving on most conventional methods while remaining below the benchmark of expert human tutoring (Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). For Islamic education, where cohorts in Qur'anic and Arabic studies vary widely in prior attainment, the capacity to individualise instruction addresses a longstanding difficulty, and the

systematic review of the field associates it with gains in understanding and efficiency (Amilusholihah & Ramadhan, 2025). A qualification from the general literature is nonetheless pertinent: adaptive systems may optimise for narrow performance indicators at the expense of broader objectives, a limitation that matters acutely in a domain whose aims are formative as well as cognitive (Selwyn, 2019).

Intelligent Tutoring and Conversational Agents

Conversational and tutoring systems are the most visible application within Islamic education. In Qur'anic instruction, voice-recognition systems analyse a learner's recitation, detect errors in articulation and tajwid, and return real-time correction, functions grouped in the systematic review of Qur'anic AI under recitation, memorisation, and pronunciation support (Mohd Noor et al., 2025). The deployment of large-scale platforms delivering supervised recitation and memorisation instruction across multiple languages illustrates the scale these applications have reached, and national initiatives in Malaysia have promoted comparable digital tools within *tahfiz* education (Hng, 2025). Exploratory studies report that learners, including those outside the Islamic-studies stream, make use of such applications to support memorisation, though their effectiveness depends on the accuracy of the underlying models and on appropriate supervision (Ab Alim et al., 2025). Empirically, the acceptance study of Muslim university students found that perceived ease of use significantly predicted perceived usefulness and, in turn, the intention to adopt such tools (Faizin et al., 2025). The consistent framing across this literature, and across the broader debate on whether automated systems can substitute for teachers, is that conversational agents supplement rather than supplant the teacher (Selwyn, 2019; Guilherme, 2019).

Assessment, Language Processing, and Administration

The remaining domains, though less prominent, are well attested. Automated assessment returns formative feedback rapidly and reduces the marking burden, a benefit documented in the PRISMA-based review of deep learning in Islamic religious education (Idham et al., 2026). The application of natural language processing to Arabic and to Qur'anic studies widens access to primary sources and supports translation and comprehension, though it confronts distinctive challenges arising from

the orthography, morphology, and sacred status of the text; researchers have nonetheless built Arabic models capable of machine reading comprehension over the Qur'an, and have surveyed the wider landscape of techniques and resources (Al-Shidi et al., 2025; Malhas & Elsayed, 2022; Bashir et al., 2023). Administrative automation, encompassing scheduling, records, and routine processes, improves operational efficiency and redirects staff time towards instruction, and features in case studies of digital transformation in Islamic schools (Djazilan et al., 2024). Across all three, the benefits are real but instrumental, bearing on the efficiency rather than the substance of religious formation.

Challenges and Risks

The analysed literature consistently highlights four clusters of challenge, summarised in Table 4 and elaborated below.

Table 4. Recurrent challenges of AI adoption in Islamic education.

Challenge cluster	Implication for Islamic education (selected sources)
Academic integrity and cognitive over-reliance	Plagiarism and cognitive offloading may weaken the reflection central to character formation (Zhai et al., 2024; Gerlich, 2025; Cotton et al., 2024).
Data privacy and algorithmic bias	Data collection raises issues of consent and protection, acute for minors; ML systems may reproduce cultural and religious bias (Miao & Holmes, 2023; Baker & Hawn, 2022; Naous et al., 2024).
Teacher authority and transmission	AI may unsettle the teacher's epistemic authority and the chain of transmission (<i>sanaad</i>) of religious knowledge (Tobing et al., 2025; Guilherme, 2019).
Doctrinal accuracy of content	Generated answers may be inaccurate, fabricated, or decontextualised, risking misguidance on creed, law, and practice (Mohd Noor et al., 2025; Bender et al., 2021; Atif et al., 2025).

Academic Integrity and Cognitive Over-Reliance

The most widely discussed challenge concerns the integrity of learning. Reviews of generative AI in higher education identify plagiarism among the most frequently cited concerns, and detailed treatments examine the difficulty of detection and the institutional responses required (Farrokhnia et al., 2024; Cotton et al., 2024). A subtler risk is cognitive offloading, the habitual delegation of thinking to the system, which a systematic review associates with diminished cognitive engagement (Zhai et al., 2024) and which recent empirical work has begun to measure

directly: a large mixed-methods study found a strong negative association between frequent AI use and critical-thinking performance mediated by offloading (Gerlich, 2025), an experimental study recorded weaker neural connectivity and recall among AI-assisted writers (Kosmyrna et al., 2025), and a survey of knowledge workers linked greater trust in AI to reduced critical engagement (Lee et al., 2025). In Islamic education this risk carries a formative dimension: where reflection and contemplation are themselves objectives of religious formation, their displacement by automated production threatens not merely the quality of learning but the cultivation of the disposition the tradition seeks to instil (Amilusholihah & Ramadhan, 2025). The literature consistently identifies educator mediation as the principal safeguard (Saqr et al., 2024; Yu, 2024).

Data Privacy and Algorithmic Bias

The second challenge concerns the data on which AI depends. The collection and processing of learner information raise questions of consent, confidentiality, and security that international guidance treats as central, proposing both the protection of data privacy and an age threshold for the independent use of generative systems, a provision of particular relevance to religious schools serving minors (Miao & Holmes, 2023). Beyond privacy, the wider literature documents the propensity of machine-learning systems to reproduce and amplify the biases present in their training data (Baker & Hawn, 2022; Mehrabi et al., 2021). For Islamic education the concern is twofold and now partly evidenced: systems trained predominantly on non-Islamic or non-Arabic sources may misrepresent the tradition, as studies measuring cultural bias in large language models have shown, finding that such models often default to Western and non-Muslim assumptions (Naous et al., 2024); and opaque models may render the grounds of their outputs unaccountable, in tension with the transparency that the responsible handling of a trust requires. Such biases are not incidental but reflect, as critical scholarship on the political economy of AI has argued, the relations of power and the classificatory choices embedded in the data and categories on which these systems are built (Crawford, 2021).

Teacher Authority and the Chain of Transmission

The third challenge concerns the reconfiguration of the teacher's role. In secular settings this is often framed as a shift from instructor to facilitator, and a sustained critical literature has questioned whether automated systems can replicate the social, emotional, and relational dimensions of teaching at all (Selwyn, 2019; Guilherme, 2019). In Islamic education the matter goes deeper, for it touches religious authority itself. Knowledge in this tradition has been transmitted through a personal relationship between teacher and student and authenticated through a recognised chain of transmission, with the teacher serving as both a source of reliable knowledge and a moral exemplar (al-Attas, 1995). Critical scholarship cautions that uncritical reliance on AI may unsettle this relationship and the epistemic authority bound up with it, and calls for a reconstruction of the teacher's role that preserves its authenticating and formative functions (Tobing et al., 2025). International competency frameworks converge on the same point from a different direction, recasting the classroom as a teacher–AI–learner relationship in which the teacher's judgement remains central (Miao & Cukurova, 2024). The teacher's role thus shifts towards discernment and verification rather than being displaced.

Doctrinal Accuracy and the Authority of Content

The fourth challenge, distinctive to the religious domain, concerns the reliability of machine-generated content on matters of creed, law, and worship. Generative systems are known to produce fluent but sometimes inaccurate or decontextualised output, a consequence of their modelling statistical form rather than meaning (Farrokhnia et al., 2024; Bender et al., 2021), and the review of Qur'anic AI identifies authenticity and contextual appropriateness as principal concerns (Mohd Noor et al., 2025). The point is no longer merely theoretical: recent evaluations of language models on religious questions have documented their tendency to fabricate citations and rulings and have argued that, for such questions, the responsible behaviour of a system is often to abstain rather than to answer (Atif et al., 2025). Because such systems lack the grounding in recognised scholarly authority that authenticates religious knowledge, their outputs on doctrinal matters require verification against authoritative sources, and the consequences of unverified reliance are

graver than in secular subjects, extending to the risk of misguidance in belief and practice. This challenge most directly motivates the principled subordination of AI to qualified human scholarship developed below.

Empirical Evidence

Although the literature is weighted towards conceptual work, several empirical findings are notable and are summarised in Table 5. The structural-equation study of Muslim university students found that perceived ease of use significantly affected perceived usefulness, attitudes, and behavioural intention, even as actual use remained limited (Faizin et al., 2025). A quasi-experimental study in Indonesian Islamic higher education reported gains in Qur'anic literacy from AI-assisted instruction (Ridwan, 2025). In the wider field, meta-analytic synthesis reported a medium-to-large effect of adaptive learning on cognitive outcomes (Wang et al., 2024), a scoping review confirmed consistent gains in engagement (du Plooy et al., 2024), and meta-analyses of intelligent tutoring established a moderate positive effect across domains (Steenbergen-Hu & Cooper, 2014; Ma et al., 2014). The reviewed scholarship on Islamic education reports gains in understanding, motivation, and engagement, conditional on pedagogical mediation and seldom disaggregated by specific cognitive process (Amilusholihah & Ramadhan, 2025; Idham et al., 2026). Within Qur'anic education, evaluations of recitation-recognition systems indicate useful accuracy in detecting tajwid errors while underscoring the continued necessity of qualified human oversight (Mohd Noor et al., 2025). Set against these gains, the cognitive literature reviewed above provides countervailing evidence that unguided use may depress critical thinking (Gerlich, 2025; Kosmyna et al., 2025). Case studies of Islamic schools, finally, identify teacher readiness and infrastructure as decisive constraints on effective adoption (Djazilan et al., 2024; Rapi et al., 2024). Across this body of work the evidence remains limited and, in places, preliminary. It would be premature to draw firm conclusions from it, and some caution is warranted in generalising to Islamic educational settings in particular.

ARTIFICIAL INTELLIGENCE IN ISLAMIC EDUCATION: OPPORTUNITIES,
CHALLENGES, AND AN ETHICAL FRAMEWORK GROUNDED IN MAQASID AL-
SHARI'AH

Table 5. Selected empirical and review evidence on AI adoption and effectiveness.

Source	Design	Principal finding
Faizin et al. (2025)	Survey; PLS-SEM; 224 students	Ease of use shapes usefulness, attitude, and intention to adopt AI in Islamic religious learning.
Ridwan (2025)	Quasi-experiment; Islamic HE	AI-driven tools improved Qur'anic literacy relative to conventional instruction.
Wang et al. (2024)	Meta-analysis	Adaptive learning systems yield a medium-to-large effect on cognitive outcomes.
Steenbergen-Hu & Cooper (2014)	Meta-analysis	Intelligent tutoring yields a moderate effect; below human tutoring, above most other methods.
Gerlich (2025)	Mixed methods; 666 participants	Frequent AI use negatively associated with critical thinking, mediated by cognitive offloading.
Amilusholihah & Ramadhan (2025)	Systematic review; 18 articles	Gains in understanding and motivation; ethical, privacy, and teacher-role concerns.
Djazilan et al. (2024)	Qualitative case study	AI aids learning but requires teacher mediation; readiness and infrastructure constrain adoption.

Interpreting the Findings

Read as a whole, the findings show AI offering Islamic education real value while introducing risks that cannot be reduced to technical performance. The opportunities cluster around access, personalisation, and immediacy of feedback, benefits especially salient in Qur’anic and Arabic studies where learners progress at markedly different rates (Mohd Noor et al., 2025; Amilusholihah & Ramadhan, 2025). These findings align with the wider literature, in which adaptive and conversational systems are credited with improving accessibility and producing measurable cognitive gains (Crompton & Burke, 2023; Wang et al., 2024). They diverge in one respect: the benefits reported in Islamic education are consistently framed as supplementary to, rather than substitutive of, the teacher, reflecting a recognition that the aims of religious instruction exceed the transfer of information (Faizin et al., 2025; Djazilan et al., 2024). The evidentiary base is, however, uneven; much of the Islamic education literature is conceptual or small in scale, and the more robust

quantitative findings derive from the general educational literature, so conclusions specific to Islamic settings remain provisional (Zawacki-Richter et al., 2019).

The challenges, in turn, acquire a distinctive coloration in this domain. The general concern that over-reliance may erode critical capacities, now supported by converging empirical studies (Zhai et al., 2024; Gerlich, 2025; Kosmyna et al., 2025), is intensified where the cultivation of reflection is itself a formative objective, so that cognitive offloading threatens the very disposition religious education seeks to instil (Amilusholihah & Ramadhan, 2025). The worry about inaccurate output becomes, in matters of creed and law, a concern about misguidance, since generative systems model form rather than meaning and lack the grounding in recognised authority that authenticates religious knowledge (Bender et al., 2021; Atif et al., 2025; Tobing et al., 2025). The reconfiguration of the teacher's role is weightier than in secular settings, because the teacher is a node in the chain of transmission and a moral exemplar (al-Attas, 1995). The documented gap between technological provision and teacher readiness compounds these concerns, indicating that responsible adoption is as much an institutional and human challenge as a technical one (Djazilan et al., 2024; Rapi et al., 2024). A central tension therefore runs through the literature, between well-supported pedagogical benefits and equally well-documented risks to the formative aims of religious education, and it is this tension that the framework developed below is intended to adjudicate.

Islamic Epistemological and Ethical Framework

The foregoing interpretation indicates that the adoption of AI in Islamic education cannot be governed by efficiency alone, nor by the procedural safeguards of the international frameworks reviewed earlier taken on their own, but requires a framework rooted in the Islamic intellectual tradition. Such a framework is developed here, integrating the tawhidic conception of knowledge, the higher objectives of the Shari'ah, and the cognate principles of *maslahah*, *amanah*, *karamah insaniyyah*, and *adab*, depicted as a layered structure in Figure 2, and then operationalised as a scheme of *maqasid*-compliant design and a recurring governance cycle.

The Tawhidic Paradigm and Naqli–Aqli Integration

The point of departure is *tawhid*, which integrates the transmitted and rational sciences within a single, divinely ordered reality rather than partitioning them into autonomous spheres (Bakar, 1991, 1998). On this view the rational instruments of the age, AI among them, are neither inherently licit nor inherently illicit; their standing is determined by whether they serve, and are disciplined by, the ends of revealed guidance (al-Faruqi, 1982). AI thus occupies the position of an '*aqli*' instrument that may legitimately assist the study of the *naqli* sciences provided it remains subordinate to them and to the qualified human authority that transmits them. This reframes the reconfiguration of the teacher's role noted earlier: the teacher's authority is not rendered obsolete but reaffirmed as the locus of discernment within which AI is used (al-Attas, 1995; Wan Daud, 1998). It also clarifies the response to machine-generated religious content: its subordination to verification by recognised scholarly authority, since the '*aqli*' instrument cannot authenticate the *naqli* sources it processes.

This integration is not a levelling of all knowledge to a single plane. The tradition orders the sciences, placing knowledge of God and of revealed guidance at the summit and arranging the instrumental and empirical sciences beneath it and in its service (Bakar, 1991, 1998). The unity of the sciences is, in Osman Bakar's terms, a unity under *tawhid* rather than a flat equivalence among ways of knowing. The pursuit of knowledge is itself disciplined by *adab*, for the learner owes a right disposition both towards what is studied and towards those from whom it is received (al-Attas, 1995; Wan Daud, 1998). Placed within this order, AI is an '*aqli*' instrument of real value, well able to extend the reach of the rational sciences. Its subordination, though, is normative and not merely practical. It stands below revelation, and below the human authority that transmits and interprets revelation, and no degree of technical sophistication revises that order of precedence.

Maqasid al-Shari'ah as an Evaluative Framework

The higher objectives of the Shari'ah furnish a structured means of evaluating AI in education. These are commonly expressed as the preservation of five fundamental values, namely religion (*din*), life (*nafs*), intellect ('*aql*), lineage (*nasl*), and property (*mal*), and the contemporary

literature has applied them as a systematic instrument for assessing emerging technologies (Mohadi & Tarshany, 2023; Auda, 2008; Kamali, 2008). Reading the four challenge clusters through these objectives, as set out in Table 6, converts general ethical concern into specific guidance: the accuracy of religious content bears on the preservation of religion; the risk of cognitive over-reliance bears on the preservation of intellect; the protection of learner data bears on the preservation of property as a trust; and the protection of minors and of communal bonds bears on the preservation of lineage and life. The *maqasid* thus serve not as an external imposition upon the discussion of educational AI but as a precise analytical instrument for it.

Table 6. The higher objectives of the Shari'ah (maqasid al-shari'ah) applied to AI in Islamic education (framework after Mohadi & Tarshany, 2023; Auda, 2008; Bakar, 1998)

Objective (<i>maqсад</i>)	Application to AI in Islamic education
Preservation of religion (<i>din</i>)	Machine-generated content on creed, law, and worship must be verified against recognised scholarly authority, safeguarding sound religious understanding.
Preservation of intellect (<i>'aql</i>)	Adoption should strengthen reasoning and reflection; cognitive offloading that dulls the intellect is to be resisted.
Preservation of life and dignity (<i>nafs</i>)	Tools should respect human dignity (<i>karamah insaniyyah</i>) and wellbeing and keep the human teacher and learner central.
Preservation of lineage and community (<i>nasl</i>)	Use should protect learners, especially minors, and reinforce rather than fragment the educational community.
Preservation of property and trust (<i>mal</i>)	Learner data, as a trust (<i>amanah</i>), must be handled with consent, confidentiality, and security; resources used responsibly.

ARTIFICIAL INTELLIGENCE IN ISLAMIC EDUCATION: OPPORTUNITIES,
CHALLENGES, AND AN ETHICAL FRAMEWORK GROUNDED IN MAQASID AL-
SHARI'AH

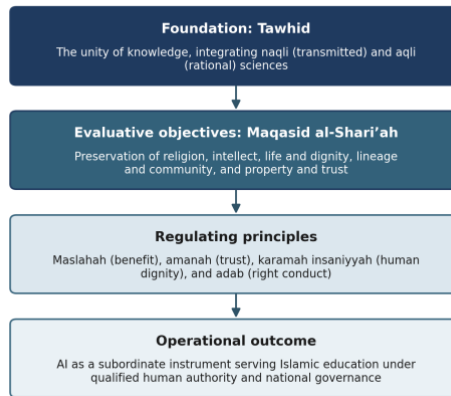


Figure 2. A layered Islamic epistemological and ethical framework for the adoption of AI in Islamic education.

Machine Authority, Sanad, and the Limits of Delegation

The advent of generative AI sharpens a question the tradition has not previously had to confront in this form. Earlier instruments of reason, such as the astrolabe, the algorithm, and the printing press, extended human capacities without claiming to produce knowledge; the generative model, by contrast, produces fluent, authoritative-seeming discourse on matters of creed and law, and thereby appears to occupy the epistemic position that the tradition reserves for the qualified knower. A tawhidic epistemology supplies the resources to locate this capacity correctly. Because knowledge in this view descends from a single divine source and is apprehended through the complementary operation of revelation and reason, the criterion of authority is not the fluency or plausibility of a statement but its grounding in the transmitted sources and its authentication by those qualified to interpret them (Bakar, 1998; al-Faruqi, 1982). Generative fluency, however sophisticated, is not itself a warrant of truth. It is a feature of linguistic form, not evidence of understanding.

A question implicit throughout therefore merits explicit treatment: whether an AI system can exercise religious authority in any meaningful sense. The Islamic tradition locates authority in knowledge transmitted

through qualified persons and authenticated by a chain of transmission (*sanad*), and in the reasoned interpretation (*ijtihad*) of those competent to undertake it. Authority is inseparable from the moral and spiritual qualification of the knower, encompassing not merely the possession of information but its right understanding, the sincerity of intention with which it is held, and accountability before God for its communication (al-Attas, 1995; Wan Daud, 1998). A generative system possesses none of these attributes: it processes text without understanding, lacks intention, and bears no accountability, a limitation that the technical literature on language models expresses in its own terms when it describes such systems as modelling linguistic form in the absence of meaning (Bender et al., 2021). It follows that the outputs of such systems cannot constitute religious authority but at most assist those who possess it, a conclusion consonant with the empirical finding that models should often abstain from religious questions rather than answer them (Atif et al., 2025; Mohd Noor et al., 2025; Tobing et al., 2025).

This has a specific implication for the institution of the *sanad*, the chain of transmission through which religious knowledge is authenticated. The *sanad* is not merely a bibliographic record but a relationship of trust and accountability linking each transmitter to the next and terminating in a living person answerable for what is conveyed (al-Attas, 1995). A generative system can reproduce the content of transmitted reports and even cite chains of narration, but it cannot stand within a *sanad*, for it is neither a witness nor a bearer of the trust that the chain embodies, and it has been shown to fabricate both citations and rulings when pressed on religious questions (Atif et al., 2025). The appropriate role of AI with respect to the *sanad* is therefore strictly instrumental: it may assist in locating, collating, and cross-referencing transmitted material, but the authentication of that material remains an act that only a qualified human authority can perform.

Three further notions sharpen what is at stake. The licence to transmit and to teach is conferred, in this tradition, through *ijazah*, a personal authorisation passed from one qualified bearer of knowledge to another, which certifies not the mere possession of information but the trustworthiness of the person who carries it. Bound up with this is *taklif*, the moral and legal accountability before God that attaches only to responsible human agents. The teacher, at his best, is a *murabbi*, one who

forms character as much as he conveys content. A generative model holds none of these. It has received no *ijazah*; it bears no *taklif*; and it cannot undertake the formation that *tarbiyah* demands. For these reasons it cannot serve as a *mufti*, whose rulings presuppose qualified judgement and answerability before God; it cannot serve as a *murabbi*, whose vocation is the cultivation of the person; and it cannot stand as a transmitter within a *sanad*, every link of which is a trust borne by a living witness. What is left to it is the more modest, and still genuinely useful, place of an instrument, subordinate to the qualified human scholarship that alone can authenticate, judge, and teach.

This conclusion has a constructive corollary. If AI cannot authenticate religious knowledge, it can nonetheless serve the scholar and the learner in defined supporting roles: the retrieval and collation of sources, the detection of recitation errors against established standards, the drafting of preliminary materials for expert revision, and the personalisation of practice. The governing criterion is clear: tasks requiring understanding, judgement, or authority remain with qualified persons, while tasks that are instrumental and verifiable may be delegated to the machine under human oversight. The relationship between the human scholar and the machine is thus best understood as one of supervised instrumentality rather than partnership between equals: the scholar sets the question, constrains the task, and bears final responsibility for the answer, while the system performs delimited, verifiable operations under that authority. This division of labour operationalises the tawhidic subordination of the *'aqli* instrument to the *naqli* sources, furnishes institutions with a workable criterion for distinguishing legitimate from illegitimate uses of AI in religious instruction, and aligns with the international consensus that AI in education should augment rather than replace the human teacher, while supplying the theological reason that consensus leaves unstated (Miao & Cukurova, 2024; Selwyn, 2019).

Maslahah, Amanah, Karamah Insaniyyah, and Adab

Four cognate principles refine the framework. The principle of public interest (*maslahah*) renders the evaluation of AI consequential and contextual: the same tool may be commendable where it deepens understanding and protects the learner, and objectionable where it fosters deception or harm, with intention (*niyyah*) bearing on its moral

quality. The principle of trust (*amanah*) governs the handling of learner data, framing privacy not merely as a legal requirement but as a religious obligation to safeguard what has been entrusted, a framing that gives the privacy and bias concerns raised earlier a specifically Islamic grounding and that converges with the emphasis on data protection in international guidance (Miao & Holmes, 2023; Mehrabi et al., 2021). The principle of human dignity (*karamah insaniyyah*) insists that AI serve rather than diminish the human person, keeping teacher and learner central and resisting their reduction to data points, in consonance with the human-centred orientation of international frameworks (Miao & Cukurova, 2024; Floridi et al., 2018). The principle of *adab*, the just ordering of conduct in accordance with knowledge, governs the disposition with which AI is used, requiring honesty, humility before qualified authority, and the preservation of the contemplative seriousness proper to religious learning (al-Attas, 1995; Wan Daud, 1998).

Towards Maqasid-Compliant AI Design

The framework as presented so far is evaluative; it remains to render it operational. If the *maqasid* are to function not merely as a lens for critique but as a design and governance instrument, each objective must be translated into concrete requirements for how AI systems are selected, configured, and used in Islamic educational settings. This section undertakes that operationalisation, mapping each of the five objectives onto specific requirements and onto the corresponding domains of the international governance frameworks reviewed earlier, as summarised in Table 7.

The preservation of religion (*hifẓ al-din*) translates into a requirement of doctrinal verification. Any machine-generated content bearing on creed, law, or worship must be subject to mandatory validation against recognised scholarly authority before instructional use; systems used for religious instruction should be configured to defer or abstain on questions exceeding their competence rather than to answer speculatively (Atif et al., 2025); and institutions should designate qualified scholars to authenticate religious content. This requirement gives substantive content to the human-oversight provisions common to the UNESCO, OECD, and EU instruments. The preservation of intellect (*hifẓ al-‘aql*) translates into a requirement of cognitive integrity: AI should be deployed to strengthen rather than supplant reasoning, with task designs

that preserve the learner's active engagement and resist the cognitive offloading associated with unguided use (Zhai et al., 2024; Gerlich, 2025), and with AI literacy and critical-verification skills cultivated as core competencies (Long & Magerko, 2020).

The preservation of life and dignity (*hifẓ al-naḥs*) translates into a requirement of human-centred oversight: systems must keep the human teacher and learner central and resist their reduction to data points (*karamah insaniyyah*), meaningful human control must be maintained over consequential decisions, and learner wellbeing must govern deployment. The preservation of lineage and community (*hifẓ al-nas*) translates into protections for the young and for the communal fabric of learning: special safeguards for minors, including the age-appropriate access controls that international guidance recommends (Miao & Holmes, 2023), and attention to whether a given use reinforces or fragments the teacher–student relationship through which formation occurs. The preservation of property and trust (*hifẓ al-māl*) translates into the treatment of data as a trust and the active mitigation of bias: learner data must be handled with consent, confidentiality, and security (*amanah*), and algorithmic bias, including the documented tendency of models trained on non-Islamic corpora to default to non-Muslim assumptions (Naous et al., 2024; Crawford, 2021), must be assessed and mitigated, so that systems neither misappropriate what is entrusted nor systematically misrepresent the tradition.

Table 7. Operationalising the maqasid as requirements for maqasid-compliant AI design and their corresponding AI-governance domains.

Objective (<i>maqasid</i>)	Design and governance requirement	Corresponding governance domain
Hifẓ al-dīn (religion)	Mandatory verification of machine-generated religious content against qualified authority; configured abstention on questions beyond competence.	Human oversight; content reliability
Hifẓ al-ʿaql (intellect)	Task designs that preserve active reasoning and resist cognitive offloading; AI literacy and critical-verification skills as core competencies.	Human oversight; education and capacity
Hifẓ al-naḥs (life and dignity)	Human-centred deployment that keeps teacher and learner central; meaningful human control over consequential decisions.	Human oversight; human-centred values

Objective (<i>maqсад</i>)	Design and governance requirement	Corresponding governance domain
Hifz al-nasl (lineage and community)	Protection of minors through age-appropriate controls; attention to the teacher–student relationship and communal cohesion.	Data protection (minors); safety
Hifz al-mal (property and trust)	Learner data treated as a trust, with consent, confidentiality, and security; assessment and mitigation of algorithmic bias.	Data governance; fairness and bias

Taken together, these requirements convert the *maqсад* from a normative aspiration into a checklist against which specific tools and practices can be assessed, procured, and audited. They also map directly onto the procedural domains of the international governance frameworks, such as human oversight, transparency, data governance, and risk management, supplying those domains with the substantive criteria that determine, in an Islamic educational context, what their satisfaction actually requires. In this way the *maqсад* operate not as a parallel and competing standard but as the substantive complement that the procedural frameworks themselves leave open.

The Maqasid-Compliant AI Governance Cycle

Operationalisation at the level of design must be complemented by a process at the level of governance, since the admissibility of an AI application is not settled once and for all but must be revisited as tools, uses, and evidence evolve. The study therefore proposes a Maqasid-Compliant AI Governance Cycle, depicted in Figure 3, which embeds the *maqсад* evaluation within a recurring institutional process. The cycle proceeds through six stages. First, AI adoption begins with a proposed application, whether a tutoring system, an assessment tool, or a content generator. Second, *maqсад* evaluation assesses the proposal against the five objectives, asking how it bears on the preservation of religion, intellect, life and dignity, lineage and community, and property and trust. Third, ethical screening applies the regulating principles of *maslahah*, *amanah*, *karamah insaniyyah*, and *adab* to the proposal’s purpose, data practices, and likely effects. Fourth, human verification establishes the mechanisms by which qualified persons validate machine-generated religious content and retain oversight of consequential decisions. Fifth,

ARTIFICIAL INTELLIGENCE IN ISLAMIC EDUCATION: OPPORTUNITIES, CHALLENGES, AND AN ETHICAL FRAMEWORK GROUNDED IN MAQASID AL-SHARI'AH

educational implementation deploys the application within these constraints, accompanied by the AI-literacy provision that critical use requires. Sixth, continuous review monitors outcomes, whether pedagogical, doctrinal, or ethical, and feeds them back into adoption decisions, so that tools failing the *maqasid* or the principles are reconfigured or withdrawn and the cycle begins again. The cyclical rather than linear form is essential: it institutionalises the recognition that responsible AI adoption is an ongoing trust (*amanah*) rather than a one-time authorisation, and it gives Islamic institutions a concrete, repeatable procedure for translating the framework into practice.

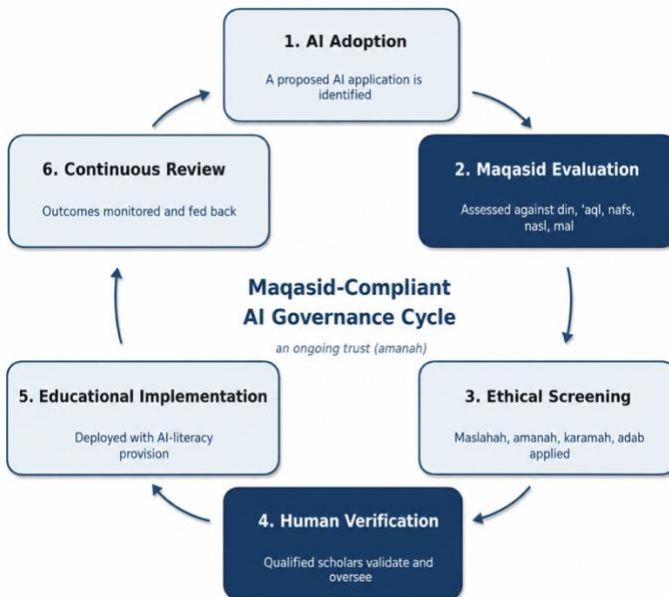


Figure 3. The Maqasid-Compliant AI Governance Cycle for Islamic educational institutions.

Theoretical Contribution

The framework advanced here constitutes the study's principal theoretical contribution, and it advances the existing literature in four specific respects. First, where prior scholarship on AI in Islamic education acknowledges ethical concern in general terms (Amilusholihah

& Ramadhan, 2025; Rosmaini, 2025), this study supplies a structured analytical instrument, the five *maqasid*, that converts diffuse concern into determinate criteria mapped directly onto the documented challenges. Second, where existing *maqasid*-based treatments of technology address AI in the abstract (Mohadi & Tarshany, 2023; Elmahjub, 2023), this study situates the objectives within a prior tawhidic foundation and a subsequent layer of regulating principles, yielding an integrated three-layered structure rather than a single-level checklist. Third, the analysis of machine authority and the *sanad* furnishes a criterion for the limits of delegation that is largely absent from the human-centred frameworks prevalent in the international literature (Miao & Holmes, 2023; Holmes et al., 2022), thereby specifying what those frameworks leave underdetermined: why and how far the human teacher and the transmitted tradition retain primacy. Fourth, by benchmarking the *maqasid* against the principal instruments of global and national AI governance, the study specifies the precise complementarity between an Islamic ethics and the international frameworks: it identifies the substantive normative content that procedural and rights-based regimes leave undetermined, and operationalises that content as a design instrument (the *maqasid*-compliant design requirements set out above) and a governance process (the governance cycle described above). The contribution is therefore not only conceptual but constructive, yielding instruments that institutions and policymakers can apply.

Policy and Practical Implications

Three governing principles connect the framework to practice. First, AI is properly positioned as a subordinate instrument that serves the formative aims of Islamic education and the transmitted tradition rather than displacing the teacher who authenticates them, consistent with the tawhidic integration of '*uqli* and '*naqli* knowledge (Bakar, 1998; al-Faruqi, 1982). Second, the evaluation of AI is governed by intention and benefit, so that the same tool is praiseworthy or blameworthy according to its purpose and effect, in line with *maslahah*. Third, responsible adoption is a shared trust (*amanah*) borne by institutions, educators, and learners, demanding guidelines, training, and oversight rather than either uncritical enthusiasm or wholesale prohibition. These principles accord with, and may be operationalised through, the governance cycle proposed above.

Implementation may be conceived as proceeding through three levels. At the level of governance, institutions require explicit policies translating the *maqasid*-based principles into rules of acceptable use, in particular the mandatory verification of machine-generated religious content (Cotton et al., 2024). At the level of capacity, educators and learners require structured development of AI and digital literacy, so that the critical appraisal of outputs becomes a habitual competence, a need that international competency frameworks and the scholarship on religious educators' competencies have made concrete (Miao & Cukurova, 2024; Miao & Shiohira, 2024; Long & Magerko, 2020; Arif et al., 2025). At the level of design, the development and selection of tools should involve scholars of the Shari'ah alongside technologists. The specific priorities that follow are set out below.

- Develop institutional guidelines defining acceptable use, addressing academic integrity, and mandating verification of machine-generated religious content against authoritative sources, in service of preserving religion (*din*).
- Cultivate AI and digital literacy among educators and learners so that outputs are appraised critically and reflection remains central, in service of preserving intellect (*'aql*).
- Treat learner data as a trust (*amanah*), with explicit practices for consent, confidentiality, and security, and deliberate mitigation of algorithmic bias.
- Reaffirm the teacher's role as facilitator, verifier, and moral exemplar, and prioritise investment in teacher capacity given the documented readiness gap.
- Foster collaboration among technologists, educators, and scholars of the Shari'ah so that tools are designed and governed in accordance with the *maqasid al-shari'ah*.

For Malaysia, these implications are directly actionable within the existing governance architecture. The National Guidelines on AI Governance and Ethics already commit the country to fairness, safety, privacy, transparency, accountability, and human benefit (Ministry of Science, Technology and Innovation, 2024); a *maqasid*-based approach extends rather than contradicts this commitment by supplying the substantive

criteria appropriate to Islamic educational institutions. The National AI Office, as the body coordinating national strategy and capacity-building and developing further ethical guidance (National AI Office, 2024), is well placed to convene scholars of the Shari'ah alongside technologists and educators and to incorporate *maqasid*-aligned guidance for Islamic institutions into its evolving frameworks. The governance cycle proposed above offers such institutions a concrete process for translating national principles into religiously grounded practice, and offers comparable Muslim-majority jurisdictions a transferable model.

Limitations and Future Research Agenda

The study is bounded by its design. As a narrative review and interpretive synthesis of published scholarship, it generates no new empirical data, and its selection of sources, while guided by explicit considerations of relevance and quality, reflects scholarly judgement rather than an exhaustive protocol; the field it surveys, moreover, remains dominated by conceptual and small-scale studies with limited representation of Malaysian higher education (Zawacki-Richter et al., 2019; Faizin et al., 2025). Its reliance on three databases and on sources in English and Malay may also have excluded relevant work in Arabic and other languages. The framework it advances, including the *maqasid*-compliant design scheme and the governance cycle, is conceptually derived and awaits empirical validation in practice.

These limitations define a research agenda that is at once empirical and constructive. First, the field's reliance on conceptual and small-scale work calls for larger empirical investigations of AI use in Islamic educational institutions, including controlled studies of learning outcomes, of teacher and learner perceptions, and of the doctrinal accuracy of machine-generated religious content (Faizin et al., 2025; Ridwan, 2025). Second, the centrality of critical verification identified above makes the development and evaluation of AI-literacy interventions for religious educators and learners a priority, including measures of whether such interventions improve the appraisal of machine-generated content (Long & Magerko, 2020; Arif et al., 2025). Third, the technical foundations merit sustained attention: the application of natural language processing to the Qur'an and its sciences raises distinctive challenges of orthography, morphology, and sacred status that remain only partially solved (Al-Shidi et al., 2025; Bashir et al., 2023; Malhas & Elsayed, 2022).

Fourth, and relatedly, the construction and evaluation of Arabic and Islamic large language models, including the measurement and mitigation of the cultural and religious bias documented in general models, is a research frontier with direct bearing on the reliability of AI in religious education; the development of explainable models, whose grounds can be inspected and contested, is especially pertinent where doctrinal accuracy is at stake (Naous et al., 2024). Fifth, the emergence of systems that generate religious rulings, sometimes described as AI fatwa or mufti systems, demands rigorous study of their reliability, their patterns of error and fabrication, and the conditions, if any, under which their output may responsibly assist qualified scholarship, including the vexed question of how, if at all, such systems might relate to the chain of transmission (*sanad*) (Atif et al., 2025). Sixth, comparative studies across Muslim-majority jurisdictions would test the portability of the framework proposed here. A study setting the Malaysian, Indonesian, Bruneian, Saudi, and Turkish approaches to governing AI in Islamic education alongside one another, for instance, would show how it fares under differing institutional and regulatory conditions, and where it would need adaptation. Pursuing this agenda would allow the framework to be refined and validated, and would allow Islamic education to engage artificial intelligence both confidently and critically.

CONCLUSION

This study examined how artificial intelligence is applied in Islamic education, evaluated the accompanying challenges, and developed an Islamic epistemological and ethical framework to guide responsible integration. In answer to the first research question, AI is applied across five domains: personalised and adaptive learning, intelligent tutoring and conversational agents for Qur'anic and Arabic instruction, automated assessment, language and manuscript processing, and administration, with reported gains in access, motivation, and engagement consistent with meta-analytic evidence of medium-to-large cognitive effects (Wang et al., 2024; Steenbergen-Hu & Cooper, 2014). In answer to the second, four challenges recur with particular force: academic integrity and cognitive over-reliance, data privacy and algorithmic bias, the reconfiguration of teacher authority, and the doctrinal accuracy of machine-generated content, each now supported by a maturing empirical

literature (Gerlich, 2025; Naous et al., 2024; Atif et al., 2025). In answer to the third, the study advanced a three-layered framework uniting the tawhidic integration of *naqli* and *aqli* knowledge, the *maqasid al-shari'ah*, and the principles of *maslahah*, *amanah*, *karamah insaniyyah*, and *adab*, and showed, through a comparison with the principal instruments of global and national AI governance, that this framework supplies the substantive ends those instruments leave undetermined.

The study makes four contributions. Theoretically, it integrates a fragmented literature, bringing the empirical evidence on effectiveness, the critical scholarship on risk, and the normative resources of the Islamic tradition into a single account. Conceptually, it articulates an explicit, layered framework, depicted in Figure 2, and a criterion for the limits of delegation derived from the analysis of machine authority and the *sanad*. Comparatively, it specifies the complementarity between the *maqasid* and the UNESCO, OECD, EU, and Malaysian governance regimes. Practically, it operationalises the framework as a *maqasid*-compliant approach to AI design and as the Maqasid-Compliant AI Governance Cycle, translating it into an institutional process connected to the policy context of Malaysia. Table 8 consolidates the argument by aligning each principal challenge with its framework-based response and governing objective.

Table 8. Alignment of principal challenges with framework-based responses and governing objectives.

Challenge	Framework-based response	Governing objective
Academic integrity and over-reliance	Guided use, verification, and cultivation of discernment; AI as supplement to reflection.	Intellect (<i>'aql</i>); <i>adab</i>
Data privacy and algorithmic bias	Treatment of data as a trust; consent, confidentiality, security, and bias mitigation.	Property (<i>mal</i>); <i>amanah</i>
Teacher authority and transmission	Reaffirmation of the teacher as verifier and exemplar; AI subordinate to human authority.	Tawhidic order; <i>karamah insaniyyah</i>
Doctrinal accuracy of content	Mandatory verification of religious content against recognised scholarly authority.	Religion (<i>din</i>)

Pursued in this spirit, the integration of artificial intelligence need not compromise the integrity of Islamic education. Governed by the *maqasid* and disciplined by qualified human authority, AI can serve as a

subordinate instrument that supports access, personalisation, and feedback while leaving intact the formative purpose at the heart of the tradition: the cultivation of knowledgeable, reflective, and upright learners. The framework, the design scheme, and the governance cycle advanced here are offered as a starting point for that endeavour, to be refined through the empirical and comparative research the field now requires.

REFERENCES

- Ab Alim, N. N. A., Wahid, K., Abd Zamani, S. N. H., & Adenan, F. (2025). Artificial intelligence as a support tool for Quran memorization: An exploratory study on the potential use of AI applications among non-Islamic studies students. *Quranica*, 17(2), 157–186. <https://doi.org/10.22452/quranica.vol17no2.7>
- Abdelgelil, M. F. M., Tahir, B. S., & Bidin, S. N. B. S. (2025). Challenges of artificial intelligence (AI) and its impact on Qur'anic qira'at: A thematic review analysis. *Quranica*, 17(2), 74–96. <https://doi.org/10.22452/quranica.vol17no2.4>
- Abubakari, M. S., Shafik, W., & Hidayatullah, A. F. (2024). Evaluating the potential of artificial intelligence in Islamic religious education: A SWOT analysis overview. In *AI-enhanced teaching methods* (pp. 216–239). IGI Global. <https://doi.org/10.4018/979-8-3693-2728-9.ch010>
- Al-Attas, S. M. N. (1995). *Prolegomena to the metaphysics of Islam: An exposition of the fundamental elements of the worldview of Islam*. International Institute of Islamic Thought and Civilization (ISTAC).
- Al-Faruqi, I. R. (1982). *Islamization of knowledge: General principles and work plan*. International Institute of Islamic Thought.
- Al-Shidi, A., Nordin, F. H., Mohamed, I. I., Younis, M. A., Elmutasim, I. E., & Abusham, E. E. (2025). A literature review on natural language processing techniques for Qur'anic studies: Challenges

- and insights. *Frontiers in Signal Processing*, 5, 1535166.
<https://doi.org/10.3389/frsip.2025.1535166>
- Amilusholihah, A., & Ramadhan, N. J. H. (2025). Exploring the implementation of artificial intelligence in Islamic education: A systematic literature review. *Ngaji: Jurnal Pendidikan Islam*, 5(1), 3–17. <https://doi.org/10.24260/ngaji.v5i1.95>
- An, A. N., Tamami, F. Q. A., Daud, Z., Salleh, N. M., Ishak, M. H., & Muthoifin, M. (2025). Understanding the integration of deep learning and artificial intelligence in Quranic education and research through bibliometric analysis. *Educational Process: International Journal*, 14, e2025012. <https://doi.org/10.22521/edupij.2025.14.12>
- Anandal, F., Dartim, D., & Yusutria, Y. (2024). The role of artificial intelligence in Islamic education for inclusive children. *Iseedu: Journal of Islamic Educational Thoughts and Practices*, 8(1), 41–46. <https://doi.org/10.23917/iseedu.v8i1.8841>
- Arif, M., Aziz, M. K. N. A., & Ma'arif, M. A. (2025). Islamic religious education teachers' competencies in the digital age: A systematic literature review. *Journal of Education and Learning (EduLearn)*, 19(2), 587–596. <https://doi.org/10.11591/edulearn.v19i2.21311>
- Atif, F., Askarbekuly, N., Darwish, K., & Choudhury, M. (2025). Sacred or synthetic? Evaluating LLM reliability and abstention for religious questions. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(1), 217–226. <https://doi.org/10.1609/aies.v8i1.36543>
- Auda, J. (2008). *Maqasid al-Shariah as philosophy of Islamic law: A systems approach*. International Institute of Islamic Thought.
- Bakar, O. (1991). *Tamhid and science: Essays on the history and philosophy of Islamic science*. Secretariat for Islamic Philosophy and Science.
- Bakar, O. (1998). *Classification of knowledge in Islam: A study in Islamic philosophies of science*. Islamic Texts Society.

- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32(4), 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Bashir, M. H., Azmi, A. M., Nawaz, H., Zaghouani, W., Diab, M., Al-Fuqaha, A., & Qadir, J. (2023). Arabic natural language processing for Qur'anic research: A systematic review. *Artificial Intelligence Review*, 56(7), 6801–6854. <https://doi.org/10.1007/s10462-022-10313-2>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). <https://doi.org/10.1145/3442188.3445922>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- 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). <https://doi.org/10.62775/edukasia.v5i2.1373>
- du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalised adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and

- engagement. *Heliyon*, 10(21), e39630.
<https://doi.org/10.1016/j.heliyon.2024.e39630>
- Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology*, 36(4), 73. <https://doi.org/10.1007/s13347-023-00668-x>
- European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union.
<http://data.europa.eu/eli/reg/2024/1689/oj>
- Faizin, N., Alfian, 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>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474.
<https://doi.org/10.1080/14703297.2023.2195846>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6.
<https://doi.org/10.3390/soc15010006>
- Guilherme, A. (2019). AI and education: The importance of teacher and student relations. *AI & Society*, 34(1), 47–54.
<https://doi.org/10.1007/s00146-017-0693-8>

- Hady, Y., Krisnawati, N. M., & Fatima, A. (2025). Artificial intelligence (AI), the erosion of tacit knowledge, and the challenges of cultivating wisdom: Epistemological and ethical implications for 21st-century Islamic education. *EDUKASI: Jurnal Penelitian Pendidikan Agama dan Keagamaan*, 23(2), 430–446. <https://doi.org/10.32729/edukasi.v23i2.2407>
- Hng, A. (2025). *Malaysia: National AI roadmap 2021–2025 and AI governance framework*. Regulations.ai. <https://regulations.ai/regulations/malaysia>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Idham, I., Rokimin, R., Akhirudin, M. T., Rahman ZH, M. H., & At Thariq, Z. Z. (2026). The application of deep learning in enhancing the effectiveness of Islamic religious education curriculum in madrasahs. *Management of Education: Jurnal Manajemen Pendidikan Islam*, 12(1), 34–48. <https://doi.org/10.18592/moe.v12i1.20058>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kamali, M. H. (2008). *Maqasid al-Shari'ah made simple*. International Institute of Islamic Thought.
- 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, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task* [Preprint]. arXiv.
<https://doi.org/10.48550/arXiv.2506.08872>
- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16).
<https://doi.org/10.1145/3313831.3376727>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918.
<https://doi.org/10.1037/a0037123>
- Malhas, R., & Elsayed, T. (2022). Arabic machine reading comprehension on the Holy Qur'an using CL-AraBERT. *Information Processing & Management*, 59(6), 103068.
<https://doi.org/10.1016/j.ipm.2022.103068>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.
<https://doi.org/10.1145/3457607>

- Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. <https://www.unesco.org/en/articles/ai-competency-framework-teachers>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. UNESCO.
- Miao, F., & Shiohira, K. (2024). *AI competency framework for students*. UNESCO. <https://www.unesco.org/en/articles/ai-competency-framework-students>
- Ministry of Science, Technology and Innovation. (2021). *Malaysia national artificial intelligence roadmap 2021–2025 (AI-RMAP)*. MOSTI.
- Ministry of Science, Technology and Innovation. (2024). *National guidelines on AI governance and ethics (AIGE)*. MOSTI.
- Mohadi, M., & Tarshany, Y. (2023). Maqasid al-Shari'ah and the ethics of artificial intelligence: Contemporary challenges. *Journal of Contemporary Maqasid Studies*, 2(2), 79–102. <https://doi.org/10.52100/jcms.v2i2.107>
- Mohd Noor, M. A. A., Wan Ramli, W. N. A., & Abdul Rahman, N. N. (2025). Artificial intelligence (AI) in Quranic education: A systematic review of pedagogical applications and challenges. *QURANICA: International Journal of Quranic Research*, 17(1). <https://ejournal.um.edu.my/index.php/quranica/article/view/64955>
- Naous, T., Ryan, M. J., Ritter, A., & Xu, W. (2024). Having beer after prayer? Measuring cultural bias in large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 16366–16393). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.acl-long.862>
- Nasr, S. H. (1993). *The need for a sacred science*. State University of New York Press.

- National AI Office. (2024). *About the National AI Office (NAIO)*. Ministry of Digital, Malaysia. <https://ai.gov.my/about-naio/>
- OECD. (2024). *Recommendation of the Council on artificial intelligence* (OECD/LEGAL/0449). Organisation for Economic Co-operation and Development. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- Rapi, M., Rusdi, M., & Idris, R. (2024). Challenges and opportunities of artificial intelligence adoption in Islamic education in Indonesian higher education institutions. *International Journal of Learning, Teaching and Educational Research*, 23(11), 423–443. <https://doi.org/10.26803/ijlter.23.11.22>
- Ridwan, R. (2025). Assessing the impact of AI-driven tools on Qur'anic literacy: A quasi-experimental study in Islamic higher education in Indonesia. *Journal of Cultural Analysis and Social Change*, 10(3), 252–259. <https://doi.org/10.64753/jcasc.v10i3.2405>
- Rosmaini. (2025). The transformation of Islamic education in the era of artificial intelligence (AI): Opportunities, challenges, and ethics of its use. *Teunuleh Scientific Journal*, 6(3), 333–347. <https://doi.org/10.51612/teunuleh.v6i3.228>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Sabzalieva, E., & Valentini, A. (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. UNESCO.
- Sagr, M., López-Pernas, S., Conde, M. A., Pavlovic, O., & Milic, M. R. (2024). The critical challenges of artificial intelligence in education. In *Proceedings of the 14th International Conference on e-Learning* (pp. 3–9).
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.
- Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students'

- academic learning. *Journal of Educational Psychology*, 106(2), 331–347. <https://doi.org/10.1037/a0034752>
- Tobing, M. A., Abrianto, D., & Harahap, M. Y. (2025). Reconstruction of Islamic religious education in the artificial intelligence era: A critical review of teacher roles, digital ethics, and epistemic authority. *Proceedings of the International Conference on Islamic Community Studies*, 7186–7192. <https://proceeding.pancabudi.ac.id/index.php/ICIE/article/view/1343>
- UNESCO. (2022). *Recommendation on the ethics of artificial intelligence*. UNESCO.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Wan Daud, W. M. N. (1998). *The educational philosophy and practice of Syed Muhammad Naquib al-Attas: An exposition of the original concept of Islamization*. International Institute of Islamic Thought and Civilization (ISTAC).
- Wang, X., Huang, R., Sommer, M., Pei, B., Shidfar, P., Rehman, M. S., Ritzhaupt, A. D., & Martin, F. (2024). The efficacy of artificial intelligence-enabled adaptive learning systems from 2010 to 2022 on learner outcomes: A meta-analysis. *Educational Technology Research and Development*, 72(3), 451–478. <https://doi.org/10.1007/s11423-024-10350-4>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995>
- Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: New demands for teachers' roles. *Heliyon*, 10(2), e24289. <https://doi.org/10.1016/j.heliyon.2024.e24289>

- 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, 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11, 28. <https://doi.org/10.1186/s40561-024-00316-7>