

Generative AI, The Authority of Religious Knowledge, and The Reorientation of Islamic Religious Education Teachers: An Islamic Epistemological Perspective

Ubaidillah

Sekolah Tinggi Agama Islam Ihyaul Ulum, Gresik, Indonesia

E-mail: obeidbahrums@gmail.com

Moh. Irfan Sulthoni

Sekolah Tinggi Agama Islam Ihyaul Ulum, Gresik, Indonesia

E-mail: irfansulthoni461@gmail.com

Abstract

The acceleration of Generative Artificial Intelligence (Generative AI) based on Large Language Models (LLM) has shifted the search for religious knowledge from the turats and talaqqī traditions towards instant algorithmic synthesis. This study aims to critically-philosophically analyze the epistemic limits of religious output generated by AI, map shifts in religious authority, and formulate a pedagogical reorientation of Islamic Religious Education (IRE) Teachers. This qualitative research uses a philosophical - analytical research library with content analysis techniques and philosophical hermeneutics. Empirical Data on the performance of ChatGPT, Claude, and Gemini were extracted from IslamicLegalBench, FiqhQA, and QIAS 2025 Shared Task, then diagnosed using tripartite Islamic epistemology (Bayani, Burhani, Irfani). The results showed three layers of limitations that are structural, not technical. First, in Bayani's reasoning, AI excels at compiling textual data and mapping fatwas, but is devoid of isnād, naqd al-isnād, and Shari'a responsibilities; the hallucination rate of 55% in some models and 21% in the best models, is not a statistical anomaly, but rather a manifestation of the absence of a living transmission relationship. Second, in Irfan's reasoning, AI experiences a total epistemic void: decreased performance in Arabic indicates AI is working through statistical patterns, not understanding the meaning of life, so it is not able to touch the tazkiyatun nafs, khasyyah, transmission, and barakah that require the presence of a conscious qalb and direct encounter (talaqqī). Third, on Burhani's reasoning, AI is able to construct comparative analyses of the 'illat law and sociology schools, but operates probabilistically without awareness of 'aql; sycophancy above 40% on 6 of the 9 models as well as the FiqhQA paradox GPT-4.0 excels in accuracy but Gemini is better at abstaining, proving the AI's output is only a crude hypothesis, not aql. Consequently, religious authority shifted from relational transmission to computational visibility giving rise to pseudo-authority. This study concluded that AI is based solely as a wasīlah (Ta'līm instrument), not a ghāyah or religious authority holder. Therefore, the role of the

PAI teacher must be reoriented from mudarris to mu'addib and murabbi through Digital Tabayyun, planting Adab al-'Ilm, and moral example (uswah hasanah).

Keywords: generative AI; authority of religious knowledge; Islamic epistemology; reorientation of Islamic religious education (PAI) teachers; digital tabayyun.

Introduction

Generative Artificial Intelligence (Generative AI) technology based on *Large Language Models* (LLM) has triggered a fundamental shift in the global education ecosystem. In classrooms and Islamic educational institutions, this technology is present not only as a data search instrument, but has transformed into an entity capable of responding to theological questions, citing fiqh texts, and constructing ethical reasoning instantly, communicatively, and persuasively. Changes in student learning patterns that increasingly rely on ready-made algorithmic answers are beginning to shift the conventional tradition of *thalabul 'ilmi* (*knowledge acquisition*), which historically relied on the study of *turats* (classical texts) and the direct guidance of an educator. In this context, Generative AI is no longer merely a matter of learning technology, but is beginning to touch on how religious knowledge is acquired, verified, and authorized. (Mahbubi, 2025; Shobri & Alfurqan, 2026) .

Viewed from the perspective of Islamic Philosophy of Education, this phenomenon gives rise to crucial epistemological tensions. In the Islamic intellectual tradition, knowledge (*'ilm*) is not merely the accumulation of cognitive data or the statistical synthesis of texts, but rather relates to truth, *etiquette*, and the process of human formation that requires verification (*tabayyun*) and the transmission of knowledge through sanad (*isnad*) and *talaqqi* (Ajwath, 2026; Al-Attas, 1991; Saidi et al., 2025) . Authoritative knowledge requires a traditional foundation, orthodoxy, and vertical transmission from the Source (of knowledge) (Nasr, 1989) . In contrast, Generative AI works through probabilistic language modeling and lacks the epistemic experience, intentions, or moral-spiritual responsibilities that humans possess (Mahbubi, 2025; Putri1 et al., 2026) . The inability of generative systems to guarantee the validity of sources and sanad, coupled with the possibility of producing erroneous or *hallucinated information*, raises questions about the limits of AI's authority when used to acquire and explain religious knowledge. This issue becomes increasingly important because answers delivered fluently and persuasively can create an authoritative impression even though the epistemic process differs from Islamic scientific traditions (Putri1 et al., 2026; Rhisky & Baharun, 2026) .

Numerous studies have discussed the use of Generative AI in education, including its potential to support learning, provide information, personalize learning, and change the role of educators (Amin, 2026; Yun, 2025) . Other studies have begun to highlight issues of accuracy, *hallucinations* , ethics, and students' dependence on AI output (Desitha Cahya et al., 2026; Qian et al., 2026) . In the context of Islamic education, studies on AI have also addressed issues of digital literacy, technology integration in learning, and the need to maintain ethical and spiritual values. However, discussions of Generative AI as a question of the authority of religious knowledge still require philosophical in-depth analysis, particularly when comparing the generative capabilities of AI with the structure of Islamic epistemology, which has different mechanisms for determining the source, validity, rationality, and spiritual experience of knowledge (Shalih, 2026) .

This gap can be seen from the lack of studies that systematically map the capabilities and limitations of Generative AI through three Islamic epistemological patterns (*Bayani*, *Irfani*, and *Burhani*), especially those formulated comprehensively by Muhammad Abed al-Jabiri. *Bayani* reasoning is a knowledge system that places text (revelation) as the source of truth, *Irfani* reasoning is an epistemological structure that bases the acquisition of knowledge on inner experience (spiritual), intuition (*kashf*), and purification of the soul (*dzaug*), not on text or reason alone, while *Burhani* reasoning is a rational-demonstrative method of thinking based on logic, the law of causality, and empirical evidence. The three provide different perspectives in understanding how knowledge is formed and validated: *Bayani* emphasizes the authority of texts, language, and the transmission of knowledge; *Irfani* is concerned with the inner dimension, spiritual experience, etiquette, and purification of the soul; while *Burhani* emphasizes rationality, argumentation, and proof (Al-Jabiri, 1991). These three perspectives are not positioned as literal characteristics of AI, but as analytical lenses to test the epistemic limits of Generative AI output (Velly et al., 2025). With this framework, the issues studied do not stop at whether AI is capable of providing textually correct answers, but also how those answers obtain validity and whether that validity aligns with Islamic epistemological principles.

These epistemic issues ultimately have direct implications for the position of Islamic Religious Education (PAI) teachers (Ali & Alistiana, 2026; Azizah et al., 2026). If the function of conveying information (*ta'lim*) is increasingly easily replicated by generative systems, then the professional legitimacy of teachers no longer lies primarily in their ability to provide information that can be found through machines (Ubaidah & Haris, 2026). The role of teachers needs to move towards a more reflective, dialogical, ethical, and transformative function as *mu'addib* and *murabbi*: forming adab, guiding the process of knowledge verification, developing *critical religious literacy skills*, and providing exemplary behavior (*uswah hasanah*) that cannot be reduced to algorithmic output (Al-Attas, 1991; al-Attas, 2014; Rhisky & Baharun, 2026). In this context, the concept of *Digital Tabayyun* becomes important as a pedagogical orientation to guide students not to accept AI output directly, but to examine the source, context, argumentation, and authority of the underlying knowledge (Sidik et al., 2026). This reorientation aligns with the discussion in this study regarding the shift in the teacher's role from *mudarris* to *mu'addib* and *murabbi*.

Based on these issues, this study aims to analyze the epistemic limits of Generative AI in the production and delivery of religious knowledge through the perspectives of *Bayani*, *Burhani*, and *Irfani*, and examine its implications for changes in the authority of religious knowledge and the role of Islamic Religious Education teachers. This study specifically asks: how the output of Generative AI can be understood within the epistemological framework of *Bayani*, *Irfani*, and *Burhani*; how the use of AI has the potential to change the pattern of authority, *isnad*, and *talaqqi* in religious learning; and how Islamic Religious Education teachers can be reoriented from the function of knowledge transmission to the function of *ta'dib* and *tarbiyah* in the AI era. Thus, the contribution of this study lies in developing a philosophical reading of Generative AI that does not solely assess technology based on accuracy or efficiency, but based on how religious knowledge obtains its epistemic legitimacy and its consequences for the pedagogical transformation of Islamic Religious Education teachers.

Methods

This study uses a qualitative approach with a philosophical-analytical library research method (Creswell et al., 2007). This approach was chosen because the main focus of the study is to epistemologically diagnose the nature and limits of religious output produced by three *Generative AI platforms* (ChatGPT, Claude, and Gemini) as documented in previous empirical studies, such as *IslamicLegalBench*, *FiqhQA*, and *QIAS 2025 Shared Taks*. Primary data sources in this study include works on classical and modern Islamic epistemology (*Al-Jabiri*, *Al-Attas*, *Nasr*) as well as empirical research reports that evaluate the performance of the three AI platforms in the Islamic domain. Secondary data sources include indexed journal articles, supplementary books, and regulatory documents relevant to the issue of AI in Islamic education, digital religious authority, and the redefinition of the role of Islamic education teachers. Data collection was carried out through documentation techniques and literature studies with four systematic stages: literature inventory through academic databases (Google Scholar, JSTOR, SINTA, DOAJ), literature categorization into three analytical focuses (AI epistemic limitations, Islamic epistemology and scientific *sanad*, and Islamic Religious Education teachers' pedagogical reorientation), selection and evaluation based on validity and novelty, and extraction of empirical findings on the output of ChatGPT, Claude, and Gemini from existing studies.

Data analysis was done using a combination of content analysis techniques and philosophical hermeneutics. The analysis procedure is carried out interactively through four steps: First, Data condensation, which encodes and groups conceptual findings and empirical findings about the output of the three AI platforms into sub-focuses of discussion; second, epistemic criticism (epistemic critique), which analyzes the empirical findings using the tripartite scalpel of Islamic epistemology (*Bayani*, *Irfani*, and *Burhani*) to uncover why the identified limitations are structural, not technical; third, the presentation of data in the form of descriptive-analytical narratives, comparison matrices, and thought schemes; and fourth, the drawing of philosophical conclusions about the limits of the validity of AI and recommendations for pedagogical reorientation of Pai teachers as *Mu'addib*. In order to guarantee academic quality and weight, the validity of the data is tested through theoretical triangulation and hermeneutical audits, that is, repeated critical examinations of textuality, thought context and philosophical inference in order to avoid subjective deviations of the researcher

Results

A. Performance Data of ChatGPT, Claude, and Gemini in the Islamic Religious Domain

1. IslamicLegalBench (718 Issues, 7 Schools of Law, 9 LLM Models)

IslamicLegalBench is the first benchmark to evaluate the ability of Large Language Models (LLMs) to understand Islamic law across seven schools of thought (Sunni and Shia) across 718 problems covering 13 tasks of varying complexity. The evaluation of nine *state-of-the-art models* revealed that the best-performing model achieved only a 68% correctness rate with a 21% hallucination rate, while several other models achieved below 35% correctness and above 55% hallucination rates. The application of *few-shot prompting* also showed relatively limited benefits, with only two of the nine models achieving performance improvements of more than 1%. On moderate-complexity tasks demanding accurate knowledge, the models

demonstrated the highest error rates, while on high-complexity tasks, indications of pseudo-competence were found through semantic reasoning abilities. Furthermore, testing for false premises revealed a risky tendency *for sycophancy*, with six of the nine models accepting misleading assumptions at a rate of more than 40%. Overall, the evaluation also showed a performance gap between *the closed-source* and *open-source models*, with a difference of 21.73 percentage points in correctness rate and 20.15 percentage points in hallucination rate (Elmahjub et al., 2026a).

2. FiqhQA (960 Question-Answer Pairs, 4 Sunni Schools, 6 LLM Models)

FiqhQA is the first benchmark to explicitly categorize Islamic fatwas based on the four main Sunni schools of thought (Maliki, Shafi'i, Hanafi, and Hanbali) in both Arabic and English, with a particular focus on *abstention*, the model's ability to recognize situations when a question should not be answered directly. The dataset consists of 960 question-answer pairs from 120 unique questions across two languages and four schools of thought. The evaluation results show that GPT-4.0 achieved the highest accuracy compared to other models, while Gemini and Fanar demonstrated better *abstention*, which is crucial for reducing the risk of providing incorrect answers with a high degree of confidence. Furthermore, all models experienced a decrease in performance when faced with questions in Arabic compared to English, indicating limitations in the LLM's reasoning and religious understanding capabilities in non-English languages (Atif et al., 2025).

3. QIAS 2025 Shared Task (Insights into ChatGPT and Gemini)

Studies at QIAS 2025 provided additional data on the performance of generative models, specifically GPT-4.0 and Gemini, on Arabic Islamic law tasks. On a *multiple-choice questions* (MCQ) task on classical Islamic knowledge, GPT-4.0 achieved 85% accuracy in a *zero-shot model* and slightly improved to 86% in a *five-shot model*. Gemini 2.5, meanwhile, showed a more substantial improvement, from 59% in a *zero-shot model* to 87% in a *few-shot model*, indicating that providing examples can help the model understand the task format and instructions. On a task related to Islamic inheritance law, o3 achieved the highest accuracy of 93.4%, followed by Gemini at 90.6%. On a more rigorous MCQ evaluation of Islamic inheritance, Gemini Flash 2.5 achieved 91.5% accuracy in its initial testing. These findings indicate that LLM performance on Islamic law tasks can differ significantly depending on the type of task, model, and *prompting strategy* used (Bouчекif et al., 2025).

4. Specific Data about Claude

Claude's specific data in the Islamic religious domain is more limited than that of GPT and Gemini. However, in the Islamic heritage task (QIAS 2025), Claude 3.7 achieved a score of 0.67 without *thinking* and 0.81 with *thinking* on the development dataset, while it achieved 0.71 on the test dataset. The 21% improvement with *thinking* indicates that structured reasoning substantially improves Claude's abilities, but its baseline score remains lower than that of GPT-4.0 and Gemini on the same task.

B. Epistemological Diagnosis Based on Empirical Data

1. Bayani's Reasoning: The Absence of *Isnad* and the Pattern of Structural Hallucinations

Empirical data from IslamicLegalBench and FiqhQA provide quantitative confirmation of AI *output*. Hallucination rates reaching 55% in some models are not statistical anomalies, but rather a manifestation of the absence of *isnad* in the LLM architecture. In *Bayani reasoning*, *isnad* plays a vital role in determining whether knowledge is directly connected to the bearer of truth, namely the Prophet Muhammad. AI can list hadith numbers and narrators' names, but

lacks a genealogy of scholarly transmission linking the speaker to an authoritative source (Elmahjub et al., 2026a) .

The finding that moderate-complexity tasks requiring accurate knowledge actually exhibit the highest errors strengthens the argument that AI operates at the level of information mapping (*ta'lim*), not epistemically binding knowledge (*'ilm syar'i*). When AI is required to remember precisely (something that in Islamic tradition requires *talaqqi* and *mudhakarah*), it consistently fails.

2. Irfani's Reasoning: An Irreparable Void

Empirical data cannot measure the intuitive dimension (*kasyf*) and purification of the soul (*tazkiyatun nafs*) which is the basis of *Irfani reasoning*, but it can show structural limitations that explain why the *Irfani dimension* is closed to AI. Consistent performance declines indicate that AI operates through statistical patterns that depend on training data, not through a living understanding of meaning (Boucekif et al., 2025) . If AI even struggles with *fiqh al-'ibadat* (ritual worship) then it clearly does not have the capacity to touch the dimension of *tazkiyatun nafs* which requires the presence of a conscious *heart*.

3. Burhani Reason: Probabilistic Without *Tarjih Authority*

Data from FiqhQA shows a very instructive pattern: GPT-4.0 excels in accuracy, but Gemini is better at *abstaining*. This means the AI is capable of producing textually “correct” answers, but lacks the awareness to know when it should not answer, something considered half of knowledge in Islamic tradition (Atif et al., 2025a) .

The finding of *sycophancy*, where 6 out of 9 models accepted the misleading premise at a rate above 40%, indicates that the AI lacks reflective *'aql* awareness. The AI constructs the logic of *ushul fiqh*, but does not internalize the *Maqasid ash-Syari'ah*; it can compare schools of thought, but does not have the authority to guide according to *shar'i*.

C. Empirical Data Recapitulation Matrix

Table 1. Empirical Data Recapitulation

Studies	Platform	Task	Key Metrics
IslamicLegalBench	9 LLMs (including GPT, Claude, Gemini)	718 Problems, 7 Schools of Thought	Best: 68% true, 21% hallucination Worst: <35% true, >55% hallucination
FiqhQA	GPT-4.0, Gemini, etc.	960 QA, 4 schools of thought	GPT-4.0: highest accuracy Gemini: best abstention
QIAS Inheritance	GPT-4.0, Gemini, Claude	Heritage MCQ	GPT-4.0: 93.4%; Gemini: 90.6% Claude 3.7: 0.71 (test)
QIAS Classic MCQ	GPT-4.0, Gemini	1,400 MCQs	GPT-4.0: 85-86%; Gemini: 59-87%

Discussion

A. The 55% Hallucination Rate Is Not a Technical Failure, But a Consequence of the Desacralization of Knowledge and the Absence of Isnad

Empirical data from *IslamicLegalBench* shows that some LLM models achieve a hallucination rate exceeding 55% in the domain of Islamic law, while even the best models are limited to a 68% correctness rate with 21% hallucination (Elmahjub et al., 2026b) . In the

epistemological framework of Al-Jabrī (Al-Jabri, 1991) and Nasr (Nasr, 1989) , this figure is not simply a "technical error" that can be eliminated through increased computational capacity or *fine-tuning* . It is a structural consequence of the process of desacralization of knowledge, where knowledge is separated from its sacred origins, existence, and divine truth and disconnected from the system of transmission chains (*isnād*).

More deeply, when viewed through the framework of Islamic educational epistemology developed by Al-Attas (Al-Attas, 1991; al-Attas, 2014) , knowledge (*'ilm*) is defined as the arrival of the meaning of something into the soul (*husūl*) or the arrival of the soul to the essence of meaning (*wuṣūl*). AI algorithms work through statistical processing and manipulation of external linguistic tokens (*bayān zāhirī*). Meanwhile, in the Islamic conception, the authority of knowledge is only held by certain people who meet strict criteria and qualifications (al-Attas, 2014) . Therefore, AI output never undergoes the process of *husūl* in the soul that is morally and spiritually conscious. When ChatGPT or Claude includes references to classical texts, it merely replicates the pattern of digital strings without carrying out narrator criticism (*naqd al-isnād*), without verifying *'adālah* and *ḍabt*, without assuming ontological and *shar'ī* responsibility for the resulting legal inferences.

The finding that " moderate complexity tasks " actually triggered the highest error rates confirms that increasing model parameters is unable to penetrate the boundaries of the nature of science. *Prompt engineering* has proven unsuccessful in addressing this issue (only 2 of 9 models improved by >1%) because the root of the problem lies in the ontological distinction between profane digital information and sacred, authoritative knowledge.

B. Sycophancy Above 40% and the Ratio Limit: Why Computing Fails to Have Intellectualus and Manners

The finding that most models accept false premises at a rate above 40% provides a solid empirical basis for epistemological criticism of *probabilistic reasoning* in AI. In Nasr's perspective (Nasr, 1989) , pure artificial intelligence operates in the realm of *ratio* (analytical-calculative reasoning disconnected from a transcendent source), not in the realm of *intellectus* (reason or *'aql* which is a divine ray capable of recognizing the Absolute). *Sycophancy* (the tendency of models to agree to false premises in order to please users) is a clear manifestation of the absence of reflective awareness and the absence of epistemic *etiquette due to the absence of intellectus*.

As also emphasized by Al-Attas (Al-Attas, 1991; al-Attas, 2014) , the modern intellectual crisis stems from the confusion of knowledge and the loss of *adab* , namely the failure to recognize the proper position of everything, including the limits of one's own abilities. AI does not have the capacity to exercise *iḥtiyāt* (caution) or say "I don't know" autonomously because it is driven solely by token probability optimization (Atif et al., 2025) . A comparison between GPT-4.0 and Gemini in *FiqhQA* , where more accurate models often have lower abstention rates, reveals a paradox: the higher the linguistic ability and artificial confidence of a system, the lower its epistemic awareness of the limits of legal validity (Shobri & Alfurqan, 2026) . This suggests that the limits of mechanistic rationality (*Burhān* AI and *ratio calculations*) are absolute and fundamental.

C. Inkhifāḍ Al-Adā' and the Limits of Sophia Perennis: Why Computation Fails to Capture the Meaning of Tradition

FiqhQA Data showed that all models experienced a decline in performance (*inkhifāḍ al-adā'*) when faced with classical Arabic language questions (Atif et al., 2025) . Within the

framework of *Irfānī epistemology* and the Islamic educational tradition, this phenomenon demonstrates the ontological limits of artificial intelligence. Arabic in Islam is not merely an instrument of semantic communication, but rather a vehicle for the transmission of meaning that is bound up with spiritual dimensions, *baraka* , and *ḥāl*. (Hikmah et al., 2026) .

Connecting these findings with the concepts of *Tradition* and *Sophia Perennis* in Nasr's view (Nasr, 1989) , religious truth is rooted in divine revelation that has a universal and permanent structure (*semantic permanence*). When AI reduces sacred religious texts to horizontal mathematical probabilities, the transcendent dimension and spiritual presence of the *talaqqī process* disappears completely. The decline in performance in Arabic indicates that AI operates on the surface of the text, isolated from the perennial wisdom and the depth of meaning that can only be absorbed through civilized human interaction.

D. The Illusion of Authority: From Relational Sanad to Mass Aggregation and Digital Pseudo-Authority

In Islam, the authority of knowledge is based on the integration between revelation , competence methodological, validation scientific , and responsible ethical responsibility (Abdullah, 2026) . Cross-model evaluations (*GPT-4.0*, *Gemini*, *Claude*) revealed performance gaps that gave rise to the phenomenon of *pseudo-authority*. High access speeds, intuitive interfaces, and syntactic fluency created the illusion of scientific legitimacy in AI, despite the models' high levels of hallucinations and minimal abstinence (Atif et al., 2025) .

In Nasr's critique (Nasr, 1989) , mass opinion or internet collectives have no authority whatsoever in matters of true knowledge; true authority demands a foundation of tradition, orthodoxy, and vertical transmission from a Supreme Source. When digital society relies on AI, which essentially aggregates popular internet discourse and data as instant religious references, a dislocation of authority occurs, leading to pseudo-authority. The following table summarizes this shift in orientation:

Table 2. Shift In The Orientation of The Authority of Knowledge

Dimensions	Traditional (Relational)	Digital Era (Computational)
Authority Base	Sanad of knowledge, orthodoxy, and moral integrity (<i>'adālah</i>)	Access speed, syntactic fluency, and mass data aggregation
Transmission Method	The tradition of <i>talaqqī</i> and vertical transmission	Token probability algorithm (<i>LLM</i>)
Source of Legitimacy	Diplomas, recognition of scholars, and <i>isnād</i>	User interface and <i>benchmark metrics</i>
Main Orientation	<i>Uswah</i> , <i>barakah</i> , <i>sophia perennis</i> , and the formation of the soul	Ease, speed and mechanistic efficiency
Main Risks	Limited physical access	<i>Pseudo-authority</i> and theological hallucinations based on mass opinion

E. Pedagogical Reorientation of Islamic Religious Education Teachers: From Mudarris to Mu'addib and Murabbī

Empirical findings regarding high rates of hallucinations (55%), *sycophancy* (>40%), and limited Arabic language comprehension demand a fundamental reorientation of the role of Islamic Religious Education (PAI) teachers. PAI teachers can no longer position themselves

merely as transmitters of cognitive information (*mudarris*), as the task of transmitting factual data has now been massively taken over by artificial intelligence systems.

Integrating Al-Attas's Islamic educational theory (Al-Attas, 1991) and Nasr's method of integrating science (Nasr, 1989), true education is *ta'dīb* —the process of cultivating recognition and acknowledgment of the proper place of everything within the order of being, which guides students in using modern instruments under the light of revelation. Therefore, the reorientation of Islamic Religious Education teachers must be placed in the following functional hierarchy:

Table 3. Pedagogical Reorientation of Islamic Religious Education Teachers

Levels	Role	Focus	Status in the AI Era
1	<i>Mudarris</i>	Cognitive information transfer (<i>ta'līm</i>)	Can be replaced by AI
2	<i>Navigator</i>	Teaching science along with understanding concepts	Some can be replaced by AI
3	<i>Mu'addib</i>	Instilling manners, ethics, and recognition of the position of everything	Cannot be replaced by AI
4	<i>Murabbi</i>	Maintenance, guidance, maturation of soul and character	Cannot be replaced by AI

(PAI) teachers in the era of digital disruption must transform into epistemic validators who guide students to critically verify AI output, train *iḥtiyāṭ* and abstain skills, and bridge digital information access into the ecosystem of *turāts* that has *adab values*. Thus, Islamic Religious Education (PAI) is not reduced to merely mechanistic-biological *tarbiyah*, but rather returns to its noble essence as the formation of civilized humans (*ta'dīb*) (Al-Attas, 1991).

Conclusion

This study concludes that the limitations of *Generative AI's religious output*, as documented by IslamicLegalBench, FiqhQA, and QIAS 2025, are structural, not technical, rooted in *next-token prediction* and *probabilistic reasoning architectures* that cannot be corrected through increased computational capacity, *fine-tuning*, or *prompt engineering*. The tripartite diagnosis (*Bayani, Irfani, Burhani*) reveals three interrelated layers of limitations. In *Bayani reasoning*, AI excels at compiling textual data and mapping fatwas, but is devoid of *isnād*, *naqd al-isnād*, and shari'a accountability; the hallucination rate of 55% in some models and 21% even in the best model is not a statistical anomaly, but rather a manifestation of the absence of a living transmission relationship between AI and authoritative sources. In *Irfani reasoning*, AI suffers from a total epistemic vacuum: consistent performance decline in Arabic indicates that AI operates through statistical patterns, not a living understanding of meaning, and thus lacks the capacity to address *the tazkiyatun nafs, khasyyah*, transmission of *ḥāl*, and *barakah* that require the presence of a conscious *heart and direct encounter (talaqqī)*. In *Burhani reasoning*, AI is able to map *the 'illat* of law and the sociology of madhhabs, but *sycophancy* above 40% in 6 of 9 models proves the absence of reflective awareness of *'aql*; the FiqhQA paradox of GPT-4.0 being superior in accuracy but Gemini being better at *abstaining* suggests that the most "accurate" AI is the least aware of the limits of its knowledge, so its outputs remain only as raw hypotheses, not *hujjah*.

The epistemological consequence of these findings is a paradigm shift in religious authority from *relational transmission* (based on sanad and *talaqqī*) to *computational visibility* (based on speed of access and ease of algorithmic interface). This shift gives rise to *pseudo-*

authority, where students tend to measure the truth of religious information based on language fluency and speed of access, rather than on sanad and methodology.

The pedagogical implication of this finding is a radical reorientation of the role of Islamic Religious Education (PAI) teachers, positioned solely as learning instruments (*wasīlah* or *ta'lim*), not as holders of religious authority (*ghāyah*). Therefore, the role of Islamic Religious Education (PAI) teachers must be fundamentally reoriented from merely conveying cognitive information (*mudarris*) to being civilized guides and guides of the soul (*mu'addib* and *murabbī*). This transformation is realized through the cultivation of *Digital Tabayyun*, the instillation of *Adab al-'Ilm*, and moral exemplars (*uswah hasanah*) that cannot be reduced by any algorithmic system. Theoretically, this study offers an epistemological diagnostic framework (Bayani, Irfani, Burhani) that goes beyond accuracy- or efficiency-based assessments, while strengthening the argument that AI's limitations are *structural*, a correction to the technological narrative that tends to be *solutionist*. However, as *library research*, this study does not produce new empirical data and is limited to three AI platforms; The effectiveness of the *Digital Tabayyun* model and the reorientation of the *mu'addib* still require field research or classroom action research to empirically test its implementation. Ultimately, the automation of cognitive knowledge transfer by AI actually emphasizes and strengthens the essential role of Islamic Religious Education teachers as *mu'addib* and *murabbi*, roles rooted in living human relationships and irreplacable by any algorithm.

References

- Abdullah, A. H. (2026). The Boundaries of Legitimate Knowledge in Islamic Thought: Authority, Pluralism, and the Politics of Interpretation. *QiST: Journal of Quran and Tafseer Studies*, 5(2), 873–902. <https://doi.org/10.23917/qist.v5i2.17051>
- Ajwath, Dr. M. R. A. (2026). Islamic Epistemology in the Age of Misinformation: Reclaiming Truth, Knowledge, and Scientific Authority. *International Journal of Research and Innovation in Social Science*, 10(6), 17769–17792. <https://doi.org/10.47772/IJRISS.2026.100601226>
- Al-Attas, M. N. (1991). *The Concept of Education in Islam: A Framework for an Islamic Philosophy of Education*. International Institute of Islamic Thought and Civilization. <https://books.google.co.id/books?id=hfxwnQAACAAJ>
- al-Attas, S. M. N. (2014). *Prolegomena To The Metaphysics Of Islam*. Penerbit UTM Press. <https://books.google.co.id/books?id=PKugBQAAQBAJ>
- Ali, M. M., & Alistiana, L. (2026). REKONSTRUKSI OTORITAS GURU DI ERA ARTIFICIAL INTELLIGENCE: ANALISIS EPISTEMIK, PEDAGOGIS, DAN MORAL DI MADRASAH IBTIDAIYAH. *Jurnal Citra Multidisiplin*, 1(4), 355–366. <https://doi.org/10.38048/jcm.v1i4.6835>
- Al-Jabiri, M. A. (1991). *Bunyah al-'aql al-'arabi: Dirasah tahliliyyah naqdiyyah li nuzhum al-ma'rifah fi al-tsaqafah al-'arabiyyah*. Al-Markaz al-Tsaqafi al-'Arabi.
- Amin, M. (2026). Ethical Use of Generative Artificial Intelligence in Islamic Religious Education: Perspectives of Madrasah Aliyah Teachers and Students in Banjarmasin. *Al-Munawwarah: Journal of Islamic Education*, 2(2), 147–161. <https://doi.org/10.38073/almunawwarah.v2i2.5654>

- Atif, F., Askarbekuly, N., Darwish, K., & Choudhury, M. (2025a). 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>
- Atif, F., Askarbekuly, N., Darwish, K., & Choudhury, M. (2025b). 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>
- Azizah, N., Al-Khoiri, F., Azhar, M. K., Irhamsyah, G., & Jordan, B. (2026). Reposisi Guru Pendidikan Agama Islam di Era Generative Artificial Intelligence: Penguatan Peran Humanistik dalam Pembelajaran. *Tadabbur: Jurnal Integrasi Keilmuan*, 5(1), 179–188. <https://doi.org/10.15408/tadabbur.v5i1.53452>
- Bouhekif, A., Rashwani, S., Sbahi, H., Gaben, S., Al-Khatib, M., & Ghaly, M. (2025). Assessing Large Language Models on Islamic Legal Reasoning: Evidence from Inheritance Law Evaluation. *Proceedings of The Third Arabic Natural Language Processing Conference*.
- Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative Research Designs: Selection and Implementation. *The Counseling Psychologist*, 35(2), 236–264. <https://doi.org/10.1177/0011000006287390>
- Desitha Cahya, Putri Ramdani, Annajmi Rauf, Andi Baso Kaswar, & M Miftach Fakhri. (2026). AI Hallucinations in AI-Assisted Educational Decision-Making and Academic Honesty Intentions Among Undergraduates. *Journal of Applied Artificial Intelligence in Education*, 48–61. <https://doi.org/10.66053/jaaie.v1i2.9>
- Elmahjub, E., Qadir, J., Mushtaq, A., Naeem, R., Ghaznavi, I., & Iqbal, W. (2026a). *IslamicLegalBench: Evaluating LLMs Knowledge and Reasoning of Islamic Law Across 1,200 Years of Islamic Pluralist Legal Traditions* (arXiv:2602.21226). arXiv. <https://doi.org/10.48550/arXiv.2602.21226>
- Elmahjub, E., Qadir, J., Mushtaq, A., Naeem, R., Ghaznavi, I., & Iqbal, W. (2026b). *IslamicLegalBench: Evaluating LLMs Knowledge and Reasoning of Islamic Law Across 1,200 Years of Islamic Pluralist Legal Traditions* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2602.21226>
- Hikmah, N., Hakiem, I. H. E., & Hakiem, D. A. H. E. (2026). THE IRFĀNĪ METHOD IN THE QIRĀ’ĀT ‘ĀṢIM TRANSMISSION NETWORK: Teacher and Student Relationship in the Path of ‘Āṣim. *Edukasi Islami: Jurnal Pendidikan Islam*, 15(02), 1125–1136. <https://doi.org/10.30868/ei.v15i02.10361>
- Mahbubi, M. (2025). Filsafat Pendidikan Islam di Era AI: Integrasi Epistemologi dan Aksiologi Islam. *An-Nuha*, 5(1), 37–45. <https://doi.org/10.24036/annuha.v5i1.591>
- Nasr, S. H. (1989). *Knowledge and the Sacred*. State University of New York Press. <https://books.google.co.id/books?id=DsooAQAAMAAJ>
- Putri1, S. A. A. Z., Muslimah, K. C., & Fauzia, S. R. (2026). ARTIFICIAL INTELLIGENCE (AI) DALAM PEMBELAJARAN AKIDAH AKHLAQ: SEBUAH TINJAUAN FILOSOFIS-EPISTEMOLOGIS. *Ilmuna: Jurnal Studi Pendidikan Agama Islam*, 8(1), 530–546. <https://doi.org/10.54437/ilmuna.v8i1.2942>
- Qian, K., Liu, S., Li, T., Raković, M., Li, X., Guan, R., Molenaar, I., Nawaz, S., Swiecki, Z., Yan, L., & Gašević, D. (2026). Towards reliable generative AI-driven scaffolding: Reducing hallucinations and enhancing quality in self-regulated learning support. *Computers & Education*, 240, 105448. <https://doi.org/10.1016/j.compedu.2025.105448>

- Rhisky, M., & Baharun, H. (2026). Knowledge Authority and Value Degradation in the AI Era Rethinking Islamic Education. *DAAR EL-IDARAH: Journal of Islamic Education Management*, 2(01), 138–152. <https://doi.org/10.66931/jmpi.v2i01.430>
- Saidi, N. S. B., Mutathahirin, M., Zulkefli, N. A. B., & Dasrizal, D. (2025). Students' Pursuit of Knowledge in Islam: Framework for Character Education from Quran and Hadith. *Muaddib: Journal of Islamic Teaching and Learning*, 1(3), 64–78. <https://doi.org/10.67055/x7gdk766>
- Shalih, B. (2026). From Kitab Kuning To Chatgpt: Authority, Authenticity, And Pesantren Pedagogy In The Age Of Generative Artificial Intelligence. *Hunafa: Jurnal Studia Islamika*, 23(1), 23–41. <https://doi.org/10.24239/jsi.v23i1.952>
- Shobri, F., & Alfurqan, A. (2026a). From Individual Competence to Distributed Capacity: How Islamic Religious Education Teachers Verify Generative-AI Content in Padang, Indonesia. *FIKROTUNA: Jurnal Pendidikan Dan Manajemen Islam*, 15(02), 707–736. <https://doi.org/10.32806/jf.v15i02.2227>
- Shobri, F., & Alfurqan, A. (2026b). From Individual Competence to Distributed Capacity: How Islamic Religious Education Teachers Verify Generative-AI Content in Padang, Indonesia. *FIKROTUNA: Jurnal Pendidikan Dan Manajemen Islam*, 15(02), 707–736. <https://doi.org/10.32806/jf.v15i02.2227>
- Sidik, A., Saputro, A. D., Sudarmadi, S., & Syam, A. R. (2026). From Hoaxes to Wisdom: A Tabayyun-Based Digital Literacy Model for Junior High Islamic Education Via Canvas LMS. *Progresiva: Jurnal Pemikiran Dan Pendidikan Islam*, 15(01), 1–14. (ahmadsidik@alfacentauri.sch.id). <https://doi.org/10.22219/progresiva.v15i01.43954>
- Ubaidah, H. H., & Haris, A. (2026). From Ta'lim to Ta'dib: Reorienting Scientific Authority in Islamic Education in the Era of Artificial Intelligence. *Intiqad: Jurnal Agama Dan Pendidikan Islam*, 18(1), 300–320. <https://doi.org/10.30596/intiqad.v17i2.29856>
- Velly, V., Rahman, A., Alharbi, A., & Awan, N. (2025). *Islamic Education 4.0: Rethinking Moral and Religious Learning for a Socially Conscious Generation*. <https://core.ac.uk/outputs/705490057/#1>
- Yun, S. (2025). Paradigm Shift in Education in the Era of Generative AI and the Redefinition of Teachers' Roles. *Korean Journal of Religious Education*, 80, 109–126.