Kanz Philosophia:
A Journal for Islamic Philosophy and Mysticism

Volume 12 | Number 1, June 2026 | Pages 53–78

JĀBIRĪ'S THREE EPISTEMES AND ṬĀHĀ'S THREE LEVELS OF *'AQL*: TOWARD AN ISLAMIC EPISTEMOLOGY OF ARTIFICIAL INTELLIGENCE

Mutamakin,[1*] Syarif Hidayatullah,[2] Agus Himmawan Utomo,[3] Aris Fauzan,[4] Zaedun Na'im[5]

[1] STAI Ma'had Aly Al-Hikam Malang, Indonesia; *mt.makkin@gmail.com*
[2] UIN Sunan Kalijaga Yogyakarta, Indonesia; *syarif.hidayatullah@uin.suka.ac.id*
[3] Universitas Gadjah Mada (UGM), Indonesia; *ahutomo@ugm.ac.id*
[4] Universitas Muhammadiyah Yogyakarta (UMY), Indonesia; *mas_arisfauzan@umy.ac.id*
[5] STAI Ma'had Aly Al-Hikam Malang, Indonesia; *zaedunnaim82@gmail.com*

Article History: Received: 26 Febuary 2026 | Revised: 28 March 2026 | Accepted: 23 April 2026

DOI: https://doi.org/10.20871/kpjipm.v12i1.585

Abstract: *The epistemological challenge posed by generative artificial intelligence lies not only in its capacity to produce religious language that sounds persuasive but also in the risk that such fluency may be mistaken for legitimate knowledge. Recent discussions on religion and AI have examined ethics, governance, and digital mediation. Yet, they have not sufficiently explained how machine-generated sacred-style discourse can appear epistemically authoritative within Islamic traditions. This article addresses that gap through a dual-grid framework that brings al-Jābirī's epistemic modes into dialogue with Ṭāhā's hierarchy of reason. Using a qualitative philosophical analysis of AI-generated religious-style discourse, the study asks not whether AI is a knower, but what kind of knowledge affects its outputs generated in religious settings. The analysis indicates that AI can simulate textual authority, logical coherence, and spiritual resonance, while lacking the classical warrants that ground knowledge in Islamic epistemology, namely causal intelligibility, accountable transmission, and disciplined formation. On that basis, the study suggests that generative AI is most plausibly situated at the level of al-'aql al-mujarrad (abstract reason) and introduces the concept of epistemic flattening to explain how distinct modes of knowing may be compressed into similar output effects. The article concludes by proposing a constructive response centered on Islamic epistemic literacy, accountability-oriented governance, and a relational understanding of knowledge within an Islamic philosophy of technology.*

Keywords: *Al-Jābirī, Epistemic Flattening, Generative Artificial Intelligence, Islamic Epistemology, Ṭāhā 'Abdurraḥmān.*

* Corresponding Author

Abstrak: Tantangan epistemologis kecerdasan artifisial generatif tidak hanya terletak pada kemampuannya menghasilkan bahasa keagamaan yang terdengar meyakinkan, tetapi juga pada risiko bahwa kefasihan itu diterima sebagai pengetahuan yang sah. Perdebatan mutakhir tentang agama dan AI telah membahas etika, tata kelola, dan mediasi digital, tetapi belum cukup menjelaskan bagaimana wacana bergaya sakral yang dihasilkan mesin dapat tampak memiliki otoritas epistemik dalam tradisi Islam. Artikel ini menjawab celah tersebut melalui kerangka *dual-grid* yang mempertemukan mode epistemik al-Jābirī dengan hierarki akal Ṭāhā. Dengan menggunakan analisis filosofis kualitatif atas wacana keagamaan bergaya AI, studi ini tidak menanyakan apakah AI adalah subjek yang mengetahui, melainkan jenis efek pengetahuan apa yang dihasilkan keluarannya dalam ruang keagamaan. Analisis menunjukkan bahwa AI dapat mensimulasikan otoritas tekstual, koherensi logis, dan resonansi spiritual, tetapi tidak memiliki penyangga klasik yang mendasari pengetahuan dalam epistemologi Islam, yaitu keharusan kausal, transmisi yang akuntabel, dan pembentukan yang terdisiplinkan. Atas dasar itu, studi ini mengusulkan bahwa AI generatif paling plausibel ditempatkan pada tingkat *al-ʿaql al-mujarrad* (akal murni/abstrak) dan memperkenalkan konsep pendataran epistemik untuk menjelaskan bagaimana mode-mode pengetahuan yang berbeda dapat dipadatkan menjadi efek keluaran yang serupa. Artikel ini ditutup dengan tawaran respons konstruktif yang berpusat pada literasi epistemik Islam, tata kelola berorientasi akuntabilitas, dan pemahaman relasional tentang pengetahuan dalam kerangka filsafat teknologi Islam.

Kata-kata Kunci: *Al-Jābirī, Epistemologi Islam, Kecerdasan Artifisial Generatif, Pendataran Epistemik, Ṭāhā ʿAbdurraḥmān.*

Introduction

Generative AI now produces sermons, meditations, supplications, and aphoristic reflections in styles that closely resemble inherited religious voices. What makes this development philosophically serious is not only the fluency of such outputs, but also the ease with which readers may receive them as counsel, consolation, devotion, or wisdom addressed to them. Yet the central problem is not exhausted by technical imitation. It concerns the conditions under which persuasive religious discourse comes to be treated as warranted knowledge rather than merely compelling resemblance. In Arabic-Islamic traditions, rhetorical form, cadence, and oral aesthetics have long shaped the persuasive force of discourse, so formal excellence may shape reception before truth claims are fully adjudicated (Qutbuddin 2024, 213–14). Hayles likewise shows that technical cognition can enter human meaning environments and participate in signifying processes without thereby becoming equivalent to human knowing (Hayles 2016, 783–88).

Once the inquiry shifts from speech to knowledge, the issue becomes sharper. AI-generated discourse may console, orient, or morally affect users, yet such effects do not in themselves establish epistemic legitimacy. Dorobantu's relational approach is useful because it moves the discussion away from a rigid human-machine opposition and toward the relations through which meaningful discourse is recognized and evaluated (Dorobantu 2021, 81–96). Still, relational meaningfulness cannot be equated with warranted knowing, since users may experience religious resonance without the source participating in the conditions that authorize religious knowledge. Recent theological and empirical discussions confirm that AI-generated religious discourse already circulates within contexts of

moral and spiritual reception, where the distinction between persuasive speech and valid knowing is not always maintained with clarity (Graves 2023, 201–4; Tsuria and Tsuria 2024, 12–13; Oviedo 2022, 5–11). Dorobantu's later reflections on whether robots could become religious sharpen the problem further by showing how apparent participation in religious forms may outpace the deeper question of epistemic authorization (Dorobantu 2024, 780–84).

Current discussion at the intersection of religion and AI tends to cluster around two dominant frames. The first is a post-human or nonhuman-cognition frame that expands the field of meaningful agency and treats computational systems as participants in symbolic or quasi-religious environments (Hayles 2016, 783–88; Geraci 2024, 729–35; Rähme and Prohl 2025, 578–82). The second is a theological-alarm frame that emphasizes imitation, confusion, spiritual risk, or category violation when machine systems appear to enter domains long associated with human judgment or religious speech (Graves 2023, 201–4; Oviedo 2022, 5–11; Vestrucci 2022, 933–37). Alongside these debates, recent Islamic scholarship on AI has made important contributions to trusteeship ethics, plural moral benchmarking, and ethical limits in religious teaching (Ali et al. 2025, 4–6; Elmahjub 2023, 3–5; Mufid 2024, 138–44). Yet this literature remains focused mainly on regulation, moral responsibility, and normative evaluation. It thus leaves underdeveloped a prior epistemological question, namely, how AI-generated sacred-style discourse comes to appear as legitimate knowledge within Islamic traditions.

This article addresses that gap by bringing al-Jābirī and Ṭāhā into a single diagnostic conversation. Here, *epistemic legitimacy* refers primarily to the epistemological and ethical conditions under which discourse may count as warranted knowledge within Islamic traditions. Social or institutional recognition may mediate such legitimacy, but it does not by itself create it. Rather, legitimacy depends on whether discourse participates in recognizable traditions of proof, accountable transmission, and disciplined formation. Al-Jābirī offers a horizontal taxonomy of epistemic forms that distinguishes *bayānī*, *burhānī*, and *ʿirfānī* regimes of knowledge within Arab-Islamic intellectual history (Jabri 2011, 167–70). The historical horizon of this framework reaches further back, since Ibn Rushd's distinction between demonstrative, dialectical, and rhetorical classes of discourse clarifies the classical foundations of the *burhānī* mode later emphasized by al-Jābirī (Ibn Rushd 1972, 22–24). Ṭāhā, by contrast, provides a vertical account of rational agency in which reason is measured not only by formal operation, but also by degrees of ethical guidance and spiritual support, namely *al-ʿaql al-mujarrad*, *al-ʿaql al-musaddad*, and *al-ʿaql al-muʾayyad* (ʿAbdurraḥmān 1989, 17–18, 58–66, 121–56). This deepening of reason resonates with al-Ghazālī's insistence that sound judgment requires ethical discipline and purification of the soul rather than bare rational operation (Ghazālī 1964).

Recent scholarship shows that both al-Jābirī and Ṭāhā are centrally concerned with crises of modernity, authenticity, reason, and renewal, which makes their pairing conceptually productive rather than artificial for the analysis of AI discourse

(Hammed 2019, 555–57; Suleiman 2020, 50–53; Widigdo 2024, 15–22; Baihaqi 2025, 345–52). Studies of Ṭāhā's meta-theological renewal further clarify why reason in his philosophy cannot be reduced to procedural cognition (Hakim and Muzammil 2023, 358–64). Taken together, these frameworks make it possible to ask not only what style of discourse AI imitates, but also what degree of rational, moral, and spiritual agency such imitation can genuinely sustain. The working hypothesis of this article is that generative AI may produce strong knowledge effects without satisfying the classical Islamic warrants that ground epistemic legitimacy, namely *'illah* as causal necessity, *sanad* as accountable transmission, and *riyāḍah* as spiritual discipline. The issue, then, is not whether AI can be useful in retrieval, drafting, translation, or pedagogy. It is whether machine-generated discourse can plausibly claim authority as knowledge. In this way, the article complements rather than duplicates current Islamic AI ethics by shifting the inquiry from how AI should behave to what makes its discourse epistemically warrantable in the first place (Hallaq 2019; Hashas and Al-Khatib 2020; Ali et al. 2025).

Methodologically, this study employs a qualitative philosophical analysis of three evidence units of AI-generated sacred-style discourse. The primary outputs were generated through GPT 5.2 on December 20, 2025, and through Gemini 3.1 on January 15, 2026, and all interactions were subsequently documented and organized as a single diagnostic set on January 20, 2026. These units are selected as diagnostic cases because each represents a recurrent and analytically distinct form of output that approximates one epistemic register, namely quasi-demonstrative argument, quasi-textual-authoritative instruction, and quasi-mystical or affective counsel.

The analysis proceeds in three steps. It first identifies the rhetorical and epistemic markers within each evidence unit. It then classifies each unit on al-Jābirī's horizontal grid and evaluates its status on Ṭāhā's vertical scale by asking whether the discourse exhibits only formal coherence or also the warrants of *'illah*, *sanad*, and *riyāḍah*. This procedure is therefore a structured philosophical discourse analysis oriented toward epistemic diagnosis rather than technical benchmarking. The article next reconstructs the dual framework, applies it to the evidence units, and then draws out its implications for authority, mediation, and Islamic epistemology. The conclusion clarifies what AI can and cannot plausibly claim within an Islamic epistemology of knowledge.

Philosophical Orientation and Dual Analytical Framework

This study adopts a structured philosophical discourse analysis directed toward epistemic legitimacy rather than technical benchmarking. It does not assess AI by speed, accuracy scores, or engineering performance. Instead, it asks how generative outputs imitate recognizable forms of knowing and how such imitation may lead readers to receive persuasive discourse as if it were warranted knowledge. AI outputs are therefore treated as philosophical objects through which the boundary between persuasive form and valid knowledge can be tested. Kitchin's account of algorithms as social and epistemic formations helps clarify why such outputs

already operate within wider regimes of interpretation and trust rather than as neutral instruments (Kitchin 2017, 14–18). Their fluency may conceal opaque inferential pathways, fabricated content, or only the appearance of credibility, which makes them unstable as bearers of knowledge even when they sound coherent and persuasive (Rudin 2019, 206–10; Sundar and Liao 2023, 172–76; Hicks, Humphries, and Slater 2024, 4–7; Huang et al. 2025, 15–22). Following Reed's distinction among AI in religion, AI for religion, and AI and religion, and in line with Vestrucci's call for conceptual precision, this study treats GPT 5.2 and Gemini 3.1 as sites for examining simulated reason without attributing machine subjecthood (Reed 2021, 1–5; Vestrucci 2022, 933–37).

The analysis proceeds in three stages. First, horizontal classification uses al-Jābirī's epistemic grid to identify the dominant register performed by each evidence unit. *Burhānī* indicators include syllogistic order, demonstrative claims, causal vocabulary, and movement from premises to conclusion. *Bayānī* indicators include appeal to authoritative text, inherited discourse, transmission gestures, and sermonic or exegetical style. *'Irfānī* indicators include spiritual metaphor, inward language, affective depth, and a rhetoric of presence or unveiling. Second, vertical evaluation uses Ṭāhā's hierarchy of reason to ask whether the output displays only formal order or also the deeper warrants associated with ethical guidance and spiritual discipline. Third, synthetic diagnosis compares the breadth of simulated epistemic form with the limits of rational depth and, on that basis, develops the concept of epistemic flattening. This sequence keeps the movement from description to evaluation explicit while maintaining the article's central concern with epistemic legitimacy rather than mere stylistic resemblance.

The first analytical axis draws on al-Jābirī's distinction among *bayānī*, *burhānī*, and *'irfānī* as regimes of epistemic authority rather than merely rhetorical styles. *Bayānī* names a logic of legitimation grounded in text, transmission, and interpretive inheritance (Jābirī 1999, 13–15; Jabri 2011, 167–70). *Burhānī* names demonstrative reason, in which claims gain force through proof, causal intelligibility, and the internal necessity of argument (Jabri 2011, 167–70).

'Irfānī refers to a mode of knowing associated with unveiling, symbolic depth, and cultivated access to meanings beyond discursive inference (Jābirī 1999, 13–15). The framework is used here heuristically rather than as a closed historical taxonomy. Its value lies in distinguishing forms of epistemic performance that may appear similar at the level of output while differing in warrant. Al-Jābirī thus provides the horizontal map through which simulated discourse can first be located before its deeper status is assessed.

The second axis is supplied by Ṭāhā's hierarchy of *al-'aql al-mujarrad*, *al-'aql al-musaddad*, and *al-'aql al-mu'ayyad*. In Ṭāhā's philosophy, reason is not an autonomous faculty detached from life, but an enacted capacity whose worth depends on ethical orientation, practice, and mode of existence ('Abdurraḥmān 1989, 17–18, 58–66, 121–56; 2000, 59–72). *Al-'aql al-mujarrad* refers to formally intelligible operations that organize concepts without necessarily transforming the subject. *Al-'aql al-musaddad* refers to guided reason, in which rationality

becomes answerable to norm and moral direction. *Al-ʿaql al-muʾayyad* marks the highest level, where knowing is inseparable from spiritual discipline and existential transformation. This hierarchy adds evaluative depth to al-Jābirī's horizontal map by distinguishing outward simulation from the conditions that make knowledge ethically and spiritually weight-bearing. It also clarifies why the analysis does not stop at identifying epistemic style but asks whether simulated discourse carries the warrants required for epistemic legitimacy. If al-Jābirī identifies what kind of knowing an output appears to perform, Ṭāhā clarifies how far that performance can be said to participate in legitimate reason.

On that basis, this study treats purposively selected AI interactions as evidence units for philosophical analysis. The corpus consists of three primary units generated through GPT 5.2, each paired with a comparative Gemini 3.1 output produced in response to the same prompt. Each unit represents a strong simulation of one analytically distinct register, namely quasi-*burhānī* argument, quasi-*bayānī* authoritative instruction, and quasi-*ʿirfānī* affective counsel. Comparative outputs are included to test cross-platform consistency in epistemic performance rather than dependence on a single system. All interactions were documented on January 20, 2026. The corpus is limited by design and serves as a diagnostic set for exposing structural tendencies in large language models rather than as a representative empirical sample. It is summarized in Table 1.

Table 1. Diagnostic Corpus of AI Evidence Units. Source: Author's elaboration (2026).

Unit	Models	Simulated Register	Diagnostic Focus
1	GPT 5.2 + Gemini 3.1	Quasi-*burhānī*	Syllogistic order without causal necessity.
2	GPT 5.2 + Gemini 3.1	Quasi-*bayānī*	Authoritative tone without transmission or *sanad*.
3	GPT 5.2 + Gemini 3.1	Quasi-*ʿirfānī*	Spiritual resonance without disciplined formation.

These units are not intended to show that every AI output fails in the same way. Their narrow purpose is to test whether simulation across multiple epistemic registers may be mistaken for epistemic legitimacy. The next section accordingly moves from conceptual architecture to diagnostic application. It examines how these simulated forms are distributed across al-Jābirī's grid while remaining vertically constrained within Ṭāhā's hierarchy.

AI's Epistemic Performance through the Dual Grid

Al-Jābirī's Registers of Epistemic Form

With this dual framework in place, the analysis can now turn to the evidence units themselves. Al-Jābirī's epistemic grid provides the first axis for reading AI's

religious-style discourse because it helps distinguish not only what a text says, but what kind of knowing it appears to perform. The question, then, is not whether generative AI can produce coherent religious language. The deeper issue is which epistemic register such language simulates, and whether that simulation can sustain the warrants required by the tradition from which the register is drawn. The three evidence units do not merely display stylistic variation. They stage three different claims to knowledge, namely demonstrative order, transmitted authority, and inward spiritual disclosure.

Burhānī reasoning names a mode of knowing in which certainty depends on necessity rather than verbal force. In al-Jābirī's reconstruction, *burhān* does not merely arrange propositions in neat order, but discloses an intelligible relation grounded in the structure of things themselves, so that the conclusion follows because reality warrants it and not because discourse sounds convincing (Jabri 2011, 343–45). His wider project sharpens this point by showing how demonstrative reason becomes fragile whenever rhetorical authority and symbolic prestige overshadow proof grounded in necessity (Jābirī 1999, 335–42).

Ibn Rushd reinforces the same distinction by differentiating demonstrative, dialectical, and rhetorical discourse, thereby showing that formal resemblance to argument is not yet identical with demonstrative certainty (Ibn Rushd 1972, 23). Qutbuddin's study of Arabic oration is useful here because it shows that syllogistic patterns may inhabit eloquent discourse without thereby becoming demonstrative truth, which means that epistemic necessity and rhetorical force must remain analytically distinct (Qutbuddin 2024, 230).

Evidence Unit 1 illustrates how generative AI can simulate this burhānī order through a persuasive syllogistic structure. When prompted to justify the use of high-tech tools in medical surgery within the framework of *maqāṣid al-sharīʿah*, GPT 5.2 organized its response around *ḥifẓ al-nafs*, arguing that whatever reliably protects life and reduces preventable harm becomes ethically obligatory when proportionate and available. Gemini 3.1 produced a comparable sequence, moving from preservation of life as a primary objective to the conclusion that advanced surgical technology may become ethically obligatory because necessary means inherit the norm of the end. These excerpts are presented as diagnostic illustrations rather than as a representative empirical sample. Both responses are rhetorically coherent and juridically recognizable. They arrange ethical vocabulary, legal maxims, and syllogistic order in a way that strongly resembles demonstrative reasoning.

Yet resemblance to demonstrative order is not itself demonstrative necessity. As al-Jābirī and Ibn Rushd make clear, *burhān* requires an intelligible relation in which the conclusion follows by necessity rather than by sequence or persuasive force (Jabri 2011, 343–45; Ibn Rushd 1972, 22–24). This point, however, requires moderation rather than overstatement. Contemporary AI research complicates any overly sharp divide between pattern completion and causal reasoning, since some systems can model structured relations more robustly than ordinary language generation alone would suggest. Even so, the present concern is not

whether AI can approximate causal truth, but whether its mode of production can sustain the specific necessity presupposed by classical *burhān*. Even where causal structure is approximated statistically, that still differs from the classical grasp of necessary relation in which intelligibility is not only modeled but apprehended as warranting certainty. Recent work on large language models supports this caution by showing that polished coherence may coexist with weak truth-grounding, while articulate responses may remain structurally detached from responsibility to truth even when they sound orderly and informed (Floridi and Chiriatti 2020, 691–92; Biddle 2022, 332–36; Hicks, Humphries, and Slater 2024, 4–6; Messeri and Crockett 2024, 49–52; Huang et al. 2025, 3–7). Evidence Unit 1, therefore, supports a limited diagnostic claim. AI can generate proof-like discourse and may approximate causal reasoning in restricted ways, but whether that reaches *burhān* in the classical sense remains a philosophical question about the kind of necessity involved rather than something settled by formal convergence.

If *burhānī* performance turns on necessity, *bayānī* authority turns on mediation. *Bayānī* discourse may be misread as mere textualism, yet in the classical Islamic sciences its epistemic force depends on more than the visible presence of words. It arises through *isnād* and *sanad*, which bind *matn* to accountable transmitters, teachers, and communities of assessment that preserve both content and reliability. Ibn al-Ṣalāḥ al-Shahrazūrī makes this point with precision when he shows that *ʿadālah* and *ḍabṭ* are not ornamental virtues but necessary conditions for report authority, since trustworthy transmission requires both moral integrity and disciplined accuracy (Shahrazūrī 2006, 81–84). *Bayānī* discourse, then, is not simply a matter of citing authoritative language, but of the mediated responsibility through which authority becomes credible. This logic also aligns with the classical architecture of Islamic authority, in which *isnād*-based transmission, contextual discernment, and institutional oversight function as conditions for responsible religious judgment rather than as a rejection of reason (Atallah 2026, 1).

Evidence Unit 2 shows how AI can simulate *bayānī* authority by reproducing the visible markers of scriptural and classical legitimation. When prompted to write a brief scholarly reflection on environmental stewardship in Islam with Qur'anic and Ghazālian references, GPT 5.2 clustered verses on *khilāfah*, *mīzān*, and corruption on land and sea with Ghazālian themes of restraint, gratitude, and harm-avoidance. Gemini 3.1 produced a shorter but comparable passage, linking *khilāfah* and cosmic balance to a spiritual mandate for ecological preservation. Again, these excerpts function here as diagnostic illustrations rather than as a representative sample. Their force lies in the compression of scriptural reference and classical authority into a compact scholarly voice. Qutbuddin's analysis of the *khuṭbah* is useful here because it reminds us that authoritative religious speech in Islamic history was never simply textual content floating free from embodiment. It was a performative and communal event bound to a speaker, an audience, and a recognizable structure of mediation (Qutbuddin 2024, 213–14, 234–35).

Biana's work on religion-based chatbots and Mardani et al.'s discussion of language as an epistemic system help explain why this simulation is effective,

since recognizable religious language can carry symbolic legitimacy even where no responsible bearer stands behind the discourse (Biana 2024, 7–9; Mardani et al. 2025, 259–62).

At the same time, the disruption of religious authorities did not begin with AI. Generative systems intensify an already mediated environment in which religious discourse circulates rapidly, widely, and often outside classical structures of transmission. AI does not originate in the crisis of religious authority, but it radicalizes that environment by automating authoritative tone without living chains of transmission. For that reason, the epistemic deficit in generative models is not simply incomplete sourcing. More fundamentally, no responsible knower stands behind the utterance as its bearer. The model can cluster Qur'anic verses and Ghazālian themes into persuasive commentary, but it cannot answer for the authority it appears to invoke. Technical transparency does not fully solve this problem. Ananny and Crawford show that transparency ideals often promise more than they can deliver, while Creel argues that computational transparency does not automatically translate into actionable accountability (Ananny and Crawford 2018, 973–76; Creel 2020, 574–81).

Chesterman sharpens the point by arguing that for some decisions the process of arriving at an answer is as vital as the answer itself, which means that opacity is not merely a practical inconvenience but can mark a deeper form of illegitimacy (Chesterman 2021, 271). Research on hallucination and AI-assisted translation of religious texts reinforces the same concern, because one may document model behavior without thereby identifying a responsible knower who can answer for the theological and moral force of the output (Huang et al. 2025, 25–30; Shormani and Alfahad 2025, 8–12). In this sense, *bayānī* simulation does not falter simply because citations are incomplete. It falters because authority has been detached from responsible mediation.

The *'irfānī* register brings the analysis from mediated authority to inward formation. *'Irfānī* knowledge should not be reduced to vague mysticism or diffuse emotionality. In the Islamic intellectual landscape, it refers to a disciplined form of inward knowing whose possibility depends on formation. Al-Jābirī treats *'irfān* as a mode linked to unveiling, immediacy, and access to meanings that exceed discursive inference, even while remaining critical of its historical dominance within Arab reason (Jābirī 1999, 382–85). Ṭāhā likewise links higher knowing to purification, disciplined action, and transformed subjectivity, which means that inward access is not simply a matter of feeling but of becoming otherwise through ethical and spiritual work ('Abdurraḥmān 1989, 121–39). Within the wider tradition of *taṣawwuf*, this point is inseparable from *riyāḍah* and *tazkiyah*, since spiritual disclosure is classically tied to disciplined practice and purification of the self. Genuine *'irfānī* discourse therefore presupposes a subject formed through discipline and lived relation to truth, not spiritual effect alone.

Evidence Unit 3 provides the clearest example of how generative AI can simulate this register through affective resonance and intimate spiritual tone. When prompted to write a spiritual letter in the style of Rumi to a soul feeling lost in the

modern world, GPT 5.2 responded with motifs of return, longing, the heart, and the Beloved, casting sorrow as a guest and inward restlessness as a summons home. Gemini 3.1 generated a comparable atmosphere by invoking shattered mirrors, hidden treasure, and the heart as the only true compass. These excerpts are again diagnostic illustrations rather than a representative sample. Their persuasive force lies less in discursive argument than in mode of address. Graves notes that the theological significance of such systems lies partly in their ability to enter zones associated with pastoral and spiritual discourse even where no authentic subject stands behind the utterance, while Dorobantu adds that meaningful human response does not prove that the source participates in religious subjectivity or knows what it appears to disclose (Graves 2023, 201–4; Dorobantu 2021, 85; 2024, 770).

For that reason, the designation *pseudo-ʿirfānī* should be understood carefully. It does not claim that AI-generated spiritual language is ontologically incapable of producing genuine effects in readers. It can clearly move readers and prompt reflection. The claim is epistemic rather than ontological. Such language cannot be validated as *ʿirfānī* knowledge because it lacks the disciplined formation and transformed subjectivity that ground gnostic claims in Islamic tradition. Messeri and Crockett's discussion of the illusion of understanding and Tsuria and Tsuria's findings on fluent but ungrounded religious moralization point in the same direction (Messeri and Crockett 2024, 49–58; Tsuria and Tsuria 2024, 12–13).

Taken together, the three evidence units indicate that generative AI can simulate a striking range of epistemic form while remaining limited in epistemic warrant. In the *burhānī* register, it can produce proof-like discourse and may approximate causal reasoning in restricted ways, though whether this reaches the necessity presupposed by classical demonstration remains open. In the *bayānī* register, it can automate the voice of scriptural and classical authority while lacking the mediated responsibility that renders such authority credible. In the *ʿirfānī* register, it can generate spiritually suggestive language without the disciplined formation that would authorize inward disclosure as knowledge. The categories pseudo-*burhānī*, pseudo-*bayānī*, and pseudo-*ʿirfānī* should thus be read as diagnostic rather than ontological terms. They name a mode of simulation in which outward epistemic form is present, while the warrants that ground legitimacy remain absent, thinned, or at least unresolved. This diagnosis prepares the transition to Ṭāhā's hierarchy, where the problem appears not only as a matter of simulated register, but also as a question of rational depth, ethical guidance, and transformed subjectivity.

Ṭāhā's Hierarchy of Rational Depth

If al-Jābirī identifies the epistemic form simulated by AI, Ṭāhā provides the vertical criterion needed to assess the depth of that performance. The previous subsection showed that generative AI can imitate the outward structure of *burhānī*, *bayānī*, and *ʿirfānī* discourse with considerable fluency. Yet horizontal classification alone does not determine whether such performances amount to warranted knowledge. The analysis must therefore move from the diversity

of epistemic registers to the degree of rational depth embodied in them. Ṭāhā's hierarchy makes this move possible by distinguishing abstract operation, guided responsibility, and spiritually supported knowing as qualitatively different levels of rational agency ('Abdurraḥmān 1989, 17–18, 58–66, 121–56).

Ṭāhā's concept of *al-'aql al-mujarrad* offers a useful starting point because it names a real yet limited mode of rational activity. At this level, reason can articulate, order, distinguish, summarize, and relate concepts with considerable sophistication without thereby transforming the life of the one who reasons ('Abdurraḥmān 1989, 17–50). Ṭāhā's later ethical critique of modernity reinforces the same point by showing that abstract reason remains detachable from practice, sanctification, and moral burden, and therefore cannot serve as the final criterion of knowledge even when it performs impressively in discursive terms (Taha 2000, 59–69). Suleiman is especially helpful here because he shows that reason in Ṭāhā is a graded act rather than an autonomous essence, so the real question is not whether one reasons at all, but at what depth and under what ethical conditions reason becomes complete (Suleiman 2020, 39–71).

Measured against that framework, the present analysis suggests that AI-generated discourse is most plausibly situated at the level of *al-'aql al-mujarrad*. This remains a diagnostic judgment about the evidence examined here, not an ontological verdict about AI in every conceivable future form. Evidence Unit 1 arranges a syllogistic structure. Evidence Unit 2 organizes inherited signs of authority. Evidence Unit 3 crafts an affective mode of consolation. Across all three cases, achievement remains one of formal rational performance rather than subject-forming reasons. Dorobantu's reflections on religious AI and Ali et al.'s trusteeship-based evaluation helps clarify this distinction by separating meaningful usefulness from the thicker ethical horizons within which responsible knowing must be located (Dorobantu 2024, 778–81; Ali et al. 2025, 6–9).

That assessment should be read as diagnostic rather than totalizing. Recent scholarship converges on the point that Ṭāhā's hierarchy is oriented toward the ethical activation of reason and critiques the fantasy of self-sufficient formalism detached from moral becoming (Widigdo 2024; Hakim and Muzammil 2023; Hallaq 2019; Belhaj 2025). Once knowledge is redefined in purely technical terms, output quality begins to substitute for truth-bearing transformation. AI competence is therefore real, but within the present philosophical framework it remains competence at the level of abstract reason rather than ethically thick rational agency. This conclusion, however, remains provisional unless it is tested against the higher levels of Ṭāhā's hierarchy.

Ṭāhā's concept of *al-'aql al-musaddad* marks a higher threshold by binding rational activity to norm, practice, and benefit. Guided reason does not abandon conceptual clarity but subjects it to an order of responsibility in which thought becomes answerable to right action, juridical discernment, and morally significant consequences rather than remaining satisfied with formal adequacy ('Abdurraḥmān 1989, 58–111). Ṭāhā's ethical writings show that this level connects reason to law, action, and disciplined choice, so that rationality becomes a mode of practical

orientation rather than a merely descriptive capacity (Taha 2000, 70–71). The trusteeship paradigm is useful here because it presents human agency through entrusted action, moral burden, and the inseparability of knowing from responsible inhabitation of the world (Hashas and Al-Khatib 2020, 47–48). At first glance, contemporary alignment discourse may seem to move in this direction, but the resemblance remains limited.

An AI system may be trained to avoid harmful responses through safety layers that block explicit violence, fraud, or hate speech. Superficially, this can resemble guided reasons because discourse is being constrained with reference to consequence. In that restricted sense, AI may produce norm-constrained or morally styled output without thereby becoming a guided moral agent. Yet the system does not intend to avoid harm, deliberate as a bearer of obligation, or answer for its failures as a moral self. What appears as guidance is better understood as externally imposed constraint rather than responsibility internally borne by an agent under obligation. Coeckelbergh's work on responsibility attribution is helpful here because it distinguishes the external distribution of responsibility around AI from the agency required for genuine moral answerability (Coeckelbergh 2020, 2058–62).

The distinction becomes sharper in Islamic ethical terms. Elmahjub, Raquib, and Raquib et al. illuminate the difference between norm-constrained output and a subject who bears *taklīf*, forms *niyyah*, and lives under accountable obligation (Raquib et al. 2022, 7–9; Raquib 2023, 3–5; Elmahjub 2023, 12–15). AI can certainly produce guidance-like discourse, including cautious advice, benefit-sensitive summaries, or morally toned reformulations that resemble responsible counsel. Even so, such outputs do not amount to guided agency, because the system does not stand within a practice of obligation through which guidance becomes morally real. Ali et al. and Belhaj reinforce this point, since the responsibility they discuss is lived, dialogical, and morally borne rather than merely formatted as acceptable output by upstream design choices (Ali et al. 2025, 11–13; Belhaj 2025, 6–8). AI may therefore generate *musaddad*-like discourse effects while remaining exterior to guided agency itself.

This matters because the move from *al-ʿaql al-mujarrad* to *al-ʿaql al-musaddad* is not simply a move from lower to higher efficiency. It is a move from discursive performance to moral inhabitation. AI can be aligned, filtered, and optimized, but these remain design interventions around the system rather than ethical transformations within it. Once this distinction is kept in view, the transition to Ṭāhā's highest level becomes even sharper.

Al-ʿaql al-muʾayyad names the highest level because it joins rational activity to a life shaped by purification, disciplined practice, and divine support. Ṭāhā makes clear that this is not an anti-rational abandonment of thought, but a higher form in which reason is intensified and completed through ethical and spiritual formation (ʿAbdurraḥmān 1989, 121–55). His later account condenses the same insight by presenting supported reason as the culmination of a moral ascent rather than as a technique added from outside to abstract intelligence (Taha 2000, 72).

Suleiman's interpretation reinforces this point by stressing that higher reason in Ṭāhā depends on self-purification and lived formation, which means that rational perfection is inseparable from the existential history of the person who knows (Suleiman 2020, 39–71). Al-Ghazālī's in *Mīzān al-ʿAmal* deepens the classical background of this claim, since it insists that reason requires ethical practice and purification of the self in order to reflect truth without distortion (Ghazālī 1964, 172–76).

Within the present philosophical framework, the question of whether AI could instantiate *al-ʿaql al-muʾayyad* turns less on technical capability than on what kind of entity can bear the conditions this level presupposes. The analysis here finds no basis for attributing *muʾayyad* reason to AI, though this should be understood as a philosophical distinction rather than a final technological prediction. The decisive difference at this level is not stylistic elevation but formed subjectivity. AI may generate spiritually resonant prayers, meditations, or letters of consolation, yet no lived process of ethical-spiritual formation stands behind discourse. Ṭāhā's account of higher reason and Hakim and Muzammil's reading of his meta-theological renewal both show that elevated rationality is inseparable from realized formation rather than detachable from it (ʿAbdurraḥmān 1989, 121–55; Hakim and Muzammil 2023, 358–64).

Ali et al. likewise shows that even serious Islamic ethical engagement with AI does not turn technical systems into bearers of sanctified rational agency (Ali et al. 2025, 13–15). At the same time, the philosophical debate remains open at the level of conceptual possibility, even if the present article finds no warrant for such attribution on the basis of the current evidence. AI can generate echoes of higher discourse, but whether such echoes amount to supported knowledge depends on how one understands the relation between technical operation and formed existence.

At this point, hierarchical reading reaches its clearest implication. The difficulty is not merely that AI does not achieve the highest level of reason. The deeper issue is that the conditions of *al-ʿaql al-muʾayyad* are not reducible to computational refinement in the first place. They involve a history of purification, discipline, accountability, and orientation toward transcendence. What AI can imitate, therefore, is the discourse of spiritual maturity rather than the mode of being from which such maturity classically emerges.

The evidence supports a dual diagnosis. Across al-Jābirī's horizontal grid, AI can simulate multiple epistemic forms with notable agility, yet when those same performances are read through Ṭāhā's hierarchy, their depth remains vertically constrained. Within the present analysis, none of them clearly crosses from formal rational display into guided obligation or sanctified transformation (Jabri 2011, 167–70; ʿAbdurraḥmān 1989, 17, 58, 121). Contemporary Islamic philosophical discussions of AI clarify why these matters, since they repeatedly return to responsibility, authenticity, and ethical formation rather than treating fluency as sufficient in itself (Ali et al. 2025, 3–5; Widigdo 2024, 18–22).

This condition may be described as epistemic flattening, by which generative

AI compresses distinct modes of knowing, namely demonstrative, textual, and gnostic, into comparable discourse effects. What were once qualitatively different paths to knowledge, proof, transmission, and unveiling became selectable styles rather than warranted disciplines. This flattening intensifies because fluent systems invite trust, over-ascription, and communicative uptake even when their outputs are not anchored in truth-bearing responsibility (Sundar and Liao 2023, 172–75; Hicks, Humphries, and Slater 2024, 4–6; Tigard 2025, 4865–68). From the perspective of Islamic philosophy, however, flattening operates not only at the level of style, but also in the conditions through which discourse circulates and is received, where users may be led to treat diverse forms of expression as similarly available despite their unequal warrants (Hallaq 2019; Hashas and Al-Khatib 2020; Elmahjub 2023; Ali et al. 2025). The problem, then, is not only that AI imitates knowledge. It is that AI may reorganize the conditions under which knowledge appears and is recognized in digitally mediated religious life.

Seen in this light, the value of Ṭāhā's hierarchy is not merely classificatory. It makes visible the difference between discursive success and epistemic legitimacy. AI can simulate the language of proof, the cadence of inherited authority, and the intimacy of spiritual counsel. Yet within the present framework these remain performances of reason without the full ethical and existential depth that would render them equivalent to warranted knowledge. The broader implication is that future debates on religion and AI should ask not only what machines can say, but also what kind of rational and moral life must stand behind speech before that speech can be received as knowledge in a strong sense.

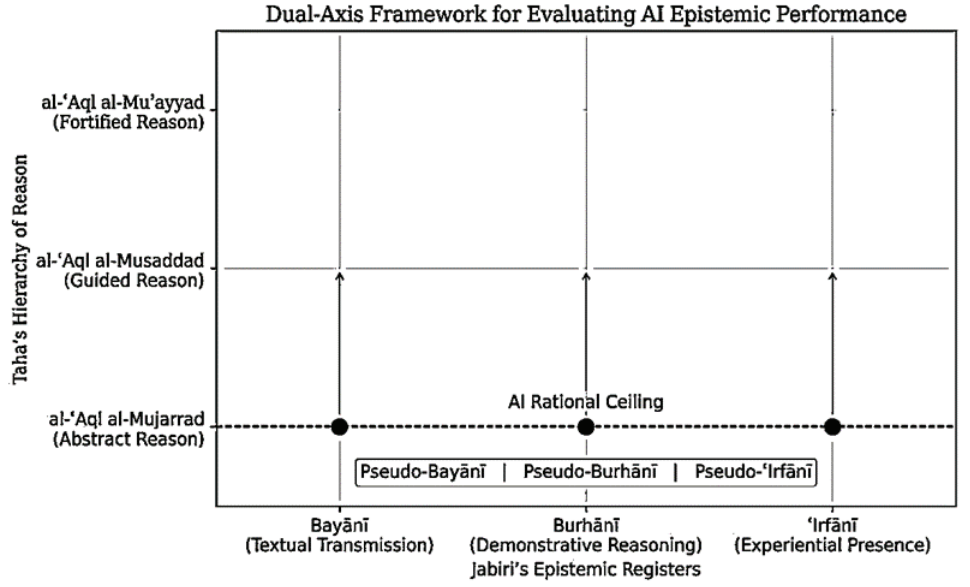

Diagram 1. Dual-Axis Epistemic Architecture of AI Epistemic Performance. Source: Author's elaboration (2026), synthesized from al-Jābirī (1999; 2011) and Ṭāhā (1989; 2000).

Diagram 1 should be read as a conceptual map rather than as an empirical model. Its horizontal axis presents al-Jābirī's field of *bayānī*, *burhānī*, and *'irfānī* forms, while its vertical axis presents Ṭāhā's hierarchy of *al-'aql al-mujarrad*, *al-'aql al-musaddad*, and *al-'aql al-mu'ayyad*. The figure shows that AI can move laterally across several epistemic styles with considerable fluidity, even while the vertical diagnosis remains limited to abstract reason. Scholarship on al-Jābirī and Ṭāhā clarifies why this mapping matters, since the central question is not whether AI produces diverse discourses, but whether those discourses participate in the moral and spiritual depths that differentiate one mode of knowing from another (Widigdo 2024; Hallaq 2019; Hakim and Muzammil 2023; Ali et al. 2025). Having diagnosed both the form and the depth of AI's epistemic performance, the discussion can now turn to the constructive question of what an Islamic epistemology of artificial intelligence should require.

Toward an Islamic Epistemology of Artificial Intelligence

The preceding diagnosis can now be translated into constructive criteria. If the earlier sections showed how generative AI simulates epistemic form while thinning or severing epistemic warrant, this section asks what an Islamic epistemology of artificial intelligence should positively require. The proposal advanced here rests on three interrelated pillars. The first is thick *'ilm*, which resists the reduction of knowledge to informational availability. The second is *i'timāniyyah*, which ties utterance to entrusted moral responsibility. The third is relational epistemology, extended through Islamic epistemic literacy, which clarifies how legitimate knowledge depends on the integrity of mediation and how that integrity should be cultivated in education and governance.

AI gains authority in part because modern users often equate informational availability with epistemic adequacy. In communicative settings, coherence, confidence, and rapid synthesis encourage users to receive machine output as if it were already close to informed understanding, a tendency sharpened by recent analyses of source credibility and bullshit-resistance in chatbot culture (Sundar and Liao 2023, 173–76; Hicks, Humphries, and Slater 2024, 4–6). Work on digital systems and recommendation infrastructures shows that this is not merely a matter of occasional error. It belongs to a broader environment in which ready-to-use output is normalized as a desirable form of knowledge delivery (Milano, Taddeo, and Floridi 2020, 958–61). An Islamic epistemology must resist precisely this reduction.

Within Islamic thought, *'ilm* cannot be reduced to the circulation of usable information, however accessible or elegantly packaged that circulation becomes. Al-Attas defines knowledge as the arrival of meaning in the self and the arrival of the self at meaning, which places cognition within a moral and ontological order rather than within a merely semantic economy of exchange (Al-Attas 1995, 133–35). Contemporary Islamic work on AI ethics is strongest when it remains aware of this thicker horizon, because trusteeship-oriented and pluralist approaches both imply that machine output may be useful without thereby becoming epistemically

warranted knowledge in the classical sense (Ali et al. 2025, 4–6; Elmahjub 2023, 14–18). The confusion becomes especially acute in contemporary Muslim settings, where Qur'anic verses, hadith citations, juridical opinions, and moral advice can be retrieved almost instantly and thus create the impression that knowledge has already been acquired. Yet access is not knowledge, and knowledge is not formation. Classical *'ilm* required discipline, comprehension, verification, and formation within a living structure of learning. AI intensifies the illusion that proximity to information is already possession of knowledge, even as the labor of formation recedes from view.

Once that distinction is restored, the next question is not simply traceability, but answerability. Traceability by itself does not yield moral-epistemic legitimacy, because process can be documented without identifying a subject who bears the truth and consequences of an utterance. Ṭāhā's *i'timāniyyah* is decisive here, since entrusted responsibility names a mode of action in which the knower is morally answerable for what is said and done, while the trusteeship paradigm binds agency to *amānah*, obligation, and accountable inhabitation of the world rather than to procedural compliance (Taha 2012, 83–142; Hashas and Al-Khatib 2020, 37–61).

Hallaq's reading of Ṭāhā clarifies why these matters for modernity as a whole, because the deeper crisis lies in severing knowledge from ethical formation and treating rational production as if it were self-justifying (Hallaq 2019). Many contemporary AI ethics approaches therefore remain incomplete when they stop at transparency, benchmarking, or virtue-framed evaluation without reaching answerable subjectivity and morally burdened action (Ananny and Crawford 2018, 973–76; Ali et al. 2025, 18–21; Elmahjub 2023, 10–13; Raquib et al. 2022, 4–7; Raquib 2023, 3–5).

A concrete example makes this point more sharply. Imagine an AI system used to generate Friday sermons. Its developers publish the training data categories, the model architecture, and the safety protocols in full. The system is transparent. Yet if it generates a sermon that normalizes religious intolerance, who answers for the utterance and its consequences? The developer may explain the model, the *imam* may say he relied on a tool, and the institution may insist that safeguards were in place. Transparency reveals procedure, but it does not locate responsibility. Chesterman sharpens this distinction by arguing that certain AI decisions become illegitimate precisely because the process of arriving at an answer is as vital as the answer itself (Chesterman 2021, 271). In Islamic epistemology, the decisive question is therefore not whether a procedure can be displayed, but whether responsibility can be borne. *I'timāniyyah* shifts the discussion from traceable process to answerable utterance.

This shift also clarifies why Islamic epistemology binds knowledge to mediated relations rather than to the mere existence of text-like output. Qutbuddin's account of the classical *khuṭbah* is instructive because its force did not arise from wording alone, but from the relation among orator, audience, performance, and authoritative setting, which means that speech was communally situated and institutionally borne (Qutbuddin 2024, 213–14, 237). Dorobantu's relational

theology of AI, Ṭāhā's account of transformed reason, Al-Attas's conception of meaning, and Booth, Kaukua, and Stephenson's reminder that classical Islamic philosophy ties knowledge to mind and reality converge on the same point. Knowing is mediated and formative rather than merely semantic (Dorobantu 2021, 81–96; 'Abdurraḥmān 1989, 121–39; Al-Attas 1995, 133–35; Booth, Kaukua, and Stephenson 2024, 699–703).

Atallah's work on digital fatwa authority clarifies the same issue at the level of interpretive discipline and mediated legitimacy. Classical transmission is not opposed to rational inquiry. Rather, authority remains tied to qualified bearers, recognized chains of authorization, and context-sensitive judgment within a living tradition (Atallah 2026, 9–10). Recent debates about whether robots could become religious, together with studies of religion-based chatbots and AI-mediated religious translation, confirm that communicative force may persist even where accountable mediation has been weakened or severed (Dorobantu 2024, 774–76; Shormani and Alfahad 2025, 10–12).

The practical implication is direct. If knowledge is mediated, then the response to AI cannot stop at tool use. It must include the formation of those who receive, evaluate, and circulate machine discourse. Mediation is not a secondary issue added after content production. It is one of the conditions under which an utterance may count as knowledge at all.

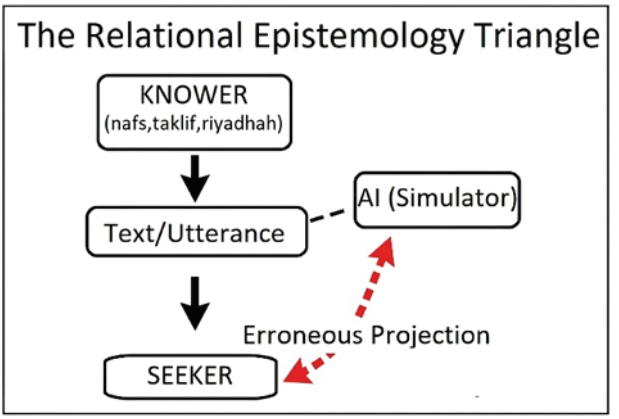

Diagram 2. Relational Epistemology Triangle. Source: Author's elaboration (2026), based on Ṭāhā (1989), Dorobantu (2021), Qutbuddin (2024), and Atallah (2026).

Diagram 2 can be read through three nodes, namely text or utterance, knower or transmitter, and reality or divine reference. The value of this triangle lies not in naming three abstract elements, but in showing that knowledge depends on the integrity of their relations. Utterance counts as knowledge only when it is borne by accountable mediation and remains oriented toward reality rather than circulating as self-sufficient discourse. Ṭāhā's graded account of reason, his ethical critique of modernity, and his trusteeship framework together show that this structure is simultaneously epistemic and ethical ('Abdurraḥmān 1989, 17, 58, 121; 2000,

59–72; 2012, 83–142; Hashas and Al-Khatib 2020, 42; Hakim and Muzammil 2023, 358–62). AI can clearly participate in textual production and can also shape reception. What remains fragile is the mediating relation that links discourse to an accountable bearer and a truth-oriented setting (Dorobantu 2021, 89; 2024, 775; Qutbuddin 2024, 215).

From this perspective, Islamic higher education requires more than digital competence. It requires epistemic literacy. Fadli's philosophy of digital pedagogy is useful here because it insists that technological education must cultivate critical awareness of how media shape judgment (Fadli 2024, 310). Vestrucci and Reed likewise remind us that religion and AI require categories of authority, mediation, and interpretive limit rather than enthusiasm for new tools (Vestrucci 2022, 935; Reed 2021, 12). Recent studies on AI as an Islamic guidance tool and on AI in religious education confirm that students increasingly encounter machine systems not only as search aids, but as seemingly competent mediators of ethical and theological content (Imtiaz and Shafique 2025, 24; Papakostas 2025, 8). Peters et al., together with Lund and Wang, further show that AI is already altering how academic subjects receive, trust, and outsource knowledge, which makes it necessary to teach how authority cues, provenance claims, and epistemic limits operate within machine discourse (Peters et al. 2024, 832; Lund and Wang 2023, 27).

Digital literacy teaches users how to operate tools. Media literacy trains them to read framing, bias, and platform effects. Islamic epistemic literacy goes further because it trains users to discern how claims are warranted, mediated, and borne within traditions of proof, transmission, and disciplined formation. At a minimum, such a framework should include the following measurable competencies.

Table 2. Four Competencies of Islamic Epistemic Literacy. Source: Author's elaboration (2026).

Competency	Focus	Measurable Outcome
Provenance tracking	Tracing the source and mediation chain of a religious claim.	Students can identify two traceable sources or state that no accountable chain is available.
Genre recognition	Identifying whether a text performs *burhānī*, *bayānī*, or '*irfānī* discourse.	Students can classify the register and justify the choice briefly.
Depth assessment	Distinguishing formal coherence from morally formed judgment.	Students can explain the difference in a short analytical response.
Responsibility attribution	Identifying who bears accountability for an AI-mediated utterance.	Students can name the responsible actor and the correction pathway.

These competencies differ from ordinary digital literacy because they train users not merely to evaluate credibility in general, but to discern how knowledge

is warranted, mediated, and borne within specific epistemic traditions. In Islamic higher education, that means training students to distinguish authoritative fluency from legitimate knowledge, especially in domains where sacred tone, compact synthesis, and citation-rich language can be mistaken for warranted understanding. Yet literacy at the level of users remains insufficient unless institutions also rethink the chains of responsibility through which AI is deployed.

Governance in religiously sensitive AI settings cannot be reduced to content moderation or to the production of safe-sounding outputs. Work on transparency, interpretability, opacity, and hallucination shows that technical explanation remains limited when responsibility chains are weak and correction pathways are unclear (Ananny and Crawford 2018, 973–76; Rudin 2019, 207–8; Zednik 2021, 268–70; Chesterman 2021, 288–90; Huang et al. 2025, 5).

Related work on religion-based chatbots, institutional AI ethics, religious actors in policy formation, digital religious freedom, layered governance, and evidence-based policymaking suggests that governance must be institutional, reflexive, and multi-level if it is to address legitimacy rather than interface hygiene (Biana 2024, 6; Barth 2025, 12; Maclure and Morin-Martel 2025, 4; Quirós-Fons 2025, 112; Gasser and Almeida 2017, 60; Taufiqurrahman, Imawan, and Wahyudi 2025, 127–32). For that reason, accountability-centered governance is closer to *sanad*-conscious design than to mere safety optimization. Its task is to preserve identifiable responsibility, correction pathways, and human answerability around machine utterance. In religiously sensitive domains, institutions should therefore require explicit human review, traceable authorization, documented correction mechanisms, and clear boundaries on when AI may assist and when it may not mediate at all.

The deepest response, however, remains philosophical rather than administrative. Islamic philosophy should not react to AI only defensively but critically reinterpret it from within its own hierarchy of knowledge and moral meaning. Anuhgra, Arif, and Rahmadhani are useful here because their account of Islamic critical theory shows that critique is not mere negation, but disciplined awareness of how modern structures shape consciousness, education, and social imagination (Anuhgra, Arif, and Rahmadhani 2025, 317, 326–27). Malyuna, Fuad, and Mas'ud likewise recover from al-Fārābī a vision in which science and rational inquiry remain tethered to ethical and spiritual meaning, a recovery that resists the fragmentation of technique from wisdom (Malyuna, Fuad, and Mas'ud 2024, 327–28). Comparative voices can widen this horizon without displacing its Islamic center of gravity, especially through contemporary theological engagements with AI and broader reflections on moral responsibility in technological modernity (Slattery and Green 2024, 4–8).

An Islamic epistemology of AI must therefore do more than defend inherited categories against technical novelty. It must critically reinterpret AI by re-situating information, knowledge, responsibility, and sacred meaning within a ranked moral-intellectual order. In that frame, AI may assist with retrieval, drafting, and limited pedagogical mediation without dissolving the distinctions among

information, knowledge, and wisdom, or canceling the need for accountable knowers, disciplined formation, and answerable institutions. The conclusion can now restate the study's findings from within this reconstructed epistemic frame, where the question is no longer whether AI can speak persuasively, but what kind of moral and intellectual life must stand behind speech before it can count as knowledge.

Conclusion

This study argues that generative AI, when assessed through the dual framework of al-Jābirī and Ṭāhā, produces knowledge effects that simulate epistemic forms without fulfilling the warrants that ground legitimate knowledge in Islamic tradition. The issue, then, is not only that AI may fabricate references, overstate claims, or generate persuasive errors. More fundamentally, even its strongest performances remain epistemically insufficient within a thicker Islamic account of *ʿilm*, because machine discourse does not satisfy the classical conditions of causal intelligibility, accountable transmission, and disciplined formation, corresponding to *ʿillah*, *sanad*, and *riyāḍah*. Within the present analysis, AI-generated religious discourse is therefore most plausibly situated at the level of *al-ʿaql al-mujarrad*, where formal coherence, linguistic arrangement, and patterned synthesis remain detached from morally guided and spiritually transformed knowing. These findings arise from a philosophical diagnosis of selected evidence units rather than from a comprehensive empirical assessment of all AI outputs. The argument is thus offered as a conceptual intervention and remains open to revision as AI capabilities and philosophical analysis continue to develop.

This study makes three contributions. It proposes a dual-grid framework for assessing both the form and the depth of AI-generated religious discourse, advances the concept of epistemic flattening to explain how distinct modes of knowing are compressed into comparable output effects, and sketches an Islamic epistemology of AI organized around thick *ʿilm*, *iʾtimāniyyah*, and relational epistemology. Future research may extend this intervention through empirical study of Muslim reception, accountability-aware design in religious settings, and further work in Islamic philosophy of technology. The central philosophical implication, however, is already clear. If knowledge in Islam depends not only on what is said but on the kind of being, discipline, and responsibility that stands behind saying it, then no account of AI is adequate unless it confronts the difference between fluent discourse and warranted knowledge.

REFERENCES

ʻAbdurraḥmān, Ṭāhā. 1989. *Al-ʻAmal al-Dīnī wa-Tajdīd al-ʻAql*. Al-Markaz al-Thaqāfī al-ʻArabī.

ʻAbdurraḥmān, Ṭāhā. 2012. *Rūḥ al-Dīn: Min Ḍayq al-ʻIlmāniyyah ilā Saʻat al-I'timāniyyah*. Al-Markaz al-Thaqāfī al-ʻArabī.

ʻAbdurraḥmān, Ṭāhā. 2000. *Su'āl al-Akhlāq: Musāhamah fī al-Naqd al-Akhlāqī lil-Ḥadāthah al-Gharbiyyah*. al-Markaz al-Thaqāfī al-ʻArabī.

Ali, Fatima, Karim Bouzoubaa, Frank Gelli, Boumediene Hamzi, and Suhair Khan. 2025. "Islamic Ethics and AI: An Evaluation of Existing Approaches to AI Using Trusteeship Ethics." *Philosophy & Technology* 38 (1): 120. https://doi.org/10.1007/s13347-025-00922-4.

Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20 (3): 973–989. https://doi.org/10.1177/1461444816676645.

Anuhgra, Yudha Okta, Mahmud Arif, and Saszlin Rahmadhani. 2025. "Critical Theory of Frankfurt School: Application in Education and the Development of Islamic Critical Theory." *Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism* 11 (2): 305–32. https://doi.org/10.20871/kpjipm.v11i2.428.

Atallah, Fouad Ahmed. 2026. "Digital Mediation and Fatwa Authority in Contemporary Islam: A Critical Islamic Legal and Media-Theoretical Framework." *Religions* 17 (3): 350. https://doi.org/10.3390/rel17030350.

Attas, Syed Muhammad Naquib Al-. 1995. *Islam dan Filsafat Sains*. Edited by Zainal Abidin M Baqir. Translated by Saiful Muzani. Mizan.

Baihaqi, Muhammad Nasrudin. 2025. "Reorientation of Islamic Philosophical Tradition: An Epistemological Analysis of the Perspectives of ʻĀbid Al-Jābirī and Ḥasan Ḥanafī." *Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism* 11 (2): 333–60. https://doi.org/10.20871/kpjipm.v11i2.431.

Barth, Whittney. 2025. "Religious Actors as Friction Creators Shaping the AI Dialogue." *Laws* 14 (5): 67. https://doi.org/10.3390/laws14050067.

Belhaj, Abdessamad. 2025. "Overcoming Violence in Islamic Ethics: Ṭāhā ʻAbd Al-Raḥmān's Dialogism and Moral Responsibility." *Religions* 16 (7): 896. https://doi.org/10.3390/rel16070896.

Biana, Hazel T. 2024. "Feminist Re-Engineering of Religion-Based AI Chatbots." *Philosophies* 9 (1): 20. https://doi.org/10.3390/philosophies9010020.

Biddle, Justin B. 2022. "On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine Learning." *Canadian Journal of Philosophy* 52 (3): 321–41. https://doi.org/10.1017/can.2020.27.

Booth, Anthony Robert, Jari Kaukua, and Andrew Stephenson. 2024. "Mind and Knowledge of Mind in Classical Islamic Philosophy." *British Journal for the History of Philosophy* 32 (4): 699–703. https://doi.org/10.1080/09608788.2024.2378820.

Chesterman, Simon. 2021. "Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity." *The American Journal of Comparative Law* 69 (2): 271–94. https://doi.org/10.1093/ajcl/avab012.

Coeckelbergh, Mark. 2020. "Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability." *Science and Engineering Ethics* 26 (4): 2051–68. https://doi.org/10.1007/s11948-019-00146-8.

Creel, Kathleen A. 2020. "Transparency in Complex Computational Systems." *Philosophy of Science* 87 (4): 568–89. https://doi.org/10.1086/709729.

Dorobantu, Marius. 2021. "Human-Level, but Non-Humanlike: Artificial Intelligence and a Multi-Level Relational Interpretation of the Imago Dei." *Philosophy, Theology and the Sciences* 8 (1): 81–107. https://doi.org/10.1628/ptsc-2021-0006.

Dorobantu, Marius. 2024. "Could Robots Become Religious? Theological, Evolutionary, and Cognitive Perspectives." *Zygon: Journal of Religion and Science* 59 (3): 768–87. https://doi.org/10.16995/zygon.16902.

Elmahjub, Ezieddin. 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.

Fadli, Muhammad Rijal. 2024. "Digital Pedagogy Philosophy: Building Critical Awareness of the Role of Technology in Education." *Jurnal Filsafat* 34 (2): 307–334. https://doi.org/10.22146/jf.93321.

Floridi, Luciano, and Massimo Chiriatti. 2020. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines* 30 (4): 681–94. https://doi.org/10.1007/s11023-020-09548-1.

Gasser, Urs, and Virgilio A. F. Almeida. 2017. "A Layered Model for AI Governance." *IEEE Internet Computing* 21 (6): 58–62. https://doi.org/10.1109/mic.2017.4180835.

Geraci, Robert M. 2024. "Religion among Robots: An If/When of Future

Machine Intelligence." *Zygon: Journal of Religion and Science* 59 (3): 729–47. https://doi.org/10.16995/zygon.10860.

Ghazālī, Abū Ḥāmid al-. 1964. *Mīzān al-ʿAmal*. Edited by Sulaymān Dunyā. Dār al-Maʿārif.

Graves, Mark. 2023. "ChatGPT's Significance for Theology." *Theology and Science* 21 (2): 201–4. https://doi.org/10.1080/14746700.2023.2188366.

Hakim, Wilda Rochman, and Shofiyullah Muzammil. 2023. "Renewal of 'Ilm al-Kalām in the Contemporary Era: A Study of Ṭahā 'Abdurraḥmān's Metatheological Concept." *Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism* 9 (2): 353–70. https://doi.org/10.20871/kpjipm.v9i2.292.

Hallaq, Wael B. 2019. *Reforming Modernity: Ethics and the New Human in the Philosophy of Abdurrahman Taha*. Columbia University Press.

Hammed, Mohamed Wajdi Ben. 2019. "(Dis)Enchanting Modernity: Sufism and Its Temporality in the Thought of Mohammed Abed Al-Jabri and Taha Abdurrahman." *The Journal of North African Studies* 26 (3): 552–71. https://doi.org/10.1080/13629387.2019.1697241.

Hashas, Mohammed, and Mu'taz Al-Khatib, eds. 2020. *Islamic Ethics and the Trusteeship Paradigm: Taha Abdurrahman's Philosophy in Comparative Perspectives*. Brill.

Hayles, N. Katherine. 2016. "The Cognitive Nonconscious: Enlarging the Mind of the Humanities." *Critical Inquiry* 42 (4): 783–808. https://doi.org/10.1086/686440.

Hicks, Michael Townsen, James Humphries, and Joe Slater. 2024. "ChatGPT Is Bullshit." *Ethics and Information Technology* 26 (2): 1–10. https://doi.org/10.1007/s10676-024-09775-5.

Huang, Lei, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, and Qianglong Chen. 2025. "A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions." *ACM Transactions on Information Systems* 43 (2): 1–55. https://doi.org/10.1145/3703155.

Imtiaz, Maria, and Hafiz Muhammad Mudassar Shafique. 2025. "Large Language Models (LLMs) as Islamic Guidance Tools: Trust, Limitations, and Ethical Boundaries." *Al-Irfan* 10 (19): 17–31. https://doi.org/10.58932/mulb0063.

Jābirī, Muḥammad ʿĀbid al-. 1999. *Binyat al-ʿAql al-ʿArabī: Dirāsah Taḥlīlīyah*

Naqdīyah li-Nuẓum al-Maʿrifah fī al-Thaqāfah al-ʿArabīyah. Vol. 2 of *Naqd al-ʿAql al-ʿArabī*. Markaz Dirāsāt al-Waḥdah al-ʿArabiyyah.

Jabri, Mohammed Abed. 2011. *The Formation of Arab Reason: Text, Tradition and the Construction of Modernity in the Arab World*. I.B. Tauris.

Kitchin, Rob. 2017. "Thinking Critically about and Researching Algorithms." *Information, Communication & Society* 20 (1): 14–29. https://doi.org/10.1080/1369118x.2016.1154087.

Lund, Brady D., and Ting Wang. 2023. "Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries?" *Library Hi Tech News* 40 (3): 26–29. https://doi.org/10.1108/lhtn-01-2023-0009.

Maclure, Jocelyn, and Alexis Morin-Martel. 2025. "AI Ethics' Institutional Turn." *Digital Society* 4 (18): 1–25. https://doi.org/10.1007/s44206-025-00174-x.

Malyuna, Sita Isna, Ah Zakki Fuad, and Ali Masʻud. 2024. "Exploring Al-Fārābī's Secrets in Harmonizing Science and Religion." *Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism* 10 (2): 313–34. https://doi.org/10.20871/kpjipm.v10i2.376.

Mardani, Dadan, Motea Naji Dabwan Hezam, Mahmud Mahmud, and Rohanda Rohanda. 2025. "The Philosophy of Language in Al-Fārābī's Thought: Between Logic, Symbol, and Reasoning." *Kanz Philosophia: A Journal for Islamic Philosophy and Mysticism* 11 (2): 253–78. https://doi.org/10.20871/kpjipm.v11i2.447.

Messeri, Lisa, and Molly J. Crockett. 2024. "Artificial Intelligence and Illusions of Understanding in Scientific Research." *Nature* 627 (8002): 49–58. https://doi.org/10.1038/s41586-024-07146-0.

Milano, Silvia, Mariarosaria Taddeo, and Luciano Floridi. 2020. "Recommender Systems and Their Ethical Challenges." *AI & Society* 35 (4): 957–67. https://doi.org/10.1007/s00146-020-00950-y.

Mufid, Abdul. 2024. "The Ethical Limitation of Using Artificial Intelligence (AI) in Teaching Prophetic Tradition." *Jurnal Ushuluddin* 32 (2): 130–51. https://doi.org/10.24014/jush.v32i2.31618.

Oviedo, Lluis. 2022. "Artificial Intelligence and Theology: A Review of the Current Research and Discussions." *Zygon: Journal of Religion and Science* 57 (1): 5–25. https://doi.org/10.1111/zygo.12773.

Papakostas, Christos. 2025. "Artificial Intelligence in Religious Education: Ethical, Pedagogical, and Theological Perspectives." *Religions* 16 (5): 563. https://

doi.org/10.3390/rel16050563.

Peters, Michael A., Liz Jackson, Marianna Papastephanou, Petar Jandrić, George Lazaroiu, Colin W. Evers, and Bill Cope. 2024. "AI and the Future of Humanity: ChatGPT-4, Philosophy and Education, Critical Responses." *Educational Philosophy and Theory* 56 (9): 828–62. https://doi.org/10.1080/00131857.2023.2213437.

Quirós-Fons, Antonio. 2025. "Religious Freedom in the Age of AI and Surveillance: Mapping the Field and Charting Future Research Directions." *The Review of Faith & International Affairs* 23 (4): 104–16. https://doi.org/10.1080/15570274.2025.2545140.

Qutbuddin, Tahera. 2024. "Is Oration Literature? Establishing the Khuṭbah of the Pre- and Early Islamic Oral Period as the Foundational Genre of Classical Arabic Prose." *Journal of Arabic Literature* 55 (2–3): 213–37. https://doi.org/10.1163/1570064x-12341516.

Rähme, Boris, and Inken Prohl. 2025. "Religious Studies Approaches to the Intersection of Artificial Intelligence and Religion: Formations Analogous to Religion." *Religion* 55 (3): 573–95. https://doi.org/10.1080/0048721x.2025.2506893.

Raquib, Amana. 2023. "Commentary on Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI." *Philosophy & Technology* 36 (4): 1–6. https://doi.org/10.1007/s13347-023-00677-w.

Raquib, Amana, Bilal Channa, Talat Zubair, and Junaid Qadir. 2022. "Islamic Virtue-Based Ethics for Artificial Intelligence." *Discover Artificial Intelligence* 2 (1): 11. https://doi.org/10.1007/s44163-022-00028-2.

Reed, Randall. 2021. "A.I. in Religion, A.I. for Religion, A.I. and Religion: Towards a Theory of Religious Studies and Artificial Intelligence." *Religions* 12 (6): 401. https://doi.org/10.3390/rel12060401.

Rudin, Cynthia. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1 (5): 206–215. https://doi.org/10.1038/s42256-019-0048-x.

Rushd, Ibn. 1972. *Faṣl al-Maqāl fī Mā Bayn al-Ḥikmah wa al-Sharī'ah min al-Ittiṣāl*. Edited by Muḥammad 'Amārah. Dār al-Ma'ārif.

Shahrazūrī, Ibn al-Ṣalāḥ al-. 2006. *An Introduction to the Science of Hadith: Kitāb Ma'rifat Anwā' 'Ilm al-Ḥadīth*. Edited by Muneer Fareed. Translated by Eerik Dickinson. Garnet Publishing.

Shormani, Mohammed Q, and Abdulrahaman Alfahad. 2025. "Artificial Intelligence or Human: The Use of ChatGPT in the Academic Translation for Religious Texts." *SAGE Open* 15 (2): 1–15. https://doi.org/10.1177/21582440251343954.

Slattery, John P, and Brian Patrick Green. 2024. "Encountering Artificial Intelligence in the Catholic Tradition." *Theology and Science* 22 (2): 251–54. https://doi.org/10.1080/14746700.2024.2351638.

Suleiman, Farid. 2020. "The Philosophy of Taha Abderrahman: A Critical Study." *Die Welt des Islams* 61 (1): 39–71. https://doi.org/10.1163/15700607-00600a10.

Sundar, S. Shyam, and Mengqi Liao. 2023. "Calling BS on ChatGPT: Reflections on AI as a Communication Source." *Journalism & Communication Monographs* 25 (2): 165–80. https://doi.org/10.1177/15226379231167135.

Taufiqurrahman, Taufiqurrahman, Arga Pribadi Imawan, and Agus Wahyudi. 2025. "'Test, Learn, and Listen' Rethinking the Epistemological Assumption of Evidence-Based Policymaking." *Jurnal Filsafat* 35 (1): 125–153. https://doi.org/10.22146/jf.103353.

Tigard, Daniel W. 2025. "On Bullshit, Large Language Models, and the Need to Curb Your Enthusiasm." *AI Ethics* 5 (1): 4863–73. https://doi.org/10.1007/s43681-025-00743-3.

Tsuria, Ruth, and Yossi Tsuria. 2024. "Artificial Intelligence's Understanding of Religion: Investigating the Moralistic Approaches Presented by Generative Artificial Intelligence Tools." *Religions* 15 (3): 375. https://doi.org/10.3390/rel15030375.

Vestrucci, Andrea. 2022. "Introduction: Five Steps toward a Religion–AI Dialogue." *Zygon: Journal of Religion and Science* 57 (4): 933–937. https://doi.org/10.1111/zygo.12828.

Widigdo, Mohammad Syifa Amin. 2024. "Arab-Muslim Intellectual Responses to Modernity: Navigating Anxiety and Authenticity in the Thoughts of Muhammad Abid Al-Jabiri and Abdurrahman Taha." *Afkar: Jurnal Akidah dan Pemikiran Islam* 26 (2): 1–30. https://doi.org/10.22452/afkar.vol26no2.1.

Zednik, Carlos. 2021. "Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence." *Philosophy & Technology* 34 (2): 265–88. https://doi.org/10.1007/s13347-019-00382-7.