https://ojs.lakaspia.org/index.php/chronicles

Citation Al-Hafi, Aban, and Rahmad Hidayat Ajrul Iman. "Source Authorization and the Limits of Delegating Qur'anic Interpretive Authority in AI Mediation." *Chronicles of Qur'anic Studies* 1, no. 1 (2026): 33–50.

Chronicles of Qur'anic Studies

Source Authorization and the Limits of Delegating Qur'anic Interpretive Authority in AI Mediation

Abstract

This article sets out to define the conceptual boundary between authorized source mediation and the delegation of interpretive authority. It employs an integrative conceptual synthesis, drawing purposively on scholarship in Qur'anic exegesis, digital religion, human–machine communication, automation, and AI accountability. The analysis compares units of claim, distinguishes levels of evidence, and tests the resulting classification against hypothetical comparative cases offered strictly as illustrations. Three distinctions follow: the legitimacy of a source is not the same as the legitimacy of an answer; computational discretion is not the same as an interpretive mandate; and influence over users is not the same as institutional delegation. Delegation is framed as a relationship that requires an identifiable mandating party, room for interpretive decision-making, and recognition of the output's standing on behalf of that party. Control and correction are treated as accountability dimensions to be assessed on their own terms. The article proposes a claim-type evaluation that separates verse retrieval, reproduction of exegesis, synthesis of views, and normative application. Its contribution is an evidentiary framework that keeps technical capability, user trust, and religious legitimacy from being conflated. The framework is the product of conceptual analysis; applying it to a particular institution or application will require operational and reception evidence not tested here.

Keywords: ***Qur'anic interpretation; artificial intelligence; religious authority; tafsir.***

Authors:

Aban Al-Hafi[1]*
[1] Universitas PTIQ Jakarta, Indonesia
* Corresponding author:
E-mail:
abanalhafi@mhs.ptiq.ac.id

Rahmad Hidayat Ajrul Iman[2]
[2] Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia
E-mail:
rahmathidayatai98@gmail.com

Article history Received 6 April 2026 · Revised 17 April 2026 · Accepted 29 May 2026 · Published 7 June 2026
https://doi.org/10.xxxxx/cqs.v1i.000
License © 2026 the Author(s). Published under a Creative Commons Attribution–ShareAlike 4.0 International (CC BY-SA 4.0) license.

Introduction

The shift from searching for references toward conversational answers changes how readers encounter Qur'anic interpretation. A user may receive an explanation that has already been selected and assembled before ever examining the works on which it rests. The institutional relevance of this issue is visible in a 2024 LPMQ publication describing plans to integrate tahlili, wajiz, and thematic tafsir into Chat Qurani; the same publication also notes that development and data population remain unfinished.[1] This description signals that the service is oriented toward the institution's scholarly products, but it does not establish the generative architecture, the quality of the answers, or the procedures used to authorize them. On the technical side, Eltanbouly and colleagues tested how language models retrieve portions of the Qur'an in a search task, so what they measure differs from any recognition of the authority to interpret.[2] Where these two domains meet, a basic problem comes into view: the availability of technology that can answer questions and the availability of recognized sources still do not tell us who is responsible for the meaning of an answer. Research on authority needs to examine that relationship before concluding that the use of AI has displaced the position of the mufassir.

The relevant literature offers complementary tools that nonetheless operate at different levels of analysis. Pink shows that contemporary tafsir can be distinguished by mode of authorship and purpose—including works by scholars, institutional works, and works that popularize explanation—while Saleh underscores how historiography and publication history shape any reading of the exegetical tradition.[3] In the field of digital religion, Campbell draws attention to questions of authority online, while Campbell and Evolvi situate the interplay of online and offline practice—including algorithms and bots—within the field's development.[4] Cheong emphasizes the role of communication in reshaping

---

[1] Bagus Purnomo, "LPMQ Susun Roadmap Kajian Al-Qur'an dan Pengembangan Aplikasi Lima Tahun Ke Depan," Lajnah Pentashihan Mushaf Al-Qur'an, 19 September 2024, https://lajnah.kemenag.go.id/info-lpmq/berita-dan-artikel/berita/lpmq-susun-roadmap-kajian-al-qur-an-dan-pengembangan-aplikasi-lima-tahun-ke-depan.html.

[2] Sohaila Eltanbouly et al., "Can LLMs Directly Retrieve Passages for Answering Questions from Qur'an?," in *Proceedings of The Third Arabic Natural Language Processing Conference* (Suzhou, China: Association for Computational Linguistics, 2025), 203–210, https://doi.org/10.18653/v1/2025.arabicnlp-main.16.

[3] Johanna Pink, "Tradition, Authority and Innovation in Contemporary Sunnī Tafsīr: Towards a Typology of Qur'an Commentaries from the Arab World, Indonesia and Turkey," *Journal of Qur'anic Studies* 12, no. 1–2 (2010): 56–82, https://doi.org/10.3366/jqs.2010.0105. Walid A. Saleh, "Preliminary Remarks on the Historiography of Tafsīr in Arabic: A History of the Book Approach," *Journal of Qur'anic Studies* 12, no. 1–2 (2010): 6–40, https://doi.org/10.3366/jqs.2010.0103.

[4] Heidi Campbell, "Who's Got the Power? Religious Authority and the Internet," *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 1043–1062, https://doi.org/10.1111/j.1083-6101.2007.00362.x. Heidi A. Campbell and Giulia Evolvi, "Contextualizing Current Digital Religion Research on Emerging Technologies," *Human Behavior and Emerging Technologies* 2, no. 1 (2020): 5–17, https://doi.org/10.1002/hbe2.149.

religious authority, while Guzman and Lewis offer a framework for understanding AI technology as part of a communicative relationship.[5] This body of work dispels the assumption that technology is merely a neutral container with no bearing on knowledge. Even so, the effect of mediation on the structure of communication should not be equated outright with the right to authorize an interpretation. What is needed is an analytical bridge that links the history of the source, the operation of the system, and the standing of the output without collapsing the differences among the three.

The gap this article addresses lies in the evidentiary threshold that applies when the term authority is used to describe AI-based tafsir services. The automation literature distinguishes the use of technology from inappropriate reliance on it, while human–machine communication scholarship examines how systems come to be treated as communicators.[6] Both perspectives help explain the interaction, but neither on its own settles whether an institution has surrendered interpretive authority. At the same time, work on AI in the study of religion is a reminder that technology has many uses—including classifying research material—that need not produce a new form of religious authority.[7] The problem is not a shortage of theory about digital religion or AI, but the risk of a category jump: endorsing a book is treated as endorsing an answer, computational latitude as a mandate, and user compliance as institutional legitimacy. This jump can produce two opposite errors—overstating a transfer of authority, or assuming technology has no effect so long as a human remains listed as responsible. This article treats keeping the categories distinct as a precondition for testing both explanations in the open.

This study asks: under what conditions can AI mediation of exegetical sources be distinguished from the delegation of interpretive authority, and what evidence is needed to recognize the difference? Its central argument is that source authorization, the performance of interpretive work, and recognition of an answer's standing are relationships that can intertwine without becoming identical. Delegation, therefore, is not adequately proven by an official label, the ability to compose an answer, or the presence of a human overseer. The article develops three elements that constitute a delegation relationship—mandate, interpretive discretion, and the standing of the output on behalf of the mandating party—while the quality of control and correction is placed under accountability. The framework of

---

[5] Pauline Hope Cheong, "The Vitality of New Media and Religion: Communicative Perspectives, Practices, and Changing Authority in Spiritual Organization," *New Media & Society* 19, no. 1 (2017): 25–33, https://doi.org/10.1177/1461444816649913. Andrea L Guzman and Seth C Lewis, "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda," *New Media & Society* 22, no. 1 (2020): 70–86, https://doi.org/10.1177/1461444819858691.

[6] Raja Parasuraman and Victor Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39, no. 2 (1997): 230–253, https://doi.org/10.1518/001872097778543886. Guzman and Lewis, "Artificial Intelligence and Communication."

[7] Randall Reed, "A.I. in Religion, A.I. for Religion, A.I. and Religion: Towards a Theory of Religious Studies and Artificial Intelligence," *Religions* 12, no. 6 (2021): 401, https://doi.org/10.3390/rel12060401.

meaningful human control helps address this last element, but it is not treated as a definition of interpretive authority.[8] The intended contribution is an analytical specification that later research can examine and contest. The article does not rule on the theological fitness of AI as a mufassir, nor does it report that delegation has occurred at LPMQ; it provides the tools to assess such claims with a commensurate kind of evidence.

Method

This study adopts a conceptual-article design based on integrative literature synthesis[9] to examine the relationships among source authorization, interpretive discretion, recognition of outputs, and accountability in AI-mediated Qur'anic explanation. Following Jaakkola and Snyder, literature was selected purposively according to its analytical relevance rather than through an exhaustive systematic-review protocol.[10] The corpus draws on studies of tafsir and religious authority, digital religion and human–machine communication, and automation, organizations, and AI governance. Sources were identified and cross-checked through thematic searches, reference lists, publisher pages, DOI/Crossref records, and academic repositories, with source verification conducted through 1 May 2026. Institutional publications were used only to establish documented factual context and not to infer technical operations, interpretive mandates, or oversight practices that they did not explicitly report.

The analysis compares the scope and relationships of the principal constructs by distinguishing source legitimacy from answer fidelity, computational discretion from delegated authority, user recognition from institutional recognition, and the existence of delegation from the adequacy of oversight. Documentary claims were interpreted within their evidentiary limits, consistent with Bowen's caution that institutional documents do not by themselves establish actual implementation.[11] Conceptual illustrations were used to test the boundaries of these distinctions and were not treated as observational evidence of actual AI applications.

Results

The first result is a distinction between the legitimacy of the corpus and the legitimacy of the output. Selecting recognized works of tafsir determines which material is

[8] Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account," *Frontiers in Robotics and AI* 5 (2018): 15, https://doi.org/10.3389/frobt.2018.00015.

[9] Hannah Snyder, "Literature Review as a Research Methodology: An Overview and Guidelines," *Journal of Business Research* 104 (2019): 333–339, https://doi.org/10.1016/j.jbusres.2019.07.039

[10] Elina Jaakkola, "Designing Conceptual Articles: Four Approaches," *AMS Review* 10, nos. 1–2 (2020): 18–26, https://doi.org/10.1007/s13162-020-00161-0

[11] Glenn A. Bowen, "Document Analysis as a Qualitative Research Method," *Qualitative Research Journal* 9, no. 2 (2009): 27–40, https://doi.org/10.3316/QRJ0902027.

fit to cite, but it does not determine whether a summary or a new answer represents that material accurately. The basis for this distinction emerges when scholarship on the exegetical tradition is read alongside research on language generation: works of tafsir carry different styles and relations of authority, while text-generating systems can produce statements that are unfaithful to the source.[12] From this relationship, the article derives the principle that endorsing an input is not transitive to every output. The principle holds even when the corpus comes from a trusted institution, because the operations of selecting, excerpting, and reassembling are additional steps. What must be checked is not only whether the source is legitimate, but whether the output's claim can be tied to the source within the same scope. The corpus's reputation thus becomes one condition of assessment, not a substitute for scrutiny of the answer.

The next distinction concerns four types of work that generate different evidentiary demands: verse retrieval, reproduction of a source's explanation, synthesis of views, and normative application. Retrieval is judged by how well the located passage fits the question; reproduction by the fidelity of a paraphrase or quotation; synthesis by the accuracy of the relationships drawn among views; and normative application by the sufficiency of the reasons for connecting meaning to a particular situation. Eltanbouly and colleagues operate on the first type, so their results cannot serve as evidence of success across all the others.[13] This classification is an analytical product of the article, not a ladder of capabilities that must be climbed in sequence. A single answer may contain several types of work at once. Consequently, an evaluation that assigns only one overall score can hide where the problem lies: the verse located may be relevant, yet the application added to it may lack sufficient grounding. The more appropriate unit of scrutiny is the relationship between each claim and the work that produced it.

Source authorization is therefore framed as a bounded relationship among the party that selects the corpus, the works selected, and the purpose of their use. Endorsement for general educational purposes, for instance, is not by itself an endorsement for resolving an individual's normative question; this example is hypothetical. The publication history and categorization of tafsir also warrant attention so that the corpus is not treated as a collection of explanations without a history.[14] The data documentation proposed by Gebru and colleagues offers a technical point of convergence for recording the provenance, composition, and usage context of material.[15] Synthesizing these two concerns yields a requirement that any description of an exegetical corpus specify the works and editions, the

[12] Pink, "Tradition, Authority and Innovation in Contemporary Sunnī Tafsīr." Ziwei Ji et al., "Survey of Hallucination in Natural Language Generation," *ACM Computing Surveys* 55, no. 12 (2023): 1–38, https://doi.org/10.1145/3571730.

[13] Eltanbouly et al., "Can LLMs Directly Retrieve Passages for Answering Questions from Qur'an?"

[14] Saleh, "Preliminary Remarks on the Historiography of Tafsīr in Arabic."

[15] Timnit Gebru et al., "Datasheets for Datasets," *Communications of the ACM* 64, no. 12 (2021): 86–92, https://doi.org/10.1145/3458723.

style of explanation, the purpose of selection, and the limits of its representativeness. Documentation, however, is not a certificate of theological truth. Its function is to make visible the choices that shape the space of possible answers and to let a reviewer identify which parts of the tradition are available, unavailable, or deliberately excluded. Corpus transparency aids assessment without removing the need for human interpretation.

The second result is the separation of computational discretion from an interpretive mandate. Computational discretion refers to the room for choice within the process of producing outputs—for instance, ranking passages of text or stringing together several explanations. An interpretive mandate refers to the relationship that gives those choices the standing of work permitted on behalf of a particular party. The two differ logically: a system can choose technically without acquiring the right to authorize an explanation, and an institution can assign a task even where the technology's room for choice is narrow. Burrell's work on opacity shows that the difficulty of understanding a system takes several forms; the invisibility of an operation therefore cannot be used as a measure of the extent of authority.[16] The resulting framework calls for separate scrutiny of what the system can do, what the operator permits, and the standing of the output produced. A delegation classification becomes weak if only one of these questions is answered. The distinction also keeps the AI label from being treated as a full account of an institutional relationship that in fact still needs to be investigated.

Mandate, interpretive discretion, and the standing of the output are the three elements that together form this article's working definition of delegating interpretive authority. A mandate requires an identifiable authorizing party and an identifiable scope of tasks. Interpretive discretion requires room to select, connect, or apply explanations within the limits of that task. The standing of the output requires that the output acquire a certain status on behalf of the mandating party, rather than appearing merely as conversational text. This definition is an analytical proposal, not terminology claimed to be settled within Qur'anic studies. Nor does it require that every answer be legally binding: an educational explanation can hold institutional standing without resembling a fatwa. Conversely, the normative character of a sentence does not prove that an institution has endorsed it. Nyholm's discussion of agency attribution and the location of responsibility helps preserve the distinction between a system's role in joint work and accountability for an action.[17] Delegation in this framework therefore denotes a sociotechnical relationship, not an acknowledgment that a language model possesses moral personhood.

---

[16] Jenna Burrell, "How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* 3, no. 1 (2016): 2053951715622512, https://doi.org/10.1177/2053951715622512.

[17] Sven Nyholm, "Attributing Agency to Automated Systems: Reflections on Human–Robot Collaborations and Responsibility-Loci," *Science and Engineering Ethics* 24, no. 4 (2018): 1201–1219, https://doi.org/10.1007/s11948-017-9943-x.

Testing through comparative cases yields three classifications that must be kept distinct. A hypothetical service that merely redisplays official explanations, with no room to compose new decisions, falls under authorized source mediation. A service that produces interpretations and is trusted by users without an institutional mandate demonstrates interpretive influence or user recognition, but does not yet meet the definition of institutional delegation. A service that is tasked with composing explanations within set limits, holds discretion, and issues outputs with standing on behalf of the mandating party meets the working definition of delegation; its normative fitness still requires a separate assessment. These three illustrations are not the product of observing applications. Their use is to show that similar outputs can arise from different relations of authority. Conversely, the same provider can offer functions with different statuses. The unit of classification is thus not the chatbot brand or the institution as a whole, but the function, the type of claim, and the authorizing relationship that can be traced in a particular use.

The third result is the separation of influence, trust, and institutional recognition. Influence refers to a change in judgment or behavior; trust refers to an attitude that can underpin reliance; institutional recognition refers to the standing granted to an output within an organizational relationship. The literature on trust in automation shows that reliance requires judgment calibrated to the character of the system and its context.[18] Krügel, Ostermaier, and Uhl show that ChatGPT's moral advice can influence participants' judgments in their experiment; this finding is not evidence about interpretive authority or the behavior of Indonesian Muslims.[19] The conceptual implication is that a system can be influential without a mandate, while an institutionally endorsed service is not necessarily trusted by its users. Usage data or agreement with answers is therefore insufficient to identify the source of legitimacy. Research seeking to assess a change in authority must specify the form of recognition under study, who confers it, and how that recognition relates to compliance and verification practices.

The existence of delegation must also be separated from the quality of its accountability. An organization can, in principle, delegate interpretive work while providing weak oversight; that case is a possible instance of unaccountable delegation, not an automatic reason to declare that no delegation exists. This separation prevents a circular definition that recognizes only the ideal case. Santoni de Sio and van den Hoven characterize meaningful control through responsiveness to relevant reasons and traceability to humans;

[18] J. D. Lee and K. A. See, "Trust in Automation: Designing for Appropriate Reliance," *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, no. 1 (2004): 50–80, https://doi.org/10.1518/hfes.46.1.50_30392.

[19] Sebastian Krügel, Andreas Ostermaier, and Matthias Uhl, "ChatGPT's Inconsistent Moral Advice Influences Users' Judgment," *Scientific Reports* 13, no. 1 (2023): 4569, https://doi.org/10.1038/s41598-023-31341-0.

Elish shows the risk of loading responsibility onto an operator whose control is limited.[20] From these two concerns, the article derives that any accountability check must include a reviewer's capacity to understand the task, correct the output, and influence the procedure. Listing an expert on the team does not prove that capacity. What is needed is the relationship among competence, information, the opportunity to intervene, and the authority to make corrections. Human oversight thus becomes an object of substantive testing, not an administrative mark that closes the evaluation.

Taken together, the three results yield a different evidentiary threshold for each type of claim. A source-authorization claim requires the identity of the corpus and of the party that establishes its use. A discretion claim requires evidence about the system's room for choice. A delegation claim requires the linkage of mandate, interpretive choice, and the standing of the output. An accountability claim requires evidence about control, traceability, and correction. A social-acceptance claim requires data from those who use or recognize the service. Model documentation can help explain purpose, evaluation, and limits of use, but it does not substitute for the full body of that evidence.[21] Where an examined document does not address a given element, the correct status is undocumented in that material. This status must be distinguished from an explicit denial and from evidence that a mechanism does not exist. The resulting framework lets researchers withhold conclusions when the evidence is insufficient, without assuming that technology has no effect. It likewise leaves room to revise the classification if new operational sources or reception data reveal a different relationship.

Table 1. Evidentiary thresholds by type of claim.

[20] Santoni de Sio and van den Hoven, "Meaningful Human Control over Autonomous Systems." Madeleine Clare Elish, "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction," *Engaging Science, Technology, and Society* 5 (2019): 40–60, https://doi.org/10.17351/ests2019.260.

[21] Margaret Mitchell et al., "Model Cards for Model Reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2019), 220–229, https://doi.org/10.1145/3287560.3287596.

Type of claim	Evidence required	What it does not establish
Source authorization	Identity of the corpus and of the party that sets its use	That every output faithfully represents the source
Computational discretion	Evidence of the system's room for choice in producing outputs	That the choices carry an institutional mandate
Interpretive delegation	Linkage of mandate, interpretive choice, and standing of the output	That the oversight of that delegation is adequate
Accountability	Evidence of control, traceability, and correction	The mere presence of a human on the team
Social acceptance	Data from those who use or recognize the service	The institutional source of legitimacy

Discussion

Distinguishing source authorization from delegation refines the discussion of authority in digital religion precisely at the point where material passes into an answer. Campbell, and Campbell and Evolvi, provide the context that authority is inseparable from religious relationships that unfold across media.[22] This article adds a narrower question: at what stage does recognition of a source become recognition of what has been made from that source? The question matters because a conversational interface can fuse citation, summary, and advice into a single voice that appears seamless. Unity of presentation does not prove unity of origin or of epistemic standing. The framework's contribution is not to assert that a new authority arises with AI, but to make the assumed transformation more amenable to examination. Analysis may reveal continuity in corpus endorsement alongside change in access arrangements, or find delegation in one function without generalizing it to the whole service. Such a reading lets variation in authority relationships emerge from evidence rather than being fixed in advance by a technology label.

A dialogue with Qur'anic studies shows why fidelity to a source is more complex than matching words. Pink's typology captures differences in mode of authorship and commentarial orientation, while Saleh's history-of-the-book approach ties any reading of the tradition to processes of historiography and publication.[23] On that basis, the article infers that digitization must preserve the identity of a view, not merely fragments of its content. Two explanations addressing the same verse cannot necessarily be fused into a single position without a marker. In a hypothetical illustration, a system might quote reasoning from one work and then conclude with a position closer to another. The problem

[22] Campbell, "Who's Got the Power? Religious Authority and the Internet." Campbell and Evolvi, "Contextualizing Current Digital Religion Research on Emerging Technologies."

[23] Pink, "Tradition, Authority and Innovation in Contemporary Sunnī Tafsīr." Saleh, "Preliminary Remarks on the Historiography of Tafsīr in Arabic."

is not always a fabricated quotation, but the loss of the connection among the explanation, the author's purpose, and the position within a debate. The representation of ikhtilāf should therefore be judged by clarity of attribution, the scope of agreement, and the limits of synthesis. This proposal does not treat the diversity of tafsir as an error to be eliminated; that diversity is itself part of the knowledge that must be explained responsibly to the reader.

The mediatization framework poses a productive challenge to this separation. Hjarvard explains how media can act as an agent of religious change, while Cheong situates communicative practice as part of the formation of authority.[24] Both remind us that limiting a formal mandate does not mean limiting all of technology's influence. Even so, the context and reach of these theories must be respected; the relationship between media and religion is not identical across every society or organization. In a tafsir service, the ordering of answers, the choice of language, and the ease of asking questions can become sites of change worth studying, even where the institution retains its formal authority. This article therefore distinguishes the shaping of influence through the interface from the granting of standing to an output. The distinction does not diminish the importance of mediation; it shows that the mechanism of change can lie outside delegation. Research may find institutional reinforcement at the level of source provision alongside a new dependence on answer composition, without forcing a single, sweeping narrative of authority transfer.

The human–machine communication perspective clarifies the possibility that the form of delivery shapes users' attributions. Guzman and Lewis discuss the functional and relational dimensions and the boundary among humans, machines, and communication.[25] For the question of tafsir, this framework directs attention to how an answer is perceived: as a library citation, an institutional explanation, or the response of an agent taken to understand the user's needs. Yet communicative perception does not settle the question of justification. Floridi and Chiriatti's discussion of GPT-3 and Bender and colleagues' critique of language-model development are reminders to distinguish convincing language production from broader claims about understanding and responsibility.[26] This article does not use the limitations of earlier models to fix the capabilities of all present-day systems. The implication more resistant to technical change is that greater linguistic fluency still does not prove a mandate. Technology may become ever better at explaining material, while who

---

[24] Stig Hjarvard, "The Mediatization of Religion: A Theory of the Media as Agents of Religious Change," *Northern Lights: Film & Media Studies Yearbook* 6, no. 1 (2008): 9–26, https://doi.org/10.1386/nl.6.1.9/1. Cheong, "The Vitality of New Media and Religion."

[25] Guzman and Lewis, "Artificial Intelligence and Communication."

[26] Luciano Floridi and Massimo Chiriatti, "GPT-3: Its Nature, Scope, Limits, and Consequences," *Minds and Machines* 30, no. 4 (2020): 681–694, https://doi.org/10.1007/s11023-020-09548-1. Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2021), 610–623, https://doi.org/10.1145/3442188.3445922.

authorizes the explanation and bears its consequences remains a separate institutional question.

Separating the types of work also curbs the tendency to treat a technical achievement as scholarly legitimacy. Ji and colleagues survey the problem of hallucination in language generation, while Eltanbouly and colleagues focus their evaluation on the retrieval of Qur'anic passages.[27] Both sources give reason to design evaluations to fit the work, not reason to judge every tafsir service by a single right-or-wrong measure. Checking a quotation requires matching it against the source edition; checking a paraphrase requires fidelity of meaning; checking a synthesis requires assessing the relationships among views; and normative application requires contextual reasoning and a party authorized to judge it. This framework does not assume a particular architecture, such as retrieval-augmented generation, in an application that has not been documented. Even where source retrieval can be demonstrated, the linkage between the answer and the source still needs to be checked. The methodological implication is to separate citation errors, distortions of a view, and inferences that exceed the evidence. Improvement in one dimension must not hide weakness in another or be promoted as an endorsement of the interpretation as a whole.

A comparison with organizational research helps explain why the division of labor between humans and AI cannot be reduced to a choice between assisting and replacing. Raisch and Krakowski discuss the relationship between automation and augmentation, Jarrahi takes up human–AI collaboration in decision-making, and Faraj, Pachidi, and Sayegh place learning algorithms within questions of work and organization.[28] Its relevance to this article is conceptual: adding assistive functions can change how work is distributed without automatically eliminating the human role. In an illustrative tafsir service, an expert may remain the final party responsible, yet the material they read has already been selected and summarized by the system. That situation should be studied as a change in the arrangement of work, not immediately labeled a loss of authority. The next empirical question is whether the reviewer still has access to the source context, time to examine it, and the opportunity to reject a recommendation. Continuity of formal structure thus does not preclude epistemic change, while a change in workflow does not by itself prove that authority has been delegated to the system.

---

[27] Ji et al., "Survey of Hallucination in Natural Language Generation." Eltanbouly et al., "Can LLMs Directly Retrieve Passages for Answering Questions from Qur'an?"

[28] Sebastian Raisch and Sebastian Krakowski, "Artificial Intelligence and Management: The Automation–Augmentation Paradox," *Academy of Management Review* 46, no. 1 (2021): 192–210, https://doi.org/10.5465/amr.2018.0072. Mohammad Hossein Jarrahi, "Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organizational Decision Making," *Business Horizons* 61, no. 4 (2018): 577–586, https://doi.org/10.1016/j.bushor.2018.03.007. Samer Faraj, Stella Pachidi, and Karla Sayegh, "Working and Organizing in the Age of the Learning Algorithm," *Information and Organization* 28, no. 1 (2018): 62–70, https://doi.org/10.1016/j.infoandorg.2018.02.005.

Even evidence of a formal mandate must be supplemented by a reading of control practices. Kellogg, Valentine, and Christin's review situates algorithms as part of a contested field of organizational control, while Kitchin directs algorithm research toward critical scrutiny of the object and its context.[29] This article uses that concern to raise a counter-possibility to its own framework: an institution may state that AI is merely an assistive tool, while operational dependence grants the system far greater influence. The present manuscript does not establish this possibility. To test it, subsequent research would need to compare written procedures, reviewers' access to sources, and correction decisions in actual work. If only public statements are available, conclusions must stop at the articulation of roles. If practice shows an authority that is routinely exercised and recognized even though it is not explicitly formulated, the definition of mandate would need to be broadened to account for authorization institutionalized through practice. In this way, the framework becomes neither overly formalistic nor immune to evidence that runs against it.

The debate over transparency shows why a list of sources alone is not enough to ensure accountability. Burrell distinguishes forms of opacity, while Ananny and Crawford criticize the equation of being able to see a system with being able to know and control it.[30] Diakopoulos's work on algorithmic accountability adds attention to scrutiny of the processes that produce information.[31] The implication for tafsir services is the need for task-appropriate transparency: readers need an account of the sources and the status of an answer, while reviewers need sufficient access to evaluate how claims are assembled. Displaying the name of a book can aid attribution but does not show whether the passage cited supports the conclusion. Conversely, disclosing a model's technical details does not necessarily help a user tell a paraphrase from a new inference. The proposed framework treats transparency as a means of enabling corrective action. Its value is judged by what the relevant parties can examine and fix, not by the volume of information published or the impression of openness the interface conveys.

The question of responsibility reveals another limit of the slogan that a human remains in oversight. Matthias discusses the responsibility gap in learning systems, while Nyholm

[29] Katherine C. Kellogg, Melissa A. Valentine, and Angèle Christin, "Algorithms at Work: The New Contested Terrain of Control," *Academy of Management Annals* 14, no. 1 (2020): 366–410, https://doi.org/10.5465/annals.2018.0174. Rob Kitchin, "Thinking Critically About and Researching Algorithms," *Information, Communication & Society* 20, no. 1 (2017): 14–29, https://doi.org/10.1080/1369118x.2016.1154087.

[30] Burrell, "How the Machine 'thinks'." Mike Ananny and Kate Crawford, "Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability," *New Media & Society* 20, no. 3 (2018): 973–989, https://doi.org/10.1177/1461444816676645.

[31] Nicholas Diakopoulos, "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures," *Digital Journalism* 3, no. 3 (2015): 398–415, https://doi.org/10.1080/21670811.2014.976411.

considers the attribution of agency in human–robot work.[32] Neither addresses tafsir directly, but both supply a question about who can be held to account when action is dispersed among actors and technology. Elish's critique further shows the risk of making the nearest human the bearer of blame despite limited capacity for control.[33] Building on that debate, this article proposes that responsibility for a tafsir service should follow the decisions that are genuinely controllable: the selection of the corpus, the design of the task, the approval of its use, the checking of outputs, and the handling of complaints. This allocation does not require one person to master every layer. What is needed is a chain of accountability across layers that does not stop at the final reviewer. The presence of an exegesis expert becomes meaningful when their scholarly judgment can actually change an answer or a procedure, and is not used merely to lend an appearance of legitimacy to decisions already made by others.

At the user level, the goal of governance should be appropriate reliance, not merely increased trust. Parasuraman and Riley distinguish use, misuse, disuse, and abuse of automation; Lee and See connect trust with context-appropriate reliance.[34] Krügel and colleagues' experiment gives further reason not to assume that disclosing a chatbot's identity is always enough to prevent the influence of its advice.[35] Its application to tafsir remains a research agenda, not a direct generalization from a moral experiment. User evaluation can examine whether people can recognize the origin of a claim, locate its references, tell disputed views apart, and know when to ask for an expert explanation. Measures of satisfaction or convenience do not substitute for those capacities. A service's success is thus determined not only by acceptance of its answers, but also by users' ability to maintain informed judgment. This framework leaves room for AI that is both useful and open to question, without treating user caution as a failure of the technology.

Any governance proposal must connect general principles with work that can be examined. Jobin, Ienca, and Vayena identify both convergence and divergence in AI ethics guidelines; Mittelstadt cautions about the limits of principles without institutional support; Hagendorff evaluates the gap between guidelines and their implementation.[36] The AI4People framework foregrounds opportunities, risks, and ethical principles, while datasheets and

[32] Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata," *Ethics and Information Technology* 6, no. 3 (2004): 175–183, https://doi.org/10.1007/s10676-004-3422-1. Nyholm, "Attributing Agency to Automated Systems."

[33] Elish, "Moral Crumple Zones."

[34] Parasuraman and Riley, "Humans and Automation." Lee and See, "Trust in Automation."

[35] Krügel and Ostermaier and Uhl, "ChatGPT's Inconsistent Moral Advice Influences Users' Judgment."

[36] Anna Jobin, Marcello Ienca, and Effy Vayena, "The Global Landscape of AI Ethics Guidelines," *Nature Machine Intelligence* 1, no. 9 (2019): 389–399, https://doi.org/10.1038/s42256-019-0088-2. Brent Mittelstadt, "Principles Alone Cannot Guarantee Ethical AI," *Nature Machine Intelligence* 1, no. 11 (2019): 501–507, https://doi.org/10.1038/s42256-019-0114-4. Thilo Hagendorff, "The Ethics of AI Ethics: An Evaluation of Guidelines," *Minds and Machines* 30, no. 1 (2020): 99–120, https://doi.org/10.1007/s11023-020-09517-8.

model cards provide forms of data and model documentation.[37] This article translates that convergence into documentation requirements at the level of a tafsir function: the service's purpose, the type of sources, the type of work, the status of answers, the party responsible for corrections, and the limits of use. This recommendation is not a claim that an institution has already implemented it or certainly lacks it. Its practical value is to supply questions that can be answered with evidence, including whether an output is a quotation, a synthesis, or contextual advice. Conformity with technical documentation is likewise not treated as a religious endorsement; scholarly judgment still requires domain competence and a clear space of accountability.

The main limitation of this framework is its conceptual character and its purposive selection of literature. It has not tested how mandate, discretion, and recognition operate in a particular service, nor does it offer a representative picture of digital tafsir in Indonesia. Cross-disciplinary comparison helps build categories but can obscure the particularities of educational traditions, teacher–student relationships, and differences of religious view. An alternative explanation that remains open is that some services are adequately understood as a continuation of library aids, while others gain influence through user reception without delegation. The framework would need revision if evidence shows institutional recognition that the three proposed elements cannot explain, or if users consistently form authority through relationships not centered on the institution. The most decisive follow-up research is not merely to accumulate development news, but to bring together assignment documents, tracing of claims in outputs, observation of corrections, and reception studies. That combination can test the boundary among authorization as declared, work as performed, and the standing of answers as actually recognized.

Table 2. Three relationships that similar outputs can conceal

Relationship	Defining condition	Evidence signature
Authorized source mediation	Redisplays official explanations with no room to compose new decisions	A legitimate corpus and a defined purpose of use
Interpretive influence / user recognition	Produces interpretations that users trust, absent an institutional mandate	Reception and reliance data from users, not an institutional grant
Institutional delegation	Tasked to compose explanations within limits, holds discretion, issues outputs with standing on behalf of the mandating party	Linkage of mandate, interpretive choice, and recognized standing of the output

[37] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines* 28, no. 4 (2018): 689–707, https://doi.org/10.1007/s11023-018-9482-5. Gebru et al., "Datasheets for Datasets." Mitchell et al., "Model Cards for Model Reporting."

Conclusion

This article addresses the question of the limits of delegation by separating three relationships that are often conflated in discussions of AI and tafsir. Endorsing a source does not automatically endorse everything made from it; a system's ability to select and compose explanations does not automatically prove a mandate; and influence over users does not automatically indicate institutional recognition. On this basis, authorized source mediation is an adequate description when the evidence shows only the establishment of material and a service function. Delegating interpretive authority requires additional evidence of a mandating party, room for interpretive decision-making, and the standing of outputs on behalf of that party. This conclusion does not settle the actual capabilities of any application, nor does it claim that the authority of the mufassir has shifted or cannot shift. It clarifies the kind of relationship that must be demonstrated before the term delegation is used as an empirical explanation.

The study's conceptual contribution lies in separating the existence of delegation from the quality of its accountability. That separation allows an analysis to recognize weakly supervised delegation, influence without a mandate, and the use of official sources unaccompanied by any authority to fix meaning. The methodological implication is evaluation by type of claim and type of work, so that success in retrieval is not made a substitute for fidelity of synthesis or the soundness of normative application. The practical implication is the need for documentation that links the corpus, the operations, the standing of answers, and the authority to make corrections. Accountability takes concrete form when the party said to be responsible has the information, competence, and ability to change a decision. The framework thus turns attention away from a technology label and toward relationships of knowledge and responsibility that can be examined.

The article's findings should be read as the product of conceptual synthesis, bounded by the selection of literature, source access, and the absence of field testing. Further development requires examination of well-defined service functions, not generalization from a platform name or an institutional identity. Evidence about a mandate must be brought together with a system's operations, its correction processes, and user recognition to determine whether the proposed relationship has in fact formed. Where such evidence is unavailable, withholding a classification is a more defensible conclusion than affirming a transfer of authority. By keeping the categories open to proof and rebuttal, the study of AI in Qur'anic interpretation can engage real change while preserving the distinction among access to knowledge, the composition of explanations, and the authority to authorize meaning.

Declaration

Use of Artificial Intelligence (AI)

The authors used ChatGPT/Codex by OpenAI as a technical aid for language editing, manuscript structural refinement, and sentence reformulation. AI was not used to generate research data. All references, substantive descriptions, and interpretations presented in the manuscript were reviewed and independently verified by the authors. AI is not listed as an author; full responsibility for the accuracy, originality, academic integrity, and final content of the manuscript rests with the authors.

Funding

This research received no specific funding from any public, commercial, or non-profit funding agency.

Conflict of Interest

The authors declare that they have no conflicts of interest related to the research, authorship, and/or publication of this manuscript.

References

Ananny, Mike, and Kate Crawford. "Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20, no. 3 (2018): 973–989. https://doi.org/10.1177/1461444816676645.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. New York: Association for Computing Machinery, 2021. https://doi.org/10.1145/3442188.3445922.

Bowen, Glenn A. "Document Analysis as a Qualitative Research Method." *Qualitative Research Journal* 9, no. 2 (2009): 27–40. https://doi.org/10.3316/QRJ0902027.

Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3, no. 1 (2016): 2053951715622512. https://doi.org/10.1177/2053951715622512.

Campbell, Heidi. "Who's Got the Power? Religious Authority and the Internet." *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 1043–1062. https://doi.org/10.1111/j.1083-6101.2007.00362.x.

Campbell, Heidi A., and Giulia Evolvi. "Contextualizing Current Digital Religion Research on Emerging Technologies." *Human Behavior and Emerging Technologies* 2, no. 1 (2020): 5–17. https://doi.org/10.1002/hbe2.149.

Cheong, Pauline Hope. "The Vitality of New Media and Religion: Communicative Perspectives, Practices, and Changing Authority in Spiritual Organization." *New Media & Society* 19, no. 1 (2017): 25–33. https://doi.org/10.1177/1461444816649913.

Diakopoulos, Nicholas. "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures." *Digital Journalism* 3, no. 3 (2015): 398–415. https://doi.org/10.1080/21670811.2014.976411.

Elish, Madeleine Clare. "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction." *Engaging Science, Technology, and Society* 5 (2019): 40–60. https://doi.org/10.17351/ests2019.260.

Eltanbouly, Sohaila, Salam Albatarni, Shaimaa Hassanein, and Tamer Elsayed. "Can LLMs Directly Retrieve Passages for Answering Questions from Qur'an?" In *Proceedings of the Third Arabic Natural Language Processing Conference*, 203–210. Suzhou, China: Association for Computational Linguistics, 2025. https://doi.org/10.18653/v1/2025.arabicnlp-main.16.

Faraj, Samer, Stella Pachidi, and Karla Sayegh. "Working and Organizing in the Age of the Learning Algorithm." *Information and Organization* 28, no. 1 (2018): 62–70. https://doi.org/10.1016/j.infoandorg.2018.02.005.

Floridi, Luciano, and Massimo Chiriatti. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines* 30, no. 4 (2020): 681–694. https://doi.org/10.1007/s11023-020-09548-1.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4 (2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. "Datasheets for Datasets." *Communications of the ACM* 64, no. 12 (2021): 86–92. https://doi.org/10.1145/3458723.

Guzman, Andrea L., and Seth C. Lewis. "Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda." *New Media & Society* 22, no. 1 (2020): 70–86. https://doi.org/10.1177/1461444819858691.

Hagendorff, Thilo. "The Ethics of AI Ethics: An Evaluation of Guidelines." *Minds and Machines* 30, no. 1 (2020): 99–120. https://doi.org/10.1007/s11023-020-09517-8.

Hjarvard, Stig. "The Mediatization of Religion: A Theory of the Media as Agents of Religious Change." *Northern Lights: Film & Media Studies Yearbook* 6, no. 1 (2008): 9–26. https://doi.org/10.1386/nl.6.1.9/1.

Jaakkola, Elina. "Designing Conceptual Articles: Four Approaches." *AMS Review* 10, nos. 1–2 (2020): 18–26. https://doi.org/10.1007/s13162-020-00161-0.

Jarrahi, Mohammad Hossein. "Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organizational Decision Making." *Business Horizons* 61, no. 4 (2018): 577–586. https://doi.org/10.1016/j.bushor.2018.03.007.

Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys* 55, no. 12 (2023): 1–38. https://doi.org/10.1145/3571730.

Jobin, Anna, Marcello Ienca, and Effy Vayena. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1, no. 9 (2019): 389–399. https://doi.org/10.1038/s42256-019-0088-2.

Kellogg, Katherine C., Melissa A. Valentine, and Angèle Christin. "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals* 14, no. 1 (2020): 366–410. https://doi.org/10.5465/annals.2018.0174.

Kitchin, Rob. "Thinking Critically about and Researching Algorithms." *Information, Communication & Society* 20, no. 1 (2017): 14–29. https://doi.org/10.1080/1369118x.2016.1154087.

Krügel, Sebastian, Andreas Ostermaier, and Matthias Uhl. "ChatGPT's Inconsistent Moral Advice Influences Users' Judgment." *Scientific Reports* 13, no. 1 (2023): 4569. https://doi.org/10.1038/s41598-023-31341-0.

Lee, J. D., and K. A. See. "Trust in Automation: Designing for Appropriate Reliance." *Human Factors* 46, no. 1 (2004): 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Matthias, Andreas. "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata." *Ethics and Information Technology* 6, no. 3 (2004): 175–183. https://doi.org/10.1007/s10676-004-3422-1.

Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. "Model Cards for Model Reporting." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. New York: Association for Computing Machinery, 2019. https://doi.org/10.1145/3287560.3287596.

Mittelstadt, Brent. "Principles Alone Cannot Guarantee Ethical AI." *Nature Machine Intelligence* 1, no. 11 (2019): 501–507. https://doi.org/10.1038/s42256-019-0114-4.

Nyholm, Sven. "Attributing Agency to Automated Systems: Reflections on Human–Robot Collaborations and Responsibility-Loci." *Science and Engineering Ethics* 24, no. 4 (2018): 1201–1219. https://doi.org/10.1007/s11948-017-9943-x.

Parasuraman, Raja, and Victor Riley. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39, no. 2 (1997): 230–253. https://doi.org/10.1518/001872097778543886.

Pink, Johanna. "Tradition, Authority and Innovation in Contemporary Sunnī Tafsīr: Towards a Typology of Qur'an Commentaries from the Arab World, Indonesia and Turkey." *Journal of Qur'anic Studies* 12, nos. 1–2 (2010): 56–82. https://doi.org/10.3366/jqs.2010.0105.

Purnomo, Bagus. "LPMQ Susun Roadmap Kajian Al-Qur'an dan Pengembangan Aplikasi Lima Tahun Ke Depan." Lajnah Pentashihan Mushaf Al-Qur'an, September 19, 2024. https://lajnah.kemenag.go.id/info-lpmq/berita-dan-artikel/berita/lpmq-susun-roadmap-kajian-al-qur-an-dan-pengembangan-aplikasi-lima-tahun-ke-depan.html.

Raisch, Sebastian, and Sebastian Krakowski. "Artificial Intelligence and Management: The Automation–Augmentation Paradox." *Academy of Management Review* 46, no. 1 (2021): 192–210. https://doi.org/10.5465/amr.2018.0072.

Reed, Randall. "A.I. in Religion, A.I. for Religion, A.I. and Religion: Towards a Theory of Religious Studies and Artificial Intelligence." *Religions* 12, no. 6 (2021): 401. https://doi.org/10.3390/rel12060401.

Saleh, Walid A. "Preliminary Remarks on the Historiography of Tafsīr in Arabic: A History of the Book Approach." *Journal of Qur'anic Studies* 12, nos. 1–2 (2010): 6–40. https://doi.org/10.3366/jqs.2010.0103.

Santoni de Sio, Filippo, and Jeroen van den Hoven. "Meaningful Human Control over Autonomous Systems: A Philosophical Account." *Frontiers in Robotics and AI* 5 (2018): 15. https://doi.org/10.3389/frobt.2018.00015.

Snyder, Hannah. "Literature Review as a Research Methodology: An Overview and Guidelines." *Journal of Business Research* 104 (2019): 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039.