CONVERGENCE: JOURNAL OF ISLAMIC AND INTERDISCIPLINARY STUDIES
Vol. 1, No. 1, April 2026, pp. 1-10
ISSN 3165-1909, DOI: 10.68831/cjiis.v1i1.2026.4

Artificial Intelligence-Generated Fatwas and Religious Chatbots: Epistemological Risks to the Authority of Islamic Law

Marlina[1], Sri Yanti[2]

[1] Department of Islamic Education Management, Faculty of Tarbiyah, Institut Agama Islam Negeri Takengon, Aceh, Indonesia
[2] Department of Islamic Religious Education, Universitas Muhammadiyah Palembang, Palembang, Indonesia

Article Info

Keywords:
Artificial intelligence
Fatwa
Chatbot
Ijtihad
Maqasid al-shariah

ABSTRACT

Generative artificial intelligence (AI) chatbots are increasingly used to answer religious questions, giving rise to what may be called artificial intelligence-generated fatwas. This conceptual study analyzes the epistemological risks that such outputs pose to the authority of Islamic law. Using a literature-based, conceptual-analytical approach, it compares fatwas issued by muftis or fatwa institutions with outputs of large language models along five dimensions: source of knowledge, juristic reasoning, accountability, sensitivity to context, and scholarly transmission (sanad). The analysis identifies five interrelated risks: opaque sources, hallucination, bias, loss of context, and an appearance of authority. Their common root is the separation between the form of a fatwa and the epistemic warrant that customarily accompanies it. These risks may weaken the gatekeeping role of the ulama, distort lay legal understanding, and affect ritual and social practice. The article proposes mitigation and governance principles grounded in usul al-fiqh and maqasid al-shariah: functional classification of outputs, source traceability, human oversight, disclosure of uncertainty, independent evaluation, and religious-digital literacy. This study is conceptual and does not empirically test any system.

This is an open access article under the CC BY-SA 4.0 License

Corresponding Author:
Marlina
Department of Islamic Education Management, Faculty of Tarbiyah, Institut Agama Islam Negeri Takengon, Aceh, Indonesia
ORCID: 0000-0001-5892-5020 - *Email: marlinasabil@gmail.com*

1. INTRODUCTION

Chatbots based on large language models (LLMs) have become a new interface for seeking information. Users no longer merely receive a list of links but narrative answers composed fluently, including for questions traditionally put to ulama, such as the validity of an act of worship, the status of a transaction, or the resolution of a family dispute. This phenomenon continues a long pattern in which new media have changed the way Muslims access religious knowledge and blurred the boundary between established authority and new voices [1]–[4]. Generative chatbots, however, bring a qualitative change: such systems do not merely display sources but *generate* new text that can take the form and tone of a fatwa [5], [6].

In Indonesia, institutional responses have already emerged. The National Conference of Ulama of Nahdlatul Ulama (NU) in September 2023 decided that asking artificial intelligence (AI) about religious matters is permissible, but that treating its answers as guidance to be put into practice is forbidden because their correctness is not guaranteed [7]. In 2025, an official of the Indonesian Ulama Council (MUI) cautioned that artificial intelligence should not be positioned as a mujtahid or a giver of fatwas [8], and MUI was reported to be preparing guidelines on the use of artificial intelligence for religious reference [9]. On the state's side, Circular Letter of the Minister of Communication and Informatics Number 9 of 2023 contains ethical guidelines for artificial intelligence that apply generally to business actors and electronic system operators [10]. These decisions and guidelines are normative; this article seeks to set out their analytical foundations from the standpoint of the epistemology of Islamic law.

The academic literature shows a spectrum of positions. Ahmed assesses, through a literature review, that chatbots offer access and consistency but are less sensitive to context, so that the role of the human mufti remains necessary [11]. Jan compares fatwas generated by ChatGPT with fatwas issued by humans and argues that such programs cannot attain the standing of a giver of fatwas [12]. Other studies discuss the issuance of fatwas with artificial

intelligence in Dubai and Egypt [13], the technical side of automatically generating Islamic legal opinions [14], the application of artificial intelligence to fatwas in the context of Islamic finance [15], the challenges artificial intelligence poses to ijtihad in the Shia tradition [16], and the challenges artificial intelligence poses to Islamic thought in general [17]. In the Indonesian context, research highlights how young scholars use online sources in the production of knowledge and fatwas [18], as well as fatwa authority and the transmission of NU's *bahtsul masail* in the digital era [19]. At the same time, the credibility of human-made online fatwas has also been questioned [20], so the comparison must be fair to both sides.

As far as the sources reviewed in this article go, existing contributions tend to be descriptive-normative (permitting or prohibiting), technical, or comparisons of particular examples. Few have systematically mapped the risks of chatbots onto the categories of the epistemology of Islamic law, namely the qualifications of the mufti, *istinbat*, *sanad*, and *maqasid al-shariah*, while avoiding the equation of classical categories with technology. This article fills that gap. Its main argument is that the epistemological risk of artificial intelligence-generated fatwas lies not merely in the possibility of wrong answers, since human muftis too can err, but in *the separation between the form of a fatwa and the epistemic warrant* that traditionally accompanies it, namely qualification, method, responsibility, and connection to the scholarly community.

The article is guided by three research questions: (1) what are the epistemological differences between a fatwa issued by a mufti or fatwa institution and a religious answer generated by a chatbot; (2) what epistemological risks arise when chatbot outputs are treated as fatwas, and what is their impact on the authority of the ulama, legal understanding, and religious practice; and (3) what mitigation and governance principles are consistent with the Islamic scholarly tradition. The aim is to answer these three questions through conceptual analysis based on a literature review. Because the problem is also one of religious learning, the article closes its analysis by drawing implications for Islamic education (Section 4.4.4). Section 2 sets out the conceptual framework, Section 3 the method, Section 4 the results and discussion, and Section 5 the conclusion.

2. CONCEPTUAL FRAMEWORK: AUTHORITY AND EPISTEMOLOGY IN ISLAMIC LAW

2.1. Authority and epistemology in Islamic law

The epistemology of Islamic law rests on *usul al-fiqh*, the discipline that governs how practical law (*fiqh*) is derived from evidence through methods that can be held accountable [21], [22]. *Al-Risala* by al-Syafi'i is regarded as one of the pioneering works in the formulation of those methods [23]. Two important distinctions should be noted. First, the distinction between the evidence and the law produced from it: on many questions of detail, the result of ijtihad is a strong presumption (*zanni*), not certainty (*qatʿi*), so that diversity of opinion (*ikhtilaf*) is a methodological consequence, not an anomaly [21], [22]. Second, the distinction between knowledge and authority: what the community accepts is not merely a true proposition, but a proposition produced by a subject deemed qualified through a recognized procedure.

Authority in Islamic law is therefore both epistemic and moral. Abou El Fadl distinguishes between authority that is *authoritative*, that is, trusted voluntarily because of competence and integrity and remaining open to questions, and authority that is *authoritarian*, which demands obedience without room for scrutiny, and stresses that speaking in God's name demands special caution [24]. Hallaq shows that continuity and change in Islamic law proceed through the interaction between legal theory, fatwa practice, and community acceptance [25]. Zaman shows that contemporary ulama continue to negotiate their role as guardians of tradition amid changing institutions and media [26].

2.2. Fatwa, mufti, and mustafti

A fatwa is a legal opinion given by a mufti in response to a questioner (*mustafti*). Unlike a judge's ruling (*qada'*), a fatwa is in principle not legally binding, but it is morally and religiously binding for those who believe in it [27]–[29]. The literature on classical mufti ethics, for example the works of al-Nawawi, al-Qarafi, and Ibn al-Qayyim, places the fatwa in a personal and moral relationship: a mufti must be learned, of integrity, and cautious, and must understand the questioner's circumstances [30]–[33]. Ibn al-Qayyim stresses that a fatwa may change according to changes in time, place, circumstances, intention, and custom [32], while al-Qarafi distinguishes the fatwa from the judge's ruling and pays attention to the role of local custom [31]. These two features, personal responsibility and sensitivity to context, serve as the main points of comparison for chatbots in this article.

At the institutional level, contemporary fatwas often arise from collective processes. MUI and NU's *bahtsul masail* in Indonesia, for example, produce fatwas or decisions through forums, commissions, and determination procedures, with broad socio-political influence even though they are generally not binding under state law [19], [34]–[36]. Authority is thus not only individual but also institutional: there are parties who can be asked for an explanation, corrected, and, in the religious sense, held accountable. Cholil Nafis affirms that a mufti is responsible for his fatwa before Allah, so artificial intelligence is not fit to be given responsibility for deciding [8].

2.3. Ijtihad, taqlid, and the limits of analogy

Ijtihad is the exertion of effort to derive law from evidence while meeting certain scholarly requirements, among them mastery of Arabic, knowledge of the sources of law, and methods of reasoning [21], [22]. *Taqlid* technically means following another's opinion without demanding the evidence for oneself. In the classical framework, laypeople are in fact required to ask those who are learned, so requesting a fatwa is a legitimate form of

taqlid, not negligence [22], [25]. The notion that "the gate of ijtihad is closed" has been criticized historiographically; Hallaq shows that ijtihad continued within the legal school tradition [37]. Fatwa theory also distinguishes the mujtahid mufti from the mufti who follows a school of law, who within certain limits may legitimately convey the school's opinion [28].

This mapping is important to avoid careless analogy. Calling a chatbot a "digital mujtahid" or a "digital muqallid" involves a *category mistake*, since both categories presuppose a subject who is burdened with obligations (*mukallaf*), capable of intention, and able to be held accountable. The argument that a chatbot merely "conveys the school's opinion" like a muqallid mufti also needs to be tested. Conveying a school's opinion still requires verification of the validity of transmission, selection of the opinion that serves as the school's reliable position (*muʿtamad*), and application to the concrete case (*tahqiq al-manat*). These three functions are not performed by token prediction alone (Section 4.1).

2.4. Sanad and scholarly transmission

Sanad (or *isnad*) in the science of hadith is a chain of transmitters whose credibility is assessed, among other things in terms of integrity (*ʿadalah*) and accuracy (*dabt*), so that the validity of a report depends on the quality of the people in the chain, not only on the content of the text [38]. In legal education, a similar pattern appears in the tradition of *ijazah*, a teacher's permission that marks a person's fitness to teach and issue fatwas [39]. Scholarly transmission is thus personal, graded, and accountable: knowledge is received from known teachers, tested, and recognized. It should be noted that source citations or reference lists, which an artificial intelligence system can produce, are not the equivalent of sanad. The former points to a document, whereas the latter guarantees the quality of the people in the chain of transmission.

2.5. Maqasid al-shariah as an evaluative framework

Maqasid al-shariah are the objectives of Islamic law which, in the classical formulation, are associated with the protection of religion, life, intellect, lineage, and property, and are used to weigh benefit and harm [22], [40]. Auda proposes a reading based on a systems approach that extends the scope of maqasid beyond the classical list [41]. In this article maqasid is used not as a license to justify technology on the basis of benefit, but as an evaluative framework for weighing risks and designing safeguards, especially the protection of religion (so that teachings are not distorted) and of intellect (so that understanding is not obscured). Because maqasid without methodological discipline easily becomes ad hoc justification, the governance principles in Section 4.5 are tied to more specific usul al-fiqh categories. Figure 1 summarizes the conceptual framework used.

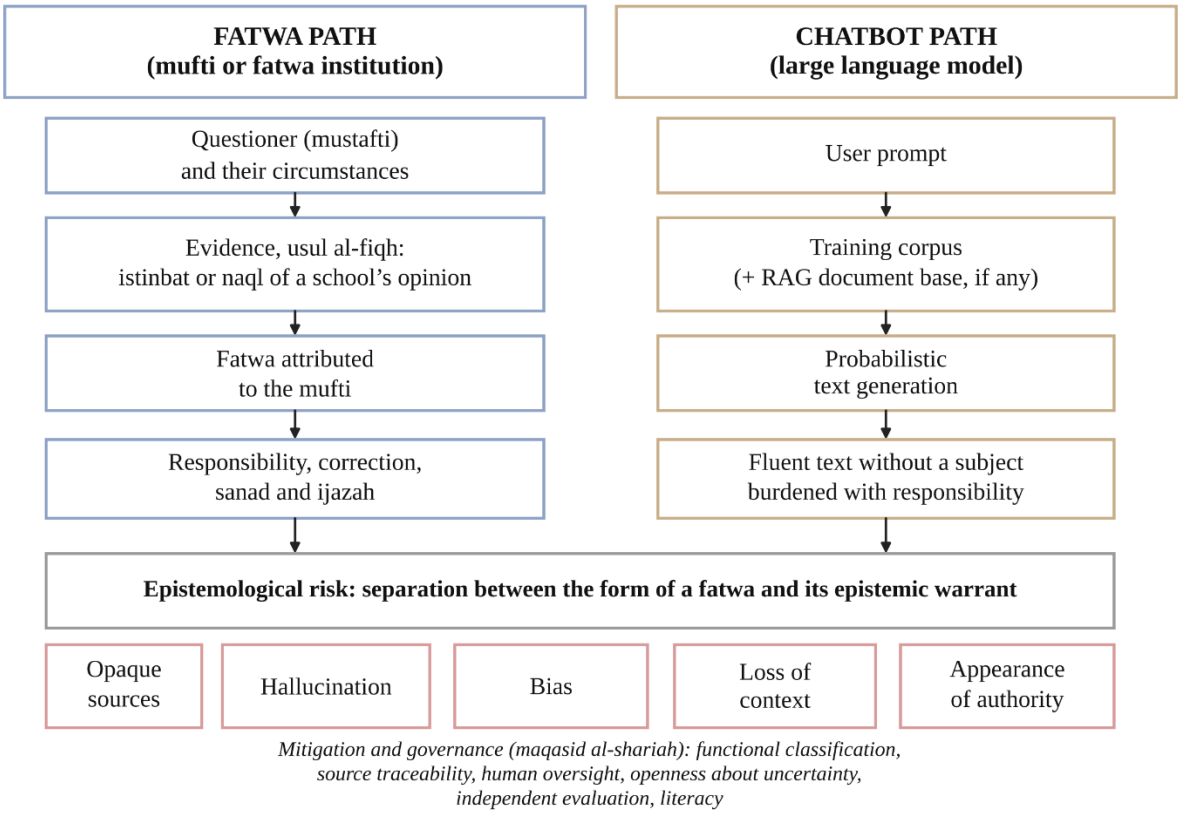

Figure 1. Conceptual framework: the fatwa path, the chatbot path, and epistemological risk points

3. METHOD

This research is a literature review with qualitative conceptual analysis. This approach was chosen because the research questions concern relations between concepts (authority, epistemology, fatwa, ijtihad, sanad, maqasid) and the characteristics of artificial intelligence systems, not the frequency or distribution of user behavior. This article conducts no interviews, surveys, or tests of chatbot systems, and does not claim primary empirical data.

The study materials consist of four groups: (a) classical primary sources of Islamic law on the methods and ethics of fatwa; (b) academic literature on usul al-fiqh, fatwa, and religious authority, including studies on the Indonesian context; (c) technical and critical literature on large language models; and (d) policy documents and decisions of religious institutions, namely the decisions of the 2023 NU National Conference as reported by the organization's official media, ministerial circular letters, and international artificial intelligence governance frameworks. Sources were selected based on their relevance to the research questions and the verifiability of their bibliographic data; the search for references was carried out until October 2026.

The analysis was conducted in four steps: (1) constructing a conceptual framework from groups (a) and (b) to obtain the characteristics of a valid fatwa; (2) identifying how chatbots work and their limitations from group (c); (3) comparing the two conceptually along five dimensions, namely source of knowledge, reasoning process, accountability, sensitivity to context, and scholarly transmission, and then deriving epistemological risks from the points of mismatch; and (4) synthesizing mitigation principles by weighing the findings against maqasid al-shariah and the governance frameworks in group (d).

This research has several limitations. First, the literature search did not follow a systematic review protocol and therefore does not claim completeness. Second, because the author did not test any system, statements about chatbot failures rest on the literature and apply at the level of structural possibility, not as frequency estimates for any particular product; system capabilities also change rapidly. Third, the usul al-fiqh framework used rests primarily on the Sunni tradition; other traditions, such as ijtihad in Shia thought, have different concepts of authority and are discussed in separate literature [16]. Fourth, the findings are conceptual and need to be tested empirically in further research.

4. RESULTS AND DISCUSSION

This section presents the findings of the analysis according to the three research questions: the epistemological differences between fatwas and chatbot outputs (4.1), the relevant workings and limitations of chatbots (4.2), epistemological risks (4.3), their impacts, including implications for Islamic education (4.4), and mitigation and governance principles (4.5).

4.1. The mufti's fatwa and the chatbot's answer: epistemological differences

A conceptual comparison shows that the main difference lies not in the content of the answer, which may well be similar, but in the structure of its warrant (Table 1). A valid fatwa is produced by a qualified subject who weighs the evidence and the questioner's circumstances, can be questioned again, and is accountable religiously and, in an institutional context, institutionally [27], [30], [31]. Chatbot output is text generated statistically from patterns in a training corpus [42], [43], which is then aligned with human preferences so as to appear helpful and polite [44].

Table 1. Conceptual comparison of fatwas and chatbot outputs

Dimension	Fatwa by a mufti or fatwa institution	Religious answer by a chatbot
Subject	A qualified mufti or committee that is known and can be contacted	A statistical system without shar'i qualifications; developers and operators do not automatically hold such qualifications
Source of knowledge	Evidence and scholarly opinions that are explicitly cited and can be checked	Patterns from the training corpus; references, if any, may be missing or wrong
Reasoning process	*Istinbat* or conveying a school's opinion with verification of transmission and *tahqiq al-manat*	Probability-based text generation; no legal reasoning procedure that can be audited as in usul al-fiqh
Questioner's context	Can ask follow-up questions and weigh the questioner's circumstances, custom, and welfare	Limited to information in the *prompt*; sensitivity to context is unstable
Accountability	The mufti is accountable before Allah and society; the institution can issue corrections	There is no *mukallaf* subject; responsibility is dispersed among operators, developers, and users
Scholarly transmission	Sanad and ijazah; membership in a scholarly community	No personal sanad; the origin of training data is generally not transparent to users
Status of output	A legal opinion attributed to the giver of the fatwa	Information not yet equivalent to a fatwa unless curated, verified, and attributed

Categorically, chatbot output is better understood as unverified religious information than as a fatwa. It can function as a tool, for example summarizing, translating, or searching texts, but it becomes a fatwa only if an authoritative party examines it and attributes it. The stance of the 2023 NU National Conference, which permits asking but prohibits treating artificial intelligence answers as guidance to be put into practice [7], can be read as consistent with the distinction between information and guidance for action. This distinction does not depend on the quality of the answer: a chatbot that is often right does not automatically possess a structure of warrant.

A reasonable objection is that human muftis too can err, be biased, and be non-transparent, and that the credibility of online fatwas has been questioned [20]. This objection is correct, but it does not invalidate the distinction above. A mufti's errors occur within a system that provides correction mechanisms, such as re-asking, ikhtilaf, collective forums, and personal responsibility, whereas a chatbot's errors often have no clear address of responsibility.

4.2. How chatbots work and the relevant limitations

4.2.1. A brief account of how they work

Modern LLMs are generally based on the *Transformer* architecture [42], trained to predict the next token on text corpora of very large scale so that they are able to perform a variety of language tasks [43]. Further stages such as instruction tuning and learning from human feedback make the model more compliant with instructions and more preferred by users [44]. Some systems are equipped with *retrieval-augmented generation* (RAG), namely the retrieval of documents from an external knowledge base to serve as context for the answer, which allows answers to be linked to source documents [45]. What is produced remains a statistically plausible continuation of text, not the result of an istinbat procedure that can be checked step by step.

4.2.2. Relevant limitations

First, fluency is not the same as reliability. Some researchers argue that models trained only on linguistic form do not by themselves acquire meaning [46] and highlight the risk of large language models as producers of convincing text with no guarantee of truth [6]. Claims about model "understanding" remain debated, so this article does not rely on them, but on the narrower fact that fluency does not guarantee accuracy. Second, models can produce content that is unfaithful to sources or non-factual (*hallucination*) [5]; on the benchmarks tested, models can also imitate common errors in human text [47]. Third, models can reproduce bias from data, design, and use [48]. Fourth, models do not always use long contexts evenly [49]. Fifth, models tend to adjust their answers to the user's views (*sycophancy*) [50]. Sixth, the ability of models to estimate their own uncertainty is still being studied and cannot be assumed to be reliable in all domains [51].

These limitations are general in nature. How often they occur on questions of Islamic law must be studied empirically. For positive law, there are already indications of high rates of hallucination on verifiable questions [52]; this finding can only serve as an analogy, not as evidence about fiqh.

4.3. Epistemological risks

From the mismatches in Table 1, five interrelated risks are derived. Their mapping onto aspects of authority and mitigation is summarized in Table 2 (Section 4.5).

4.3.1. Opaque sources

In the fatwa tradition, the strength of an opinion depends on its evidence and on who states it [30]. A chatbot without RAG cannot show a checkable basis for an opinion, because its knowledge is dispersed in the model's parameters [45]. Even when sources are cited, there are three problems: references can be fictitious or mismatched [5]; quotations of evidence can be truncated or detached from the context of their sentence while the status of the text is not examined; and a list of references is not a sanad (Section 2.4). As a result, users cannot do what students of knowledge classically do, namely trace, weigh, and test.

4.3.2. Hallucination

Hallucination in a fatwa differs in weight from hallucination in a news summary. It can fabricate evidence, attribute opinions to ulama or schools of law that did not hold them, or claim a consensus (*ijmaʿ*) that does not exist. A false claim of consensus is especially dangerous because it closes off the legitimate space of ikhtilaf [22]. Hallucination is a documented problem in text generation [5]. RAG can help link answers to documents [45], but the quality of the result still depends on the quality of the corpus and on the accuracy of the interpretation of the retrieved documents.

4.3.3. Bias

Bias can originate from data, design, and use [48]. A study of GPT-3 documented persistent negative associations with Muslims [53], which shows that training corpora are not neutral toward representations of Islam. In the domain of fatwas, bias can take the form of language dominance (popular English-language online sources versus Arabic-language classical books and local sources), the dominance of certain opinions over others, or the framing of issues following user preferences. Potential bias is among the reasons reported in news coverage of the 2023 NU National Conference decision [54]. As far as reviewed, such intra-Islamic bias has not been systematically documented; it is therefore formulated as a structurally plausible risk that needs testing, not as an empirical finding.

4.3.4. Loss of context

A good fatwa is contextual. Ibn al-Qayyim and al-Qarafi stress the questioner's circumstances and local custom [31], [32]; a mufti can ask back, weigh conditions, and even postpone or decline to answer. A chatbot obtains context only from the text the user types, while lay users often do not know which facts are legally relevant, for example the details of a contract, marital status, or health condition. Systems also do not always use long contexts evenly [49] and tend to answer rather than hold back. There is also collective context: contemporary fatwas respond

to social change and the conditions of the local community [34], whereas generic answers across jurisdictions obscure it.

4.3.5. Appearance of authority

This risk may be the most dangerous because it operates on perception. Fluency, speed, a confident tone, the format of evidence and citations, and a personal conversational interface produce an appearance of authority out of proportion to the epistemic warrant behind it. Using Abou El Fadl's framework, such a system can function, in perception, as an *authoritarian* authority, that is, giving a final answer without opening room for scrutiny, even without intending to [24]. The tendency to adjust answers to the user [50] opens the way to "fatwa shopping" by altering the *prompt*. Studies of online religious authority show that authority is negotiated through practices of recognition, not only through office [2]; users' recognition of a chatbot can form a de facto authority even though its developers do not claim it.

4.4. Impacts on the authority of the ulama, legal understanding, and religious practice

4.4.1. Authority of the ulama

The first impact concerns the shift in the gatekeeper function. The ulama are no longer the only node of access, continuing a trend of new media that has given rise to alternative figures and platforms of authority [1], [2], [4]. Chatbots accelerate this shift because they offer personal, instant answers available at any time. The risk is that authority moves from "who is learned and accountable" to "the system that is easiest to access". The impact is not singular. Ulama and institutions can use artificial intelligence as a tool for literature searching and initial drafting; a literature review judges that artificial intelligence can broaden access and consistency while retaining the human element in determining rulings [11]. The 2023 NU National Conference itself recommended that PBNU develop artificial intelligence with content from authoritative parties [55], and its commission was reported to regard the development of religious artificial intelligence as fardhu kifayah [56].

4.4.2. Legal understanding

For lay users, the main risk is simplification. Islamic law can be reduced to a "permissible" or "not permissible" answer without understanding of method, ikhtilaf, and conditions of application. When a chatbot presents a single answer as "Islamic law", the diversity of schools of law can be erased; conversely, when all opinions are presented without a framework, users may choose wrongly. In the long run, dependence can reduce engagement in the learning process that in the tradition takes place through teachers and community (Section 2.4). Conversely, if designed to direct users to sources and explain ikhtilaf, a chatbot can broaden legal literacy; this needs to be tested empirically.

4.4.3. Religious practice

In the realm of worship and social transactions, an erroneous output can affect the validity of worship, the rights of others (contracts, inheritance, marriage), and health. The higher the maqasid implications, namely the protection of life, lineage, and property, the greater the consequences of error. Social risks also arise: different answers from a chatbot or different *prompt*s can trigger fragmentation of references among the congregation, and answers that justify sensitive decisions, such as declaring others to be unbelievers or violence, require special handling in the form of escalation to human authorities.

4.4.4. Implications for Islamic education

The risks above are also educational. Students and young Muslims increasingly meet religious questions first through chatbots, and online sources have already become part of how Islamic knowledge is produced and received [18]. In classical learning, a student acquires an answer together with a way of reaching it: a teacher who is known, a method that can be examined, and a community that corrects. A chatbot supplies the answer without that formation. The educational risk is therefore not only that a student receives a wrong ruling, but that the student comes to treat fluent output as knowledge and stops asking who is responsible for it and on what evidence it rests (Section 2.4).

Islamic education institutions, including madrasahs, pesantren, and Islamic higher education, are thus well placed to respond. Three implications follow. First, teachers and lecturers need to distinguish clearly between artificial intelligence as a learning aid (summarizing, translating, locating texts) and artificial intelligence as a source of religious rulings, in line with the distinction between information and guidance for action in Section 4.1. Second, curricula in usul al-fiqh, fiqh, and Islamic studies can use chatbot answers as material for critical exercises, for example asking students to trace the claimed evidence, check attributions to schools of law, and identify where ikhtilaf has been erased. Third, educators need their own preparation, because the shift in the gatekeeper function discussed in Section 4.4.1 reaches classrooms as well as fatwa councils [2], [11]. These implications are conceptual and need to be tested through empirical studies of teaching practice.

4.5. Mitigation and governance principles that respect the Islamic scholarly tradition

The following principles are derived from the distinctions in Section 4.1 and the risks in Section 4.3, then weighed against maqasid, especially the protection of religion and intellect, and existing artificial intelligence governance frameworks. All seven are conceptual proposals, not standards that have been tested.

First, functional classification and honest naming. The system does not call its output a "fatwa", distinguishes general information, search assistance, and fatwa, and directs requests for a legal ruling on a personal

case to a mufti or fatwa institution. This principle is consistent with the distinction in Section 4.1 and with the demand for transparency that users are interacting with an artificial intelligence system [57], [58].

Second, source traceability and corpus curation. RAG can be run on a corpus curated by scholarly institutions, for example reference books and official fatwa collections, by displaying checkable source texts, marking the school of law and differences of opinion, and declining to answer when sources are inadequate [45], [59]. This step reduces the document problem, but it does not replace sanad.

Third, human oversight and a clear address of responsibility. High-implication cases (worship, family, inheritance, finance, health, declarations of unbelief, and violence) need to be escalated to a mufti. Fatwa institutions need to designate who is responsible for a system that bears their name. This principle is consistent with the emphasis on human oversight and determination and on accountability [10], [58] and with the affirmation that a mufti bears responsibility for his fatwa [8]. An artificial intelligence-assisted fatwa should be attributed only after it has been verified by an authorized mufti.

Fourth, openness about uncertainty and limits. The system states the limits of its knowledge, brings out ikhtilaf instead of a single final answer, and does not claim consensus without verification. Calibration of model uncertainty is not a guarantee [51], so this openness needs to be supported by interface design and testing.

Fifth, independent evaluation and bias audits. Panels of ulama and experts can compile benchmarks of fiqh questions across schools of law and languages, and test for hallucinated evidence, bias, and the tendency to please users, accompanied by documentation of the intended use and limitations of the system [50], [53], [59], [60].

Sixth, religious digital literacy. Users need to be equipped with an understanding of when artificial intelligence is appropriate to use (learning, finding references) and when they need to ask the ulama (putting into practice), as well as the importance of scholarly sanad. The role of educational institutions and religious organizations is important here, consistent with UNESCO's attention to awareness and literacy [58] and with the finding that online sources have become part of Islamic knowledge production [18]. In practice, this literacy can be built through teacher training, a short module on evaluating artificial intelligence-generated religious answers within existing fiqh and Islamic studies courses, and guidance for students on when to take a question to a teacher or fatwa institution.

Seventh, layered and risk-based governance. National ethical guidelines [10], risk management frameworks [60], transparency provisions [57], and the internal guidelines of religious institutions can complement one another. MUI's plan to draw up guidelines on the use of artificial intelligence for religious reference [9] indicates this direction; the content and final status of those guidelines need to be traced in official MUI documents. Ulama need to be involved from the design stage, and maqasid is used to weigh benefit and harm, not as a tool of justification.

Table 2. Mapping of epistemological risks, affected aspects of authority, and mitigation

Risk	Main mechanism	Affected aspect of authority	Main mitigation
Opaque sources	Knowledge dispersed in model parameters; references missing, wrong, or detached from context	Evidence and transmission (sanad)	Corpus curation; RAG with checkable source texts; declining to answer when sources are inadequate
Hallucination	Plausible text that is unfaithful to sources or facts	Truthfulness of evidence and attribution of opinions	Expert verification of attribution; testing for hallucinated evidence; escalation of high-risk cases
Bias	Training data, design, and use that are not neutral	Fairness in the representation of schools of law, languages, and groups	Bias audits; cross-school corpus; documentation of limits of use
Loss of context	Context limited to the user's text; generic answers across jurisdictions	*Tahqiq al-manat* and sensitivity to custom and circumstances	Clarifying questions; directing personal cases to a mufti; jurisdiction tagging
Appearance of authority	Fluency, confident tone, format of evidence, and the tendency to please users	Community recognition and responsibility	Labeling as not a fatwa; communicating uncertainty and ikhtilaf; user literacy

Two extreme positions are equally problematic. A total prohibition ignores the reality of use and the potential benefits, such as broader access and literacy, whereas unconditional permission ignores the structure of warrant behind a fatwa. The stance of the 2023 NU National Conference, namely that asking is permitted but treating answers as guidance to be put into practice is forbidden [7], can be read as a middle path that separates the function of information from the function of guidance for action; the principles above seek to make that middle path operational. The implementation challenges are real: the capacity and cost of curation for institutions, differences between schools of law and organizations, the absence of specific binding regulation, a global product market beyond the control of local institutions, and the possibility that "ulama-labeled artificial intelligence" may actually strengthen the appearance of authority if its oversight is weak.

4.6. Limitations and agenda for further research

Because this study is conceptual, it needs to be complemented by empirical research, for example audits of the answers of various systems to fiqh questions across schools of law and languages, studies of Indonesian users' perceptions of chatbot authority, evaluation of RAG on curated corpora of classical books, and legal studies on the allocation of responsibility among developers, operators, religious institutions, and users. Comparative studies across legal traditions, including Shia [16], are also needed so that this framework is not limited to one tradition.

5. CONCLUSION

This article answers the three research questions through conceptual analysis based on a literature review. First, fatwas and chatbot answers differ epistemologically primarily in the structure of their warrant: a fatwa is produced by a qualified subject through usul al-fiqh methods, is sensitive to the questioner's context, is accountable, and is connected to scholarly transmission, whereas chatbot output is probabilistic text with no subject burdened with responsibility and no personal sanad. For this reason, chatbot output is better understood as unverified religious information, not a fatwa.

Second, the epistemological risks include opaque sources, hallucination, bias, loss of context, and an appearance of authority, with a common root in the separation between the form of a fatwa and its epistemic warrant. The impacts may take the form of a shift in the gatekeeping function of the ulama, a simplification of legal understanding, and consequences for religious practice, especially in high-implication domains. They also reach Islamic education, where students may come to treat fluent output as knowledge, so educators need to distinguish artificial intelligence as a learning aid from artificial intelligence as a source of rulings. These impacts are not singular: artificial intelligence can be a tool for ulama and fatwa institutions if placed in an appropriate function.

Third, mitigation principles consistent with the Islamic scholarly tradition include functional classification and honest naming, source traceability, human oversight with a clear address of responsibility, openness about uncertainty, independent evaluation, religious digital literacy, and layered, risk-based governance. This conclusion is conceptual and is limited by the absence of empirical testing and by the rapid pace of technological change. Further research in the form of system audits, studies of user perceptions, and studies of legal responsibility is needed to test and refine the proposed framework.

AUTHOR CONTRIBUTIONS

Marlina: conceptualization, methodology, investigation, writing – original draft, and project administration. Sri Yanti: literature review, validation, writing – review and editing, and supervision. Both authors have read and approved the final version of the manuscript.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

DATA AVAILABILITY

No new empirical data were generated or analyzed in this study. All sources analyzed are cited in the References.

ACKNOWLEDGMENTS

The authors thank the Faculty of Tarbiyah, Institut Agama Islam Negeri Takengon, and Universitas Muhammadiyah Palembang for institutional support, and the colleagues who commented on earlier drafts of this article.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

REFERENCES

[1] D. F. Eickelman and J. W. Anderson, Eds., *New Media in the Muslim World: The Emerging Public Sphere*, 2nd ed. Bloomington, IN, USA: Indiana Univ. Press, 2003.

[2] H. Campbell, "Who's got the power? Religious authority and the Internet," *J. Comput.-Mediat. Commun.*, vol. 12, no. 3, pp. 1043–1062, 2007.

[3] G. R. Bunt, *iMuslims: Rewiring the House of Islam*. Chapel Hill, NC, USA: Univ. North Carolina Press, 2009.

[4] G. R. Bunt, *Hashtag Islam: How Cyber-Islamic Environments Are Transforming Religious Authority*. Chapel Hill, NC, USA: Univ. North Carolina Press, 2018.

[5] Z. Ji *et al*., "Survey of hallucination in natural language generation," *ACM Comput. Surv.*, vol. 55, no. 12, Art. no. 248, 2023.

[6] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?," in *Proc. ACM Conf. Fairness, Accountability, Transparency (FAccT)*, 2021, pp. 610–623.

[7] "Munas NU 2023: Bertanya ke AI boleh, tapi haram dijadikan pedoman untuk diamalkan," *NU Online*, Sep. 19, 2023. Accessed: Nov. 4, 2026. [Online]. Available: https://nu.or.id/nasional/munas-nu-2023-bertanya-ke-ai-boleh-tapi-haram-dijadikan-pedoman-untuk-diamalkan-HvEJ3

[8] "Kiai Cholil Nafis ingatkan AI tak bisa jadi mujtahid atau pemberi fatwa," *NU Online*, Aug. 13, 2025. Accessed: Nov. 4, 2026. [Online]. Available: https://www.nu.or.id/nasional/kiai-cholil-nafis-ingatkan-ai-tak-bisa-jadi-mujtahid-atau-pemberi-fatwa-wo0YX

[9] "MUI siapkan pedoman pemanfaatan AI untuk rujukan keagamaan," *NU Online*, Nov. 18, 2025. Accessed: Nov. 4, 2026. [Online]. Available: https://nu.or.id/nasional/mui-siapkan-pedoman-pemanfaatan-ai-untuk-rujukan-keagamaan-DHf7z

[10] Menteri Komunikasi dan Informatika Republik Indonesia, *Surat Edaran Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial*. Jakarta, Indonesia, Dec. 19, 2023. Accessed: Nov. 4, 2026. [Online]. Available: https://jdih.komdigi.go.id/produk_hukum/view/id/883/t/surat+edaran+menteri+komunikasi+dan+informatika+nomor+9+tahun+2023

[11] S. Ahmed, "Fatwā and AI: A literature-based assessment on how AI impacts the role of the muftī," *Journal for Islamic Studies*, vol. 42, no. 1–2, 2024, doi: 10.25159/2957-9163/14400.

[12] A. Jan, "Digital authority and Islamic legal reasoning: A comparative study of AI-generated and human-issued fatāwā," *Online J. Res. Islamic Stud.*, vol. 13, no. 1, pp. 1–16, 2026, doi: 10.22452/ris.vol13no1.1.

[13] S. Tsourlaki, "Artificial intelligence on Sunni Islam's fatwa issuance in Dubai and Egypt," *Islamic Inquiries*, vol. 1, no. 2, pp. 107–125, 2022, doi: 10.22034/IS.2022.339182.1082.

[14] A. A. Munshi, W. H. AlSabban, A. T. Farag, O. E. Rakha, A. A. Al Sallab, and M. Alotaibi, "Automated Islamic jurisprudence legal opinions generation using artificial intelligence," *Pertanika J. Sci. Technol.*, vol. 30, no. 2, pp. 1135–1156, 2022, doi: 10.47836/pjst.30.2.16.

[15] A. Polat, S. Khan, and U. Bashir, "An inquiry into the application of artificial intelligence on fatwā," in *Digital Transformation in Islamic Finance: A Critical and Analytical View*, Y. Suzuki and M. D. Miah, Eds. London, U.K.: Routledge, 2022, pp. 273–286, doi: 10.4324/9781003262169-18.

[16] H. Latifi, "Challenges of using artificial intelligence in the process of Shiʿi *ijtihad*," *Religions*, vol. 15, no. 5, p. 541, 2024, doi: 10.3390/rel15050541.

[17] S. A. Malik, "Artificial intelligence and Islamic thought: Two distinctive challenges," *J. Islamic Muslim Stud.*, vol. 8, no. 2, pp. 108–115, 2024, doi: 10.2979/jims.00020.

[18] M. S. Hasyim, "A new Islamic knowledge production and fatwā rulings: How Indonesia's young Muslim scholars interact with online sources?," *J. Indones. Islam*, vol. 14, no. 2, pp. 499–518, 2020, doi: 10.15642/JIIS.2020.14.2.499-518.

[19] I. Yahya and A. A. Junaidi, "Authority of fatwā in the digital era: Methodology and transmission of Nahdlatul Ulama's Bahtsul Masail in Indonesia," *Webology*, vol. 18, no. 2, 2021, doi: 10.14704/WEB/V18I2/WEB18345.

[20] S. Yucel and A. Ismail, "Are online fatwas credible? A critical analysis of fifty fatwas on organ donation and transplantation," *Religions*, vol. 12, no. 10, p. 902, 2021, doi: 10.3390/rel12100902.

[21] W. B. Hallaq, *A History of Islamic Legal Theories: An Introduction to Sunnī Uṣūl al-Fiqh*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[22] M. H. Kamali, *Principles of Islamic Jurisprudence*, 3rd ed. Cambridge, U.K.: Islamic Texts Society, 2003.

[23] Al-Shāfiʿī, *Islamic Jurisprudence: Shāfiʿī's Risāla*, M. Khadduri, Trans. Baltimore, MD, USA: Johns Hopkins Press, 1961.

[24] K. Abou El Fadl, *Speaking in God's Name: Islamic Law, Authority and Women*. Oxford, U.K.: Oneworld, 2001.

[25] W. B. Hallaq, *Authority, Continuity, and Change in Islamic Law*. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[26] M. Q. Zaman, *The Ulama in Contemporary Islam: Custodians of Change*. Princeton, NJ, USA: Princeton Univ. Press, 2002.

[27] M. K. Masud, B. Messick, and D. S. Powers, "Muftis, fatwas, and Islamic legal interpretation," in *Islamic Legal Interpretation: Muftis and Their Fatwas*, M. K. Masud, B. Messick, and D. S. Powers, Eds. Cambridge, MA, USA: Harvard Univ. Press, 1996, pp. 3–32.

[28] W. B. Hallaq, "*Iftāʾ* and *ijtihād* in Sunnī legal theory: A developmental account," in *Islamic Legal Interpretation: Muftis and Their Fatwas*, M. K. Masud, B. Messick, and D. S. Powers, Eds. Cambridge, MA, USA: Harvard Univ. Press, 1996, pp. 33–43.

[29] J. Skovgaard-Petersen, *Defining Islam for the Egyptian State: Muftis and Fatwas of the Dār al-Iftā*. Leiden, The Netherlands: Brill, 1997.

[30] Y. al-Nawawī, *Adab al-Fatwā wa al-Muftī wa al-Mustaftī*. Damascus, Syria: Dār al-Fikr, 1988.

[31] Shihāb al-Dīn al-Qarāfī, *al-Iḥkām fī Tamyīz al-Fatāwā ʿan al-Aḥkām wa Taṣarrufāt al-Qāḍī wa al-Imām*. Aleppo, Syria: Maktab al-Maṭbūʿāt al-Islāmiyyah, 1995.

[32] Ibn Qayyim al-Jawziyyah, *Iʿlām al-Muwaqqiʿīn ʿan Rabb al-ʿĀlamīn*. Beirut, Lebanon: Dār al-Kitāb al-ʿArabī, 1998.

[33] S. A. Jackson, "The second education of the mufti: Notes on Shihāb al-Dīn al-Qarāfī's tips to the jurisconsult," *Muslim World*, vol. 82, no. 3–4, pp. 201–217, 1992, doi: 10.1111/j.1478-1913.1992.tb03553.x.

[34] M. B. Hooker, *Indonesian Islam: Social Change through Contemporary Fatāwā*. Honolulu, HI, USA: Univ. Hawaiʿi Press, 2003.

[35] N. Hosen, "Behind the scenes: Fatwās of Majelis Ulama Indonesia (1975–1998)," *J. Islamic Stud.*, vol. 15, no. 2, pp. 147–179, 2004.

[36] M. N. Ichwan, "Ulama, state and politics: Majelis Ulama Indonesia after Suharto," *Islamic Law Soc.*, vol. 12, no. 1, pp. 45–72, 2005.

[37] W. B. Hallaq, "Was the gate of *ijtihād* closed?," *Int. J. Middle East Stud.*, vol. 16, no. 1, pp. 3–41, 1984.

[38] J. A. C. Brown, *Hadith: Muhammad's Legacy in the Medieval and Modern World*. Oxford, U.K.: Oneworld, 2009.

[39] G. Makdisi, *The Rise of Colleges: Institutions of Learning in Islam and the West*. Edinburgh, U.K.: Edinburgh Univ. Press, 1981.

[40] Abū Isḥāq al-Shāṭibī, *al-Muwāfaqāt fī Uṣūl al-Sharīʿa*, M. H. Āl Salmān, Ed. Al-Khobar, Saudi Arabia: Dār Ibn ʿAffān, 1997.

[41] J. Auda, *Maqāṣid al-Sharīʿah as Philosophy of Islamic Law: A Systems Approach*. London, U.K.: International Institute of Islamic Thought, 2008.

[42] A. Vaswani *et al.*, "Attention is all you need," in *Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 30, 2017, pp. 5998–6008.

[43] T. B. Brown *et al.*, "Language models are few-shot learners," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 1877–1901.

[44] L. Ouyang *et al.*, "Training language models to follow instructions with human feedback," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 35, 2022, pp. 27730–27744.

[45] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 9459–9474.
[46] E. M. Bender and A. Koller, "Climbing towards NLU: On meaning, form, and understanding in the age of data," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2020, pp. 5185–5198.
[47] S. Lin, J. Hilton, and O. Evans, "TruthfulQA: Measuring how models mimic human falsehoods," in *Proc. 60th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, vol. 1, 2022, pp. 3214–3252.
[48] R. Navigli, S. Conia, and B. Ross, "Biases in large language models: Origins, inventory, and discussion," *ACM J. Data Inf. Qual.*, vol. 15, no. 2, Art. no. 10, 2023.
[49] N. F. Liu *et al.*, "Lost in the middle: How language models use long contexts," *Trans. Assoc. Comput. Linguistics*, vol. 12, pp. 157–173, 2024.
[50] M. Sharma *et al.*, "Towards understanding sycophancy in language models," arXiv:2310.13548, 2023.
[51] S. Kadavath *et al.*, "Language models (mostly) know what they know," arXiv:2207.05221, 2022.
[52] M. Dahl, V. Magesh, M. Suzgun, and D. E. Ho, "Large legal fictions: Profiling legal hallucinations in large language models," *J. Legal Anal.*, vol. 16, no. 1, pp. 64–93, 2024.
[53] A. Abid, M. Farooqi, and J. Zou, "Persistent anti-Muslim bias in large language models," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc. (AIES)*, 2021, pp. 298–306.
[54] "Kecerdasan buatan jadi panduan fatwa, Munas NU: 'Haram!'," *ANTARA Kaltim*, Sep. 19, 2023. Accessed: Nov. 4, 2026. [Online]. Available: https://kaltim.antaranews.com/berita/195243/kecerdasan-buatan-jadi-panduan-fatwa-munas-nu-haram
[55] M. Syakir NF, "Munas NU 2023 rekomendasikan PBNU bikin AI," *NU Online*, Sep. 20, 2023. Accessed: Nov. 4, 2026. [Online]. Available: https://nu.or.id/nasional/munas-nu-2023-rekomendasikan-pbnu-bikin-ai-mluw3
[56] "Komisi BM Waqi'iyah Munas NU: Pengembangan AI keagamaan bersifat fardhu kifayah," *NU Online*, Sep. 19, 2023. Accessed: Nov. 4, 2026. [Online]. Available: https://nu.or.id/nasional/komisi-bm-waqi-iyah-munas-nu-pengembangan-ai-keagamaan-bersifat-fardhu-kifayah-YQeus
[57] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," *Off. J. Eur. Union*, vol. L, Art. no. 2024/1689, Jul. 12, 2024.
[58] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*. Paris, France: UNESCO, 2021.
[59] M. Mitchell *et al.*, "Model cards for model reporting," in *Proc. Conf. Fairness, Accountability, Transparency (FAT*)*, 2019, pp. 220–229.
[60] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, Gaithersburg, MD, USA, 2023, doi: 10.6028/NIST.AI.100-1.