



PLANNING MALAYSIA:

Journal of the Malaysian Institute of Planners

VOLUME 24 ISSUE 3 (2026), Page 228 – 242

AI CONTENT FABRICATION AND DECEPTION: AN ISLAMIC ETHICAL AND SOCIETAL ANALYSIS

Fareed Ahmad Malik¹, Mohd Shukri Hanapi²

^{1,2} *Centre for Islamic Development Management Studies (ISDEV),*
UNIVERSITI SAINS MALAYSIA

Abstract

The pervasive increase of AI-generated content, such as deepfakes and forged religious texts has raised serious concerns about deception, disinformation, trust breaches, ethical accountability, and more. These issues concern both secular and religious spheres. Although research on the ethics of AI is growing, the Islamic scholarly works that deal specifically with ethical frameworks of AI-generated misinformation and deception are inadequate. This study employs a qualitative interpretive approach rooted in ethical reasoning and normative analysis, involving hermeneutical engagement with primary Islamic texts, classical legal opinions, along with contemporary fatwas, literature, and illustrative real-life case studies. This study formulates an Islamic ethical framework rooted in principles such as *ṣidq* (truthfulness), *amānah* (trust), and *maṣlahah* (public interest), which serve as foundation for a profound digital trust infrastructure. This framework tackles the epistemic and ontological dilemmas of AI, while maintaining human ethical accountability of AI content. The findings reveal that the concoction of content is forbidden (*ḥarām*) in the Shariah due to its capacity to cause societal harm and division in the society. Effective public-information governance necessitates Interdisciplinary collaboration of scholars and institutions, technologists, educators and policy makers to create effective verification and digital literacy protocols. This framework provides a structured approach that combines Islamic principles along with technology development to enhance ethical AI oversight and awareness.

Keywords: Islamic Ethics, Artificial Intelligence, Content Fabrication, Deepfakes, Techno-Ethical Governance

² Corresponding author. Email: hshukri@usm.my

INTRODUCTION

The global expansion of AI technologies can be seen as the beginning of an era full of innovation (Islami et al., 2025; Yakub et al., 2021) and ethical and socio-cultural uncertainty (Momeni, 2024). AI generating false content is an increasingly contentious matters, posing significant dangers as anyone can create hyper-realistic and false digital content (Popa et al., 2025). At present, anyone can create a hyper-realistic but false digital object for any purpose, using deep-fake or voice-impersonating technologies (Popa et al., 2025). It affects politics, spreads disinformation, religion, personal attacks, and other far-reaching implications (Kharvi, 2024). Pira (2023) argues that this constitutes a technology revolution that impacts how societies establish truth and trust institutions. Although secular ethical frameworks highlight the importance of transparency and accountability, their impact, particularly on fake information, fatwas, fabricated hadiths, and impersonation of scholars, is limited (Abdulrahman & Walusimbi, 2024; Bhojani & Schwarting, 2023). In Muslim societies where religious knowledge is generally disseminated by scholars (*'ulamā'*), religious institutions and authoritative texts, the impact of AI-generated fake content shows its own repercussions. A case of this is the jeopardy of AI content fabrication in Muslim societies. The proliferation of AI-driven deepfakes on religion-related matters creates more confusion among the digitally engaged population, a phenomenon increasingly explored in academic discourse (Al-Arja, 2025; Metz & Blumenthal, 2019). Such deceptive content, capable of inciting extremism and falsely justifying illegal acts, poses a big threat to community resilience and social cohesion. In the era of digitalization, the misinformation created by AI led to epistemic moral anarchy, which is forbidden in Islam (Hate, 2024; Zhang et al., 2025). The revealed text of Islam promotes truthfulness (*ṣidq*) and the preservation of peace (*sulh*). The Prophet Muhammad (peace be upon him) said, “whoever deceives us is not one of us” (Sahih Muslim, 102). So, any sort of AI-generated deception can be construed as a moral violation with serious spiritual consequences.

The study examines the use of AI-generated content specifically in religious texts and to explore it from an Islamic ethical and social understanding. The study adopts basic principles of Islam and applies them to real-life case studies of deepfake fatwas, fabricated hadiths, and politically motivated deepfake videos to formulate an ethical framework. This study is guided by the following key research questions.

1. How do Islamic ethical and juristic values assess deceitful content generated by AI, particularly in the fabrication and dissemination of religious or political content?

2. What are the implications of AI-generated misinformation on social and religious life of people in Muslim societies? And how it impacts on the community's admiration for religious authority, public trust, and harmony?
3. How Islamic values including *maqāṣid al-shariah*, *amānah* and *ṣidq* can support the ethical governance of AI technology while enhancing public awareness and institutional regulations?

The proposed Islamic ethical framework for AI governance aims to inform robust policy responses. This includes drawing up institutional regulations for AI development and deployment, alongside mandates for public awareness and digital literacy programs to prevent digital harm.

LITERATURE REVIEW

This section reviews the key themes arising from AI ethics and Islamic moral philosophy, delineates their areas of convergence and differences, and emphasizes the pressing necessity for contextually linked Islamic ethical responses.

Recent Empirical Studies on AI-generated Misinformation, Deepfakes, and Synthetic Media

The recent statistical data of OECD (2026) that reports a tenfold increase in AI content incidents (Gaudiaut, 2026; OECD, 2026), reveals a remarkable increase in AI-generated misinformation and its societal repercussions. A global survey conducted by Gillespie et al. (2025) reveals that 70% of the respondents find it difficult to trust information on the internet due to the uncertainty regarding AI-generated content. 64% are concerned about the ability of the AI content to meddle with the electoral process and the issue of democracy losing its integrity. Only 47% of people are confident in spotting AI-generated misinformation. There is strong public support for regulation of AI. 88% advocate for laws to curb AI misinformation, while 86% support enhanced fact checking by news and social media organizations, alongside mechanisms to identify AI-generated content across text, images, audio, and video (See Table 1).

Table 1: Impact, and Responses to AI-Generated Misinformation

Impact Description	% Agree*	Number of Incidents	Source
Media-reported incidents of AI-generated content and deepfakes (Jan 2026)	—	~500 monthly cases	(Gaudiaut, 2026; OECD, 2026)
Media-reported incidents baseline (early 2020)	—	~50 monthly cases	
Escalation Factor	—	10x increase	

Impact Description	% Agree*	Number of Incidents	Source
Difficulty trusting online information due to AI-generated content uncertainty	70%	—	(Slotta, 2025)
Difficulty discerning AI-generated content (A Global Survey Report, 2025)	70%	—	(Gillespie et al., 2025)
Concern about AI-generated content manipulating electoral processes	64%	—	
US adults concerned about deepfake videos and audio	60%	—	(Statista, 2026)
Global online users regarding deepfakes and AI-assisted scams as threats	69%	—	(Zandt, 2024)
Confidence in ability to detect AI-generated misinformation	47%	—	(Gillespie et al., 2025)
Support for enhanced fact-checking protocols	86%	—	
Support for laws to curb dissemination of AI-generated misinformation	88%	—	

*% Agree = ‘Somewhat agree’, Agree’, and ‘Strongly agree’

Source: Compiled by the authors based on empirical studies and data extracted from: (Gaudiaut, 2026), (Gillespie et al., 2025), (Slotta, 2025), (Statista, 2026), and (Zandt, 2024).

In the United States, 60% of adults express concern about AI-produced deepfake videos and audio, while 69% of global online users consider deepfakes and AI-assisted scams significant threats (Statista, 2026; Zandt, 2024). The European Union's Artificial Intelligence Act, implemented in August 2024, addresses AI-driven misinformation risks (Gillespie et al., 2025; Stoll, 2025). Nonetheless, data on AI-generated religious misinformation remains limited, with statistics primarily focusing on political and social impacts. The evidence shows we need comprehensive strategies to counter this, which combines legal frameworks to impose accountability on AI developers, technological safeguards to detect and prevent misinformation, and public education.

Major AI Ethical Concerns in Secular Perspective

Within the secular AI ethics space, we can find the Asilomar AI Principles (Future of Life Institute, 2025), OECD Principles on AI (OECD, 2025), and Ethics Guidelines from the EU's High-Level Expert Group on AI (Lemonne, 2021), based on norms of transparency, accountability, fairness, explainability, and non-maleficence. Nonetheless, these elements do not encompass certain fundamental components present in the social and theocentric framework of Islamic ethics (Huang et al., 2022). Mainstream AI ethics lacks religious perspectives, which is problematic for Muslim societies and globally significant,

emphasizing the need for inclusive ethical governance in AI development (Chaudhary, 2020; Roche et al., 2022; Elmahjub, 2023). With religion influencing global morals, any AI ethics excluding religious traditions like Islam is incomplete. Islam's legal reasoning and ethical governance apply to digital issues. The lack of Islamic values in AI ethics seems to assert a foreign hegemony over Muslim societies. Legal maxims (*qawā'id fiqhiyyah*) must be employed to engage Islamic ethical values. According to Al-Zuhaili (2006), the maxim *ḍarar yuzāl* (harm must be avoided) concerns AI safety, while *al-umūr bi-maqāṣidihā* (actions are judged by intentions) relates to AI accountability.

Islamic Moral Values on *Ṣidq* (Truth), *Amānah* (Trust), and *Ḍarar* (Harm)

Islamic ethics, rooted in the Qur'an and Prophetic traditions, provide moral terms for addressing deception. Key ethical values include truthfulness (*ṣidq*), trust (*amānah*), and preventing *ḍarar* (harm). The Qur'an commands believers to uphold truth and avoid falsehood. In the Qur'an (49:6), it states "*O you who have believed, if there comes to you a disobedient one with information, investigate, lest you harm a people out of ignorance and become, over what you have done, regretful.*" This principle is emphasized in Prophetic traditions, including the hadith narrated by Muslim that whoever deceives us is not one of us. The principles of *amānah* require honesty in speech, trustworthiness in knowledge dissemination, and integrity in responsibilities. These principles are violated when AI technologies deceive people through impersonating scholars or spreading false religious content. The legal maxim '*al-ḍarar yuzāl*' (harm must be avoided) forbids spreading synthetic news in Islam. These principles are embedded in Islamic legal objectives (*maqāṣid*), including protection of religion (*ḥifẓ al-dīn*), life (*ḥifẓ al-naḥs*), intellect (*ḥifẓ al-'aql*), lineage (*ḥifẓ al-nasl*), and wealth (*ḥifẓ al-māl*). AI-generated content can undermine these objectives, making this issue significant for Islamic legal deliberations.

Contemporary Islamic Scholarship and Institutional Responses to AI Technologies

Recent advancements in Islamic scholarship show increasing recognition of ethical dilemmas posed by AI. Scholars and institutions actively discuss AI ethics issues. Scholars like Shaykh Abdallah bin Bayyah and Dr. Jasser Auda advocate reinterpreting *maqāṣid al-sharī'ah* for digital ethics, emphasizing Islamic legal theory's adaptation to AI. Discussions focus on moral accountability (*taklīf*) of AI developers and users, exploring principles like *amānah* (trust) and *ṣidq* (truthfulness) in AI content. Scholars examine *ijtihād* in the digital era to authenticate AI-produced religious content. Islamic institutions like Dar al-Ifta', Majelis Ulama Indonesia, and the International Islamic Fiqh Academy have issued fatwas on AI ethics, denouncing deepfakes and false claims. They emphasize verification and labelling to counter AI misinformation. Collaborative

frameworks uniting scholars, technologists, and ethicists are needed for AI guidelines based in Islamic jurisprudence. Universities in Muslim-majority countries are establishing AI ethics research units integrating Islamic principles with governance models. Islamic digital literacy curricula evaluate AI content through principles like *ḥifẓ al-dīn* and *ḍarar yuzāl*.

Although the literature on AI ethics is growing, but scholarship highlighting the gaps from an Islamic perspective is scant. Lack of comprehensive Islamic frameworks to evaluate the misinformation generated by AI, especially disinformation affecting spiritual authority. There are not enough case studies demonstrating how AI manipulates data to fool Muslims using deepfake fatwas and AI-generated impersonations. Most fatwa bodies and Islamic leaders ignore procedures for assessing AI technology risks. The promotion of Islamic digital literacy is insufficient, particularly among the youth who are susceptible to misinformation through mobile applications and social media. The global discourse on AI governance is largely dominated by secular frameworks and lacks Islamic perspectives. The holes necessitate an integrated approach that considers the ethical issues of AI disinformation and provides solutions based on Islamic philosophy. This research aims to contribute towards offering a framework for assessing AI-generated content, resolving strategy issues including real-life case studies, and ideas in an Islamic perspective.

METHODOLOGY

This research employs a qualitative interpretive method based on Islamic ethical reasoning and normative analysis to critique the AI content fabrication and disinformation, particularly among Muslim societies. Using hermeneutical engagement with primary Islamic texts and classical works of jurisprudence, it develops normative ethical principles while normative legal reasoning with approaches such as *qiyās*, *maqāṣid al-shariah*, and *qawā'id fiqhiyyah*. This framework allows for a principled evaluation of AI-based misinformation using ethical concepts such as preventing harm, intention-based judgement, and safeguarding the public interest. The present study uses a variety of primary and secondary data that are relevant to the research questions.

A triangular data source model (as shown in **Figure 1**) is applied to the present study. The data sources are based on a range of normative, jurisprudential and social responses to the research questions. The approach involves case studies of incidents involving AI misinformation, triangulation of secondary sources, including fatwas and academic literature, and socio-technical contextualization from digital media and AI governance. Quantifying the impact of AI-generated religious misinformation with empirical and statistical data, A dynamic approach would adjust the ethical framework to change in context and assists in creating policy recommendations to alleviate the issues of AI deception in Muslim societies.

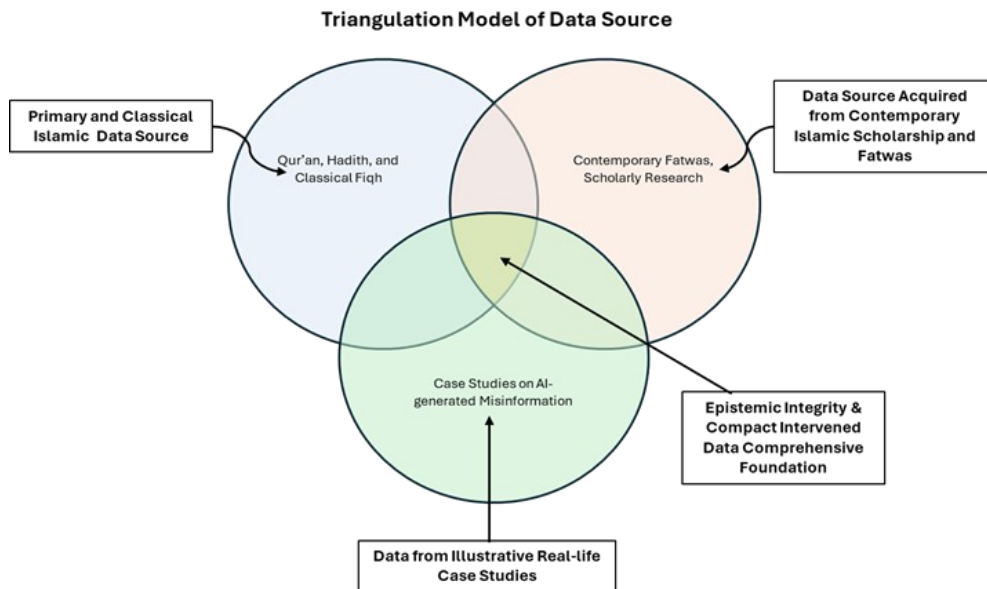


Figure 1: Triangulation Data Source Model
Source: Created by the Author

CASE STUDIES

This section investigates three different illustrative case studies that have raised ethical, epistemic and social issues regarding AI-generated content from within the Muslim majority regions. These case studies were chosen for their widespread impact and high social-religious stakes. These cases serve as critical real-life illustrations, providing empirical justification for the urgent development of robust planning and governance policies to avoid AI-driven harms. To achieve these objectives, the study aims to showcase how *maqāṣid al-shariah*, *qawā'id fiqhīyyah*, and epistemic integrity can help provide diagnoses and measures to tackle the cases of AI disinformation by applying Islamic ethics principles and juristic reasoning to these cases.

Case 1: Deepfake Fatwa by a Prominent Scholar

In June 2024, the Media Center of the Egyptian Dar al-Iftaa affirmed that Minister of Endowments Dr. Mohamed Mokhtar Gomaa did not endorse dubious individuals or disseminate spurious hadiths. Dar al-Iftaa exonerates the Minister of Endowments in the AI-generated deepfake video. The center stated that the video of the Minister of Endowments circulating online was fabricated and deceptive, generated by artificial intelligence, and emphasized that it served as a promotional campaign for an electronic gaming application. The center further clarified that Dr. Mokhtar Gomaa does not endorse dubious individuals. The

Media Center highlighted the peril of employing AI to distort facts and create deceptive videos that endorse dubious products or applications aimed at defrauding individuals, asserting that all legal actions would be pursued in this matter (Ahl Masr News, 2024; OECD.AI, 2024).

Impact and Analysis

The public's response was outrage and confusion, with calls for investigation. Religious authorities were thrown into disarray and had to issue clarifications and takedown demands. However, this episode made it such that authorities would have to discuss the issue. With a sermon penned, local muftis and preachers addressed Friday prayer service, warning against unauthenticated religious content.

From a jurisprudential standpoint, the incident is concerned with the relationship between the *‘āmmī* (layperson) and the *‘ālim* (scholar). The Islamic legal tradition allows *taqlīd* (the following of a scholar's opinion) when one does not have skills for *ijtihād* (the independent reasoning). *Taqlīd* demands an understanding of the scholar's identity, background, and credibility. A fatwa made by a digital device that imitates the appearance of a scholar is termed as false attribution (*iftirā*) and is a violation of the maxim *‘al-umūr bi-maqāsidihā* (actions are judged by intentions). Reforming trust infrastructure requires a well-defined accountability, as AI-generated content without *niyyah* (pure intention) and human oversight loses legal status. Second, it is the responsibility (*amānah*) of institutions to implement authentication protocols, thereby restoring epistemic integrity and public trust.

Case 2: AI-Generated Hadiths in Islamic Apps

In June 2025, observations of numerous Islamic mobile applications using generative AI have revealed instances where some apps produced weak and forged Islamic content (Patel, 2025). This case illustrates a prevalent issue rather than a single or isolated event. An unbiased appraisal undertaken by Islamic scholars showed that some of the apps produced weak (*da‘īf*) and completely fabricated (*mawḍū‘*) hadiths and false information. There were no citations or grading of authenticity for most of the content. Many users, including youth and people learning about Islam, allegedly widely circulated the content on social media thinking it was true (Patel 2025).

Repercussions and Ethical Judgment from Hadith Sciences and Epistemology

This scenario shows the incident of epistemological crisis in digital propagation (*da‘wah*) of religious content. Though these apps may have been developed with the intent of democratizing access to religious knowledge, they inevitably become misinformation platforms by their very design. Using such apps, the

youth are making themselves more vulnerable by relying on mobile apps instead of conventional sources such as scholars and textbooks.

The propagation of unauthenticated narrations wrongly portrays the sacred assets of the knowledge of Prophetic traditions, and it misleads the people to false beliefs and practices. Some fake content presented as Prophetic traditions may incite superstitions or extremist behaviours which have absolutely nothing to do with the very essence and balanced teachings of Islam.

The hadith sciences, from its classical development, has always had principles and rigorous criteria for assessing narrations to consider their authenticity. This includes a consideration of the transmitters (*rijāl al-ḥadīth*) and the chains of narration (*isnād*). Generative AI cannot produce works that are religiously authoritative because it is not produced by any of these filters. The use of fabricated hadith is a direct violation of the Qur'anic command, "*And do not say about what your tongues assert of untruth...*" (Qur'an 16:116). It also transgresses the moral principle *ṣidq* and the epistemic duty of verifying information before transmitting it. This breach is also mentioned in another verse of the Qur'an where Allah states: "*O you who have believed, if there comes to you a disobedient one with information, investigate, lest you harm a people out of ignorance and become, over what you have done, regretful.*" (Qur'an 49:6).

Religious texts written by AI cannot be accepted unless validated by qualified scholars. Applications that are created for such purpose must have an authorizing system. Those who fail to do so align with the promoters of disinformation that warns that causing an error is like committing it.

Case 3: Deepfake Incitement in Conflict Zones

In April 2024, a picture of burning a building went viral on Facebook in Sudan. According to the captions that circulated widely with the image stated, this building had belonged to Al-Jazeera University in Wad Madani and had been bombed by the Sudanese army. Many politicians and influential personalities became its victims and promulgated it. But that was not an isolated incident. As fighting continues between the national army and rebel paramilitary Rapid Support Forces, the social media arena has turned into another battleground. Both sides are using AI-driven deepfakes to spread disinformation and gain support. Such incidents pose considerable challenges to the Northeast African nation that desperately needs a sturdy information ecology, a concern widely discussed in studies on disinformation in conflict zones (Jareh, 2025; Suliman, 2024).

Impact and the relevance of Islamic legal perspective

This case demonstrates how AI can be deployed as a tool for factional provocation in conflict zones. Although forensic experts later proved that the footage was fake, the damage was already done. People became increasingly

estranged and there was no more dialogue between the two groups. What's worse, political leaders and other important people fell victim to it and shared it.

In an Islamic juristic perspective, such cases are governed by multi-layer *qawa'id fiqhiyyah* (legal maxims):

Firstly, for policy making, the Islamic legal maxim *Darar yuzāl* (harm must be removed) mandates immediate removal of false content and the adoption of preventive strategies, while holding platforms accountable.

Secondly, *Dar' al-mafāsīd muqaddam 'alā jalb al-maṣāliḥ* (preventing harm takes precedence over acquiring benefit) implies that AI technologies, despite potential benefits, require rigorous scrutiny and regulation in volatile contexts to prevent digital harm. Thirdly, the legal maxim *Al-maṣlaḥa al-'amma muqaddama 'alā al-maṣlaḥa al-khāṣṣa* focuses on public interest by taking precedence over private interest. To maintain harmony (*sulh*) and avoid discord (*fitnah*), public policies are introduced even if it curtails the right to freedom of expression to a certain extent of the content creator.

Beyond condemning digital incitement, Islamic ethics advocates for strong institutional coordination among scholars, government, and technical platforms, as essential strategies for improving community resilience against the harmful impact of AI deception. The Qur'an states: "*Indeed, those who like that immorality should be spread [or publicized] among those who have believed will have a painful punishment in this world and the Hereafter...*" (Qur'an 24:19).

DISCUSSION AND RECOMMENDATIONS

The rapid increase of fake AI content, especially content that impersonates religious authority, fabricates hadiths, or instigates religious sectarianism, raises critical legal and ethical questions for Islamic jurisprudence as to who bears the liability. Technology may be advanced, but the rulings given in the past have ample precedent on false religious literature, false testimonies, and misrepresentation of Islam.

Legal Maxims (*Qawā'id Fiqhiyyah*) as Ethical Anchors

Islamic Legal maxims apply Shariah to new areas. These principles provide ethical clarity and legal guidance beyond the classical texts. Two maxims about the ethics of AI-generated content are significant: *Al-Darar Yuzāl* (Harm Must Be Removed) and *Al-Umūr Bi-Maqāṣidihā* (Actions Are Judged by Intentions). The first maxim ranks amongst the highest in Islam jurisprudence. It holds that AI content causing misinformation, deviation from true religion, or social discord is harmful. Through disinformation, it can cause epistemic harm, mislead believers, and incite sectarianism. According to this advice, designers and users of AI should foresee and warn against dangers, which will ensure human intervention when needed. The second maxim relates to the ethical status and the intention (*niyyah*). Machines themselves have no intent, but the designers, users

and distributors do. Policy design for digital harm prevention must establish clear responsibility frameworks, ensuring liability falls on human designers and platform providers when AI undertakes tasks without supervision. This necessitates regulatory assessment of both the *maqāṣid* (objectives) and *niyyah* (intention) of the human actors involved in the AI content lifecycle.

Developers and Users: *Taklīf* and *Dhimmah*

The ethical and legal aspects of Islamic thought focus on the notion of moral agency and accountability. The term *mukallaḥūn* refers to the human condition. It is an Arabic word that relates to people who are bound by obligations. These are obligations that will make them responsible according to Shariah. In the context of AI, we may accentuate *mukallaḥūn*, which refers both to the developer and user.

Developers of generative AI systems producing religious content bear considerable responsibility. In the absence of any verification mechanisms or disclaimers in the system, where the data sources are not curated, and the AI is imitating known scholars without any consent, they will have to take more responsibility. The acts in question entail either an intentional misrepresentation (*tadlīs*) or negligent causation (*tasabbub*), resulting in binding the action and the agent with a *dhimmah*. They take part in epistemic harm with potential. According to *Surah Al-Hujrāt*, verse 6, the Qur'an instructs to verify information. Those who disseminate AI-generated content without checking will be liable. The duty to spread verified religious content lies with the users. This is because the Prophet (peace be upon him) said: "It is enough falsehood for a man to speak of everything that he hears" (Sahih Muslim, 31:6). Islamic ethics requires Muslims to practice good digital discernment, especially with matters of faith.

***Tazwīr* and Fabricated Fatwas in Fiqh Tradition**

Although there are new challenges with AI, Islamic law has been used in the past to address forgery and false knowledge of religion. Classical texts examine forgery, fake hadith and false fatwas. *Tazwīr* involves forging documents and faking signatures and seals. Jurists object to forgers since they threaten textual truth. The scholarship on false hadiths, as presented by the scholars like Ibn al-Jawzī (d. 1201) and al-Dhahabī (d.1348), explains the spiritual and legal consequences of following false narrations. People without proper knowledge who contested law were incorrect because of religious opinion faced very strong opposition. These judgments apply to AI systems that cause such issues in the digital environment. A false fatwa generated by AI impersonating a scholar is same as a fabricated fatwa in history. Like how Hadith scholars create chains of transmission (*isnād*) for verification, digital authentication must be made to preserve epistemic sanctity.

Recommendations for Framework on Integrating Islamic Ethics with Contemporary AI Technologies

The ethical challenges caused by AI-generated content in Muslim communities require a revised framework relating traditional Islamic ethics with modern AI technology. Although Islamic Law has principles such as truthfulness, trust, and avoidance of harm, applying these principles to the scale, autonomy, and complexity of AI poses various challenges.

An expansion of traditional Islamic legal maxims and their connection to ethical theories of AI, algorithmic decision-making and data bias is required. The principle of *maqāṣid al-sharī'ah* can be used anew for digital justice and accountability and ensure, at a minimum, the compatibility of AI systems with the public interest while safeguarding values.

The AI generating religious content raises several pertinent questions. The impact of machine-generated knowledge on information trustworthiness must be addressed by Islamic ethics. There is a need for such a criterion to evaluate the AI Output. There should be an Integration of the Islamic principles of *ijtihād* and *taqlīd* in the digital setting. Also, there should be human oversight of the religious content.

The framework must define ethical accountability in creating and sharing AI content. Islamic principles concerning *taklīf* (moral responsibility) apply to AI developers, platforms, and users in AI ecosystems. Although AI systems do not have intent, it is the humans designing and deploying the systems who are ethically responsible, which aligns with “actions are judged by intentions.”

To address the socio-cultural implications of AI, an Islamic framework for technology ethics must integrate insights from AI ethics, digital media studies, and social sciences. In Muslim communities, how AI will impact community harmony, religious leadership and identity, it is necessary for scholars, technologists and policymakers to create a framework for the development of AI that coordinates with Islamic framework through digital literacy programs and governance.

As AI develops, Islamic ethical governance must evolve from legal maxims to contemporary *ijtihād* and through the establishment of institutional capacity which is equipped to assess AI. Not only that, but ethical discussions must also be promoted along with the advancement of technological foresight into legal processes.

Implications

This framework recognizes the common ethical challenges that people experience such as preserving the integrity of religious knowledge, promoting communal harmony (*sulh*), and managing digital epistemic trust in the wake of AI-driven misinformation. It suggests:

- Formation of technology-oriented Islamic ethical rulings for AI that incorporate ethical plurality.
- Advancing digital literacy among people to protect them from the harmful aspects of AI.
- Fostering collaboration to create AI governance frameworks that cater to specific cultural contexts.
- Assistance for an international conversation on AI ethics that respects Islamic values and people's sentiments.

CONCLUSION

This study concludes with the fact that any promulgation of AI-generated disinformation such as fabricated fatwas, forged hadiths, and deepfake impersonations is categorically prohibited in Islamic law. This judgment proceeds from its direction violation of *ṣidq* and *amānah*, and its potential to inflict societal harm. The study used a qualitative interpretive method that brings together Islamic jurisprudence with real-life examples. This approach has helped to produce a clear ethical framework for addressing various issues on AI. It emphasizes the ethical accountability of AI developers, service providers, and users. It argues that while AI itself lack intent, human actors bear ethical and legal responsibility consistent with the legal maxims of Islam. The study recommends development of AI-specific ethics, propagation of digital awareness programs which are conducted by interdisciplinary and transdisciplinary scholars. The study highlights the essentiality of conducting such programs to safeguard epistemic integrity, harmony, and uphold the higher objectives of *maqāṣid* in rapidly changing digital world.

ACKNOWLEDGEMENT

This research was supported by the Post-doctoral Fellowship [FPD 2024.143] at Universiti Sains Malaysia (USM). I would like to express my sincere appreciation to my supervisor, Associate Professor Dr. Mohd Shukri Hanapi [from Centre for Islamic Development Studies (ISDEV), USM], for his guidance, valuable ideas, and continuous support.

REFERENCES

- Abdulrahman, M. M., & Walusimbi, A. H. M. (2024). Examining the role of artificial intelligence (GPT-3.5) in issuing fatwas for Islamic family cases: A comparative analysis. *Journal of Contemporary Islamic Law*, 9(2), 30–38. <https://doi.org/10.26475/jcil.2024.9.2.04>
- Ahl Masr News. (2024, June 14). The Fatwa Council acquits the Minister of Endowments of the artificial intelligence video. *Ahl Masr*. Retrieved June 17, 2025, from <https://ahlmassrnews.com/news/egyptnews/13294469/%D8%AF%D8%A7%D8%B1-%D8%A7%D9%84%D8%A5%D9%81%D8%AA%D8%A7%D8%A1>

- Al-Arja, A. (2025, May 18). Weaponized artificial intelligence: The unseen threat to factchecking. *AI Jazeera Media Institute*.
<https://institute.aljazeera.net/en/ajr/article/2861>
- Al-Zuhailī, M. (2006). *Al-qawā'id al-fiqhīya wa-taṭbīqātuhā fī 'l-maḍāhib al-arba'a* (aṭ-Ṭab'a al-ūlā). Dār al-Fikr.
- Bhojani, A.-R., & Schwarting, M. (2023). Truth and regret: Large language models, the Quran, and misinformation. *Theology and Science*, 21(4), 557–573.
<https://doi.org/10.1080/14746700.2023.2255944>
- Chaudhary, M. Y. (2020). Initial considerations for Islamic digital ethics. *Philosophy & Technology*, 33(4), 639–657. <https://doi.org/10.1007/s13347-020-00418-3>
- Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology*, 36(4).
<https://doi.org/10.1007/s13347-023-00668-x>
- Future of Life Institute. (2025, June 16). Asilomar AI principles.
<https://futureoflife.org/ai-principles/>
- Gaudiaut, T. (2026, February 17). AI content incidents skyrocket: A growing threat in the digital age. *Statista Daily Data*. <https://www.statista.com/chart/35846/ai-incidents-involving-content-generation/>
- Hate, C. for C. D. (2024). The double-edged sword of AI: How generative language models like Google Bard and ChatGPT pose a threat to countering hate and misinformation online. *Harvard Data Science Review*.
<https://doi.org/10.1162/99608f92.be4e28f0>
- Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, 4(4), 799–819.
<https://doi.org/10.1109/tai.2022.3194503>
- Islami, R. C., Hatmoko, J. U. D., Hermawan, F., Sulaiman, L., & Wulandari, M. (2025). Issues on implementation of artificial intelligence (AI) in Indonesia construction projects. *Planning Malaysia*, 23. <https://doi.org/10.21837/pm.v23i38.1781>
- Jareh, A. (2025). An in-depth assessment of the acceptability and effectiveness of AI tools in identifying misinformation in media: Literature review. *SAGE Open*, 15(4).
<https://doi.org/10.1177/21582440251381272>
- Kharvi, P. L. (2024). Understanding the impact of AI-generated deepfakes on public opinion, political discourse, and personal security in social media. *IEEE Security & Privacy*, 22(4), 115–123. <https://doi.org/10.1109/msec.2024.3405963>
- Kubaisi, A. A. S. H. A. (2024). Ethics of artificial intelligence a purposeful and foundational study in light of the Sunnah of Prophet Muhammad. *Religions*, 15(11), 1300. <https://doi.org/10.3390/rel15111300>
- Lemonne, E. (2021, March 5). Ethics guidelines for trustworthy AI. FUTURIUM European Commission.
<https://ec.europa.eu/futurium/en/aiallianceconsultation.1.html#:~:text=The%20Ethics%20Guidelines%20for%20Trustworthy,strategy%20announced%20earlier%20that%20year>
- Lundberg, E., & Mozelius, P. (2024). The potential effects of deepfakes on news media and entertainment. *AI & Society*. <https://doi.org/10.1007/s00146-024-02072-1>
- Metz, C., & Blumenthal, S. (2019, June 9). How A.I. could be weaponized to spread disinformation. *The New York Times*.

- <https://www.nytimes.com/interactive/2019/06/07/technology/ai-text-disinformation.html>
- Momeni, M. (2024). Artificial intelligence and political deepfakes: Shaping citizen perceptions through misinformation. *Journal of Creative Communications*. <https://doi.org/10.1177/09732586241277335>
- OECD. (2025). AI principles. Retrieved July 17, 2025, from <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>
- OECD. (2026). Trends in AI incidents and hazards reported by the media. *OECD Artificial Intelligence Papers*, No. 53. OECD Publishing. <https://doi.org/10.1787/4f5ff43c-en>
- OECD.AI. (2024, June 14). Egyptian authorities warn against AI-generated deepfake videos. <https://oecd.ai/en/incidents/2024-06-14-b642>
- Patel, M. (2025, June 30). The danger of relying on AI for fatwa writing: A caution for scholars and muftis. Whitethread. <https://www.whitethread.org/the-danger-of-relying-on-ai-for-fatwa-writing/>
- Pira, F. (2023). Disinformation a problem for democracy: Profiling and risks of consensus manipulation. *Frontiers in Sociology*, 8. <https://doi.org/10.3389/fsoc.2023.1150753>
- Popa, C., Pallath, R., Cunningham, L., Tahiri, H., Kesavarajah, A., & Wu, T. (2025). Deepfake technology unveiled: The commoditization of AI and its impact on digital trust. <https://doi.org/10.48550/ARXIV.2506.07363>
- Roche, C., Wall, P. J., & Lewis, D. (2022). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, 3(4), 1095–1115. <https://doi.org/10.1007/s43681-022-00218-9>
- Slotta, D. (2025, December 17). Topic: Misinformation on social media. Statista. <https://www.statista.com/topics/9713/misinformation-on-social-media/>
- Statista. (2026, January 8). U.S. adults worry about AI-generated deepfakes 2023–2025. <https://www.statista.com/statistics/1470834/us-adults-fear-of-artificial-intelligence-deepfakes/>
- Stoll, J. (2025, December 17). Themenseite: Digitale Desinformation. Statista. <https://www.statista.com/themen/12212/digitale-desinformation/>
- Suliman, M. (2024, November 7). The deepfake is a powerful weapon in the war in Sudan. African Arguments. <https://africanarguments.org/2024/10/the-deepfake-is-a-powerful-weapon-in-the-war-in-sudan/>
- Yakub, A. A., Ali, H. M., Achu, K., Jalil, R. A., & AO, S. (2021). An integrated approach based on artificial intelligence using ANFIS and ANN for multiple criteria real estate price prediction. *Planning Malaysia*, 19(17). <https://doi.org/10.21837/pm.v19i17.1005>
- Zandt, F. (2024, April 17). What are the biggest perceived dangers of AI? *Statista Daily Data*. <https://www.statista.com/chart/32112/most-problematic-ai-scenarios/>
- Zhang, J., Song, W., & Liu, Y. (2025). Cognitive bias in generative AI influences religious education. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-99121-6>

Received: 6th April 2026. Accepted: 20th April 2026