ILSIIS
Islamic Law and Social Issues in Society

Vol. 2, No. 2 (November, 2026), 153-174
P-ISSN: 3109-6344 | E-ISSN: 3110-0147
DOI: https://doi.org/10.64929/ilsiis.v2i2.69
https://ejournal.tuah.or.id/index.php/ilsiis/index

| Research Article |

Algorithmic Accountability through *Maqāṣid al-Sharīʿah*: Toward an Islamic Juridical Framework for Evaluating Artificial Intelligence Harms

Mohamed Cheikh Banane,*[1] Abderrazzak Sbihi,[2] Zinah Ghanim Younus,[3] Jaber Ali Ahmed Qitoon Al Shahri[4]

[1] Université Ibn Zohr, Morocco
[2] Mohammed V University, Rabat, Morocco
[3] College of Law, Ninevah University, Iraq
[4] Dhofar University, Oman

* Correspondence: isbanane1978@gmail.com

Received: April 27, 2026 | Accepted: October 28, 2026 | Published: November 23, 2026

Abstract: Algorithmic systems increasingly shape access to credit, healthcare, employment, and governance in Muslim-majority societies, yet their evaluation remains largely grounded in Euro-American ethical frameworks. This study addresses the absence of an operational Islamic juridical approach by mapping contemporary algorithmic harms onto the framework of *maqāṣid al sharīʿah*. Using a doctrinal analytical method rooted in *uṣūl al fiqh* and informed by recent AI ethics scholarship from 2018 to 2025, the study identifies five major categories of harm: bias, opacity, surveillance, labour displacement, and autonomous decision making. These harms are then mapped onto the five classical objectives and assessed across the three levels of *ḍarūriyyāt*, *ḥājiyyāt*, and *taḥsīniyyāt*. The findings reveal a structural mismatch between existing Islamic institutional guidance, which largely operates at the level of refinements and needs, and the reality that many algorithmic harms function at the level of necessities, posing risks to the protection of life, intellect, property, religion, and progeny. To address this gap, the study proposes a four-stage juridical review procedure consisting of classification, harm identification, *maqāṣid* mapping, and juridical determination. This framework contributes to Islamic legal theory and AI ethics by translating classical juristic reasoning into an operational and context-sensitive evaluative model. It offers a practical tool for *sharīʿah* advisory bodies, regulators, and ethicists to assess algorithmic systems while preserving the normative foundations of Islamic jurisprudence in contemporary technology governance.

Keywords: *Maqāṣid al-Sharīʿah*, Artificial Intelligence, Algorithmic Accountability, AI Ethics, Islamic Legal Theory.

Copyright © 2026 by Author(s)
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Introduction

Artificial intelligence has moved from the margins of computer science into the systems that support modern governance. Credit scoring, healthcare triage, predictive policing, content moderation, labor matching, and judicial case management now rely on algorithmic systems whose internal logic is opaque to the people affected by their decisions. The speed of the transition has outpaced most normative frameworks. A study of eighty-four public and private AI ethics guidelines found broad convergence around five principles, namely transparency, justice, non-maleficence, responsibility, and privacy, but substantial differences in how those principles are interpreted or enforced (Jobin et al., 2019). Later work has questioned whether principle-based ethics is adequate at all, given the absence of shared professional aims, validated ways to put principles into practice, and the accountability mechanisms that govern medicine and law (Mittelstadt, 2019). Researchers have identified a gap between principles and practice because high-level commitments fail to govern the granular decisions made by engineers, procurement officers, and regulators (Floridi et al., 2018; Mittelstadt et al., 2016).

The origins of these frameworks matter when they are applied in other settings. Much of the dominant discourse on AI ethics comes from institutions in Western Europe and North America, as well as a small number in East Asia. Its vocabulary, examples, and underlying assumptions about human nature reflect a particular philosophical heritage (Mohamed et al., 2020). Two problems follow for societies organized around other traditions. First, imported frameworks can miss harms specific to local contexts, such as the erosion of *waqf*-based (charitable endowment) welfare functions or the algorithmic distortion of religious authority. Second, indigenous normative resources, including *maqāṣid al-sharīʿah*, *qawāʿid fiqhiyyah* (legal maxims), and the ethical vocabulary of *akhlāq* (moral character) and *amānah* (trust), remain underused in the jurisdictions where their persuasive force is greatest (Chaudhary, 2020; Elmahjub, 2023; al-Najjār, 2008).

Muslim-majority states have begun to respond. The Organisation of Islamic Cooperation, through its Independent Permanent Human Rights Commission, adopted the Jeddah Declaration on the Guiding Principles of Artificial Intelligence in 2024, anchoring AI ethics in the Islamic concepts of free will, trusteeship, and the promotion of virtue (OIC-IPHRC, 2024). The Islamic World Educational, Scientific and Cultural Organization coordinated the Riyadh Charter on AI Ethics in the Islamic World, endorsed unanimously by fifty-three member states during the 45th Executive Council session (ICESCO, 2024). The OIC-15 Dialogue Platform issued the Tehran Declaration on Trustworthy and Ethical AI for Sustainable Development in 2025 (OIC-15 Dialogue Platform, 2025). At the national level, Dār al-Iftāʾ al-Miṣriyyah (the Egyptian Fatwa Authority) has issued rulings on the use of AI for fatwa generation and on the performance of religious rites through immersive digital environments (Dār al-Iftāʾ al-Miṣriyyah, 2024a, 2024b). These instruments establish broad

commitments. They articulate aspirations and prohibitions without specifying how classical Islamic juristic methods can be used to assess particular algorithmic harms and weigh competing interests.

Academic work on Islamic AI ethics has taken several approaches over the past five years. Raquib, Channa, Zubair, and Qadir (2022) proposed a virtue-based framework grounded in the five classical *maqāṣid*, arguing that contemporary technologies reshape cultures and redefine harms in ways that require a reinterpretation of the *ḍarūriyyāt* for the digital age. Mohadi and Tarshany (2023) examined how AI affects privacy and human dignity and creates opportunities for manipulation. They argued for incorporating ethical safeguards during design. Elmahjub (2023) advanced a pluralist ethical benchmarking approach rooted in *maṣlaḥah* (public welfare), and Ghaly (2024) traced the moral world of work in the Islamic tradition with specific attention to AI-mediated labor. Ali, Bouzoubaa, Gelli, Hamzi, and Khan (2025) adapted the *iʾtimānī* or trusteeship framework of Ṭāhā ʿAbd al-Raḥmān as an alternative ethical model, while Chaudhary (2020) sketched the philosophical groundwork for an Islamic approach to digital ethics. Regional scholarship in Southeast Asia has begun to engage AI through frameworks such as *wasaṭiyyah* or religious moderation (Kosasih et al., 2024).

A specific gap remains. Existing Islamic contributions either address general principles or focus on single applications such as fatwa issuance, autonomous vehicles, or marketing. None has yet produced a systematic juridical mapping that relates the harm categories identified in technical and regulatory scholarship to the five *maqāṣid* and evaluates them across the three classical thresholds of *ḍarūriyyāt*, *ḥājiyyāt*, and *taḥsīniyyāt*. Without such a grid, Islamic legal reasoning cannot perform the evaluative and prioritizing function that it performs in Islamic finance or bioethics, where classical juristic architecture is regularly translated into operational rules by *sharīʿah* boards and regulators.

This article responds to that gap. It aims to translate classical *maqāṣid* reasoning into a practical tool for evaluating algorithmic accountability and to demonstrate its use through three scenarios in Muslim-majority jurisdictions. The argument is organized in nine sections. Section 2 sets out the doctrinal-analytical method. Section 3 develops the conceptual foundation in *uṣūl al-fiqh* and contemporary *maqāṣid* theory. Section 4 conducts the core analysis by mapping five harm categories onto the five classical objectives. Section 5 compares three institutional instruments. Section 6 synthesizes the findings into a four-stage juridical review procedure. Section 7 discusses theoretical and practical implications, Section 8 identifies limitations, and Section 9 concludes. The contribution is a structured method for applying the *maqāṣid* that *sharīʿah* advisory bodies, regulators, and Muslim technology ethicists can use to assess specific algorithmic systems and trace their assessments to the underlying evidence.

Method

The study uses a qualitative doctrinal-analytical method grounded in the *uṣūl al-fiqh* tradition, supplemented by comparative textual analysis of Islamic institutional instruments and current AI ethics scholarship. Doctrinal research of this kind proceeds by identifying the authoritative sources of a legal system, reconstructing how jurists reason within that system, and applying their methods to a problem the classical sources did not directly address. The approach is well established in Islamic legal studies, where it has been used to evaluate biomedical technologies, innovations in Islamic finance, and environmental questions (Auda, 2008; Kamali, 2008; Elmahjub, 2021; al-Raysūnī, 1995).

The study draws on three groups of sources. The first group includes classical works of *uṣūl al-fiqh* and *maqāṣid* theory, including the writings of al-Ghazālī (1993), al-Shāṭibī (2004), and Ibn ʿĀshūr (2001), together with contemporary reconstructions by al-Yūbī (1998), al-Khādimī (2001), and al-Qaraḍāwī (2006). These sources supply the conceptual vocabulary of the five necessary objectives, the three thresholds of *ḍarūriyyāt*, *ḥājiyyāt*, and *taḥsīniyyāt*, and the juristic tools used to identify and weigh competing interests. The second group consists of contemporary Islamic ethical scholarship on AI published in peer-reviewed venues between 2018 and 2025, including works indexed in Scopus, Web of Science, and the Directory of Open Access Journals. Core contributions include Chaudhary (2020), Raquib et al. (2022), Mohadi and Tarshany (2023), Elmahjub (2023), Ghaly (2024), Ali et al. (2025), Benmamoun and Sobh (2024), and regional work such as Kosasih et al. (2024). The third group consists of institutional instruments adopted by transnational Islamic bodies and national fatwa authorities, namely the Jeddah Declaration (OIC-IPHRC, 2024), the Riyadh Charter (ICESCO, 2024), the Tehran Declaration (OIC-15 Dialogue Platform, 2025), and selected fatwas of Dār al-Iftāʾ al-Miṣriyyah (2024a, 2024b).

The analysis proceeded in three stages. The first was doctrinal reconstruction, which restated the classical objectives and thresholds in contemporary policy language while preserving their meaning. The second was harm mapping. Five categories of algorithmic harm identified by the AI ethics literature, namely bias, opacity, surveillance, labor displacement, and autonomous decision-making, were assessed to determine which of the five classical objectives each harm threatens and at which threshold (Mittelstadt et al., 2016; Jobin et al., 2019; O'Neil, 2016). The third stage was institutional comparison. Three OIC-affiliated declarations and the Dār al-Iftāʾ fatwas were analyzed for the level at which they address AI harms, using a content-analytic protocol informed by the coding procedure of Miles, Huberman, and Saldana.

The comparison has a defined scope. The article does not attempt a comprehensive country comparison of national AI regulations. It compares three institutional instruments operating at the transnational Islamic level, together with illustrative fatwas from one authoritative national body, to identify recurring patterns

in how these Islamic institutions emphasize particular jurisprudential concerns. This bounded scope is appropriate for a conceptual article whose primary contribution is theoretical rather than empirical. The study thus follows the tradition of conceptual Islamic legal scholarship represented by Auda (2008) and Kamali (2008), while adding contemporary institutional analysis. Validity rests on triangulation across the three groups of sources and on explicit identification of points of convergence and divergence among the institutional instruments.

Conceptual Framework

The framework rests on three interlocking ideas drawn from *uṣūl al-fiqh* and contemporary *maqāṣid* theory: the five classical objectives, the three thresholds, and the juristic distinction between means and ends. The analysis briefly explains each idea before combining all three in the harm-mapping matrix developed in Section 4.

The five classical objectives, formulated by al-Ghazālī in *al-Mustaṣfā* (al-Ghazālī, 1993) and consolidated by al-Shāṭibī in *al-Muwāfaqāt* (al-Shāṭibī, 2004), are the preservation of religion (*ḥifẓ al-dīn*), life (*ḥifẓ al-nafs*), intellect (*ḥifẓ al-ʿaql*), progeny (*ḥifẓ al-nasl*), and property (*ḥifẓ al-māl*). Together, they summarize the interests that Islamic rulings systematically protect. Ibn ʿĀshūr (2001), al-Yūbī (1998), and al-Raysūnī (1995) have argued that the five are best read as a system rather than a list, since the protection of each entails the protection of others. More recently, some scholars have proposed the addition of environmental preservation (*ḥifẓ al-bīʾah*) as a sixth objective, because environmental harms meet the *ḍarūrī* threshold under all five classical objectives at once (Raquib et al., 2022; al-Najjār, 2008). The present article retains the classical five for analytical clarity, while noting that the framework can accommodate future revisions to the objectives.

The three thresholds distinguish Islamic rulings by their urgency and protective force. *Ḍarūriyyāt* denotes necessities, interests whose destruction would threaten the existence or basic integrity of the protected objective. *Ḥājiyyāt* denotes needs, interests whose absence would cause hardship without destroying the objective. *Taḥsīniyyāt* denotes embellishments or refinements, interests whose fulfillment improves the quality of an objective without being indispensable. The thresholds form a hierarchy. Rulings at the *ḍarūrī* level take precedence over those at the *ḥājī* level, and rulings at the *ḥājī* level take precedence over those at the *taḥsīnī* level when the categories come into conflict (Kamali, 2008; al-Khādimī, 2001). Beyond classifying interests, the hierarchy helps jurists weigh competing interests when no single ruling can protect every objective.

The third idea is the juristic distinction between means and ends. Classical scholars held that the ruling that applies to an end also applies to the means of reaching it: if an end is obligatory, the means necessary to achieve it are obligatory, and if an end is prohibited, the means that lead to it are also prohibited. Two concepts

from *uṣūl al-fiqh* illustrate how this principle applies to AI governance: *sadd al-dharāʾiʿ* (blocking the means) and *fatḥ al-dharāʾiʿ* (opening the means) (Ibn Qayyim al-Jawziyyah, 1991; Ibn ʿAbd al-Salām, 1991). Algorithmic systems are instruments whose moral evaluation depends on the objectives they serve, the harms they may cause, and the alternatives available. Ali et al. (2025) have reformulated the principle within a trusteeship-based ethical framework, arguing that the flexibility of means is compatible with fixed ends, which allows new technologies to be adopted when they serve legitimate objectives and prohibited when they undermine them.

The analysis combines these ideas in a two-dimensional grid. The horizontal axis lists the five classical objectives, and the vertical axis lists the three thresholds. Each cell of the resulting fifteen-cell matrix asks a juristic question: Does the harm under consideration threaten this objective at this threshold? The grid allows harms to be classified, compared, and prioritized in a manner consistent with classical juristic reasoning but applicable to contemporary algorithmic systems. It also exposes the regulatory mismatch identified in the analysis below, in which existing institutional instruments address AI at the level of refinement even when the harms threaten necessities.

The framework draws on ideas shared by other recent proposals and gives them a distinct structure for evaluating AI harms. Raquib et al. (2022) developed a virtue-based approach that derives essentials from Islamic tradition without building an explicit grid organized by threshold. Elmahjub's (2023) pluralist approach to ethical benchmarking is grounded in *maṣlaḥah* and operates at a higher level of abstraction. Mohadi and Tarshany (2023) identified categories of AI harm against the five objectives without stratifying them across the three thresholds. Ghaly (2024) focused on the specific domain of AI-mediated work. Ali et al. (2025) advanced the trusteeship alternative. The present framework complements these contributions by producing an explicit mapping by threshold that can be used to compare specific algorithmic systems.

Mapping Algorithmic Harms onto the *Maqāṣid* Grid

This section applies the evaluative grid to five categories of algorithmic harm that are well established in the technical and regulatory literature. For each harm, the analysis identifies the objectives most directly threatened and the threshold the threat reaches. The overall mapping is summarized in Table 1, and the synthesis that follows informs the comparative analysis in Section 5.

1. Algorithmic Bias and the Protection of Progeny and Property

Algorithmic bias refers to systematic error in the outputs of machine-learning systems that disadvantages particular demographic groups. The Gender Shades study documented that commercial facial analysis systems misclassified darker-skinned women at rates of up to 34.7 percent, compared with rates as low as 0.8 percent for

lighter-skinned men (Buolamwini & Gebru, 2018). Comparable patterns have been observed in credit scoring, hiring, healthcare triage, and predictive policing. The consistent finding is that biased systems reproduce and amplify existing social inequalities rather than correcting them (O'Neil, 2016; Mittelstadt et al., 2016).

On the *maqāṣid* grid, algorithmic bias threatens the preservation of progeny (*ḥifẓ al-nasl*) by perpetuating disadvantage across generations through overlapping barriers to credit, education, and housing, and threatens the preservation of property (*ḥifẓ al-māl*) by distorting access to economic opportunities. Where biased triage systems delay or deny care, bias also threatens the preservation of life (*ḥifẓ al-nafs*). The threshold is *ḍarūrī* rather than *ḥājī*, because systematic exclusion from credit, employment, and healthcare destroys rather than merely inconveniences the protected interests. The juristic implication is direct: algorithmic bias at this threshold threatens necessities, so rulings that mitigate it take precedence over rulings aimed at efficiency or competitive advantage (Ibn ʿAbd al-Salām, 1991). This aligns with the position articulated in the Tehran Declaration, which commits signatories to trustworthy AI for sustainable development and to addressing the inequities generated by technological systems (OIC-15 Dialogue Platform, 2025).

2. Opacity and the Protection of Intellect and Religion

Opacity refers to the inability of affected parties, and often of developers themselves, to explain why a model produced a particular output. In contemporary deep-learning systems, opacity is structural because the representations learned by such systems are high-dimensional and not interpretable without post-hoc methods. The ethical literature has long identified opacity as a distinct harm, separate from the question of correctness, because it prevents people from contesting decisions, holding decision-makers accountable, and giving informed consent (Mittelstadt et al., 2016; Jobin et al., 2019; Floridi et al., 2018).

On the *maqāṣid* grid, opacity threatens the preservation of intellect (*ḥifẓ al-ʿaql*) by depriving affected persons of the epistemic conditions required for rational assent. It also threatens the preservation of religion (*ḥifẓ al-dīn*) when opaque systems generate religious guidance or mediate acts of worship. Dār al-Iftāʾ al-Miṣriyyah has addressed this religious concern by prohibiting reliance on AI to issue fatwas because such systems lack the juristic qualifications and contextual awareness required for authoritative religious guidance (Dār al-Iftāʾ al-Miṣriyyah, 2024a). The opacity of automated religious outputs threatens the preservation of religion at the *ḍarūrī* level when automated systems replace the authority of qualified *muftīs* and at the *ḥājī* level when the systems assist qualified *muftīs* without replacing them. This stratified reading is compatible with the broader argument advanced by Taddeo and Floridi (2018) that AI can function as a force for good only under conditions that keep humans meaningfully in control.

3. Surveillance and the Protection of Honor and Progeny

Pervasive algorithmic surveillance, including biometric identification, behavioral tracking, and predictive profiling, threatens several objectives at once. The Islamic tradition has long protected the private sphere through the concept of *ḥifẓ al-ʿirḍ*, sometimes grouped with *ḥifẓ al-nasl* and sometimes treated as a quasi-independent objective. Classical jurists prohibited *tajassus*, or spying, and contemporary scholars have extended this prohibition to digital surveillance (Chaudhary, 2020; Elmahjub, 2023; al-Qaraḍāwī, 2006).

Mass surveillance threatens the preservation of progeny and honor by creating lasting records of people's private lives that can be used to coerce, exclude, or shame individuals and their descendants, and threatens the preservation of intellect by chilling free deliberation and religious inquiry. The threshold depends on the intensity and integration of surveillance. Isolated, targeted surveillance with judicial authorization and for a narrowly defined purpose may operate at the *ḥājī* level, whereas systematic, untargeted population-scale surveillance operates at the *ḍarūrī* level because it destroys the conditions of private moral agency that the objectives protect. The Riyadh Charter addresses this dimension in general terms by calling for AI systems that respect human dignity and privacy (ICESCO, 2024), but the operational threshold still requires case-specific juristic review.

4. Labor Displacement and the Protection of Property and Life

AI-driven labor automation threatens the preservation of property by disrupting livelihoods. It also threatens the preservation of life less directly by putting access to health insurance, housing security, and mental health support at risk, since these often depend on employment (Ghaly, 2024). Classical Islamic ethics treats legitimate work as both permissible and morally meritorious, and the tradition has historically developed mechanisms such as *waqf*, *zakāt*, and communal obligations to mitigate economic disruption. The juristic question raised by labor automation is not whether new technologies are permitted, but whether institutions protect displaced workers' interests when these technologies are deployed.

The threshold for harm from labor displacement depends on the scale and pace of automation. Gradual automation that preserves transition pathways for workers operates at the *ḥājī* or *taḥsīnī* level. Rapid displacement across occupational sectors without compensation operates at the *ḍarūrī* level. The juristic implication is that the permissibility of AI-driven automation depends on whether *zakāt*-compatible or equivalent redistributive mechanisms are in place. This conditional permissibility reflects the classical juristic tool of *al-iḥtijāj bi-al-maṣlaḥah*, by which a ruling is calibrated to the net public benefit under specified conditions (Kamali, 2008; al-Yūbī, 1998).

5. Autonomous Decision-Making and the Preservation of Religion, Life, and Property

Autonomous decision-making refers to the delegation of consequential decisions to algorithmic systems without meaningful human review. Autonomy differs from opacity because an autonomous system may be fully interpretable and still produce morally unacceptable outputs. It also differs from bias because an autonomous system may be statistically fair and still displace human moral agency. Autonomous decision-making threatens the preservation of religion where it substitutes for juristic reasoning, the preservation of life in clinical and military contexts, and the preservation of property in financial and insurance decisions (Elmahjub & Qadir, 2023; Raquib et al., 2022).

The threshold is context-dependent but generally *ḍarūrī*. Dār al-Iftāʾ al-Miṣriyyah has held that fully autonomous AI cannot perform the functions of *iftāʾ* (fatwa issuance) (Dār al-Iftāʾ al-Miṣriyyah, 2024a). It has addressed the performance of ritual obligations through immersive digital substitutes in similarly restrictive terms (Dār al-Iftāʾ al-Miṣriyyah, 2024b). In Islamic finance and healthcare, the juristic analogy suggests that autonomous systems may be deployed as decision-support tools but not as final decision-makers without qualified human oversight. The Tehran Declaration explicitly endorses a human-centered approach that preserves meaningful human control over consequential AI outputs (OIC-15 Dialogue Platform, 2025), and Latifi (2024) has raised parallel concerns in the *shīʿī ijtihād* tradition.

Taken together, the analysis of the five categories reveals a recurring juridical pattern. Under conditions of scale and integration, the harms most frequently generated by contemporary AI systems operate at the *ḍarūrī* level of the *maqāṣid* grid. The threshold varies by harm and context, so each case requires careful analysis. The recurring pattern provides a basis for the comparative analysis of institutional instruments that follows.

Table 1. Mapping of five algorithmic harms onto the five *maqāṣid al-sharīʿah*.

Algorithmic harm	**The five *maqāṣid***				
	Ḥifẓ al-dīn (Religion)	*Ḥifẓ al-nafs* (Life)	*Ḥifẓ al-ʿaql* (Intellect)	*Ḥifẓ al-nasl* (Progeny)	*Ḥifẓ al-māl* (Property)
Bias	*	**	**	***	***
Opacity	***	**	***	*	**
Surveillance	**	**	*	***	*
Labor displacement	*	***	**	**	***
Autonomous decision-making	***	***	***	**	**

Table 1. Cell entries indicate the threshold at which each harm engages each objective: *** = *ḍarūrī* (necessity); ** = *ḥājī* (complementary need); * = *taḥsīnī* (refinement). The table summarizes the analysis in Sections 4.1–4.5.

Comparative Analysis of Institutional Instruments

The analysis in Section 4 identifies a largely consistent pattern across the categories of harm on the *maqāṣid* grid. The most consequential harms operate at the *ḍarūrī* level across multiple objectives, not merely at the *taḥsīnī* level within a single objective. This section examines three institutional instruments adopted by Islamic institutions to determine which thresholds they address and where their approaches converge or diverge. Table 2 summarizes the comparison.

The Jeddah Declaration on the Guiding Principles of Artificial Intelligence, adopted by the OIC Independent Permanent Human Rights Commission in 2024, anchors AI ethics in the Islamic concepts of free will and trusteeship, and calls on member states to develop a comprehensive strategy for AI ethics that aligns Islamic values with global ethical standards (OIC-IPHRC, 2024). The declaration addresses human dignity, fairness, privacy, and the protection of family values, and highlights the role of the International Islamic Fiqh Academy and related research institutions. It addresses harms primarily at the *ḥājī* and *taḥsīnī* levels. The declaration sets out shared principles but gives no criteria for determining whether a particular AI system causes harm at the *ḍarūrī* level.

The Riyadh Charter on Artificial Intelligence in the Islamic World, adopted by 53 ICESCO member states during the 45th Executive Council session, was described by its drafters as a comprehensive strategic framework for AI development in the Islamic world (ICESCO, 2024). The Charter identifies ethical principles that align with Islamic values and with international AI governance, and proposes a guiding framework that member states can use to keep pace with technological developments. Like the Jeddah Declaration, the Riyadh Charter addresses harms primarily at the *ḥājī* and *taḥsīnī* levels. Its strength lies in the breadth of its endorsement and its explicit alignment with the educational and cultural mandate of ICESCO, but it does not provide a grid organized by threshold to support case-specific juristic review.

The Tehran Declaration on Trustworthy and Ethical AI for Sustainable Development, adopted at the Second OIC-15 Ministerial Meeting in 2025, commits ten technologically advanced OIC member states, including Indonesia, Malaysia, Pakistan, Saudi Arabia, Türkiye, and Qatar, to cooperation on AI education, research, ethics, and governance (OIC-15 Dialogue Platform, 2025). The declaration is notable for its explicit link between AI cooperation and sustainable development goals and for its endorsement of human-centered AI. Like the other declarations, it sets out general principles without classifying harms across the three classical thresholds. Its

focus on sustainable development provides a basis for further juristic analysis of *ḥifẓ al-bīʾah* (al-Najjār, 2008).

The fatwas of Dār al-Iftāʾ al-Miṣriyyah provide more specific guidance. Fatwa No. 22255 addresses the use of AI applications to obtain legal rulings and concludes that reliance on AI for fatwa generation is impermissible, in part because AI cannot satisfy the epistemic, juristic, and contextual conditions of authoritative religious guidance (Dār al-Iftāʾ al-Miṣriyyah, 2024a). The fatwa on the performance of acts of worship through immersive technology prohibits the substitution of AI-mediated virtual experiences for the physical performance of *ḥajj* and *ʿumrah*, while permitting the use of such technologies for educational purposes (Dār al-Iftāʾ al-Miṣriyyah, 2024b). These rulings engage AI harms at the *ḍarūrī* level, because they concern the integrity of religion and acts of worship, and offer practical guidance for Muslim practitioners.

The comparison reveals three patterns. First, the transnational declarations state broader principles than the national fatwas, consistent with their distinct institutional functions but producing the regulatory mismatch identified in Section 4. Second, all four instruments converge on a principle of human-centered AI, although they describe the principle in varying degrees of detail. Third, the Dār al-Iftāʾ fatwas show that jurists can address harms at the *ḍarūrī* level when the issue is clearly defined. This supports the practical applicability of the proposed *maqāṣid* grid. Southeast Asian scholarship on AI governance, including the *wasaṭiyyah*-based approach proposed by Kosasih et al. (2024), complements these findings by showing how regional jurisprudence can mediate between transnational principles and local regulatory practice.

The comparative evidence supports the core claim of this article. Existing Islamic institutional instruments tend to articulate AI governance at the levels of refinement (*taḥsīnī*) and need (*ḥājī*), while the harms they address often operate at the *ḍarūrī* level. The mismatch does not invalidate the instruments, but it indicates that they need to be supplemented by an evaluative grid organized by threshold if they are to perform the protective function for which the *maqāṣid* framework was originally developed (Ibn ʿĀshūr, 2001; al-Raysūnī, 1995). Section 6 proposes a procedure that links the broad principles of declarations with the case-specific guidance of fatwas.

Table 2. Comparative Analysis of Six Institutional Instruments Engaging AI Governance.

Instrument	Scope	Harms addressed	Threshold engaged	Key gap
Jeddah Declaration (OIC-IPHRC, 2024)	OIC member states	General principles; virtue and trusteeship	*Ḥājī* – *taḥsīnī*	No harm-specific guidance
Riyadh Charter (ICESCO, 2024)	53 ICESCO states	AI ethics principles; cultural heritage	*Ḥājī*	No enforcement mechanism
Tehran Declaration (OIC-15, 2025)	OIC-15 platform	Trustworthy AI; sustainable development	*Ḥājī*; partial ḍarūrī	Broad aspirational scope
Fatwa No. 22255 (Dār al-Iftāʾ, 2024a)	Egypt; iftāʾ domain	AI-generated religious guidance	*Ḍarūrī*	Single-domain focus
Acts of Worship (Dār al-Iftāʾ, 2024b)	Egypt; ritual practice	Immersive digital substitutes	*Ḍarūrī*	Limited to ritual domain
UNESCO Recommendation (2021)	UN member states	Comprehensive AI harms	N/A (non-Islamic frame)	Not grounded in Islamic jurisprudence

Note: The table classifies each instrument by its scope, the harms it addresses, the *maqāṣid* threshold at which it operates, and its main limitation. The UNESCO Recommendation is included as a non-Islamic benchmark for comparison.

Integrative Framework: A Four-Stage Juridical Review Procedure

The findings of the comparative analysis support a four-stage juridical review procedure that integrates the *maqāṣid* grid into a repeatable assessment process (Figure 1). The procedure is designed for use by *sharīʿah* advisory bodies, national fatwa authorities, regulators, and Muslim technology ethicists, and is compatible with both transnational declarations and national regulatory instruments.

The first stage is classification. The system under review is described in neutral technical terms, including its training data, model architecture, deployment context, affected populations, and known error patterns. Technical experts, regulators, and *sharīʿah* scholars prepare the description together, following the collaborative model used by *sharīʿah* advisory boards in Islamic finance. The output is a standardized

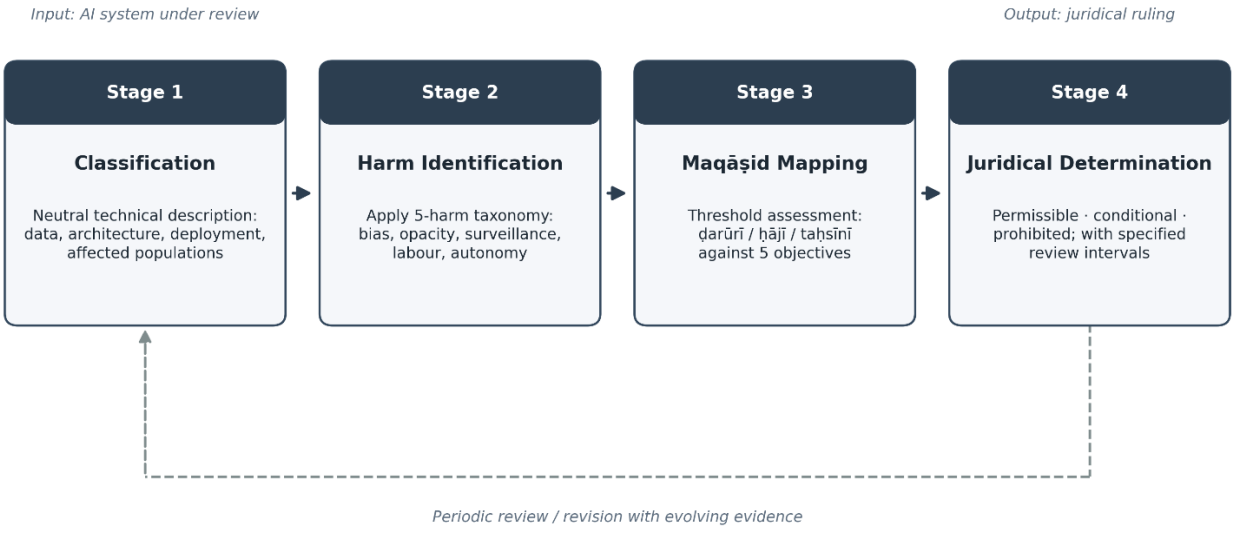
description that scholars without specialized AI expertise can assess and trace to the original technical documentation.

The second stage is harm identification. Using the five categories defined in Section 4, scholars identify which harms the system causes or could cause and assess their severity. The taxonomy can accommodate additional categories. New harm categories, such as environmental footprint, fabrication, or distortion of religious authority, can be added as the technology evolves. The output is a harm profile that records the presence, absence, or borderline status of each category for the system under review.

The third stage is *maqāṣid* mapping. Each identified harm is assessed against the five classical objectives using the grid developed in Section 3, and a threshold is assigned: *ḍarūrī*, *ḥājī*, or *taḥsīnī*. Scholars justify the assigned threshold by citing classical sources, contemporary scholarship, and institutional instruments. They can revise it as new information becomes available. The output is a threshold-indexed evaluative matrix specific to the system under review.

The fourth stage is juridical determination. Drawing on the threshold-indexed matrix, scholars determine whether the system is permissible, conditionally permissible, or prohibited. Conditional permissibility is the most common outcome, because most algorithmic systems produce a mixture of benefits and harms that requires mitigation rather than outright prohibition. Scholars state the conditions explicitly, citing the thresholds the harms reach and specifying review intervals. The output is a juridical determination that can be communicated to regulators, operators, and affected populations (al-Khādimī, 2001; al-Qaraḍāwī, 2006).

Figure 1. The Four-Stage Juridical Review Procedure.



Note: The procedure moves from technical classification (Stage 1) through harm identification (Stage 2) and *maqāṣid* threshold mapping (Stage 3) to juridical

determination (Stage 4). The dashed feedback loop indicates periodic review as new evidence about the system becomes available.

The integrative framework shares structural similarities with the Ethical Impact Assessment procedure proposed by UNESCO (2021), but differs in three respects. First, it is grounded in classical Islamic juristic categories rather than in universal ethical principles alone, which allows it to operate with the authority that Islamic legal reasoning carries in Muslim-majority jurisdictions. Second, it is threshold-stratified rather than flat, which allows it to prioritize harms that threaten the existence of protected interests over harms that cause hardship. Third, it is designed to produce juridical determinations rather than ethical recommendations, which aligns with the institutional roles of *sharīʿah* boards and fatwa authorities. The framework is complementary to existing secular instruments such as those developed by Floridi et al. (2018), Mittelstadt (2019), and Taddeo and Floridi (2018), and can be used alongside them in pluralist regulatory settings.

Discussion

The proposed framework has implications in three areas: Islamic legal theory, AI ethics scholarship, and comparative technology governance. For Islamic legal theory, the framework demonstrates that classical *maqāṣid* reasoning can be applied in areas classical jurists did not anticipate, provided that the links between classical concepts and contemporary terms are clear and carefully justified. The framework does not claim to replace the juristic deliberation of qualified scholars, but to make it transparent, reproducible, and easier to communicate across linguistic and institutional boundaries (Auda, 2008; Kamali, 2008; Elmahjub, 2021; al-Raysūnī, 1995).

For AI ethics scholarship, the framework offers an example of how a non-Western normative tradition can translate its classical principles into practical rules for algorithmic accountability. Recent scholarship has called for pluralist approaches to AI ethics that incorporate cultural and religious traditions beyond the Euro-American mainstream (Mohamed et al., 2020; Elmahjub, 2023). The present framework answers that call by showing how institutions can translate those principles into procedures and thresholds. It also contributes to the debate on the principle-practice gap identified by Mittelstadt (2019) and Jobin et al. (2019) by proposing a procedure organized by threshold that can be incorporated into *sharīʿah* advisory review.

For comparative technology governance, the framework illustrates how religious-legal architecture can contribute to the regulation of algorithmic systems in plural societies. Muslim-majority jurisdictions that combine state law with classical juristic review, such as Malaysia, Saudi Arabia, the United Arab Emirates, and Indonesia, already have the institutional infrastructure to adopt the framework without radical reform. The framework is also relevant to Muslim-minority contexts

where *sharīʿah* advisory functions operate within private Islamic finance, halal certification, and community-level ethical review. Regional scholarship such as Kosasih et al. (2024), together with institutional literature from Malaysian and Indonesian fatwa authorities, shows that such integration is already underway, though on a case-by-case basis.

The framework has several practical applications. *Sharīʿah* advisory bodies in Islamic fintech, which evaluate products such as algorithmic credit scoring, robo-advisory services, and *sharīʿah*-compliant neobanks, can use the threshold grid to prioritize harms and to structure conditions for permissibility. Halal certification bodies, which are increasingly asked to certify AI-enabled products, can use the framework to distinguish refinements from necessities. National fatwa authorities, whether in Egypt, Indonesia, Malaysia, or the Gulf, can use the framework to produce juridical determinations that are structurally comparable across jurisdictions, which would facilitate transnational coordination without requiring uniformity. Secular regulators in Muslim-majority states can use the framework to anticipate how religious advisory bodies are likely to evaluate proposed regulations, and to structure consultation accordingly.

The framework does not claim universality. Its premises are grounded in Sunni *uṣūl al-fiqh* and *maqāṣid* theory as developed by al-Ghazālī, al-Shāṭibī, Ibn ʿĀshūr, and their contemporary interpreters (al-Ghazālī, 1993; al-Shāṭibī, 2004; Ibn ʿĀshūr, 2001; al-Yūbī, 1998). *Shīʿī* juristic methodologies share many features with Sunni *uṣūl* but differ on specific questions of authority, *ijtihād*, and the sources of law, and a parallel framework grounded in *shīʿī* jurisprudence would need to be developed separately (Latifi, 2024). *Maqāṣid*-based reasoning is also not the only Islamic ethical framework for AI. Alternative proposals, including the trusteeship approach of Ali et al. (2025), the virtue-based approach of Raquib et al. (2022), and the *tawḥīd*-based (divine unity) proposals explored in recent Islamic business ethics literature (Benmamoun & Sobh, 2024), can complement one another within Islamic intellectual traditions.

The framework can be refined as Islamic legal scholarship and AI technologies develop. The harm taxonomy developed in Section 4 can be expanded as AI technologies give rise to new forms of harm, while the proposed threshold assignments remain subject to reassessment by qualified scholars. This capacity for refinement reflects the classical principle of *taghayyur al-aḥkām bi-taghayyur al-zamān wa-al-makān*, or the variation of rulings with time and place, through which Islamic legal reasoning accommodates changing circumstances while preserving its foundations (Ibn Qayyim al-Jawziyyah, 1991).

Four limitations warrant further research. First, the study's conceptual claims about regulatory mismatch require empirical validation through systematic analysis of institutional instruments and case decisions, including comparative application of the grid to national AI regulations and *sharīʿah* advisory opinions. Second, expanding

the five-harm taxonomy requires addressing its predominantly Euro-American foundations, which risk reproducing the biases it seeks to challenge. Future research should therefore recognize harms arising in Muslim-majority contexts, including the algorithmic distortion of religious authority, automated halal certification, and pilgrimage surveillance, as distinct categories. Third, the framework's grounding in Sunni jurisprudence calls for fuller engagement with *shīʿī*, *ibāḍī*, and other legal traditions, examining how procedural differences affect institutional adoption and whether threshold assignments converge across traditions (Latifi, 2024). Finally, its translation into regulatory practice requires standardized documentation, review intervals, and appeal procedures, developed through collaboration among fatwa authorities, Islamic finance regulators, and technology companies. Existing *sharīʿah* advisory bodies could draw on their experience with technical frameworks in Islamic finance.

Conclusion

The article develops a juridical framework based on *maqāṣid al-sharīʿah* for evaluating the harms of algorithmic systems. It argues that existing instruments adopted by Islamic institutions, including the Jeddah Declaration, the Riyadh Charter, the Tehran Declaration, and the fatwas of Dār al-Iftāʾ al-Miṣriyyah, tend to operate at the levels of refinement and need, while the harms they address often operate at the level of necessity. The resulting regulatory mismatch undermines the protective function of *sharīʿah*-based review and leaves Muslim-majority jurisdictions without an operational grid that connects specific algorithmic harms to the classical threshold structure. The article proposes such a grid by mapping five harm categories onto the five classical objectives and stratifying them across the three thresholds, and synthesizes the results into a four-stage juridical review procedure.

The main finding is that bias, opacity, surveillance, labor displacement, and autonomous decision-making each threaten multiple classical objectives at the *ḍarūrī* level under contemporary conditions of deployment, and that juridical review of these harms must therefore address them at the level of necessity. The framework translates *maqāṣid* reasoning into an explicit procedure for assessing these harms by threshold and complements existing virtue-based, trusteeship, and pluralist proposals in the Islamic AI ethics literature without displacing them.

The article makes three contributions. It presents the first structured mapping of AI harms onto the full three-tier *maqāṣid* grid. It then synthesizes the findings into a four-stage juridical review procedure that can be used immediately by *sharīʿah* advisory bodies and national fatwa authorities. Its comparison of institutional instruments also shows how classical Islamic legal reasoning can be applied to contemporary technology governance. Further research should extend the framework through empirical case analysis, incorporate harm categories specific to Muslim-

majority jurisdictions, articulate parallel procedures within non-Sunni traditions, and collaborate with regulatory bodies to translate the framework into standardized documentation. The article argues that classical Islamic jurisprudence can help contemporary societies govern algorithmic systems in ways that protect fundamental interests and guide institutional change.

Declarations

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

Conflict of interest. The authors declare that they have no competing interests.

Ethical Statement. The study was conducted in accordance with the principles of voluntary participation, confidentiality, anonymity, and respect for participants' rights. The researchers obtained informed consent before collecting data. The authors declare that the study complies with applicable ethical research principles.

AI Usage Declaration. The authors used Claude AI to assist with initial reference searches and the formulation of theoretical concepts. The authors reviewed, verified, and thoroughly edited all generated output. The authors bear full responsibility for all published content.

References

Ab Rahim, S. F., Ab Rahman, M. F., Abdullah Thaidi, H. A., Nik Mohd Azimi, N. N. M. A., & Jailani, M. R. (2025). Artificial intelligence for fatwa issuance: Guidelines and ethical considerations. *Journal of Fatwa Management and Research*, 30(1), 76–100. https://doi.org/10.33102/jfatwa.vol30no1.654

Abdelnour, M. G. (2022). *The higher objectives of Islamic theology: Toward a theory of maqāṣid al-ʿaqīdah*. Oxford University Press. https://doi.org/10.1093/oso/9780197648636.001.0001

Al-Ghazālī, Abū Ḥāmid Muḥammad ibn Muḥammad. (1993). *Al-Mustaṣfā min ʿilm al-uṣūl* (M. ʿA. al-S. ʿA. al-Shāfī, Ed.). Dār al-Kutub al-ʿIlmiyyah.

Al-Khādimī, N. M. ibn Mukhtār. (2001). *ʿIlm al-maqāṣid al-sharʿiyyah*. Maktabat al-ʿUbaykān.

Al-Najjār, ʿA. al-M. (2008). *Maqāṣid al-sharīʿah bi-abʿād jadīdah*. Dār al-Gharb al-Islāmī.

Al-Qaraḍāwī, Y. (2006). *Dirāsah fī fiqh maqāṣid al-sharīʿah bayna al-maqāṣid al-kulliyyah wa-al-nuṣūṣ al-juzʾiyyah*. Dār al-Shurūq.

Al-Raysūnī, A. (1995). *Naẓariyyat al-maqāṣid ʿinda al-Imām al-Shāṭibī*. al-Maʿhad al-ʿĀlamī lil-Fikr al-Islāmī.

Al-Shāṭibī, Abū Isḥāq Ibrāhīm ibn Mūsā. (2004). *Al-Muwāfaqāt fī uṣūl al-sharīʿah* (ʿA. Darrāz, Ed., Vols. 1–4). Dār al-Kutub al-ʿIlmiyyah.

Al-Yūbī, M. S. ibn Aḥmad. (1998). *Maqāṣid al-sharīʿah al-islāmiyyah wa-ʿalāqatuhā bi-al-adillah al-sharʿiyyah*. Dār al-Hijrah.

Ali, F., Bouzoubaa, K., Gelli, F., Hamzi, B., & Khan, S. (2025). Islamic ethics and AI: An evaluation of existing approaches to AI using trusteeship ethics. *Philosophy & Technology*, 38(3), 120. https://doi.org/10.1007/s13347-025-00922-4

Alkhiri, T. A. A. (2022). Islamic ethical foundations of AI and its modern applications. *International Journal of Computer Science and Network Security*, 22(5), 741–746. https://doi.org/10.22937/IJCSNS.2022.22.5.101

Auda, J. (2008). *Maqāṣid al-sharīʿah as philosophy of Islamic law: A systems approach*. International Institute of Islamic Thought. https://doi.org/10.2307/j.ctvkc67tg

Benmamoun, M., & Sobh, R. (2024). An Islamic perspective on AI ethics. *Journal of Business Ethics Education*, 21, 231–244. https://doi.org/10.5840/jbee20242114

Bunt, G. R. (2018). *Hashtag Islam: How cyber-Islamic environments are transforming religious authority*. University of North Carolina Press. https://doi.org/10.5149/northcarolina/9781469643168.001.0001

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

Cath, C. (2018). Governing artificial intelligence: Ethical, legal, and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A*, 376(2133), 20180080. https://doi.org/10.1098/rsta.2018.0080

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

Dār al-Iftāʾ al-Miṣriyyah. (2024a). *Fatwa No. 22255: Using AI applications to obtain fatwas*. https://www.dar-alifta.org/en/fatwa/details/22255/using-ai-applications-to-obtain-fatwas

Dār al-Iftāʾ al-Miṣriyyah. (2024b). *Acts of worship in the digital age*. https://www.dar-alifta.org/en/fatwa/details/20709/acts-of-worship-in-the-digital-age

Elmahjub, E. (2021). Islamic jurisprudence as an ethical discourse: An enquiry into the nature of moral reasoning in Islamic legal theory. *Oxford Journal of Law and Religion*, 10(1), 16–42. https://doi.org/10.1093/ojlr/rwab003

Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology*, 36(4), 73. https://doi.org/10.1007/s13347-023-00668-x

Elmahjub, E., & Qadir, J. (2023). How to program autonomous vehicle (AV) crash algorithms: An Islamic ethical perspective. *Journal of Information, Communication and Ethics in Society*, 21(4), 452–467. https://doi.org/10.1108/JICES-02-2023-0015

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Gazali, H. M., Jumadi, J., Ramlan, N. R., Abd Rahmat, N., Uzair, S. N. H. M., & Mohid, A. N. (2020). Application of artificial intelligence (AI) in Islamic investments. *Journal of Islamic Finance*, 9(2), 70–78. https://doi.org/10.31436/jif.v9i2.485

Ghaly, M. (Ed.). (2016). *Islamic perspectives on the principles of biomedical ethics*. World Scientific.

Ghaly, M. (2024). What makes work "good" in the age of artificial intelligence (AI)? Islamic perspectives on AI-mediated work ethics. *The Journal of Ethics*, 28, 429–453. https://doi.org/10.1007/s10892-023-09456-3

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hamsin, M. K., Anwar, S., Hidayati, M., & Hasanudin, H. (2024). Unveiling ethical implications: AI robot accountability in Islamic context. *Jurnal Media Hukum*, 30(1), 117–135. https://doi.org/10.18196/jmh.v30i2.18524

Ibn ʿAbd al-Salām, ʿIzz al-Dīn. (1991). *Qawāʿid al-aḥkām fī maṣāliḥ al-anām* (Vols. 1–2). Maktabat al-Kulliyyāt al-Azhariyyah.

Ibn ʿĀshūr, M. al-Ṭ. (2001). *Maqāṣid al-sharīʿah al-islāmiyyah* (M. al-Ṭ. al-Maysāwī, Ed.). Dār al-Nafāʾis.

Ibn Qayyim al-Jawziyyah, M. ibn Abī Bakr. (1991). *Iʿlām al-muwaqqiʿīn ʿan rabb al-ʿālamīn* (M. ʿA. al-S. Ibrāhīm, Ed., Vols. 1–4). Dār al-Kutub al-ʿIlmiyyah.

ICESCO. (2024). *Riyadh Charter on artificial intelligence ethics in the Islamic world*. Islamic World Educational, Scientific and Cultural Organization.

Ismail, A. M., & Baseeruddin, A. S. (2022). Moderation in fatwā and ijtihād. *AHKAM: Jurnal Ilmu Syariah*, 22(1). https://doi.org/10.15408/ajis.v22i1.22299

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kamali, M. H. (2008). *Maqāṣid al-sharīʿah made simple* (Occasional Papers Series No. 13). International Institute of Islamic Thought.

Kosasih, E., Islamy, M. R. F., & Wiwaha, R. S. (2024). Artificial intelligence in the era of Society 5.0: Compromising technological innovation through the wasaṭiyyah

approach within the framework of Islamic law. *Al-Istinbath: Jurnal Hukum Islam*, 9(2), 519–540. https://doi.org/10.29240/jhi.v9i2.9596

Latifi, H. (2024). Challenges of using artificial intelligence in the process of shīʿī ijtihād. *Religions*, 15(5), 541. https://doi.org/10.3390/rel15050541

Mahmood, S. F., Fikry, A., Hamzah, M. I., Khalid, M. M., Rosyad, A. M., & Bhari, A. (2023). Fiqh robotic for artificial intelligence in humanoids used for therapy, services and other social activities: An integration of AI and maqāṣid sharīʿah. *Journal of Fatwā Management and Research*, 28(2), 527. https://doi.org/10.33102/jfatwa.vol28no2.527

Malik, S. A. (2024). Artificial intelligence and Islamic thought: Two distinctive challenges. *Journal of Islamic and Muslim Studies*, 8(2), 108–115. https://doi.org/10.2979/jims.00020

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). https://doi.org/10.1177/2053951716679679

Mohadi, M. (2023). Interview with Jasser Auda. *Journal of Contemporary Maqasid Studies*, 2(1), 159–174.

Mohadi, M., & Tarshany, Y. (2023). Maqāṣid al-sharīʿah and the ethics of artificial intelligence: Contemporary challenges. *Journal of Contemporary Maqasid Studies*, 2(2), 79–102. https://doi.org/10.52100/jcms.v2i2.107

Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Munshi, A. A., AlSabban, W. H., Farag, A. T., Rakha, O. E., Al Sallab, A. A., & Alotaibi, M. (2022). Automated Islamic jurisprudence legal opinions generation using artificial intelligence. *Pertanika Journal of Science and Technology*, 30(2), 1135–1156.

Nashwan, A. J. (2025). Data endowment as a digital waqf: An Islamic ethical framework for AI development. *Philosophy & Technology*, 38 (74). https://doi.org/10.1007/s13347-025-00905-5

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press. https://doi.org/10.1126/science.abm5861

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishers.

Organisation of Islamic Cooperation Independent Permanent Human Rights Commission [OIC-IPHRC]. (2024). *Jeddah Declaration on the guiding principles of artificial intelligence*. OIC-IPHRC.

Organisation of Islamic Cooperation-15 Dialogue Platform [OIC-15]. (2025). *Tehran Declaration on trustworthy and ethical artificial intelligence for sustainable development*. OIC-15 Second Ministerial Meeting.

Qadir, J., Islam, M. Q., & Al-Fuqaha, A. (2022). Toward accountable human-centered AI: Rationale and promising directions. *Journal of Information, Communication and Ethics in Society*, 20(2), 329–342. https://doi.org/10.1108/JICES-06-2021-0059

Raquib, A. (2015). *Islamic ethics of technology: An objectives (maqāṣid) approach*. The Other Press.

Raquib, A. (2023). Commentary on Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology*, 36(74). https://doi.org/10.1007/s13347-023-00677-w

Raquib, A., Channa, B., Zubair, T., & Qadir, J. (2022). Islamic virtue-based ethics for artificial intelligence. *Discover Artificial Intelligence*, 2(1), 11. https://doi.org/10.1007/s44163-022-00028-2

Sardar, Z. (1988). *Information and the Muslim world: A strategy for the twenty-first century*. Mansell.

Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751–752. https://doi.org/10.1126/science.aat5991

Tsourlaki, S. (2022). Artificial intelligence and fatwā issuance: A case study of Dubai and Egypt. *Islamic Inquiries*, 1(2), 107–125.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

Wahid, S. H., & Kususiyanah, A. (2020). The relevance of the usage of artificial intelligence and machine learning in legal analysis: An analysis of legal provisions in the National Sharia Council fatwā and Financial Services Authority regulation. In *Proceedings of the 1st International Conference on Research in Social Sciences and Humanities (ICORSH 2020)*.