AI and The Future of Muslim Societies: Toward A Maqāṣid-Based, Human-Centered, and Socially Beneficial AI Governance

Suparno
Universitas Islam Negeri K.H. Abdurrahman Wahid, Pekalongan, Indonesia
Email: suparno.sabila@gmail.com

Abstract

The rapid acceleration and global proliferation of Artificial Intelligence (AI) technologies present deep ontological, ethical, and socio-legal challenges to contemporary Muslim societies. Existing global AI governance frameworks are predominantly anchored in Eurocentric legal positivism and secular bioethics, often failing to address the distinct normative, cultural, and spiritual imperatives of Islamic jurisprudence. This study proposes an integrative framework for AI governance rooted in Maqāṣid al-Sharīʿah (the higher objectives of Islamic law), synthesizing classical Islamic legal philosophy with contemporary human-centered AI principles. Using a qualitative normative-philosophical and comparative legal approach, this article examines how the traditional protection of religion (ḥifẓ al-dīn), life (ḥifẓ al-nafs), intellect (ḥifẓ al-ʿaql), lineage (ḥifẓ al-nasl), and wealth (ḥifẓ al-māl) can be operationalized to mitigate algorithmic bias, data commodification, automated decision-making risks, and existential threats to human agency. The research demonstrates that a Maqāṣid-based AI governance model offers an ethically robust, socially beneficial, and human-centric alternative to prevailing market-driven regulatory models. Ultimately, the study advocates for proactive legal and institutional reform across Muslim-majority nations to ensure that AI development aligns with social justice, human dignity, and collective well-being (maṣlaḥah).

Keywords: artificial intelligence, AI governance, maqāṣid al-sharīʿah, human-centered AI, Islamic jurisprudence

Introduction

The global emergence of Artificial Intelligence (AI) represents one of the most transformative socio-technological shifts of the twenty-first century. Advanced machine learning models, autonomous systems, and generative algorithms are fundamentally reshaping public administration, healthcare, economic structures, and legal adjudication. For Muslim societies, this technological revolution is not merely a technical upgrade; it is a complex socio-legal phenomenon that touches upon every aspect of social, moral, and spiritual life. As digital transformation accelerates across the Islamic world, institutions are increasingly forced to

grapple with systems that make autonomous decisions impacting human livelihoods, legal rights, and moral agency.

The integration of AI into Muslim-majority nations offers immense potential for social development, financial inclusion through Islamic fintech, and enhanced administrative efficiency. However, it also introduces unprecedented risks regarding algorithmic surveillance, algorithmic bias, and the erosion of traditional community values. The uncritical adoption of western-centric AI architectures risks importing embedded ideological assumptions that may conflict with Islamic ethical norms. Consequently, Muslim scholars, legal practitioners, and policymakers must proactively engage with AI technologies to ensure that their deployment serves genuine human flourishing rather than technological determinism or corporate hegemony.

At the core of this challenge lies the tension between rapid technological acceleration and the slower pace of legal adaptation. Traditional legal systems in many Muslim-majority countries struggle to address the novel legal personalities, liability issues, and data sovereignty concerns raised by advanced AI applications. Without a clear ethical and normative anchor, AI adaptation risks degenerating into either passive acceptance or total resistance. Therefore, establishing a forward-looking legal and philosophical framework is essential to navigate this paradigm shift gracefully and effectively.

Muslim societies stand at a critical historical juncture where technology must be governed rather than merely consumed. The future of Muslim societies in the digital age depends on their capacity to formulate normative standards that honor their rich jurisprudential heritage while actively contributing to global discourse. Achieving this balance requires moving beyond reactionary fatwas toward a proactive, systemic governance structure. Such a structure must seamlessly integrate technological innovation with immutable ethical principles.

Ultimately, the goal is not to resist technological advancement, but to humanize and moralize it. AI must be oriented toward serving social justice, poverty alleviation, and communal solidarity rather than exacerbating socioeconomic inequalities. By establishing a robust governance vision, Muslim societies can ensure that technology serves humanity, keeping moral responsibility (taklīf) firmly at the center of automated operations.

Contemporary global AI governance is largely dominated by frameworks originating from North America and the European Union, such as the EU AI Act and various OECD guidelines. While these frameworks provide valuable legal instruments regarding data privacy, risk classification, and algorithmic transparency, they are fundamentally rooted in secular legal positivism, utilitarian ethics, or neoliberal market principles. These epistemological foundations prioritize individual autonomy, risk mitigation, and economic efficiency, often overlooking broader communal, spiritual, and socio-ecological obligations.

In secular regulatory models, ethics is frequently reduced to technical compliance checklists, risk management protocols, or corporate self-regulation. This minimalist approach often fails to address deeper existential questions regarding what it means to be human, the nature of moral accountability, and the holistic preservation of community structure. Furthermore, Western AI models frequently carry implicit biases that reflect the cultural priorities, linguistic dominance, and historical experiences of the Global North, thereby marginalizing non-Western worldviews and normative frameworks.

For Muslim societies, relying exclusively on imported governance models creates a normative disconnect. European or American regulatory instruments do not account for the

spiritual dimensions of human agency (khilāfah) or the moral imperative to uphold public welfare (maṣlaḥah) over profit maximization. When secular frameworks are applied without adaptation in Muslim contexts, they often create regulatory gaps in addressing issues such as algorithmic distortion of religious authority, digital public morality, and the ethical limits of human enhancement.

Moreover, market-driven AI governance models often struggle to curb predatory technological practices, such as surveillance capitalism, algorithmic manipulation of user behavior, and the systemic exploitation of digital labor. These practices run counter to the core tenets of Islamic social justice, which strictly prohibit harm (ḍarar), exploitation, and unjust enrichment. A purely risk-based or market-led governance regime lacks the philosophical depth needed to counter these systemic structural harms effectively.

Therefore, relying solely on secular AI governance is inadequate for Muslim-majority nations seeking to construct a digitally sovereign and morally coherent future. There is an urgent need to supplement global AI regulations with a comprehensive, value-laden framework that derives its authority from divine guidance and centuries of legal philosophy. Such a framework must bridge universal human rights with specific Islamic ethical requirements, offering a rich model for equitable governance.

To resolve the limitations of existing regulatory approaches, this study proposes a Maqāṣid-based framework for AI governance. Maqāṣid al-Sharīʿah, or the higher objectives of Islamic law, provides a dynamic, flexible, and comprehensive ethical framework designed to promote human welfare and prevent harm. By anchoring AI governance in the core objectives of protecting religion, life, intellect, lineage, and wealth, legal theorists and technologists can establish a holistic framework for evaluating the moral legitimacy and social utility of emerging technologies.

A Maqāṣid-based approach transforms AI governance from a reactive mechanism of error-correction into a proactive methodology for technology design and deployment. It requires that algorithms be constructed not only to optimize technical performance, but also to uphold fundamental human values and social justice. This human-centered perspective aligns naturally with Islamic theology, which views humanity as God's stewards (khulafāʾ) on Earth, entrusted with the moral responsibility to build a just, balanced, and sustainable world order.

Furthermore, a Maqāṣid-based governance model offers a principled way to reconcile technological innovation with traditional Islamic legal jurisprudence (fiqh). Rather than evaluating each algorithmic update through isolated legal rulings, Maqāṣid provides systemic principles that can guide software engineering, data management, and policy formulation. It allows legal scholars and computer scientists to converse in a shared normative language centered on human dignity, public benefit (maṣlaḥah), and the mitigation of systemic harm (rafʿ al-ḥaraj).

By placing the human being at the center of technological development, this framework explicitly rejects technological determinism, the notion that technological progress is an uncontrollable force shaping society. Instead, it asserts that technology must always remain an instrument subject to human moral judgment and legal accountability. Automated systems must never strip humans of their moral agency, decision-making capacity, or fundamental dignity.

Ultimately, formulating a Maqāṣid-based and human-centered AI governance model is not merely an academic exercise; it is an urgent socio-legal necessity for the Muslim world.

It offers a viable roadmap for Muslim-majority countries to establish digital sovereignty, protect their citizens from technological harms, and contribute a meaningful, justice-oriented perspective to the global discourse on ethical AI.

Method

This study employs a qualitative, normative-philosophical, and comparative legal research methodology designed to investigate the intersection of Artificial Intelligence, governance, and Islamic jurisprudence. Normative legal research is uniquely suited for this inquiry as it focuses on analyzing legal principles, philosophical foundations, statutory frameworks, and doctrinal texts rather than empirical statistical metrics. By adopting a philosophical lens, this study examines the underlying values that govern AI technologies and contrasts them with the ethical imperatives derived from Islamic legal philosophy.

The methodological architecture of this research is structured around uṣūl al-fiqh (Islamic legal theory) and contemporary jurisprudence, utilizing the analytical tools of ijtihād (independent legal reasoning) to address novel technological challenges. The research bridges classical Islamic juristic principles with modern legal, regulatory, and technical standards governing AI. This dual focus allows for a comprehensive critique of contemporary secular frameworks alongside a rigorous formulation of an alternative Maqāṣid-based model.

Furthermore, a comparative legal approach is utilized to evaluate existing global AI governance regimes, such as the European Union's Artificial Intelligence Act and international ethical guidelines, against Islamic legal objectives. This comparative analysis identifies areas of convergence, such as transparency and fairness, while highlighting critical divergent philosophies regarding human agency, communal responsibility, and moral purpose.

The research pipeline follows a multi-stage conceptual model moving from epistemological data collection to normative legal synthesis, as structured below:

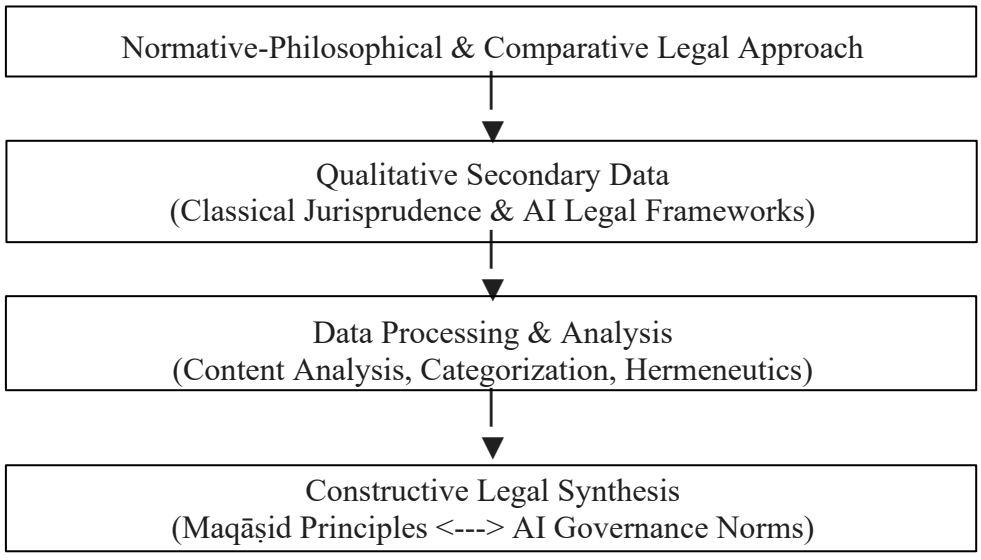


Figure 1. Multi-Stage Conceptual Model

The data utilized in this research consists exclusively of qualitative secondary data drawn from primary legal texts, classical Islamic jurisprudence, contemporary academic literature, and international legal instruments. Primary Islamic juristic sources include foundational classical works on Maqāṣid al-Sharīʿah by seminal scholars such as Al-Ghazālī,

Al-Shāṭibī, and Ibn ʿĀshūr, which provide the doctrinal basis for the protection of the five fundamental necessities (al-ḍarūriyyāt al-khams). Secondary datasets encompass modern academic literature on AI governance, machine learning ethics, algorithmic transparency, and digital law, alongside international regulatory frameworks including the EU AI Act and UNESCO's Recommendations on the Ethics of Artificial Intelligence.

Data processing and analysis were carried out using a three-stage qualitative analysis strategy comprising conceptual filtering, systematic categorization, and constructive legal synthesis. In the first stage, raw textual data collected from classical and modern sources were subjected to qualitative content analysis. In the second stage, the conceptual data were categorized using the analytical framework of the five essential objectives of Maqāṣid al-Sharīʿah (ḥifẓ al-dīn, ḥifẓ al-nafs, ḥifẓ al-ʿaql, ḥifẓ al-nasl, and ḥifẓ al-māl). The third stage employed a hermeneutic and legal harmonization method, evaluating modern AI governance principles against the normative requirements of Islamic legal maxims (al-qawāʿid al-fiqhiyyah), particularly "Harm shall be eliminated" (al-ḍarar yuzāl) and "Prevention of harm takes precedence over the attraction of benefit" (darʾ al-mafāsid muqaddam ʿalā jalb al-maṣāliḥ).

Results

Mapping the Ethical Gaps in Current AI Governance The systematic legal analysis reveals significant ethical, regulatory, and conceptual gaps within current global AI governance frameworks when evaluated against human-centered standards. While existing frameworks excel at identifying technical vulnerabilities and operational risks, they frequently fail to address deeper structural and existential harms perpetuated by automated systems. Table 1 provides a comparative analysis mapping prevailing AI governance limits against the proposed Maqāṣid-based framework.

A Comparison between Secular Governance Frameworks and the Maqāṣid al-Sharīʿah Framework

Table 1. Comparative Analysis of Governance Frameworks

No.	Comparative Dimension	Secular Frameworks	Maqāṣid Framework	Core Analytical Finding
1	**Philosophical Anchor**	Grounded in Secular Legal Positivism and Utilitarianism. Legitimacy is primarily determined by positive law, rationality, institutional authority, and calculations of utility, benefits, and costs.	Grounded in Maqāṣid al-Sharīʿah (the higher objectives of Islamic law) and Khilāfah (human stewardship). Governance is guided by moral values, divine accountability, ethical responsibility, and the preservation of human dignity.	Secular models do not necessarily establish an absolute, transcendent moral imperative. Consequently, ethical considerations may become negotiable when confronted with market pressures and economic interests. In contrast, the Maqāṣid framework positions ethics as a fundamental normative foundation that cannot be reduced to economic calculations.
2	**Primary Goal**	Prioritizes Market Efficiency,	Prioritizes Public Benefit (Maṣlaḥah), Social	Represents a paradigm shift from merely preventing

		productivity, economic optimization, institutional performance, and Risk Management.	Justice, human welfare, and the realization of comprehensive societal well-being.	technical failures and economic losses toward actively promoting holistic human flourishing and sustainable social welfare.
3	**Human Agency**	Emphasizes Individual Autonomy, individual freedom, Consumer Choice, and the protection of individual interests. Humans are primarily understood as autonomous decision-makers and rights-bearing individuals.	Emphasizes Moral Responsibility (Taklīf) and Stewardship (Khilāfah). Humans are regarded as morally accountable agents entrusted with the responsibility to govern, manage, and use technology ethically.	Repositions humans from passive data subjects or technology consumers into morally accountable agents who bear responsibility for the consequences of technological development and utilization for individuals and society.
4	**Regulatory Focus**	Focuses on Technical Compliance, legal conformity, institutional supervision, regulatory standards, and Risk Auditing.	Focuses on Systemic Harm Elimination and the principle of Darʾ al-Mafāsid (prevention and elimination of harm).	Regulatory priorities extend beyond technical compliance toward preventing structural, cultural, social, and spiritual harms. This approach emphasizes the identification and prevention of systemic risks before they generate widespread consequences.
5	**Scope of Protection**	Primarily emphasizes Individual Rights, consumer protection, civil liberties, and Data Privacy.	Integrates the protection of individuals, families, communities, society, and future generations through the principles of public benefit and the higher objectives of Islamic law.	Expands the scope of regulatory protection from individual interests toward safeguarding communal integrity, social cohesion, intergenerational justice, and the well-being of future generations.

Conceptual Comparative Synthesis

To clarify the fundamental distinctions between the two governance approaches, the following table summarizes their respective orientations and the resulting paradigm shifts.

Table 2. Conceptual Comparative Synthesis

No.	Analytical Dimension	Secular Governance Orientation	Maqāṣid Governance Orientation	Resulting Paradigm Shift
1	Philosophical Foundation	Legality, instrumental rationality, and institutional legitimacy.	Moral values, divine accountability, and human stewardship.	From legal legitimacy toward integrated legal and moral legitimacy.
2	Governance Objectives	Efficiency, productivity, and risk management.	Public benefit, social justice, and holistic human flourishing.	From efficiency-driven governance toward welfare-oriented governance.
3	Human Position	Individual autonomy, consumer rights, and personal choice.	Moral responsibility, accountability, and stewardship.	From passive consumers toward morally accountable human agents.
4	Regulatory Orientation	Technical compliance, legal standards, and risk auditing.	Prevention of systemic harm and realization of public benefit.	From compliance-based regulation toward harm-prevention governance.

5	Protective Scope	Individual rights, personal interests, and data privacy.	Individual, family, community, societal, and intergenerational protection.	From individual-centered protection toward comprehensive social protection.

Academic Interpretation and Discussion

The comparative analysis presented in Table 1 identifies five fundamental distinctions between secular governance frameworks and the Maqāṣid al-Sharīʿah-based governance framework. These distinctions demonstrate that the two approaches differ not only in their philosophical foundations but also in their governance objectives, conceptualization of human agency, regulatory priorities, and scope of protection.

Differences in Philosophical Foundations

The first distinction concerns the philosophical foundations underlying each governance framework. Secular governance frameworks generally derive their legitimacy from legal rationality, positive law, institutional authority, and utilitarian considerations. Within this perspective, governance decisions are commonly evaluated through their legality, effectiveness, efficiency, and measurable consequences.

By contrast, the Maqāṣid al-Sharīʿah framework integrates legal reasoning with moral obligations and transcendent accountability. Through the concept of Khilāfah, human beings are understood as stewards entrusted with responsibilities that extend beyond individual interests and economic benefits.

Accordingly, the Maqāṣid framework establishes a normative foundation in which ethical responsibility constitutes an integral component of governance rather than merely an additional consideration.

Differences in Governance Objectives

The second distinction relates to the primary objectives of governance. Secular frameworks frequently prioritize market efficiency, institutional performance, productivity, and risk mitigation. These objectives are particularly relevant in modern regulatory environments characterized by technological innovation, economic competition, and institutional complexity. Nevertheless, the Maqāṣid framework extends the concept of governance success beyond efficiency and risk management. Its primary orientation is the realization of Maṣlaḥah (public benefit) and social justice. This approach emphasizes that governance should not merely prevent institutional failures or economic losses but should also contribute positively to human dignity, social welfare, equitable development, and holistic human flourishing.

Differences in the Conceptualization of Human Agency

The third distinction concerns the position of human beings within governance systems. Secular governance frameworks generally emphasize individual autonomy, freedom of choice, consumer rights, and personal interests.

Although these principles provide important foundations for protecting individual liberties, they may be insufficient when governance challenges involve collective responsibility and the broader consequences of technological development.

The Maqāṣid framework introduces the concept of Taklīf, which establishes human beings as morally accountable agents. Combined with Khilāfah, this concept emphasizes

that individuals and institutions have responsibilities concerning the consequences of their decisions. Consequently, humans are repositioned from passive data subjects or technology consumers into morally responsible stewards of technological systems and their societal implications.

Differences in Regulatory Orientation

The fourth distinction concerns regulatory priorities. Secular governance frameworks generally emphasize technical compliance, legal conformity, institutional supervision, and risk auditing. These mechanisms are essential for ensuring regulatory accountability. However, technical compliance alone may not adequately address broader structural, cultural, and social harms arising from complex governance systems.

The Maqāṣid framework addresses this limitation through the principle of Darʾ al-Mafāsid, which emphasizes the prevention and elimination of harm. Under this approach, regulatory assessment should consider not only whether an institution or technological system complies with established technical standards but also whether its operations generate systemic harm, undermine human dignity, or threaten collective welfare. Therefore, the regulatory approach shifts from a predominantly compliance-oriented model toward a preventive and ethically grounded governance framework.

Differences in the Scope of Protection

The fifth distinction concerns the scope of protection. Secular governance frameworks generally emphasize individual rights, personal interests, consumer protection, and data privacy. These protections remain fundamental to modern governance systems. Nevertheless, the increasingly interconnected nature of contemporary social and technological environments requires a broader understanding of regulatory protection.

The Maqāṣid framework offers an integrated approach encompassing individuals, families, communities, society, and future generations. This broader perspective recognizes that governance decisions may generate consequences extending beyond immediate individual interests. Accordingly, regulatory protection must also safeguard communal integrity, social cohesion, intergenerational justice, and the sustainability of human welfare.

As shown in Table 1, prevailing frameworks treat ethics as an external constraint rather than an internal design requirement. This structural limitation allows technology developers to prioritize profit and speed, implementing ethical safeguards only after harms have occurred. Furthermore, secular frameworks heavily emphasize individual data privacy while remaining virtually silent on the collective, societal, and cultural disruptions caused by algorithmic deployment.

Systematizing the Five Essential Objectives (Al-Ḍarūriyyāt al-Khams) for AI Regulation A central result of this study is the operationalization of the five essential objectives of Maqāṣid al-Sharīʿah (al-ḍarūriyyāt al-khams) into a concrete normative standard for AI legal regulation. Each traditional objective directly corresponds to a specific risk vector and regulatory imperative within the contemporary AI landscape, as systematized in Table 2.

An Operational Framework for Integrating Islamic Ethical Principles into Artificial Intelligence Governance.

Table 3. Operationalization of Maqāṣid Al-Sharīʿah in AI Governance

No.	Maqāṣid al-Sharīʿah Dimension	AI-Related Risks	Governance Implications	Operational Governance Measures
1	Ḥifẓ al-Dīn (Protection of Faith and Moral Values)	Algorithmic distortion of religious information, fabricated religious narratives, manipulation of religious content, and religious deepfake materials.	Mitigating moral distortion, preserving religious integrity, and ensuring the authenticity and credibility of religious narratives disseminated through AI systems.	Implementing religious content verification mechanisms, AI-generated content disclosure, authentication protocols for religious information, and ethical oversight of AI-driven religious platforms.
2	Ḥifẓ al-Nafs (Protection of Life and Safety)	Autonomous system failures in healthcare and transportation, unsafe AI-assisted medical decisions, and the development or deployment of lethal autonomous weapons.	Ensuring human safety, preventing physical harm, strengthening system reliability, and maintaining meaningful human control over safety-critical AI applications.	Conducting rigorous AI safety audits, establishing mandatory human oversight, implementing safety testing and risk assessments, and developing accountability mechanisms for AI-related harm.
3	Ḥifẓ al-ʿAql (Protection of Intellect and Truth)	Cognitive manipulation, widespread disinformation, deepfakes, algorithmic manipulation of public opinion, and technology-driven behavioral addiction.	Protecting intellectual autonomy, promoting information integrity, preventing cognitive exploitation, and preserving the capacity for independent and informed decision-making.	Implementing misinformation detection systems, deepfake identification mechanisms, algorithmic transparency, digital literacy programs, and safeguards against manipulative recommendation systems.
4	Ḥifẓ al-Nasl (Protection of Lineage and Family)	Exploitation of children's personal data, exposure to harmful digital content, AI-enabled child exploitation, and invasion of family and domestic privacy.	Safeguarding children's digital well-being, protecting family privacy, preventing exploitation, and preserving the rights and welfare of future generations.	Establishing child-centered AI safeguards, strengthening parental and family privacy protections, restricting harmful content, implementing age-appropriate AI systems, and enforcing strict data protection standards for minors.
5	Ḥifẓ al-Māl (Protection of Wealth and Economic Equity)	Automated predatory lending, algorithmic monopolies, discriminatory pricing, financial exclusion, and exploitative automated financial decision-making.	Ensuring economic fairness, financial transparency, protection against exploitation, equitable access to financial services, and accountability in AI-driven economic activities.	Requiring transparent algorithmic decision-making, conducting financial algorithm audits, preventing discriminatory pricing, strengthening consumer protection, and ensuring fairness in automated lending and financial services.

Conceptual Synthesis: Operationalizing Maqāṣid in AI Governance

The five dimensions of Maqāṣid al-Sharīʿah provide an integrated normative foundation for identifying, assessing, and mitigating the ethical, social, economic, and institutional risks associated with artificial intelligence.

Table 4. Operationalizing Maqāṣid in AI Governance

No.	Maqāṣid Principle	Protected Interest	Primary AI Governance Objective	Expected Governance Outcome
1	Ḥifẓ al-Dīn	Faith, religious integrity, and moral values	Preventing religious misinformation and moral distortion.	Preservation of religious authenticity and ethical integrity in digital environments.
2	Ḥifẓ al-Nafs	Human life, physical safety, and well-being	Preventing AI-related physical harm and ensuring human oversight.	Safer, more accountable, and human-centered AI systems.
3	Ḥifẓ al-ʿAql	Intellectual autonomy, knowledge, and truth	Preventing cognitive manipulation and protecting information integrity.	Improved intellectual autonomy and informed decision-making.
4	Ḥifẓ al-Nasl	Lineage, family integrity, children, and future generations	Protecting children and families from digital exploitation and privacy violations.	Stronger family protection and sustainable intergenerational digital well-being.
5	Ḥifẓ al-Māl	Wealth, property, and economic justice	Preventing financial exploitation and algorithmic discrimination.	Fairer, more transparent, and accountable AI-driven economic systems.

Ḥifẓ al-Dīn: Protection of Faith and Moral Values

The principle of Ḥifẓ al-Dīn provides an ethical foundation for protecting religious integrity and moral values in AI-driven environments. The increasing capacity of generative AI to produce synthetic religious narratives, manipulate religious symbols, and disseminate fabricated religious information creates new challenges for the authenticity of religious knowledge.

Within this framework, AI governance should incorporate mechanisms for verifying religious content, identifying synthetic materials, and ensuring the credibility of digital religious information. The objective is not to restrict technological innovation but to ensure that technological development does not undermine religious integrity, moral values, and the freedom of individuals to practice their faith.

Ḥifẓ al-Nafs: Protection of Life and Safety

Ḥifẓ al-Nafs establishes the protection of human life and physical safety as fundamental governance priorities. This principle is particularly relevant to AI applications in healthcare, transportation, industrial automation, and autonomous weapon systems.

AI-related failures in these sectors may generate serious consequences, including physical injury, loss of life, and violations of human safety. Therefore, governance mechanisms should require rigorous safety assessments, continuous monitoring, human oversight, and clear institutional accountability.

The principle establishes that technological efficiency must never take precedence over the protection of human life.

Ḥifẓ al-ʿAql: Protection of Intellect and Truth

The principle of Ḥifẓ al-ʿAql emphasizes the preservation of human intellectual capacity, rational judgment, and access to reliable knowledge.

In the context of artificial intelligence, algorithmic manipulation, disinformation, deepfake technologies, and addictive recommendation systems may influence cognitive autonomy and distort public understanding.

Accordingly, AI governance should promote information integrity, algorithmic transparency, digital literacy, and effective mechanisms for identifying misleading or manipulated content.

The central objective is to ensure that AI technologies strengthen human intellectual capabilities rather than undermine independent reasoning and informed decision-making.

Ḥifẓ al-Nasl: Protection of Lineage and Family

Ḥifẓ al-Nasl extends the scope of AI governance to the protection of family integrity, children's welfare, and the interests of future generations.

The exploitation of children's personal data, exposure to harmful digital content, AI-enabled child exploitation, and unauthorized access to domestic information represent significant challenges to family protection in digital environments.

The Maqāṣid framework therefore requires child-centered safeguards, stronger privacy protection, responsible data governance, and preventive mechanisms against digital exploitation.

This principle also introduces an intergenerational perspective, recognizing that AI governance decisions may affect the welfare, dignity, and opportunities of future generations.

Ḥifẓ al-Māl: Protection of Wealth and Economic Equity

The principle of Ḥifẓ al-Māl establishes economic justice, property protection, and the prevention of financial exploitation as essential components of AI governance.

The increasing use of algorithmic decision-making in banking, lending, insurance, and digital financial services may create risks involving discriminatory pricing, predatory lending, financial exclusion, and algorithmic monopolization.

To address these challenges, governance frameworks should require transparency in automated financial decisions, independent algorithmic audits, fair treatment of consumers, and effective accountability mechanisms.

The objective is to ensure that AI-driven economic development contributes to equitable prosperity rather than reinforcing existing economic inequalities.

By systematizing the five essential objectives, this study demonstrates that Maqāṣid al-Sharīʿah provides an exhaustive, logically coherent framework for evaluating modern AI risks. It provides legal scholars and regulatory authorities with an objective, value-driven rubric to assess, license, or restrict AI technologies based on their holistic impact on society.

Comparative Legal Synthesis: Convergence and Divergence with Global Frameworks

The research produced a comparative legal mapping that highlights critical points of convergence and divergence between a Maqāṣid-based AI governance framework and prominent global legal standards, specifically the EU AI Act and UNESCO's Ethics of AI recommendations.

In terms of convergence, all three frameworks strongly emphasize the necessity of algorithmic transparency, data privacy, accountability, and the reduction of technical bias. The Maqāṣid framework fully supports the EU AI Act’s risk-based classification model, agreeing that high-risk AI deployment requires pre-market conformity assessments and continuous

human oversight. Similarly, the Maqāṣid framework concurs with UNESCO's insistence on human dignity, non-discrimination, and environmental sustainability.

However, significant divergences emerge regarding ultimate purpose and normative justifications. While secular frameworks view risk mitigation primarily as a tool to maintain market stability and user trust, the Maqāṣid framework views governance as a moral obligation (amānah) aimed at achieving holistic public benefit (maṣlaḥah) and divine stewardship (khilāfah). Furthermore, secular frameworks overwhelmingly prioritize individual rights, whereas the Maqāṣid model balances individual rights with collective communal protections.

Operationalizing "Human-Centeredness" through Khilāfah (Stewardship) The results of this study redefine the concept of "Human-Centered AI" through the foundational Islamic doctrine of Khilāfah (human stewardship on Earth). Under the framework of Khilāfah, human beings are endowed with unique intellectual, moral, and spiritual capacities that assign them complete accountability (taklīf) for their actions and inventions. Technology, including the most sophisticated artificial intelligence, is categorically classified as an instrumental tool (wāsilah) designed to assist humanity in fulfilling its stewardship obligations. Consequently, AI can never be granted moral agency, legal personhood, or autonomous rights that rival or displace human responsibility.

Operationalizing human-centeredness through Khilāfah requires establishing strict legal boundaries against the over-delegation of human judgment to automated systems. In legal adjudication, public administration, and ethical decision-making, final discretionary power must permanently reside with human actors. Algorithmic recommendations must strictly remain advisory tools (qarār inshāʾī) rather than binding legal determinations.

Institutional Architecture for Maqāṣid-Compliant AI Governance The final key result of this study is the proposal of an institutional architecture designed to operationalize Maqāṣid-based AI governance within national legal systems. Regulating complex algorithmic ecosystems requires specialized institutional bodies capable of conducting real-time legal, technical, and ethical evaluations. The proposed framework outlines a tri-partite institutional model:

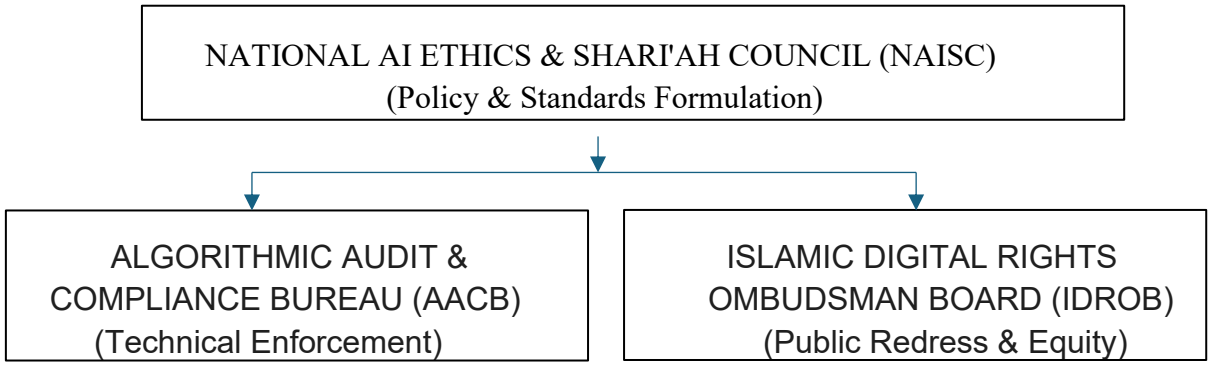


Figure 2. Framework Outlines a Tri-Partite Institutional Model

1. National AI Ethics and Sharīʿah Council (NAISC): A centralized, multidisciplinary regulatory body comprising Islamic legal jurists, computer scientists, legal scholars, and ethicists. Tasked with issuing binding policy directives and dynamic Maqāṣid compliance standards.

2. Algorithmic Audit and Compliance Bureau (AACB): An operational technical body responsible for conducting mandatory pre-market certification, algorithmic transparency checks, and bias audits.
3. Islamic Digital Rights Ombudsman Board (IDROB): An independent judicial mechanism established to handle public grievances, user complaints, and civil litigation arising from algorithmic discrimination and automated rights violations.

Discussion

Re-evaluating Algorithmic Autonomy and Moral Agency in Islamic Jurisprudence The rapid evolution of autonomous AI systems requires a rigorous jurisprudential re-evaluation of legal personality (ahliyyah) and moral agency within Islamic law. Modern computational theories often debate whether advanced deep-learning systems should be granted a form of artificial legal personality similar to corporate entities. However, from a Maqāṣid jurisprudence perspective, granting legal personality or moral agency to non-human, non-conscious software algorithms is fundamentally problematic and legally untenable.

In Islamic jurisprudence, legal capacity is bifurcated into ahliyyat al-wujūb (the capacity to acquire legal rights and obligations) and ahliyyat al-adāʾ (the capacity to execute legal actions and bear moral responsibility). The latter is strictly contingent upon ʿaql (rational intellect) and taklīf (moral accountability before God). Because AI systems operate purely through statistical pattern recognition and deterministic processing, devoid of consciousness, soul (rūḥ), or moral intent (niyyah), they can never possess genuine legal capacity or moral accountability.

Attempting to assign legal personality to AI risks creating legal tax havens and accountability voids, allowing developers, corporations, and state actors to evade liability for harm caused by their systems. Islamic legal principles dictate that liability must always trace back to human actors. Consequently, legal jurisprudence must classify AI systems strictly as instruments or agents (wakīl) operating under the absolute responsibility of their human principals. Under the classical legal maxim "The act of the agent is the act of the principal," ultimate civil and criminal liability rests with the human developers, corporate executives, or administrative officers who set the parameters of the algorithm.

Mitigation of Systemic Harm (Darʾ al-Mafāsid) versus Innovation and Economic Benefit A central debate in contemporary AI governance centers on balancing technological innovation with risk mitigation. Industry proponents frequently argue that stringent legal regulations smother innovation. However, Islamic legal philosophy offers a sophisticated resolution through core jurisprudence maxims, particularly "Prevention of harm takes precedence over the attraction of benefit" (darʾ al-mafāsid muqaddam ʿalā jalb al-maṣāliḥ).

In the context of AI governance, economic benefit (jalb al-maṣāliḥ)—such as automated productivity gains and rapid financial transactions—cannot be pursued at the expense of systemic societal harm (mafsadah). When an AI deployment poses credible risks of structural bias, mass labor displacement, or the erosion of democratic discourse, the legal imperative to mitigate harm takes absolute priority over profit generation.

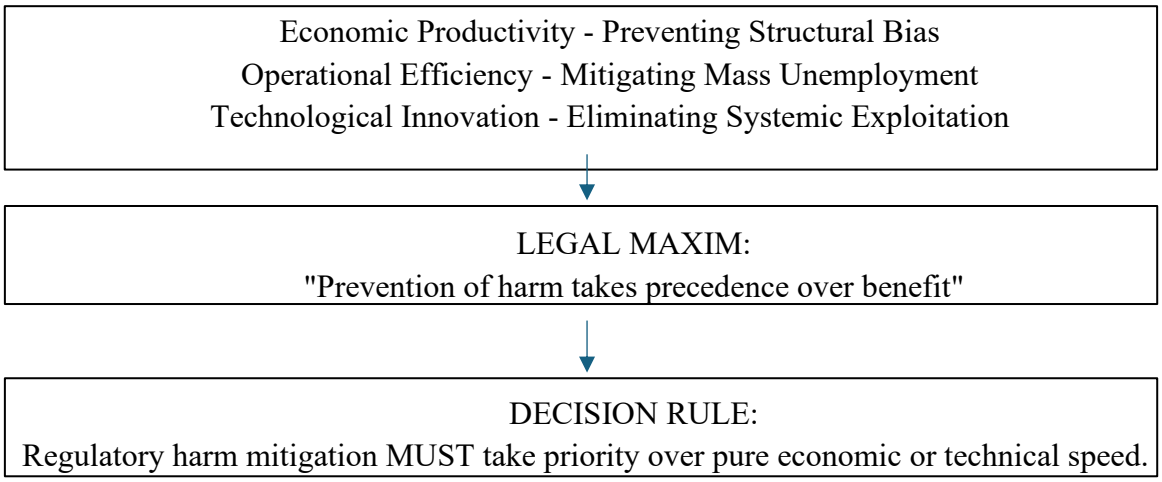


Figure 3. Jalb al-Maṣāliḥ (Attracting Benefit) Darʾ al-Mafāsid (Preventing Harm)

This methodological prioritization does not imply a total rejection of technological innovation. Rather, it establishes a precautionary principle (sadd al-dharāʾiʿ) requiring rigorous pre-market risk assessments. Furthermore, the legal maxim "Harm shall be eliminated" (al-ḍarar yuzāl) mandates that when unforeseen algorithmic harms emerge post-deployment, regulatory authorities possess the immediate legal power to suspend, modify, or recall the offending AI system.

Data Sovereignty, Algorithmic Bias, and Social Justice (ʿAdālah) Data is the fundamental fuel driving modern AI systems, yet current global data harvesting models perpetuate severe structural inequalities. Large technology conglomerates routinely harvest vast amounts of data from populations across the Global South and Muslim world. This dynamic has given rise to "digital colonialism," wherein harvested data is processed abroad to generate commercial AI products that are then resold to source populations, embedding foreign cultural norms and biases.

From an Islamic socio-legal perspective, this extractive dynamic violates social justice (ʿadālah) and public ownership of communal resources. Data generated by a society constitutes a collective resource linked to digital sovereignty and cultural identity. The uncompensated exploitation of communal data represents a contemporary form of unjust enrichment (akl amwāl al-nās bi-al-bāṭil). Furthermore, foreign-trained AI models deployed indiscriminately in Muslim societies risk distorting Islamic religious concepts and undermining the legal presumption of innocence (al-aṣl barāʾat al-dhimmah). To counter these systemic issues, a Maqāṣid-based governance model advocates for strict national data sovereignty laws and localized data infrastructure.

Preserving Cultural, Linguistic, and Religious Identity in Generative AI The proliferation of Generative AI systems presents unprecedented challenges to the preservation

of cultural, linguistic, and religious identity across Muslim societies. These engines are primarily trained on English-centric corpora, inherently imbuing them with secular worldviews. When Generative AI systems encounter Islamic jurisprudence, theology, and historical scholarship, they frequently struggle with contextual nuance, leading to hallucinatory legal rulings and the flattening of diverse scholarly traditions.

Preserving intellect (ḥifẓ al-ʿaql) and religion (ḥifẓ al-dīn) requires active legal and technological intervention to safeguard authentic sacred knowledge. Legal systems must mandate clear digital labeling and watermarking for all AI-generated media and text, preventing the deceptive dissemination of synthetic religious texts or deepfake statements attributed to scholars. Furthermore, state institutions across the Muslim world must actively collaborate to build high-quality, peer-reviewed, open-access digital repositories of Islamic jurisprudence to serve as foundational datasets for training localized AI models.

Strategic Roadmap for Policy, Legal Reform, and Technical Standardization Translating a Maqāṣid-based AI governance framework into an operational reality requires a clear, phased strategic roadmap:

1. Short-Term Milestones (1–2 Years): Establish national regulatory bodies (NAISC and AACB). Enact immediate executive orders requiring clear mandatory labeling for generative synthetic media, mandatory pre-market audits for high-risk AI applications in finance and healthcare, and strict prohibitions on automated biometric mass surveillance.
2. Medium-Term Milestones (3–5 Years): Enact comprehensive legislative reform updating civil, criminal, and administrative codes to integrate AI accountability and data sovereignty. Develop standardized technical benchmarks for assessing algorithmic bias and model explainability. Mandate the construction of sovereign cloud data centers.
3. Long-Term Milestones (5–10 Years): Harmonize AI governance standards across Organization of Islamic Cooperation (OIC) member states to establish a unified regional digital market. Fully integrate Islamic legal philosophy and Maqāṣid ethics into national computer science, data analytics, and law school curricula.

Conclusion

This study has articulated a holistic, human-centered framework for Artificial Intelligence governance grounded in the higher objectives of Islamic law (Maqāṣid al-Sharīʿah). The research demonstrates that prevailing global AI regulatory regimes are structurally ill-equipped to address the deep cultural, ethical, and existential challenges facing contemporary Muslim societies. By operationalizing the five essential objectives of Maqāṣid (ḥifẓ al-dīn, ḥifẓ al-nafs, ḥifẓ al-ʿaql, ḥifẓ al-nasl, and ḥifẓ al-māl), this study provides a coherent, value-laden framework that transforms AI governance from post-hoc technical error-correction into a proactive methodology for ethical technology design, deployment, and legal oversight.

The theoretical contribution of this article lies in bridging classical Islamic jurisprudence (uṣūl al-fiqh) with contemporary digital legal theory, offering a dynamic interpretation of human agency (khilāfah) and strict liability (ḍamān) in automated environments. Practically, the study provides a tripartite institutional model—comprising an Ethics and Sharīʿah Council, an Algorithmic Audit Bureau, and a Digital Rights Ombudsman—capable of implementing pre-

market conformity assessments, mitigating algorithmic bias, enforcing data sovereignty, and ensuring accessible judicial redress. Future empirical research is needed to evaluate the real-world performance of these institutional bodies within specific national legal systems.

References

Abū Dāwūd, S. b. al-Ashʿath. (2009). *Sunan Abī Dāwūd*. Dār al-Risālah al-ʿĀlamiyyah.

Al-Ghazālī, A. Ḥ. (1901). *Al-Mustaṣfā min ʿilm al-uṣūl*. Al-Maṭbaʿah al-Amīriyyah.

Al-Shāṭibī, A. I. (1997). *Al-Muwāfaqāt* (I. Ramaḍān, Ed.). Dār al-Fikr.

Al-Zuḥaylī, W. (1986). *Uṣūl al-fiqh al-Islāmī*. Dār al-Fikr.

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. http://data.europa.eu/eli/reg/2024/1689/oj

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). http://dx.doi.org/10.1162/99608f92.8cd550d1

Ibn ʿĀshūr, M. al-Ṭ. (2006). *Ibn Ashur: Treatise on Maqāṣid al-Sharīʿah* (M. E.-T. El-Mesawi, Trans.). International Institute of Islamic Thought.

Ibn Manẓūr, M. b. M. (1990). *Lisān al-ʿArab*. Dār Ṣādir.

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

Kamali, M. H. (2008). *Shari'ah law: An introduction*. Oneworld Publications.

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

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. http://dx.doi.org/10.4159/harvard.9780674736061

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.

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

UNESCO. (2022). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Kannike, U. M. M., & Fahm, A. O. (2025). Exploring the ethical governance of artificial intelligence from an Islamic ethical perspective. *Jurnal Fiqh, 22*(1), 134–161. http://dx.doi.org/10.22452/fiqh.vol22no1.5