Ethical and Epistemological Perspectives of Islam on Artificial Intelligence and Social Change: An Integrative Framework Based on Maqasid al-Shariah

Muchammad Fairuz Ubadi
Universitas Islam Negeri K.H Abdurrahman Wahid Pekalongan, Indonesia
E-mail: muchammad.fairuz.ubadi226006@mhs.uingusdur.ac.id

Abstract

The exponential development of Artificial Intelligence (AI) not only transforms global technological and economic landscapes but also reshapes social structures, religious authorities, and knowledge-production patterns in modern society. Despite providing extraordinary efficiencies, AI introduces fundamental ethical dilemmas that secular-utilitarian ethical paradigms and pure technocratic profit criteria fail to resolve completely. This study aims to elaborate deeply on Islamic ethical and epistemological perspectives in responding to the existential challenges of AI, as well as to construct a technology governance framework rooted in the philosophy of Maqasid al-Shariah. Employing a qualitative normative-library research method, this study analyzes classical and contemporary Islamic legal sources compared with recent digital ethics literature. The findings indicate that the integration of Maqasid al-Shariah values—specifically the preservation of religion (hifz al-din), life (hifz al-nafs), intellect (hifz al-aql), progeny/honor (hifz al-nasl), and wealth (hifz al-mal)—can cover the blind spots in modern AI governance. The principles of harm prevention (la darar wa la dirar), transcendental moral responsibility (taklif), and information verification criteria (tabayyun) are capable of mitigating algorithmic bias, personal data commercialization, and the erosion of critical human thinking. This study concludes that AI must be positioned as an instrument of stewardship (khilafah) oriented toward public welfare (maslahah mursalah), and recommends the formation of an interdisciplinary consortium among Islamic legal scholars, technologists, and policy regulators.

Keywords: artificial intelligence, islamic ethics, digital epistemology, maqasid al-shariah, social change, algorithmic bias, tabayyun

Introduction

The 21st century is characterized by massive technological acceleration, where Artificial Intelligence (AI) emerges as the primary driver of the fourth industrial revolution. The capacity of AI to process big data, recognize complex patterns, and generate autonomous content through machine learning and generative AI has permeated nearly every facet of human life. Ranging from medical diagnosis systems and financial market predictions to industrial automation, public policy formulation, and education, AI is no longer merely a mechanical assistant but an autonomous decision-making agent influencing the fate of millions (Aroud et al., 2016; Sinaga, 2025). However, behind this technological optimism lies a profound ethical and epistemological crisis.

Widespread AI adoption without a robust ethical framework has triggered alarming socio-humanitarian issues. One of the most critical challenges is the emergence of algorithmic bias and structured discrimination. Because machine learning systems train themselves on historical human data that frequently contains stereotypes, flawed habits, or racial and gender prejudices, AI-generated decisions often perpetuate these social injustices automatically (Priyatna, 2025). Furthermore, the phenomena of personal data commercialization (surveillance capitalism) and algorithmic opacity—where AI reasoning processes reside within a 'black box' that is difficult to trace—eliminate transparency and legal accountability. In socio-religious contexts, the proliferation of generative AI tools also threatens traditional religious authorities and triggers the erosion of critical thinking due to the onslaught of disinformation, hoaxes, and information manipulation (Apriansyah et al., 2026).

The dominant global discourse on AI ethics is largely dictated by modern Western philosophy centered on utilitarianism (maximizing material utility) or secular individualistic rights. This approach proves to have fundamental limitations because it tends to treat ethics merely as procedural compliance or economic efficiency cost-benefit calculations (B'Markone, 2017; Raquib et al., 2022). Secular approaches frequently lose the spiritual dimension, holistic justice, and transcendental moral commitment. Consequently, technological development often becomes isolated from the noble goals of human dignity preservation and social-ecological order maintenance.

Within this landscape, the Islamic intellectual tradition offers a rich, holistic, and comprehensive perspective. Islam as a universal teaching (rahmatan lil 'alamin) possesses normative, philosophical, and jurisprudential instruments capable of providing an ethical foundation oriented toward public welfare (Firdaus et al., 2025). Through the principles of Maqasid al-Shariah—the foundational objectives of Islamic law—technology is not viewed as a neutral entity or an end in itself, but as a means (wasilah) to achieve genuine human well-being in this world and the hereafter (Mubaraq, 2026). Although interest in AI and religious discourse has begun to grow in Indonesia, research mapping the relationship between Islamic epistemology, algorithmic governance, and social change concretely and deeply remains relatively limited.

Therefore, this study aims to elaborate comprehensively on how Islamic ethics and epistemology respond to the dynamics of AI development. Specifically, this research focuses on: (1) analyzing the root ethical and epistemological problems caused by modern AI autonomization, (2) formulating an AI governance framework based on the five pillars of Maqasid al-Shariah, and (3) establishing concrete strategies for an interdisciplinary synthesis among Islamic legal scholars, data scientists, and policy regulators to steer equitable social change.

Methods

Research Design

This research adopts a qualitative approach utilizing library research with a normative-philosophical and hermeneutic-analytical orientation. This approach is chosen because the

primary focus involves examining abstract concepts, epistemological structures, and Islamic legal principles, and dialogically mapping them onto contemporary AI technological ethics and philosophy. The study does not merely present literal texts but conducts a critical reconstruction of Islamic values so they can be operationalized within modern information technology.

Subjects and Sources of Data

Research data sources are categorized into two primary groups to maintain academic validity:

- **Primary Data:** Authoritative Islamic legal and philosophical texts, including classical usul al-fiqh literature concerning Maqasid al-Shariah (such as works by Al-Shatibi and Al-Ghazali), contemporary religious fatwas related to technology, as well as official national AI ethics regulations and policy documents (such as the Ministry of Communication and Informatics Circular Letter on AI Ethics) and international standards (such as UNESCO Recommendation on the Ethics of AI).
- **Secondary Data:** Peer-reviewed scientific journal articles (indexed in SINTA and Scopus), academic textbooks, and reliable research reports concerning digital ethics, machine learning, sociology of technology, and Islamic epistemology published between 2010 and 2026.

Research Procedures and Stages

This study was conducted through four systematic methodological stages:

- **Literature Inventory Stage:** Collecting documents, fiqh texts, and scientific journals relevant to AI topics, computer ethics, and Islamic legal philosophy.
- **Thematic Categorization Stage:** Sorting and grouping AI ethical issues (such as bias, privacy, autonomy, and hoaxes) based on their impact domains on societal life.
- **Conceptualization and Maqasid Mapping Stage:** Analyzing AI issues using the lens of the five pillars of Maqasid al-Shariah and legal maxims (qawa'id fiqhiyyah) to discover ethical resolution nodes.
- **Synthesis and Recommendation Formulation Stage:** Constructing an integrative AI governance framework and providing concrete inputs for developers and policy makers.

Data Collection and Processing Techniques

Data collection utilized documentation techniques and structured literature reviews. Data verification was achieved through source triangulation by cross-referencing views from various Islamic legal experts and technology ethics specialists. Data was processed through rigorous filtering based on relevance, issue recency, and academic depth.

Data Analysis Techniques

Data analysis was performed qualitatively using a combination of content analysis and thematic qualitative analysis. The author conducted close reading, theme coding, and deductive-inductive reasoning. The analytical framework is anchored in the five pillars of Maqasid al-Shariah—namely hifz al-din (preservation of religion), hifz al-nafs (preservation of life), hifz al-aql (preservation of intellect), hifz al-nasl (preservation of progeny/honor), and hifz al-mal

(preservation of wealth)—combined with primary legal maxims such as 'la darar wa la dirar' (no harm and no reciprocating harm) and the principle of 'tabayyun' (verification of truth).

Results

The analysis results successfully identify and map the philosophical and practical contributions of Islamic ethics and epistemology to modern AI existential challenges. Findings are categorized into three main domains: (1) mapping AI vulnerabilities through the lens of Maqasid al-Shariah, (2) formulating bias mitigation and data transparency ethics, and (3) safeguarding society's epistemic integrity from information distortion.

1. Mapping AI Vulnerabilities Based on the Five Pillars of Maqasid al-Shariah

The analysis demonstrates that threats posed by uncontrolled AI deployment affect all core dimensions of human preservation in Islam (al-daruriyyat al-khamsah). The table below outlines the mapping between technical AI vulnerabilities and Maqasid al-Shariah preservation pillars:

Table 1. Mapping of AI Ethical Issues within the Maqasid al-Shariah Framework

Maqasid al-Shariah Pillar	AI Threat / Technical Vulnerability	Islamic Ethical Solution & Principle
Hifz al-Din (Preservation of Religion)	Erosion of religious authority due to automated AI fatwa commodification, sacred text hallucinations, and digital ritual shifts.	Scholarly integrity, verification of knowledge chains (sanad), and restricting AI strictly as an assistant tool (not an autonomous mufti).
Hifz al-Nafs (Preservation of Life)	Algorithmic discrimination in healthcare access, biased security risk assessments, and the threat of lethal autonomous weapons systems (LAWS).	Principle of human dignity (karamah insaniyyah) and absolute prohibition on delegating life-death decisions to machines.
Hifz al-Aql (Preservation of Intellect)	Massive dissemination of hoaxes/deepfakes, addictive algorithm loops, and critical thinking decline due to passive dependency.	Implementation of Tabayyun (verification), cognitive independence protection, and algorithmic reasoning transparency.
Hifz al-Nasl (Preservation of Progeny/Honor)	Identity theft, spread of AI-generated non-consensual deepfake pornography, and family privacy data commercialization.	Strict privacy protection (hifz al-sirr), individual honor safeguarding, and the right to be forgotten.
Hifz al-Mal (Preservation of Wealth)	Economic monopoly by big tech conglomerates, computational financial fraud, and consumer data exploitation	Economic distributive justice (adl), prohibition of gharar (uncertainty/manipulation), and prevention of data monopoly

	without fair compensation.	(ihtikar).

2. Algorithmic Bias Mitigation and Enforcement of Accountability (Taklif)

A fundamental finding of this study is that algorithmic bias reflects historical human biases perpetuated by program codes. In the perspective of Sharia, tolerating biases that disadvantage marginalized groups violates the principle of fundamental justice ('adl) and the prohibition of oppression (zulm). Through the maxim 'la darar wa la dirar' (no harm and no reciprocating harm), AI system developers are morally and legally required to conduct ethical audits on training data to sever automated discrimination chains (Hadisaputra et al., 2026).

Furthermore, regarding algorithmic opacity (the black box problem), Islamic legal philosophy rejects the notion that autonomous technology can be exempted from legal liability. In Islam, the concept of a legal subject bearing ethical obligations is termed Taklif. Because machines or computer programs possess neither moral consciousness (dhamir) nor a soul, AI can never become a mukallaf subject. Therefore, accountability (mas'uliyyah) remains fully attached to the humans behind the technology—namely the algorithm designers, data provider companies, and operating institutions (Bin-Nashwan et al., 2023). Explaining algorithmic operational processes (explainable AI) becomes a religious-legal obligation to ensure justice.

3. Islamic Epistemology and Safeguarding Knowledge Integrity

The third finding relates to the epistemological domain. Generative AI and natural language processing (NLP) systems operate based on statistical and probabilistic data computations. This implies that AI does not comprehend 'truth' or 'meaning'; rather, it mathematically predicts the most likely word sequences. If society treats AI output as absolute truth, a dangerous epistemological shift occurs: scientific and spiritual truths degrade into mere statistical popularity data (Apriansyah et al., 2026).

Islamic epistemology affirms that genuine knowledge originates from divine revelation (wahyu), rational sound reasoning (aql), and verified sensory observation (hiss). To safeguard intellect integrity (hifz al-aql), Islam mandates the tradition of Tabayyun—the process of verification, confirmation, and critical evaluation of any information before acceptance and dissemination. Implementing tabayyun within the digital ecosystem demands designing AI systems equipped with transparency markers (watermarking), data source traceability, and human-in-the-loop confirmation mandates for critical decisions (Supriatin et al., 2025).

Discussion

The discourse surrounding these research findings highlights the critical urgency of integrating Islamic thought into global AI ethics debates. Hitherto, AI ethics frameworks developed in Western nations tend to center on individual autonomy, the precautionary principle, and formal regulatory compliance. Although these principles hold positive value, they frequently prove fragile when confronted with commercial capitalist pressures prioritizing product release velocity and market profit efficiency (Cranch et al., 1999; Nurrisma & Inayatillah, 2026).

The Islamic ethical framework based on Maqasid al-Shariah offers a counter-paradigm while complementing these deficiencies. Islam incorporates balanced social justice (tawazun) and transcendental spiritual responsibility. From the Islamic viewpoint, humanity carries a trust as vicegerent on earth (khalifah fi al-ard). As khalifah, humans are granted freedom to innovate and explore science-technology, yet this freedom is bounded by the duty to preserve social and ecological environments. Technology is not merely an efficiency tool, but an instrument of devotion to spread public welfare across creation (maslahah mursalah) (Ananda et al., 2025).

The implications of this discussion are fundamental for social change in contemporary Muslim societies. On one hand, Islam rejects technophobia (fear and aversion to technology) because technological advancement can be optimized to accelerate social welfare, poverty alleviation, and educational quality enhancement. On the other hand, Islam firmly rejects blind technocracy that exalts algorithms above human values and religious morality. In developing countries such as Indonesia, integrating Maqasid al-Shariah into national AI ethics policies can serve as a cultural defense bulwark against cybercrimes, data exploitation, and digital colonialism that erodes national sovereignty (Halim, 2024).

Research Limitations and Future Research Directions

Although this study has successfully constructed a comprehensive philosophical and normative framework, certain limitations must be acknowledged. First, this research is conceptual-qualitative based on literature, meaning it lacks empirical testing or technical algorithm auditing on operating AI systems in the field. Second, applying Islamic law to contemporary technological issues requires evolving interdisciplinary consensus (ijma) alongside rapid AI advancements. Consequently, future research is advised to venture into applied research, such as developing Maqasid-based AI ethics auditing instruments directly integrable into programming code (ethical coding/value-sensitive design) and testing public trust levels toward ethically labeled AI applications.

Conclusion

Based on the entire discussion above, this study concludes that Artificial Intelligence (AI) represents a dual-nature human innovation: it holds immense potential to advance civilization, yet simultaneously brings existential risks to ethical and social orders if developed without moral controls. The Islamic ethical and epistemological framework grounded in Maqasid al-Shariah proves capable of providing a holistic guide to steer AI development. By safeguarding the five core pillars—religion, life, intellect, progeny, and wealth—AI development can shift away from exploitative capitalist orientations toward equitable, just usage centered on human dignity enhancement.

The principle of harm prevention (la darar wa la dirar) mandates algorithmic bias mitigation, while the accountability principle (taklif) asserts that ethical and legal responsibility remains vested in human designers and operators. Furthermore, the tabayyun principle serves as an epistemological bulwark protecting society's thinking integrity from digital disinformation onslaughts. The primary recommendation of this study is the importance of building sustainable

interdisciplinary synergy and collaboration among Islamic legal scholars (fuqaha), computer/data scientists, sociologists, and government regulators. This synergy is essential to formulate technical regulations, ethical auditing standards, and technology education curricula imbued with Islamic values and national perspective.

Acknowledgments

Acknowledgments and funding information have been omitted to comply with double-blind review requirements.

Conflict of Interests

The author declares no financial conflicts of interest or personal relationships that could have influenced the objectivity of this research.

References

Abramson, M. (2014, May 11). Understanding digital ethics in modern news dissemination. The Washington Post. http://www.washingtonpost.com/dir/subdir/2014/05/11/story.html

Al-Ghazali, A. H. (1901). Al-Mustasfa min 'Ilm al-Usul. Al-Maktabah al-Amiriyyah.

Al-Shatibi, A. I. (1997). Al-Muwafaqat fi Usul al-Shari'ah (H. Diraz, Ed.). Dar Ibn 'Affan.

Ananda, A. F., Nurhidayah, S., Zahirianata, M., & Sa'adah, N. (2025). Pandangan Islam terhadap Etika Kecerdasan Buatan (Artificial Intelligence) dalam kehidupan sehari-hari. Sujud: Jurnal Agama, Sosial Dan Budaya, 1(3), 397–409. https://doi.org/10.63822/yd7xwt39

Apriansyah, A., Hasanah, U., Fauzi, M., & Kalsum, N. U. (2026). Islam, teknologi, dan artificial intelligence: Tantangan epistemologi, etika digital, dan otoritas keagamaan dalam konteks pendidikan Islam. Pendas: Jurnal Ilmiah Pendidikan Dasar, 11(2), 183–196.

Aroud, B. H., Breck, P. T., & Falkone, R. M. (2016). Algorithmic bias and social impact in automated systems. Journal of Research in Personality, 13, 254-276. http://dx.doi.org/10.1016/0032-026X.56.6.895

Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Tanggung jawab etis penggunaan artificial intelligence di tanah pendidikan: Formulasi paradigma baru untuk teknologi otonom. MASMAN Master Manajemen, 3(2), 57–83.

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.

B'Markone, S. O. (2017). Modern education and technological transformation: Questions and answers. Journal of Language and Education, 4(3), 55-79.

Cranch, B. E., Aroud, B. P., & Falkone, R. (1999). Technology in modern life: Ethical perspectives. Pearson.

Dignum, V. (2019). Responsible Artificial Intelligence: How to develop and use AI in a responsible way. Springer Nature. https://doi.org/10.1007/978-3-030-30371-6

El-Ghazali, M. (2020). Islamic ethics and modern technology: Bridging the gap. Journal of Islamic Studies and Contemporary Thought, 8(1), 45-62.

Firdaus, M. A., Syihabuddin, M., & Fuady, Z. (2025). Islam and Artificial Intelligence: Perspectives from traditionalist and modernist Muslim communities in Indonesia. Miqot: Jurnal Ilmu-Ilmu Keislaman, 49(1), 141. https://doi.org/10.30821/miqot.v49i1.1333

Floridi, L. (2013). The ethics of information. Oxford University Press.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1), 1-15. https://doi.org/10.1162/99608f92.8cd550d5

Hadisaputra, Az Zahra, S. A., Amaliyah, R., Zaliha, N., Ramadani, N. R., Alamsyah, A., & Nur, A. A. (2026). Perspektif Islam tentang Etika Akal Imitasi (AI) di Kalangan Mahasiswa Muslim: Studi Kasus Universitas Muhammadiyah Makassar. Al-Zayn: Jurnal Ilmu Sosial & Hukum, 4(2), 344–359. https://doi.org/10.61104/alz.v4i2.4158

Halim, A. (2024). Implementasi kecerdasan buatan dalam ijtihad kontemporer: Peluang dan tantangan hukum Islam di era 5.0. Scholastica: Jurnal Pendidikan dan Kebudayaan, 77.

Herbst, D. M., Griffith, N. R., & Slama, K. M. (2014). Human behavior and digital interaction norms. Journal of Rural Mental Health, 38, 20–35. http://dx.doi.org/10.1037/rmh0000008

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

Mubaraq, H. H. (2026). Artificial intelligence dalam perspektif Al-Qur'an: Telaah etika, ontologi, dan batasan kemanusiaan. Al-Widad, 4(2), 40–57.

Nurrisma, I. I., & Inayatillah. (2026). Integrasi nilai-nilai Islam dalam pemanfaatan media digital dan artificial intelligence di era teknologi modern. IMTIYAZ: Jurnal Ilmu Keislaman, 10(1), 39–49.

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

Priyatna, S. E. (2025). Penerapan AI dan machine learning dalam pendidikan Islam: Tantangan etika dan pendekatan integratif berbasis Maqāṣid al-Sharī'ah. Tarbiyatuna: Jurnal Pendidikan Ilmiah, 10(1), 119–136. https://doi.org/10.55187/tarjpi.v10i1.6236

Qadir, J., & Ahad, A. (2021). Artificial intelligence for social good: An Islamic perspective. IEEE Access, 9, 98115-98130.

Rahman, F. (1982). Islam and modernity: Transformation of an intellectual tradition. University of Chicago Press.

Raquib, A., Channa, B., Zubair, T., & Qadir, J. (2022). Islamic virtue-based ethics for artificial intelligence. Discover Artificial Intelligence, 2, 11. https://doi.org/10.1007/s44163-022-00028-0

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

Sartor, G. (2020). AI and law: Ethical boundaries and legal responsibility. European Law Journal, 26(3), 210-229.

Sinaga, F. Y. (2025). Etika penggunaan kecerdasan buatan dalam perspektif Islam: Mitigasi risiko dan optimalisasi manfaat. Scope: Jurnal TarbiyahbilQalam, 65.

Supriatin, I., Syarifah, S., Susilawati, E., Supriadi, I., & Apriani, R. (2025). Formulasi etika kecerdasan buatan (AI) dalam pendidikan Islam: Pendekatan Maqāṣid al-Sharī'ah dan tafsir tematik Al-Qur'an. Halaqa: Journal of Islamic Education, 1(2), 121–140. https://doi.org/10.61630/hjie.v1i2.28

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO Digital Library.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.