

Article history (leave this part):

Submission date: 06.02-2026

Acceptance date: 28-03-2026

Available online: 20-05-2026

Keywords:

artificial intelligence, AI ethics, Islamic ethics, maqāṣid al-sharī'ah, justice ('adl), human dignity (karāmah), surveillance and privacy, algorithmic bias

Funding:

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

Competing interest:

The author(s) have declared that no **competing interests** exist.

Cite as (leave this part):

Hanan Abufares Elkhimry; . (2024). Title. Journal of Science and Knowledge Horizons: 4(1), 283-293. <https://doi.org/10.34118/jskp.v2i02.2727>



The authors (2026). This Open Access article is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0) (<http://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial reuse, distribution, and reproduction are permitted with proper citation

Artificial Intelligence and Moral Responsibility: An Islamic Normative Approach

- 1- Dr. Salma Razzaq * (Assistant Professor, Air University, Islamabad, Pakistan), wajisuperb@gmail.com, ORCID: 0009-0008-6175-0311
- 2- Nasir Razzaq (CEO, Xpertno Research Center, Islamabad, Pakistan), rajasunny.razzaq@gmail.com, ORCID: 0000-0001-6939-2856

Abstract:

Artificial Intelligence (AI) is rapidly transforming social, economic, and religious spheres, raising ethical concerns that extend beyond technical performance to questions of justice, accountability, human dignity, and moral responsibility. This paper examines contemporary AI ethics through an Islamic moral framework, arguing that AI should be ethically evaluated rather than merely regulated or celebrated. Using a normative ethical analysis grounded in Islamic moral philosophy, the study first identifies key ethical challenges associated with AI, including algorithmic bias, privacy and surveillance, automation-driven economic disruption, autonomous decision-making, and accountability deficits. It then articulates core Islamic ethical principles, such as tawḥīd, khilāfah, amānah, 'adl, maṣlaḥah, maqāṣid al-sharī'ah, and karāmah insāniyyah, as normative criteria for ethical judgement. These principles are subsequently applied to assess the moral permissibility, limitations, and conditions governing contemporary AI practices. The analysis demonstrates that Islamic ethics permits beneficial AI applications while imposing firm constraints where technologies undermine justice, human agency, dignity, or public welfare. The study concludes that Islamic moral philosophy offers a rigorous and culturally grounded framework for ethically responsible AI governance.

* Salma Razzaq

Introduction

The rapid growth of Artificial Intelligence (AI) has led to a huge change in society and technology as intelligent technologies have become a part of almost every aspect of human life. AI systems are changing the way people converse with each other, work, worship, govern, and make decisions. They promise to make things easier, faster, and more creative than ever before. AI currently affects people's lives on both an individual and a group level, from algorithmic decision-making in jobs and economics to predictive systems in healthcare, governance, and security. But the quick growth of AI has also raised major ethical questions that need to be looked into by scholars. Algorithmic bias, data privacy, accountability, dehumanization, and autonomous decision-making have become key problems that raise important questions about justice, human dignity, and moral responsibility in societies that use technology (Bostrom et al., 2014; Russell, 2019).

One of the biggest ethical worries about AI is algorithmic bias, which happens when systems educated on old or unrepresentative data reinforce and make worse existing socioeconomic disparities. Research has shown that AI-driven decision-making can lead to unfair results that hurt people because of their ethnicity, gender, religion, or socioeconomic background, frequently while pretending to be neutral (Tang & Kamarudin, 2025). Increasing concerns about data privacy and surveillance are closely related to this problem. AI's dependence on extensive data collection engenders significant apprehensions over the safeguarding of personal information, informed consent, and the exploitation of sensitive data. These worries are especially strong in religious and cultural settings where privacy, trust, and moral responsibility are seen as important moral duties (Nuha, 2025).

In addition to concerns about fairness and privacy, academics have expressed apprehension on the potentially dehumanizing consequences of AI, particularly in areas historically reliant on human judgment, empathy, and ethical reasoning. In religious contexts, excessive dependence on AI-generated guidance jeopardizes the function of scholars ('ulamā'), diminishes personal spiritual participation, and simplifies intricate moral reasoning to algorithmic outputs (Waseu, 2025). The ethical challenges posed by AI are further intensified by the unprecedented speed and scale of its development. Advances in machine learning, deep learning, and natural language processing have enabled AI technologies to evolve at an exponential rate, outpacing existing ethical, legal, and regulatory frameworks (Anderson et al., 2007).

AI applications now span a wide range of sectors, including healthcare, finance, transportation, entertainment, and public administration, each presenting distinct ethical risks and responsibilities. Automation, while enhancing productivity and efficiency, has raised serious concerns about job displacement,

labour deskilling, and economic inequality, prompting debates about social justice and the dignity of work in increasingly automated economies (Floridi & Luciano, 2013).

Similarly, AI-driven personalization systems, though beneficial for user engagement, have intensified concerns over data exploitation, behavioural manipulation, and the erosion of individual autonomy (Bostrom et al., 2014). In governmental contexts, AI's use in predictive policing, administrative decision-making, and surveillance infrastructures has sparked urgent debates about transparency, accountability, and the potential institutionalization of bias within public systems (Marcus et al., 2019).

These changes have led to a growing agreement that AI is not just a technical issue, but also a very moral one. As AI systems progressively affect life-and-death choices, access to opportunities, and the allocation of societal goods, inquiries regarding fairness, accountability, human agency, and moral responsibility have become inescapable (Russell, 2019). While much of the existing discourse on AI ethics has been shaped by secular philosophical traditions, liberal conceptions of autonomy, and technocratic governance models, there is increasing recognition of the need to engage alternative ethical frameworks capable of addressing the moral complexity of intelligent technologies in culturally grounded ways.

Islamic civilization has a long history of thinking on ethics, knowledge, justice, and human responsibility. Islamic moral philosophy, grounded in Qur'ānic revelation and Prophetic guidance, has historically formulated intricate normative frameworks to address social, legal, and technical transformations (Nasr, 2006). Fundamental ethical principles, including fairness (ʿadl), trust (amānah), human stewardship (khilāfah), and public welfare (maṣlaḥah), provide a cohesive moral lexicon for assessing human activities and their repercussions. Modern scholars have increasingly contended that these principles offer significant ethical frameworks for addressing contemporary technological dilemmas, particularly those associated with AI (Ramadan, 2009; Mohadi & Tarshany, 2023). Recent studies have underscored the significance of maqāṣid al-sharīʿah, the higher objectives of Islamic law, in evaluating matters such as privacy, manipulation, and technological harm, promoting the incorporation of Islamic ethical principles into AI development and governance (Mohadi & Tarshany, 2023; Saputra & Asbi, 2025).

Building on this growing body of scholarship, the present study approaches AI ethics through an Islamic moral lens, not merely by demonstrating ethical compatibility, but by employing Islamic principles as evaluative criteria for judging contemporary AI practices. Rather than treating Islamic ethics as a symbolic or supplementary framework, this paper adopts a normative analytical

approach that uses Islamic moral concepts to assess the ethical legitimacy of AI systems and their societal implications.

The paper proceeds in three main sections to critically evaluate contemporary ethical challenges of Artificial Intelligence by applying foundational Islamic ethical principles as normative criteria for judging AI practices. The first section identifies and contextualizes the primary ethical challenges posed by AI in contemporary society. The second section articulates key Islamic ethical principles and their philosophical foundations relevant to technological development and moral responsibility. The third section applies these principles to evaluate specific AI-related ethical dilemmas, demonstrating how Islamic teachings can inform ethically responsible and socially grounded approaches to AI governance and practice.

Research Objectives

- To examine major ethical challenges arising from the development and deployment of Artificial Intelligence
- To articulate the key ethical principles of Islamic teachings relevant to technological development and human responsibility
- To apply Islamic ethical principles as evaluative tools in assessing the moral permissibility, limitations, and conditions governing contemporary AI practices

Research Methodology

This study adopts a normative ethical analysis grounded in Islamic moral philosophy to evaluate contemporary ethical challenges posed by Artificial Intelligence. Rather than employing empirical data collection or experimental methods, the research is conceptual and analytical in nature, focusing on the ethical assessment of AI practices through established Islamic normative principles. This approach is particularly suited to inquiries that seek to judge moral permissibility, responsibility, and human values in technological contexts where ethical reasoning, rather than measurement, is central.

The analysis proceeds through a structured, multi-stage evaluative process. First, major ethical challenges associated with the development and deployment of Artificial Intelligence are identified through a critical engagement with contemporary scholarship in AI ethics, technology studies, and applied philosophy. These challenges include, but are not limited to, algorithmic bias, data privacy, surveillance, automation and labour displacement, autonomous decision-making, accountability, and the impact of AI on human dignity and moral agency. The identification of these challenges is informed by widely cited works in AI ethics literature and reflects issues that have significant societal and moral implications.

Second, the study articulates key Islamic ethical principles relevant to technological development and human responsibility by examining foundational sources of Islamic moral thought. These include Qur'ānic teachings, Prophetic traditions, and classical as well as contemporary Islamic ethical and jurisprudential scholarship. Central concepts such as justice (‘adl), trust (amānah), human stewardship (khilāfah), public welfare (maṣlahah), human dignity (karāmah insāniyyah), and the objectives of Islamic law (maqāsid al-sharī‘ah) are treated not merely as descriptive values but as normative criteria for ethical evaluation.

Third, these Islamic ethical principles are systematically applied as evaluative tools to assess contemporary AI practices and ethical dilemmas. Each identified AI challenge is examined in relation to relevant Islamic moral principles to determine its ethical permissibility, limitations, or conditional acceptability. This stage of analysis involves explicit normative judgement, assessing whether particular AI practices align with, violate, or require modification under Islamic ethical standards. The evaluation emphasizes moral accountability, preservation of human dignity, and the protection of essential human interests as articulated within Islamic ethical thought.

By employing Islamic ethics as an active evaluative framework rather than a symbolic or supplementary reference, this methodology enables a principled assessment of AI technologies that foregrounds human agency, moral responsibility, and social justice. The approach contributes to broader ethical debates by demonstrating how religious moral frameworks can function as rigorous normative tools in contemporary discussions on AI governance and responsible technological development.

Ethical Challenges of Artificial Intelligence

The rapid integration of Artificial Intelligence into contemporary societies has produced a complex ethical landscape in which technological advancement is increasingly intertwined with moral, social, and spiritual concerns. While AI offers significant opportunities for innovation, efficiency, and problem-solving across sectors, its deployment has simultaneously generated profound ethical challenges that demand careful evaluation. These challenges are not merely technical in nature but fundamentally concern questions of justice, accountability, human dignity, and moral agency. Existing scholarship in AI ethics has highlighted that without principled ethical frameworks, AI systems risk reinforcing structural inequalities, eroding privacy, undermining human responsibility, and destabilising social and moral orders (Floridi & Luciano, 2013; Russell, 2019). From an Islamic ethical perspective, these challenges acquire additional significance, as they directly implicate values central to moral accountability, trust, and the preservation of human dignity.

a. Algorithmic Bias, Fairness, and Ethical Integrity

One of the most extensively documented ethical challenges of Artificial Intelligence concerns algorithmic bias and the erosion of fairness in decision-making systems. AI systems trained on historical or socially embedded data often reproduce and amplify existing inequalities related to race, gender, religion, and socioeconomic status, while presenting such outcomes as neutral or objective (Saffiya, 2018). These biases have been observed in domains such as hiring, credit scoring, facial recognition, and criminal justice, where algorithmic decisions can significantly affect individuals' life chances. The ethical concern extends beyond technical inaccuracies to deeper questions of justice and human worth, as individuals are reduced to statistical patterns rather than treated as morally significant agents (Vallor, 2016).

The opacity of many AI systems further intensifies this challenge. The so-called "black box" nature of complex machine-learning models obscures the reasoning processes underlying algorithmic decisions, making it difficult to identify responsibility or contest unjust outcomes (Pasquale, 2015; Mittelstadt et al., 2016). From an ethical standpoint, such opacity conflicts with principles of procedural justice and moral accountability, as affected individuals are denied meaningful explanations for decisions that shape their lives. Scholars have therefore argued that fairness in AI cannot be reduced to mathematical optimisation alone but must be grounded in broader ethical commitments that recognise human dignity and contextual judgement (Bostrom et al., 2014).

Islamic ethical scholarship has increasingly engaged with these concerns, emphasising that justice (*'adl*) and trust (*amānah*) require AI systems to avoid harm, discrimination, and unjust advantage. Studies focusing on Islamic ethical frameworks argue that algorithmic systems must be evaluated in light of *maqāṣid al-sharī'ah*, particularly the preservation of human dignity and social equity, rather than solely on technical efficiency (Ajizah et al., 2025; Saputra & Asbi, 2025). In this sense, algorithmic bias represents not only a technical flaw but a moral failure that demands normative judgement and ethical reform.

b. Privacy, Surveillance, and the Erosion of Moral Autonomy

A second major ethical challenge arises from AI's dependence on large-scale data collection, which has enabled unprecedented forms of surveillance and behavioural monitoring. AI-powered systems now track, categorise, and predict individual behaviour across digital platforms, public spaces, and institutional contexts, often without meaningful consent or transparency (Zuboff, 2019). This expansion of surveillance infrastructures raises serious concerns regarding personal autonomy, informational self-determination, and the concentration of power in the hands of those who control data and algorithmic systems (Crawford, 2021).

The ethical implications of such practices are particularly acute in religious and cultural contexts where privacy, confidentiality, and trust are considered moral obligations rather than mere legal rights. Scholars have noted that AI-mediated data practices risk violating deeply held ethical norms related to personal dignity and moral accountability, especially when sensitive information is exploited for profit, control, or behavioural manipulation (Nuha, 2025). In educational and religious domains, concerns have also been raised about the misuse of AI systems that compromise confidentiality, distort authority, or undermine trust between individuals and institutions (Arif et al., 2025; Niam, 2024).

Beyond privacy violations, the aggregation of personal data generates asymmetries of power that limit individuals' ability to contest or escape algorithmic categorisation. Continuous monitoring and automated profiling can subtly shape behaviour, erode moral autonomy, and normalise intrusive forms of governance (Floridi & Luciano, 2013). Ethical analyses therefore emphasise the need for transparency, data minimisation, and robust oversight mechanisms capable of addressing AI's dynamic and evolving risks. Without such safeguards, AI-driven surveillance threatens to transform societies into systems of continuous monitoring that undermine both individual freedom and collective trust.

c. Automation, Economic Disruption, and Social Justice

The economic consequences of AI-driven automation constitute a third major ethical challenge, particularly with respect to employment, labour dignity, and distributive justice. While automation has historically accompanied technological progress, the scale and speed of AI-induced disruption are widely regarded as unprecedented (Brynjolfsson & McAfee, 2014). AI systems increasingly replace or restructure human labour across manufacturing, services, administration, and knowledge-based professions, raising concerns about job displacement, deskilling, and long-term economic inequality.

Scholars have warned that AI-driven labour transformation risks concentrating wealth and power among those who own and control algorithmic infrastructures, while marginalising workers who lack access to reskilling opportunities (Eubanks, 2018). Algorithmic management systems further exacerbate these concerns by subjecting workers to opaque performance evaluations and automated disciplinary mechanisms that undermine human judgement and dignity. Ethical responses to these challenges increasingly emphasise the importance of reskilling, social safety nets, and inclusive economic policies that prioritise human well-being over purely profit-driven automation (Anderson et al., 2007).

From an Islamic ethical perspective, these economic disruptions raise fundamental questions about justice, responsibility, and the moral value of work. Contemporary Islamic scholarship highlights that technological advancement

must serve public welfare (maṣlaḥah) and prevent harm, rather than exacerbate inequality or undermine human dignity (Ajizah et al., 2025; Alamsyah et al., 2025). Accordingly, AI-driven economic change demands ethical evaluation not only in terms of efficiency but also in relation to fairness, shared benefit, and moral accountability.

d. Autonomy, Accountability, and the Challenge to Human Dignity

Perhaps the most profound ethical challenge posed by Artificial Intelligence concerns the rise of autonomous decision-making systems and their implications for human agency and moral responsibility. AI systems now operate in high-stakes domains such as healthcare, criminal justice, transportation, military operations, and religious guidance, where algorithmic decisions can directly affect life, liberty, and moral authority (Floridi et al., 2016). The delegation of such decisions to machines raises urgent questions about accountability, liability, and the erosion of human judgement.

Autonomous systems act at unprecedented speed and scale, allowing errors, biases, and harmful outcomes to propagate rapidly and become embedded within institutional structures (Whittlestone et al., 2019). When combined with opaque decision-making processes, this dynamic undermines individuals' ability to contest decisions or seek moral redress. The ethical challenge is further intensified by concerns over human deskilling and over-reliance on automation, which risk diminishing critical thinking and moral engagement over time (Noble, 2018).

In religious and ethical contexts, these concerns are especially significant. Scholars have cautioned that AI-mediated religious guidance and counselling may compromise doctrinal integrity, dilute moral responsibility, and weaken the role of human scholars and ethical deliberation (Arif et al., 2025; Niam, 2024). Islamic ethical discourse emphasises that moral agency, intention (niyyah), and accountability cannot be delegated to machines without undermining foundational principles of responsibility and human dignity. Consequently, autonomous AI systems pose not only regulatory challenges but existential ethical questions about what it means to remain morally responsible in an increasingly automated world.

Key Ethical Principles of Islamic Teachings

Islamic ethical thought offers a comprehensive moral framework for evaluating human action, grounded in divine guidance, moral accountability, and the pursuit of justice and public welfare. Unlike purely instrumental or consequence-based ethical systems, Islamic ethics integrates metaphysical commitments with normative principles that govern responsibility, intention, and social impact. As technological systems such as Artificial Intelligence increasingly mediate human decision-making and social organisation, these ethical foundations provide a principled basis for assessing whether technological development serves human flourishing or undermines moral responsibility.

a. Moral Authority, Accountability, and Human Stewardship

At the core of Islamic ethics lies the principle of *tawḥīd*, the absolute unity and sovereignty of God, which establishes that ultimate authority belongs to the Divine alone (Hallaq, 2009). *Tawḥīd* situates all human knowledge, power, and creativity within a framework of accountability, rejecting the notion that technological capability grants moral autonomy independent of ethical restraint. Human beings, as creators and deployers of technology, remain morally responsible for the consequences of their actions and innovations. From this perspective, AI systems are not neutral artefacts but human constructions that must be aligned with ethical purposes consistent with human accountability before God (Kamali, 2003). *Tawḥīd* thus challenges technocratic assumptions that prioritise efficiency or innovation over moral responsibility, insisting instead that technological progress must be ethically justified.

Closely linked to *tawḥīd* is the Qur'ānic concept of *khilāfah*, which defines human beings as stewards or trustees of creation (Qur'ān 2:30; 6:165). This notion affirms both human dignity and responsibility, positioning humans as moral agents entrusted with the careful management of resources, knowledge, and power. Applied to AI development, *khilāfah* implies that technology should enhance human capacity to fulfil ethical stewardship rather than displace moral judgement or agency. AI systems, therefore, should function as tools that support responsible decision-making, social justice, and environmental care, rather than mechanisms that absolve humans of ethical responsibility or concentrate unchecked power (Sachedina, 2009; Abou El Fadl, 2014).

Another foundational concept is *amānah*, which refers to trust, responsibility, and fidelity in fulfilling moral obligations. The Qur'ān describes humanity's acceptance of moral responsibility as a trust that even the heavens and earth declined to bear (Qur'ān 33:72). In the context of AI, *amānah* underscores that technological power entails ethical obligations toward individuals and communities affected by automated systems. Developers, institutions, and policymakers act as trustees responsible for ensuring that AI technologies do not exploit vulnerability, violate consent, or erode trust. This principle is particularly relevant where AI systems affect marginalised populations or operate in contexts, such as education, healthcare, or religious guidance, where ethical harm may not be immediately visible but deeply consequential (Abou El Fadl, 2014).

b. Justice, Public Welfare, and Ethical Governance

The ethical demand for justice is articulated through the principle of *'adl*, which occupies a central position in Islamic moral reasoning. The Qur'ān repeatedly commands believers to uphold justice, even when doing so conflicts with personal interest or social bias (Qur'ān 4:135). Islamic justice extends beyond procedural fairness to encompass substantive equity, requiring that outcomes themselves reflect moral balance and social responsibility.

Complementing justice is the principle of *iḥsān*, which calls for excellence, compassion, and moral integrity beyond minimal compliance (Qur'ān 16:90). Together, these principles establish that ethical systems—including AI technologies—must not only avoid harm or discrimination but actively promote fairness, dignity, and human welfare. In technological contexts, this demands heightened ethical scrutiny of systems that reproduce inequality, obscure accountability, or privilege efficiency over equity.

Islamic jurisprudence further refines ethical evaluation through the concept of *maṣlaḥah* (public interest) and the framework of *maqāṣid al-sharī'ah*, the higher objectives of Islamic law. These objectives aim to preserve essential human goods, including life, intellect, religion, lineage, and property (Auda, 2008). Rather than rigid rule-following, *maqāṣid*-based reasoning evaluates actions and systems according to their broader social impact and capacity to prevent harm. Applied to AI, this framework enables a principled assessment of whether technological systems promote human well-being or undermine core ethical goods. AI applications that threaten human life, manipulate cognition, erode moral autonomy, or concentrate wealth unjustly fail to meet *maqāṣid*-based ethical standards, while technologies that protect life, enhance knowledge, or promote equitable access align with Islamic ethical objectives (Auda, 2008; Siddique & Rauf, 2025).

Islamic ethical reasoning is not monolithic but accommodates pluralist moral reasoning, integrating both duty-based and outcome-oriented considerations within a coherent normative framework. Contemporary scholars argue that Islamic ethics supports a pluralist approach to AI governance, allowing for interdisciplinary engagement while maintaining moral boundaries rooted in divine accountability and public welfare (Elmahjub, 2023). This pluralism reinforces the need for collaboration between technologists, ethicists, and religious scholars to ensure that AI systems are evaluated not only for technical performance but for their ethical legitimacy within specific cultural and moral contexts (Nuha, 2025; Waseu, 2025).

Islamic ethics places strong emphasis on human dignity (*karāmah insāniyyah*) as an inherent and divinely bestowed status. The Qur'ān affirms that all human beings are honoured regardless of social position or ability (Qur'ān 17:70). This principle directly challenges AI practices that reduce individuals to data points, predictive profiles, or algorithmic outputs. Any technological system that undermines human agency, moral reasoning, or the capacity for ethical choice violates the foundational Islamic commitment to dignity and responsibility (Sachedina, 2009).

These principles, *tawḥīd*, *khilāfah*, *'adl*, *iḥsān*, *amānah*, *maṣlaḥah*, *maqāṣid al-sharī'ah*, and *karāmah*, constitute a coherent normative framework for

evaluating technological development. Rather than offering abstract moral ideals, Islamic ethics provides concrete evaluative criteria through which AI practices can be judged in terms of permissibility, limitation, and ethical responsibility. These principles form the foundation for the subsequent analysis, in which contemporary AI challenges are assessed through an Islamic ethical lens to determine whether technological innovation aligns with moral accountability and human flourishing.

Applying Islamic Ethical Principles as Evaluative Tools for Contemporary AI Practices

Having identified the principal ethical challenges posed by Artificial Intelligence and articulated the foundational Islamic ethical principles relevant to technological development, this section applies those principles as normative evaluative tools for assessing the moral permissibility, limitations, and conditions governing contemporary AI practices. Rather than treating Islamic ethics as a symbolic value system, the analysis employs Islamic moral concepts to judge whether specific AI applications align with justice, human dignity, accountability, and public welfare. This approach reflects a growing body of scholarship that emphasises virtue ethics, maqāṣid-based reasoning, and moral responsibility as central to Islamic engagement with emerging technologies (Raquib et al., 2022; El-Hady & Zenrif, 2024).

Islamic ethical reasoning does not reject technological advancement outright; instead, it adopts a principled stance that permits beneficial technologies while imposing moral constraints to prevent harm, injustice, and moral erosion. Contemporary scholars argue that AI systems must be evaluated not only for functional efficiency but for their capacity to cultivate ethical responsibility, protect human dignity, and serve collective well-being (Lusiana et al., 2024; Ghaly, 2023). From this perspective, Islamic ethics functions as a regulative framework that distinguishes ethically permissible uses of AI from those that require limitation, modification, or prohibition.

a. Algorithmic Bias and Justice (‘Adl)

The Islamic imperative of justice provides a robust ethical foundation for evaluating algorithmic bias in AI systems. The Qur’ān commands believers to uphold justice without compromise, declaring that justice is nearer to righteousness (Qur’ān 5:8). Applied to AI, this principle demands that algorithmic systems avoid discriminatory outcomes and ensure equitable treatment across social groups. AI practices that reproduce racial, gender, or socio-economic bias violate the ethical requirement of ‘adl and are therefore morally impermissible unless substantively reformed.

Justice in Islamic ethics extends beyond formal equality to substantive fairness, requiring that systems account for structural disadvantage rather than

merely applying uniform rules. Inclusive development processes are ethically warranted under the principle of *shūrā* (consultation), which emphasises participatory decision-making and the inclusion of affected communities in governance (Qurʾān 42:38; Esposito & Voll, 1996).

Furthermore, Islamic legal principles concerning liability and compensation for harm (*damān*) support the establishment of accountability mechanisms for algorithmic discrimination. The complementary principle of *ihsān* reinforces the obligation to move beyond harm avoidance toward the active promotion of equity, suggesting that ethically legitimate AI systems should reduce—not entrench—social injustice.

b. Privacy, Surveillance, and Moral Boundaries (Ḥalāl and Ḥarām)

Islamic ethics places strong normative limits on surveillance and data exploitation through its protection of privacy, dignity, and trust. The Qurʾān explicitly prohibits spying and unwarranted intrusion into private life (Qurʾān 49:12), and Prophetic teachings condemn violations of personal privacy as serious moral transgressions (Hasan, 1984). These teachings establish that AI-driven surveillance practices lacking meaningful consent are ethically impermissible, as they undermine human dignity and moral autonomy.

Islamic legal reasoning further refines these limits through the principles of *ṣatr* (protection of privacy) and *ḍarūrah* (necessity). While necessity may permit limited intrusion under exceptional circumstances, such intrusions must be proportionate, transparent, and strictly constrained to legitimate purposes (Kamali, 2003). Data collected through AI systems constitutes an *amānah*, imposing obligations of protection, non-exploitation, and responsible stewardship on those who control it. Consequently, Islamic ethics supports strong safeguards for consent, purpose limitation, transparency, and rights of access, correction, and deletion. AI practices that prioritise surveillance, behavioural manipulation, or profit-driven data extraction without ethical justification fail to meet these standards.

c. Autonomous Weapon Systems and the Sanctity of Life

Few AI applications raise ethical concerns as severe as autonomous weapons systems. Islamic ethical reasoning, grounded in the *maqāṣid* principle of preserving life, places strict moral limits on technologies capable of lethal force (Auda, 2008). Islamic just war theory (*ādāb al-ḥarb*) establishes conditions of discrimination, proportionality, and moral intention that govern the legitimacy of warfare (Hashmi & Sohail, 2002). Autonomous systems that select and engage targets without human deliberation violate these conditions by removing moral intention (*niyyah*) and accountability from life-and-death decisions.

Islamic law strictly prohibits harm to non-combatants and requires restraint even in armed conflict (Hasan, 1984). Given the current inability of autonomous systems to reliably distinguish combatants from civilians or assess

proportionality, their deployment remains ethically impermissible under Islamic principles. From this perspective, meaningful human control over lethal force is a moral requirement rather than a technical preference. AI technologies may assist human judgement but must not replace human moral agency where the sanctity of life is at stake.

d. Economic Justice, Labour, and Social Responsibility

AI-driven automation presents significant ethical challenges in relation to employment, economic inequality, and the dignity of work. Islamic economic ethics affirms that material resources and wealth carry social obligations, articulated through mechanisms such as zakāt and ṣadaqah, which ensure equitable circulation of wealth within society (Qur'ān 9:60; 51:19). AI practices that exacerbate unemployment, concentrate wealth, or marginalise workers without compensatory measures violate the ethical requirements of justice and public welfare.

Islamic teachings honour labour and craftsmanship as morally valuable forms of human activity, with Prophetic traditions emphasising the dignity of earning through one's own effort (Ṣaḥīḥ al-Bukhārī, vol. 3, ḥadīth 2072). From this standpoint, AI systems that displace workers without ethical safeguards threaten not only economic stability but moral dignity. Islamic ethics therefore supports redistributive mechanisms, retraining initiatives, and policies that ensure AI-generated wealth contributes to collective well-being rather than private accumulation. Practices consistent with maṣlaḥah prioritise productive use of technology to enhance human capability rather than render human contribution obsolete.

e. Transparency, Accountability, and Due Process

Islamic legal principles strongly support transparency and accountability in decision-making. The concept of bayyinah (clear evidence) requires that judgments affecting individuals be explicable and subject to scrutiny (Qur'ān 2:245–252). Applied to AI, this principle challenges opaque “black box” systems that deny individuals understanding or recourse. Ethical AI systems must therefore be interpretable, contestable, and governed by due process.

The concept of amānah further obligates developers and institutions to ensure competence, safety, and responsible deployment. Islamic jurisprudence permits exceptions under necessity but within strict ethical limits; harmful systems cannot be justified solely by potential benefits if less harmful alternatives exist (Kamali, 2003). Consequently, AI practices that obscure accountability or evade responsibility fail to meet Islamic ethical standards.

f. Preserving Human Agency and Moral Responsibility

At the heart of Islamic ethics lies the affirmation of human beings as moral agents endowed with intellect (‘aql) and free will (ikhtiyār) (Nasr, 2006). This conception has significant implications for AI deployment. Technologies should

augment human capacities rather than replace moral judgement in ethically consequential domains. Over-reliance on automation risks atrophying critical reasoning and moral engagement, undermining the very faculties that Islamic tradition seeks to cultivate (Ramadan, 2009).

Decisions involving moral weight, such as healthcare, criminal justice, and employment, require meaningful human deliberation. Islamic ethics therefore cautions against the automation of moral choice, viewing it as a form of ethical and spiritual harm. Education, ethical literacy, and critical engagement with AI are essential to preserving moral responsibility, aligning with the Qur'ānic emphasis on knowledge and reflection (Qur'ān 39:9; 96:1–5).

Through these evaluations, Islamic ethical principles function as active normative criteria for judging AI practices rather than abstract moral ideals. AI systems are ethically permissible when they uphold justice, dignity, accountability, and public welfare; ethically limited when they require safeguards and human oversight; and ethically impermissible when they undermine moral agency, violate dignity, or cause preventable harm. This evaluative framework demonstrates how Islamic ethics can contribute substantively to contemporary debates on AI governance, offering principled guidance rooted in moral responsibility and human flourishing.

Recommendations

Based on the ethical evaluation conducted through Islamic moral principles, this study proposes a set of normative and practical recommendations aimed at guiding the development, deployment, and governance of Artificial Intelligence in ethically responsible ways. These recommendations are not intended to restrict technological innovation, but rather to ensure that AI systems operate within moral boundaries that preserve human dignity, accountability, justice, and public welfare. Islamic ethics, when applied as a normative framework, provides clear guidance for shaping AI practices that serve humanity while preventing ethical harm.

i. Embedding Ethical Accountability in AI Design and Deployment

AI systems should be developed and deployed within clear frameworks of human accountability that recognise technology as a moral trust (*amānah*) rather than an autonomous authority. Institutions and developers must retain responsibility for the outcomes of AI systems, particularly in high-stakes domains such as healthcare, governance, finance, education, and security. Ethical oversight mechanisms, such as review boards, impact assessments, and continuous monitoring, should be integrated throughout the AI lifecycle to ensure that systems remain aligned with moral obligations and public interest. From an Islamic ethical perspective, accountability before God and society requires that

technological decisions remain traceable to human agents rather than obscured behind automated processes.

ii. Inclusive and Transparent Systems

To address algorithmic bias and structural inequality, AI systems should be designed with explicit commitments to justice (*'adl*) and equity. This requires systematic auditing of training data, decision logic, and outcomes to identify and mitigate discriminatory effects. Inclusive development practices, such as engaging diverse stakeholders and affected communities, are essential for preventing ethical blind spots and reinforcing participatory governance consistent with the Islamic principle of *shūrā*. Transparency in algorithmic decision-making should be prioritised to enable explanation, contestation, and redress, thereby reinforcing procedural justice and trust in AI systems.

iii. Strengthening Privacy Protections and Ethical Data Governance

Given the ethical centrality of privacy and dignity in Islamic teachings, AI systems that rely on data collection must adopt robust data governance practices grounded in consent, proportionality, and purpose limitation. Surveillance-oriented applications should be strictly regulated and ethically justified under clearly defined conditions of necessity, ensuring that intrusions into privacy remain minimal and accountable. Data controllers must treat personal information as an ethical trust (*amānah*), implementing safeguards that prevent misuse, exploitation, or commercial manipulation. Clear policies on data access, correction, and deletion should be established to preserve individual autonomy and moral agency.

iv. Maintaining Human Oversight and Moral Decision-Making

AI applications that affect life, liberty, and moral responsibility, such as autonomous weapons, judicial systems, and medical decision-making, should remain subject to meaningful human oversight. Islamic ethical reasoning strongly discourages delegating morally consequential decisions entirely to machines, as this undermines intention (*niyyah*), accountability, and human agency. AI systems may assist human judgement but should not replace it where ethical evaluation and moral responsibility are essential. Maintaining human control ensures that technology supports ethical reasoning rather than displacing it.

v. Promoting Ethical Economic Integration

In response to AI-driven automation and labour disruption, policymakers and institutions should adopt economic strategies that protect the dignity of work and promote distributive justice. Ethical deployment of AI requires proactive investment in reskilling, education, and social safety mechanisms that enable workers to adapt to technological change. Wealth generated through AI systems should be distributed in ways that serve public welfare (*maṣlahah*) rather than concentrating power and resources among elites. Islamic economic principles

support policies that prioritise collective benefit, productive use of technology, and protection of vulnerable populations.

vi. Institutionalizing Interdisciplinary and Religious Ethical Oversight

Effective ethical governance of AI requires sustained collaboration between technologists, ethicists, legal experts, and religious scholars. Engaging Islamic scholars in AI governance processes can help ensure that technologies are evaluated within relevant moral and cultural contexts, particularly in Muslim-majority societies. This interdisciplinary approach supports pluralistic ethical reasoning while maintaining normative coherence rooted in Islamic moral philosophy. Ethical review bodies and advisory councils should be empowered to guide AI policy, regulation, and innovation in alignment with shared moral commitments.

Enhancing Ethical Literacy and Public Awareness

Fostering ethical awareness and critical engagement with AI is essential for preserving moral responsibility in increasingly automated societies. Educational initiatives should promote AI literacy that enables individuals and institutions to understand the ethical implications of intelligent systems rather than engaging with them passively. Islamic tradition places strong emphasis on knowledge, reflection, and moral reasoning; these values should inform educational strategies that empower communities to evaluate technology ethically and responsibly.

Conclusion

The rapid expansion of Artificial Intelligence has intensified a central ethical tension of the contemporary era: technologies capable of extraordinary efficiency and innovation are simultaneously capable of reproducing injustice, eroding privacy, weakening accountability, and displacing forms of human judgement that are ethically and spiritually significant. As this study has shown, the ethical challenges associated with AI, especially algorithmic bias, surveillance and data exploitation, labour disruption, opaque autonomy, and the dehumanising reduction of persons to predictive profiles, are not marginal side effects of technical progress. Rather, they reflect structural risks that emerge when powerful systems operate without robust moral constraints and meaningful human oversight (Bostrom et al., 2014; Russell, 2019). The scale and speed of AI development, often outpacing regulatory and institutional capacity, further underline the urgency of grounded ethical evaluation capable of addressing both systemic harms and everyday applications (Anderson et al., 2007; Marcus et al., 2019).

In response to this urgency, the paper set out to contribute to AI ethics by engaging Islamic moral philosophy not as a symbolic complement to secular debates, but as a coherent normative framework for judgement. The analysis first

examined major ethical challenges posed by contemporary AI, highlighting how AI systems can institutionalise bias and inequality, enable pervasive surveillance infrastructures, restructure labour in ways that threaten distributive justice and the dignity of work, and introduce accountability gaps through opaque and autonomous decision-making (Floridi & Luciano, 2013; Zuboff, 2019; Brynjolfsson & McAfee, 2014). These challenges become particularly acute in religious and cultural contexts where trust, privacy, moral authority, and human responsibility are treated as ethical duties rather than optional values (Nuha, 2025; Waseu, 2025).

Building on this diagnosis, the paper articulated key Islamic ethical principles relevant to technological development and human responsibility. Concepts such as *tawhīd*, *khilāfah*, *amānah*, *‘adl*, *ihsān*, *maṣlaḥah*, *maqāṣid al-sharī‘ah*, and *karāmah insāniyyah* were presented as normative criteria that define moral authority, require justice and public welfare, and protect human dignity and agency. These principles offer a structured evaluative vocabulary capable of assessing both the social consequences of AI systems and the moral responsibilities of the humans who design, deploy, and govern them (Hallaq, 2009; Kamali, 2003; Auda, 2008; Sachedina, 2009). In this sense, Islamic ethics contributes not by merely affirming that Islam supports “good values,” but by clarifying what ethical legitimacy requires when technological power expands beyond the boundaries of human comprehension or control.

The third objective of the study applied these Islamic principles as evaluative tools to judge the moral permissibility, limitations, and conditions governing contemporary AI practices. The analysis demonstrated that Islamic ethics supports the use of beneficial technologies while imposing clear constraints against practices that violate justice, dignity, trust, and accountability. Algorithmic bias and discriminatory outcomes are ethically impermissible unless reformed to satisfy justice (*‘adl*) and equity, while surveillance practices lacking consent and necessity violate Islamic privacy norms and the moral trust embedded in *amānah* (Qur’ān 5:8; 49:12; Kamali, 2003). Autonomous weapons systems raise severe ethical objections under *maqāṣid*-based reasoning, given the sanctity of life and the requirement of moral accountability in decisions involving lethal force (Auda, 2008; Hashmi & Sohail, 2002). Similarly, AI-driven labour disruption requires ethical governance that preserves the dignity of work and ensures distributive justice, aligning technological innovation with *maṣlaḥah* rather than concentrated accumulation (Brynjolfsson & McAfee, 2014; Eubanks, 2018). Across these dilemmas, the paper reaffirmed that ethically consequential domains must retain meaningful human oversight, since moral agency, intention, and accountability cannot be delegated to machines without undermining foundational Islamic commitments to responsibility and human dignity (Nasr, 2006; Ramadan, 2009).

The recommendations proposed in this study translate these ethical evaluations into practical guidance for AI development and governance. They emphasise embedding accountability mechanisms in design and deployment, strengthening transparency and contestability, protecting privacy through purpose limitation and ethical data stewardship, safeguarding labour dignity through reskilling and just distribution, institutionalising interdisciplinary oversight including engagement with Islamic scholars, and expanding ethical literacy so that communities can critically engage AI rather than submit to it passively.

This study is conceptual and normative rather than empirical, and its contribution lies in clarifying evaluative criteria rather than measuring real-world outcomes. Future research could extend this framework through empirical case studies, particularly in Muslim-majority contexts, examining how AI systems are currently deployed in governance, education, finance, healthcare, and religious guidance, and assessing the extent to which ethical safeguards align with Islamic normative standards. Nevertheless, the present analysis establishes a foundational argument: as AI reshapes social life at scale, ethically responsible governance cannot remain exclusively within secular or technocratic paradigms. Islamic moral philosophy provides a rigorous, culturally grounded, and socially oriented ethical framework capable of evaluating AI practices, setting moral limits, and guiding the pursuit of innovation toward justice, dignity, and human well-being.

References

- Abou El Fadl, K. (2014). *Reasoning with God: Reclaiming shari'ah in the modern age*. Rowman & Littlefield.
- Ajizah, R. U. N., Su'aidi, Z., & Huda, M. (2025). Artificial intelligence in Islamic studies and academic ethics: Perspectives on development and implementation based on Islamic values. *Tarbawi Ngabar*, 6(1), 147–168. <https://doi.org/10.55380/tarbawi.v6i1.985>
- Alamsyah, M. N., Yayuli, & Remanda, A. (2025). Ethics of the use of artificial intelligence (AI) in the paradigm of Islamic law. *Solo International Collaboration and Publication of Social Sciences and Humanities*, 4(1), 91–104. <https://doi.org/10.61455/sicopus.v4i01.393>
- Anderson, M., & Anderson, S. L. (2007). *Machine ethics: Creating an ethical intelligent agent*. Oxford University Press.
- Arif, M., Al Irfan, S., Namanyane, T., & Ali, W. (2025). AI and machine learning in Islamic guidance: Opportunities, ethical considerations, and future directions. *Guidance: Jurnal Bimbingan Dan Konseling*, 21(2), 429–451. <https://doi.org/10.34005/guidance.v21i02.4449>
- Auda, J. (2008). *Maqasid al-shariah as philosophy of Islamic law: A systems approach*. International Institute of Islamic Thought.
- Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In K. Frankish & W. M. Ramsey (Eds.), *The Cambridge handbook of artificial intelligence*. Cambridge University Press.
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

- El-Hady, E. H. F., & Zenrif, M. F. (2024). Pandangan Islam terhadap etika kecerdasan buatan (artificial intelligence) dalam kehidupan sehari-hari. *Nuansa*, 21(2), 84–98. <https://doi.org/10.19105/nuansa.v21i2.16613>
- Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology*, 36. <https://doi.org/10.1007/s13347-023-00668-x>
- Esposito, J. L., & Voll, J. O. (1996). *Islam and democracy*. Oxford University Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Floridi, L. (2013). *The ethics of information*. Oxford University Press.
- Ghaly, M. (2023). What makes work “good” in the age of artificial intelligence (AI)? Islamic perspectives on AI-mediated work ethics. *The Journal of Ethics*. <https://doi.org/10.1007/s10892-023-09456-3>
- Hallaq, W. B. (2009). *Shari'a: Theory, practice, transformations*. Cambridge University Press.
- Hasan, A. (Trans.). (1984). *Abu Dawud*. Sh. Muhammad Ashraf.
- Hashmi, S. H. (Ed.). (2002). *Islamic political ethics: Civil society, pluralism, and conflict*. Princeton University Press.
- Kamali, M. H. (2003). *Principles of Islamic jurisprudence* (3rd ed.). Islamic Texts Society.
- Lusiana, M. H., Fathimah, E., & Daulay, W. I. H. (2024). Challenges and threats of artificial intelligence in maqasid sharia perspective. *KnE Social Sciences*. <https://doi.org/10.18502/kss.v9i2.15004>
- Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Vintage.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21.
- Mohadi, M., & Tarshany, Y. (2023). Maqasid al-shari'ah and the ethics of artificial intelligence. *Journal of Contemporary Maqasid Studies*. <https://doi.org/10.52100/jcms.v2i2.107>
- Nasr, S. H. (2006). *Islamic philosophy from its origin to the present: Philosophy in the land of prophecy*. State University of New York Press.
- Niam, M. F. (2024). Does artificial intelligence go beyond the limits of religious authority? An ethical review on IslamGPT. *AI' Adalah* (Jember), 27(1), 71–84. <https://doi.org/10.35719/aladalah.v27i1.477>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Nuha, A. A. (2025). Etika dan tantangan dakwah dengan artificial intelligence. *Dakwatuna: Jurnal Dakwah Dan Komunikasi Islam*, 11(1), 97–107. <https://doi.org/10.54471/dakwatuna.v11i1.3470>
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Ramadan, T. (2009). *Radical reform: Islamic ethics and liberation*. Oxford University Press.
- Raquib, A., Channa, B., Zubair, T., & Qadir, J. (2022). Islamic virtue-based ethics for artificial intelligence. *Discover Artificial Intelligence*, 2(1). <https://doi.org/10.1007/s44163-022-00028-2>
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- Sachedina, A. (2009). *Islam and the challenge of human rights*. Oxford University Press.
- Saputra, A., & Asbi, E. A. (2025). Islam, big data, and digital ethics: Challenges in religious life in the digital age. *Pharos Journal of Theology*, 106(4). <https://doi.org/10.46222/pharosjot.106.4022>
- Siddique, B., & Abdul Rauf, U. F. (2025). Artificial intelligence, bioethics, and Islamic law: A cross-disciplinary study for the modern age. *Journal of Applied Linguistics and TESOL*, 8(3), 78–88. <https://doi.org/10.63878/jalt1034>
- Tang, Q. W., & Kamarudin, S. (2025). A systematic literature review of theological reflections and institutional responses in religious engagements with AI ethics. In *Advances in computational intelligence and robotics book series* (pp. 353–384). <https://doi.org/10.4018/979-8-3373-2170-7.ch014>

-
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
- Waseu, I. (2025). Peran teknologi artificial intelligence (AI) dalam memperkuat dakwah Islam di era digital. *Dialogika : Jurnal Penelitian Komunikasi dan Sosialisasi*, 1(3), 257–265.
<https://doi.org/10.62383/dialogika.v1i3.561>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.