Islamic Ethical Perspectives on AI and Digital Transformation in the 21st Century

Uzma Waseem [a] and Abdur Rahim [b]

[a-b] Independent Writer and Researcher, Pakistan

Keywords: Artificial Intelligence, Islamic Ethics, Digital Transformation, AI Governance

Article Information			
Received	14th Oct 2025	Accepted	15th Dec 2025
Published	30th Dec 2025		

Abstract

The rapid advancement of artificial intelligence (AI) in the 21st century has spurred intense ethical debates worldwide, yet the prevailing discourse remains an ethics monoculture dominated by Western secular frameworks. This paper addresses that imbalance by proposing a novel Islamic AI ethics framework and providing a systematic mapping of Islamic principles to AI governance challenges. We demonstrate that Islamic ethics not only reinforces global principles like fairness and accountability but also introduces a critical dimension of spiritual accountability (amānah, the concept of divine trust) often absent from secular models. A qualitative, comparative methodology is applied: we review both mainstream AI ethics and Islamic ethical thought, then analyze four Muslim-majority country case studies (UAE, Saudi Arabia, Malaysia, Pakistan). Findings: Islamic ethical precepts show broad convergence with global AI ethics – for example, Islam's emphasis on justice aligns with fairness in algorithms, and the Islamic right to privacy complements data protection norms. More strikingly, Islamic ethics frames responsible AI as a moral duty and form of worship, embedding accountability before God as a motivator for ethical AI. The case studies illustrate how these nations integrate ethical AI guidelines with cultural and religious values – from the UAE's voluntary AI Ethics Principles to Saudi Arabia's Shariah-referenced data laws, Malaysia's new AI Office (NAIO) balancing multicultural norms, and Pakistan's National AI Policy aspiring to "ethical, inclusive" AI. The discussion highlights the potential of Islamic ethics to enrich the pluralistic dialogue on AI governance, ensuring technological innovation proceeds with moral responsibility, cultural inclusivity, and robust oversight (to prevent mere "ethics-washing"). We conclude with recommendations for policymakers – an Islamic AI governance framework – including the drafting of an Islamic AI Ethics Charter and the institutionalization of ethics oversight, suggesting that a cross-cultural approach to AI governance can promote globally beneficial, human-centric innovation while fulfilling the Islamic principle of stewardship (khilāfah).

Introduction

Artificial intelligence is revolutionizing economies and societies worldwide, raising profound ethical and governance questions. From algorithmic bias and privacy violations to job displacement and autonomous weapons, AI's risks are global in scope. In response, governments and organizations have advanced numerous AI ethics guidelines and even binding regulations. For instance, the European Union's Artificial Intelligence Act – the world's first comprehensive AI law – was passed in 2024. This EU AI Act adopts a risk-based approach, banning certain "unacceptable risk" AI uses and imposing strict requirements on high-risk systems. International bodies like UNESCO have also issued frameworks (e.g. the Recommendation on the Ethics of AI, 2021) to guide responsible AI development. These global efforts reflect a broad consensus on core principles such as transparency, accountability, fairness, and human safety in AI.

Despite these developments, the ethical discourse on AI has been predominantly Western and secular in orientation. Scholars note that Western theories and norms "overwhelmingly dominate" the AI ethics debate,

as most guidelines originate from Europe or North America. This has prompted calls for more diverse perspectives, including those grounded in non-Western cultural and religious traditions. In particular, there is growing interest in how Islamic ethics can contribute to the global conversation on AI governance. Islamic civilization has a rich ethical and legal heritage, and many of its core values – justice (al-ʿadl), mercy (ar-raḥmah), public interest (maṣlaḥa), human dignity (karāmah) – resonate with universal principles of good governance and social welfare. An Islamic ethical framework may offer not only a familiar moral compass for the 1.8 billion Muslims worldwide, but also novel insights and spiritual depth to augment secular approaches (e.g. a sense of higher accountability).

This paper investigates AI and digital transformation through an Islamic ethical lens, asking how Islamic moral principles can guide the development and use of AI in the 21st century. We aim to bridge the gap between global AI ethics and Islamic ethics, examining points of convergence and divergence. The study is timely as Muslim-majority countries are rapidly digitizing and deploying AI in governance, finance, education, and other sectors. Notably, nations like the United Arab Emirates and Saudi Arabia have launched ambitious AI strategies explicitly aligned with their cultural values, while Malaysia and Pakistan have recently set up national AI initiatives with ethical oversight components. These cases present an opportunity to observe Islamic ethics in action within AI policy and practice.

The structure of the paper is as follows. First, we review relevant literature: summarizing mainstream AI ethics principles and then delving into Islamic ethical concepts relevant to technology. Next, we outline our methodology for normative analysis and case study comparison. In the findings, we propose an Islamic ethical framework for AI, mapping key Islamic values to known AI ethical risks (e.g. bias, privacy, misuse of AI). We then present case studies of the UAE, Saudi Arabia, Malaysia, and Pakistan – highlighting each country's AI policies (such as the UAE's National AI Strategy 2031, Saudi Arabia's National Strategy for Data & AI 2020, Malaysia's NAIO, and Pakistan's National AI Policy 2025) and how they incorporate ethical principles rooted in Islam or local culture. In the Discussion, we consider the implications of integrating Islamic ethics into AI governance, including potential challenges (such as varying interpretations of Shariah or risks of superficial adoption) and the benefits of a pluralistic ethical framework. Finally, the paper offers a Conclusion with Recommendations for policymakers, suggesting ways to operationalize Islamic ethical guidelines in AI development and calling for greater cross-cultural collaboration in global AI governance.

Review of the Literature

- **Global AI Ethics: Principles and Frameworks**

By the early 2020s, a consensus had emerged around certain fundamental principles for ethical AI. Across dozens of guidelines proposed by governments, international organizations, and industry consortia, recurring themes include transparency, accountability, fairness/non-discrimination, privacy, human-centered design, and safety. These principles seek to ensure AI systems are understandable, do not unlawfully bias people, protect individual rights, and remain under meaningful human control. For example, the UAE's AI Ethics Principles (issued by its Ministry of AI) require Transparency – "AI systems and their decision-making processes [must be] understandable and accessible to users" – as well as Accountability and Fairness among others. Such guidelines often urge regular audits of AI systems and bias mitigation measures to uphold these values.

On the regulatory front, the EU AI Act 2024 stands out as a landmark legal framework. Adopted in 2024 and entering into force in 2025, it categorizes AI applications by risk level and bans those that pose "unacceptable risk" to fundamental rights. High-risk AI (e.g. in recruitment, credit, law enforcement) will face strict requirements on data quality, transparency, and human oversight, while lower-risk applications are largely

unregulated. This risk-based model is influencing policy worldwide. Tech superpowers like the US and China are also developing AI regulations (though with different approaches), and by 2023 more than a dozen countries had introduced national AI strategies focusing on "trustworthy AI" and governance. Internationally, the OECD AI Principles (2019) – endorsed by over 40 countries – and UNESCO's Recommendation on AI Ethics (2021) reflect a broad alignment on values such as promoting human-centered innovation, fairness, transparency, accountability, and sustainability in AI.

Notably, however, these global principles are often articulated in secular terms. They emphasize legal compliance and human rights, but typically do not invoke religious or spiritual dimensions of ethics. Some critics argue this makes the frameworks somewhat technocratic and lacking in motivational depth for certain cultures. As AI ethics matures, scholars and policymakers have started to explore how non-Western ethical systems can enrich the discourse. A pluralistic approach could address criticisms of a one-size-fits-all model and ensure ethical AI guidelines have legitimacy across different civilizations. In summary, global AI ethics provides a foundational set of norms – a "common language" of principles – that our analysis will compare and contrast with Islamic ethical perspectives.

- **Islamic Ethics and Technology: Key Principles**

Islamic ethics is rooted in the Sharīʿa (Islamic law and moral code), derived from the Qur'an and Hadith and elaborated by jurists over centuries. Unlike secular ethics which may separate law from morality, Islamic thought integrates them, emphasizing that following ethical principles is ultimately a form of worship and obedience to God. In the context of modern technology, Islamic scholars have begun interpreting classical principles to address new issues. Key concepts from the Islamic tradition relevant to AI include the Maqāṣid al-Sharīʿa (higher objectives of Islamic law), which prioritize the preservation of faith, life, intellect, lineage, and property (and by extension honor/dignity). These objectives can provide a values-based lens: for instance, any AI application that endangers life or intellect would be ethically suspect from a Maqāṣid perspective.

A central principle is justice (al-ʿadl) – the Qur'an commands "Be just; that is closer to piety" (Q.5:8). Justice in Islamic ethics entails impartiality and fairness in dealings. Applied to AI, this implies systems must not discriminate unfairly or entrench bias against any group. An AI hiring or loan algorithm should be audited for bias, as failing to treat individuals fairly violates al-ʿadl. Islamic teachings on justice would support measures like dataset bias mitigation and inclusive design, aligning with the global principle of fairness. In fact, in countries like Saudi Arabia and the UAE, "fairness" is explicitly stressed in AI frameworks, reflecting the Islamic commitment to justice for all ethnicities and faiths.

Another foundational value is benevolence and non-maleficence – doing good (iḥsān) and avoiding harm (lā ḍarar). The Prophet Muhammad (PBUH) taught, "There shall be no harm and no harassment (in Islam)". This translates to prioritizing beneficial AI uses and prohibiting clearly harmful ones. AI should ideally advance public welfare (maslaḥa) – e.g. using AI in healthcare to save lives or in education to spread knowledge reflects the Islamic value of mercy to mankind. Conversely, AI applications that cause injury or moral corruption (fasād), such as autonomous weapons targeting civilians or deepfake pornography (assaulting someone's dignity), would be categorically unethical in Islam. The principle of "no harm, no harassment" in Islamic law supports the common AI ethics maxim of "Do no harm," but with a moral force backed by religious accountability. Islamic ethics thus endorses strong caution against harmful AI, weighed through the lens of public interest (maṣlaḥa): an AI with minor harms but great benefits (e.g. a medical diagnostic AI that slightly impinges on privacy) might be permissible if it serves the greater good, whereas an AI with serious harms to life or morality should be curbed.

Responsibility and accountability (mas'ūliyyah) are also emphasized. In Islam, every individual is accountable for their actions before God and society; "Each of you is a shepherd and each of you is responsible for his flock," said the Prophet. This inculcates an ethos that developers and users of AI must take ownership of the impacts of AI systems. Deploying AI in a "value-neutral" way and then shrugging off harms as unintended side-effects is not acceptable. There is an ethical duty to foresee and prevent harm, and if harm occurs, to answer for it. Many global AI frameworks call for clear human accountability in AI (e.g. someone to be liable for an automated decision), and Islamic ethics reinforces this by viewing AI as part of one's entrusted responsibilities. The concept of amānah (trust) is particularly salient: the Qur'an (33:72) refers to a great "Trust" placed upon humanity, and in Islamic theology humans are khulafā' (stewards) of God's creation. Thus, developing or using AI is an amānah – a trust from God. Misusing it or neglecting its ethical implications is a breach of that trust. Concretely, an Islamic approach would advocate for auditability of AI (to trace decisions), liability for damage caused by AI, and independent ethical review (analogous to the historical ḥisbah institution that monitored marketplace morality). Both developers and officials are seen as stewards who must answer to the public and to God for how AI is implemented. This perspective adds a transcendent layer of accountability: even if an AI misstep escapes human punishment, one believes it cannot escape divine justice.

Truthfulness and transparency (ṣidq and wuḍūḥ) form another pillar. Honesty is a fundamental Islamic virtue; the Qur'an forbids concealing truth (Q.2:42). This underpins calls for transparency in AI – systems should be as explainable and open as possible. Users have a right to know when they are interacting with an AI (versus a human) and to understand how important decisions are made. "Black-box" AI that cannot be explained even to experts is problematic, especially if it affects individuals' rights. The Islamic mandate for clarity in dealings suggests that AI systems, particularly in governance or finance, should include explanatory features and avoid deception. For example, if an AI system denies someone social benefits, it should provide understandable reasons rather than an inscrutable output – aligning with both Islamic ethics and emerging AI regulations on explainability. Indeed, the UAE's ethical AI Toolkit and national principles call for transparency as a key tenet, which researchers note corresponds to the Islamic virtue of truthfulness. In sum, Islam would treat algorithmic transparency not just as a user right, but as part of moral honesty.

Beyond these principles, Islamic jurists also apply legal maxims (qawāʿid fiqhiyya) and specific rulings to new situations. For instance, the maxim "al-ḍarar yuzāl" (harm must be eliminated) provides a basis to forbid or regulate harmful AI. Another maxim, "al-yaqīn lā yazūlu bi-shakk" (certainty is not overruled by doubt), could inform requirements for reliable AI outputs – one should not act on AI decisions that are too uncertain when rights are at stake. While classical jurisprudence (fiqh) can guide on permissibility (halal/haram) of AI use cases (e.g., scholars issuing fatwas on AI in finance or medicine), some argue that a purely legalistic approach might miss broader ethical aspirations. For example, focusing only on whether a specific AI application is formally "allowed" under Shariah might overlook questions of intent and spiritual impact. Recent Islamic ethics scholarship therefore urges going beyond checklists of halal/haram toward a more holistic framework of values and trusteeship.

- **Synthesis: Gaps in Literature and the Trusteeship Framework**

While scholars like Elmahjub (2023) and Ali et al. (2025) have begun bridging global and Islamic AI ethics, a significant gap remains in empirically grounding these theoretical discussions. Much of the existing work is conceptual; there is a lack of comparative policy analysis examining how principles are operationalized at the national level in Muslim contexts. Our study aims to fill this gap by analyzing real-world AI strategies and policies in multiple countries. By doing so, we move beyond abstract theology or philosophy and assess practical implementation (or lack thereof) of ethical principles.

A particularly promising theoretical lens is Taha Abdurrahman's Iʾtimāniyya (trusteeship) model, introduced to AI ethics by Ali et al. (2025). This framework is built on the concept of divine trust (amānah) and articulates responsibilities in three spheres: towards God (spiritual integrity), towards knowledge (honesty and pursuit of truth), and towards creation (compassion and care for others and the environment). In essence, it reconceptualizes AI ethics as a form of trusteeship where humans are trustees of technology on God's behalf. Applying this model means that developing or deploying AI carries covenants: one must remain conscious of God (spiritual accountability), uphold truth and accuracy in data/algorithms (epistemic duty), and ensure AI serves humanity and nature rather than harms them (social/environmental duty). This trusteeship ethos provides an overarching framework that can integrate the various Islamic principles discussed above. It adds a transcendent motivation – viewing ethical AI as not just compliance with human laws or public opinion, but part of one's duty to the Creator. This approach could galvanize stronger ethical commitment, as violating AI ethics becomes akin to a moral and spiritual failure, not merely a legal one.

In our analysis, we elevate the Iʾtimāniyya model as a guiding framework. It is used to interpret findings and shape recommendations, emphasizing amānah (trust) and khalīfah (stewardship) as central tenets. By framing ethical AI governance as trusteeship, we underline that for Muslim-majority societies, ethical AI is more than a utilitarian choice – it is fulfilling a trust before God and humanity. This perspective is woven through our case studies and discussion, providing a cohesive theoretical foundation for our comparative approach.

Methodology

This research adopts a qualitative, interdisciplinary approach to analyze how Islamic ethics can inform AI governance in the modern world. The methodology comprises two main components:

Conceptual Analysis: We conducted a normative analysis of Islamic ethical sources in relation to identified AI ethical challenges. This involved reviewing scholarly literature (academic articles, religious edicts, policy papers) on Islamic perspectives of technology and AI ethics, as well as major global AI ethics frameworks for comparison. Key Islamic concepts (e.g. maqāṣid, maslaḥa, amānah) were examined and "mapped" to known AI risk domains (such as bias, privacy, autonomy, security). We employed principles of uṣūl al-fiqh (Islamic jurisprudential methodology) to ensure our derivations align with Islamic textual sources and recognized juristic methods. This mirrors approaches of recent scholars who use uṣūl al-fiqh to perform "value alignment" analysis for emerging technologies. Specifically, where the Qur'an and Hadith provide general guidance (e.g. on justice, honesty), we extrapolate those to AI contexts; where explicit guidance is lacking, we rely on analogical reasoning (qiyās) and consideration of public interest (istiṣlāḥ/maṣlaḥa) to reach ethical judgments. We also compare these Islamic principles with secular ethics frameworks to highlight overlaps or distinctive differences.

Case Studies: We selected four Muslim-majority countries – United Arab Emirates (UAE), Kingdom of Saudi Arabia (KSA), Malaysia, and Pakistan – for in-depth case studies. These countries were chosen due to their active engagement with AI at the national policy level and their diverse socio-religious contexts (Gulf Arab states with formal Islamic governance elements vs. Southeast Asian and South Asian states with more pluralistic settings). These four represent a spectrum of governance styles (from absolute monarchy to federal republic), levels of economic development (high-income to developing), and approaches to Islamic law in governance (from explicit Shariah implementation to more implicit influence), allowing for a rich comparative analysis. For each case, we gathered data from official strategy documents, government announcements, and secondary analyses regarding their AI initiatives. We used content analysis to identify references to ethical principles, values, or regulatory measures in these initiatives, with special attention to whether and how Islamic values are referenced or incorporated. For example, in the UAE and Saudi cases, we

examined national AI strategies and ethics guidelines for alignment with principles like justice (ʿadālah) or privacy (khuṣūṣiyyah). In Malaysia and Pakistan, we looked at the mandates of newly established AI governance bodies and policy statements for any culturally-specific ethical approaches. We also drew on scholarly critiques and media reports to assess each policy's strengths and gaps – for instance, expert commentary on Pakistan's AI Policy 2025 highlighting implementation issues.

Throughout, our analysis is comparative. We compare the ethical content of the case study policies with both global AI ethics standards and Islamic ethical norms, triangulating multiple sources (religious texts, academic work, policy documents, news reports) to strengthen our interpretations. All sources used are credible (peer-reviewed journals, official publications, reputable news), with an emphasis on recent materials (2022–2025) to capture the latest developments. In synthesizing findings, we aimed to formulate an integrative framework that policymakers and researchers can use as a reference for incorporating Islamic ethics into AI governance.

This methodology inevitably has limitations. The analysis is largely literature-based and qualitative; it does not include field surveys or interviews that might capture how Muslim publics or different scholars perceive AI ethics on the ground. Moreover, Islamic ethical interpretation is not monolithic – there are diverse viewpoints within the tradition that a broad study like this can only partially reflect. We acknowledge that our focus on dominant principles may not fully capture intra-Islamic debates or minority interpretations. Despite these caveats, the methodology provides a structured way to explore our research questions. By linking abstract principles to concrete policy examples, we generate insights that are both conceptually grounded and practically relevant.

Findings: An Islamic Ethical Framework for AI

Drawing together the literature insights and conceptual analysis, we propose an Islamic Ethical AI Framework that maps key Islamic values to major AI ethical risk domains. Table 1 (conceptually summarized here) illustrates how foundational Islamic ethical principles correspond to issues raised by AI. Overall, we find that Islamic ethics strongly complements global AI ethics, often reaching similar prescriptions via different reasoning. Below, we detail several core principles and how they address AI's challenges:

Justice and Fairness (al-ʿAdālah): In Islam, justice is paramount – "Be just; that is closer to piety" (Qur'an 5:8). In AI governance, this translates into a requirement for fairness in algorithms. AI systems should not discriminate or perpetuate bias against any group. The Islamic concept of justice entails impartiality and equal treatment of people of all backgrounds. Thus, an AI hiring tool or loan algorithm must be audited for bias; failing to treat individuals fairly would violate al-ʿadl. Islamic ethics would support measures to mitigate data bias and ensure inclusivity, aligning with the global principle of fairness. In practice, Saudi Arabia's and the UAE's AI frameworks both explicitly stress fairness, reflecting the Islamic commitment to justice for all ethnicities and faiths. Policies encourage that AI decisions be free of prejudice, resonating with the Quranic ideal that "no person is discriminated against except by their deeds." In practical terms, implementing this may involve algorithmic bias testing, using representative training data (including local populations, not just Western datasets), and offering redress mechanisms for unfair outcomes.

Beneficence and Non-Maleficence (Raḥmah & ʻAdam al-Ḍarar): Islamic ethics strongly promotes doing good (iḥsān) and mercy/compassion (raḥmah), while prohibiting causing harm (ḍarar) without just cause. A famous hadith states: "There shall be no harm and no harassment in Islam." For AI, this dual principle means prioritizing applications that benefit humanity and avoiding those that inflict physical or moral harm. AI should ideally be used to advance public welfare – for example, employing AI in healthcare to save lives or in education to spread knowledge – reflecting the Islamic value of mercy towards mankind. Conversely, AI use cases that are clearly harmful or corrupting – such as autonomous weapons targeting civilians, deepfake

pornography (which assaults personal dignity), or addictive algorithms exploiting human weaknesses – would be categorically unethical from an Islamic view. They violate the lā ḍarar (no harm) maxim and the duty of compassion. The principle of public interest (maṣlaḥa), as discussed earlier, guides weighing benefits and harms: an AI that entails minor harms but great benefits (e.g. a medical diagnostic AI that slightly intrudes on privacy) might be permitted if it serves the greater good, whereas one with serious harms to life or morality should be forbidden. Our analysis suggests Islamic ethics endorses the "Do no harm" principle common in AI ethics, but with a moral force backed by religious accountability. A Muslim AI engineer might thus view avoiding harmful design choices as part of answering to God for their work, not just a professional guideline.

Accountability and Trust (Mas'ūliyyah & Amānah): Islam teaches that everyone is answerable for their actions before God and society. "Each of you is a shepherd and each is responsible for his flock," said the Prophet Muhammad. This instills a sense of mas'ūliyyah (responsibility) that applies directly to AI developers, policymakers, and users – they must take ownership of the impacts of AI systems they create or deploy. In an Islamic framework, it is unacceptable to release AI in a moral vacuum and then shrug off harm as "unintended"; moral (and possibly legal) accountability is expected for AI outcomes. Accordingly, the UAE and Saudi ethical guidelines both emphasize accountability – clearly assigning who is answerable for an AI system's behavior. Moreover, the concept of amānah (trust, or sacred responsibility) reinforces this: AI is seen as a trust from God placed in human hands, so misuse of AI is a breach of that trust. In practical terms, Islamic ethics would advocate for strong accountability measures in AI governance: audit trails to trace AI decisions, liability frameworks for damages caused by AI, and ethical review processes (akin to a hisbah) to monitor AI projects. Both AI developers and government officials are viewed as stewards (khalīfa) who must answer to the public and ultimately to God for how AI is implemented. This principle supports robust accountability mechanisms like algorithmic impact assessments and oversight boards. It also connects to the trusteeship concept – decision-makers are trustees who must uphold the trust of safeguarding society's welfare when deploying AI.

Truthfulness and Transparency (Ṣidq & Wuḍūḥ): Honesty is a fundamental Islamic virtue; the Qur'an warns against hiding the truth or deceiving others. Applied to AI, this underpins the principle of transparency – AI operations should be as explainable and open as possible. Users have a right to know when they are interacting with an AI (rather than a human), and to understand the basis of significant algorithmic decisions. Black-box AI systems, especially those affecting individual rights (e.g. in criminal justice or finance), are problematic in Islam if they result in judgments without clarity or accountability. The Islamic mandate for clarity (wuḍūḥ) in transactions suggests AI systems, particularly in governance, should include explanatory features and avoid misleading users. For example, an AI system used to determine access to benefits or medical treatment should provide human-understandable reasons for its decisions, rather than an inscrutable output – this aligns with both Islamic ethics and secular calls for AI explainability. The UAE's Ethical AI Toolkit explicitly urges transparency and explainability, which scholars Gorian & Osman (2024) show correspond to the Islamic values of ṣidq (truthfulness) and trustworthiness. In sum, ṣidq implies that AI systems must not be instruments of fraud or misinformation; transparency becomes an ethical duty to uphold truth and trust.

Human Dignity and Privacy (Karāmah & Ḥurmah): Islamic teachings hold human dignity (karāmah) as inviolable, and the Qur'an and Hadith establish privacy as a right – for example, condemning spying or unwarranted intrusion into others' private lives. The concept of ḥurmat al-bayt (sanctity of the home) and ḥifẓ al-ʿirḍ (preservation of honor) indicate that personal data and one's private affairs should be protected from unnecessary exposure. In AI terms, this reinforces principles of privacy, data protection, and consent. Surveillance technologies or AI-driven data mining are ethically suspect if they violate the sphere of privacy without compelling justification (such as legitimate security needs proportionately addressed). Islam would

encourage strong data governance: limiting data collection to what is needed (data minimization) and preventing misuse or public exposure of personal information. Notably, Islamic privacy norms view certain infringements not just as legal violations but as sins unless justified, thus adding weight to protecting privacy. We observe that Saudi Arabia's data protection law, for instance, aligns with European GDPR standards but was also framed in terms of protecting personal privacy as an Islamic value. The elevation of privacy to a near-sacred status (a God-given right to personal sanctity) is a unique contribution Islamic ethics can make to AI governance debates, especially as AI enables mass surveillance and data aggregation. We will later see the tension this creates in cases like the UAE and KSA, which pursue security through AI surveillance while upholding privacy in principle.

Sharīʿa Compliance and Public Morality: Many Islamic discussions on technology focus on ensuring that new innovations do not facilitate what is religiously forbidden (ḥarām) or undermine moral norms. This involves applying specific rulings or seeking fatwas on issues like AI in finance (e.g., algorithmic trading avoiding usury), AI in art (avoiding idolatrous depictions), or AI in warfare (complying with Islamic laws of war). The maxim "al-ḥalāl bayyin wal-ḥarām bayyin" (permissible matters are clear and forbidden matters are clear) is sometimes invoked – if an AI application clearly leads to prohibited outcomes (like gambling or pornography), it should be disallowed. Islamic governments may explicitly ban certain AI uses on these grounds (for example, hypothetical bans on generative AI creating images that could be seen as blasphemous). We note that this can diverge from secular ethics which might allow something provided it does no harm in a worldly sense, whereas Islamic ethics could prohibit it due to spiritual/moral harm. However, this aspect is generally complementary to global principles – e.g., banning deepfake pornography would be supported by both Islamic morality and secular concerns about harassment and consent. The challenge is to avoid a shallow "checkbox" compliance ("Sharīʿa-compliant AI" in name only) and instead integrate the spirit of Islamic morality – promoting good and forbidding evil – in AI design. One emerging model, as noted, is trusteeship ethics, which goes beyond narrow legalism to infuse spiritual objectives into tech ethics. Our framework thus views Sharīʿa compliance not just as abiding by literal rulings, but fulfilling the higher objectives (maqāṣid) – such as justice, compassion, and protection of the vulnerable – through AI.

In summary, the proposed framework aligns closely with widely accepted AI ethics principles (fairness, "do no harm," accountability, transparency, privacy, etc.), but augments them with Islamic context and conviction. It adds layers of meaning – e.g. fairness is not just social equity, it's a divine command; privacy is not just a human right, it's part of one's God-given dignity; accountability is not only to regulators, but to a higher Judge. These added dimensions can motivate ethical conduct even when laws are weak or absent, by tapping into spiritual conscience. Table 1 (omitted here for brevity) maps specific Islamic concepts to corresponding AI ethics principles and risk areas, illustrating this convergence.

With this framework in mind, we now turn to how it manifests (or fails to manifest) in our case study countries. Each case provides a real-world snapshot of Islamic ethics interacting with AI policy. We examine each country's AI strategy and policies, the explicit or implicit ethical commitments therein, how (if at all) they link to Islamic values, and what strengths and gaps are apparent.

Case Studies

Each case study is organized under a common structure – National Strategy & Key Policies, Explicit/Implicit Ethical Commitments, Linkage to Islamic Values, and Strengths and Critical Gaps – to facilitate comparison.

❖ **United Arab Emirates (UAE)**

National Strategy & Key Policies: The UAE has positioned itself as a global leader in AI innovation, while also articulating a commitment to ethical AI consistent with its cultural and religious context. In October 2017, the UAE made headlines by appointing the world's first Minister of State for Artificial Intelligence, H.E. Omar bin Sultan Al Olama (at age 27), signaling that AI is a national priority. Soon after, the government launched the UAE National Strategy for AI 2031, a comprehensive roadmap to make the Emirates a global AI hub by 2031. The strategy outlines objectives such as developing local AI talent, attracting investment, and integrating AI across sectors like healthcare, education, transport, and government services. Institutionally, the UAE established bodies like the UAE Council for Artificial Intelligence and Blockchain (formed 2018) to coordinate AI policy across sectors. A specialized university, Mohamed bin Zayed University of AI (MBZUAI), opened in 2020 to boost R&D and has programs touching on AI policy and ethics. These initiatives illustrate a top-down, proactive approach to AI governance, framing the UAE as both a tech-forward and values-conscious state.

Explicit Ethical Commitments: As part of the National AI Strategy, the UAE government released a set of UAE AI Ethics Principles and Guidelines in 2019 (updated 2022). These guidelines form a voluntary ethical framework for AI development and deployment in the UAE. They enumerate eight key principles for AI: fairness, accountability, transparency, explainability, safety & security, human-centric design, sustainability, and privacy. Each principle is accompanied by recommended practices – e.g. under fairness, curbing algorithmic bias; under privacy, data minimization and protection by design. Notably, these Ethics Guidelines align their goals with both the UAE's national strategy and the UN Sustainable Development Goals, blending local and universal ethical commitments. The guidelines themselves are phrased in general, secular language – appropriate for an international business hub – and participation is voluntary (they are not enforced by law). However, the government actively promotes them through toolkits and encourages organizations to adopt them. For example, Digital Dubai issued an Ethical AI Toolkit providing checklists to help companies comply with the principles. The UAE even proposed an "AI Seal" quality mark to certify products as ethical AI, as an incentive for compliance. All these steps indicate an explicit commitment to ethical AI, albeit via soft governance. A critical observation is that without legal mandates, the risk of "ethics-washing" exists – i.e. companies might espouse the principles publicly but not rigorously implement them. We will discuss enforcement issues shortly.

Linkage to Islamic Values: Although the UAE's Ethics Guidelines do not overtly cite Islam (likely to maintain international and multi-cultural appeal), the influence of Islamic values is discernible upon analysis. A 2024 study by Gorian & Osman found a clear alignment between the UAE's eight AI principles and Islamic ethical teachings. For instance, the guideline of fairness corresponds to al-ʿadl (justice); accountability corresponds to mas'ūliyyah; transparency and honesty correspond to ṣidq (truthfulness) and wuḍūḥ (clarity); privacy corresponds to the Islamic injunction of ḥifẓ al-ʿawrah / iḥtirām al-khuṣūṣiyyah (protection of privacy and honor); a human-centric approach reflects raḥmah (compassion); and the notion of sustainability and responsibility toward environment echoes khilāfat al-arḍ (stewardship of the earth). The fact that an external scholarly article could map each principle to Quranic or Hadith values suggests the guideline framers (likely consulting local cultural experts) intentionally ensured consistency with Islamic ethics. Indeed, UAE officials often emphasize that the nation's tech innovation is guided by values like "community happiness" and public welfare, concepts resonating with the Islamic principle of maṣlaḥa (common good). In practice, while the policy documents use largely secular terminology, one can view the ethical AI push in the UAE as implicitly grounded in long-standing cultural and religious norms of justice, mercy, and trustworthiness.

Strengths and Critical Gaps: The UAE's approach shows strength in proactiveness and integration – it was among the first countries globally to create a dedicated AI ministry and strategy, and to publish national AI ethics guidelines. It successfully portrays ethical AI as part of its brand, combining modernization with moral responsibility. The voluntary guidelines allow flexibility and rapid iteration, and the alignment with both Islamic values and global standards makes them broadly palatable. Additionally, institutional support (AI Council, AI Office, toolkits) provides resources to implement ethics in practice. However, a critical gap is the lack of enforcement mechanisms. The ethics principles are not laws; adoption by industry is encouraged but not mandatory. This raises the risk of "ethics-washing" – i.e., companies or agencies might pay lip service to principles without robust compliance. There is, for example, no independent ethics regulator to audit AI projects in the private sector for bias or privacy violations. Another challenge is reconciling the ethics framework with some on-ground practices. Privacy is a principle in the guidelines, yet the UAE is known for extensive use of surveillance technologies (CCTV networks, facial recognition in "smart cities", etc.). Critical question: How does the UAE reconcile its expanding AI-driven surveillance (justified for security and smart governance) with the strong Islamic and ethical commitment to privacy and dignity? This tension – between public security interests (maslaḥa ʿāmma) and individual privacy rights – is something the UAE will need to navigate carefully. So far, the narrative is that surveillance is used for public safety and service improvement, not to intrude unjustly, but concrete oversight (like audits to prevent abuse) is not transparent. In summary, the UAE case demonstrates a blend of explicit ethics initiatives and implicit Islamic alignment, but it highlights the importance of moving from voluntary guidelines to accountable governance to avoid principles remaining only on paper.

- **Kingdom of Saudi Arabia (KSA)**

National Strategy & Key Policies: Saudi Arabia, as the birthplace of Islam and a G20 economy, offers a unique case of AI ethics grounded in Islamic values alongside a rapid modernization drive (Vision 2030). In 2020, Saudi Arabia formally adopted the National Strategy for Data and AI (NSDAI), aiming to make the Kingdom one of the top AI-powered nations by 2030. The NSDAI, unveiled by the Saudi Data & AI Authority (SDAIA), set ambitious targets: e.g., to attract $20 billion in AI investment, train 20,000 AI specialists, and spawn hundreds of AI startups, while applying AI across government and industry. More pertinent to our focus, the NSDAI emphasized that Saudi Arabia will pursue AI "ethically and responsibly to ensure public trust." One initiative under the strategy was to develop official AI Ethics Principles for the country.

SDAIA delivered on this by issuing Saudi Arabia's AI Ethics Principles (Version 2.0) in September 2023. These principles closely mirror global norms but also reflect local priorities. They cover fairness, transparency, accountability, privacy, security, safety, reliability, interpretability, and human-centered values. (Notably, "integrity" is mentioned, tying to honesty – an Islamic virtue). The principles were endorsed at high levels of government, signaling intent to guide AI development ethically. However, like the UAE's, they are guidelines rather than hard law, and primarily directed at government ministries and state-controlled projects. Alongside SDAIA, Saudi Arabia established a multi-tier governance framework: SDAIA as the central data/AI authority (with an explicit mandate to ensure AI aligns with national values), a National Center for AI (NCAI) focusing on research and policy including ethics training, and a Digital Government Authority (DGA) contributing to AI ethics and data regulations. This indicates a robust institutional setup to integrate ethics into AI governance.

Explicit Ethical Commitments: The NSDAI itself includes ethical principles in its text, and SDAIA's AI Ethics Principles 2.0 is the flagship ethical commitment. The content of these principles largely overlaps with universal AI ethics: e.g., fairness/non-discrimination, transparency & explainability, privacy & data protection, security & reliability of AI systems, safety (preventing misuse or accidents), and accountability

(assigning responsibility for AI outcomes). What differentiates Saudi's approach is the explicit linking of these principles to public trust and to its legal context of Shariah. For instance, official statements highlight that AI will be used in ways that "uphold Islamic ethics and serve humanity", and not in ways that contradict religious tenets. The strategy frames ethical AI as essential for societal acceptance of AI, which in Saudi's context means alignment with Islamic morals. Concrete policy measures include Saudi's Personal Data Protection Law (updated 2022) which, akin to GDPR, protects privacy but was promoted as in line with Islamic principles of not disclosing private affairs without permission. Another example is the use of the concept of integrity in AI systems – interpreted not just as data integrity but moral integrity (no deceit), reflecting Islamic expectations of honesty. The presence of an official ethics document (the SDAIA principles) is a strong commitment, though enforcement relies on SDAIA's authority. The principles are currently non-binding, but SDAIA as a central agency can insist that government and possibly large private projects adhere to them (especially in sectors of national importance).

Linkage to Islamic Values: Saudi Arabia's integration of Islamic ethics in AI is explicit. As an Islamic state with Shariah as the foundation of its legal system, there is a requirement (or at least an expectation) that new laws and policies not conflict with Shariah. In practice, Saudi officials and scholars frequently reference Islamic principles when discussing technology. For example, privacy is often justified by citing Islamic teachings about the inviolability of private life. A comparative study noted that Saudi Arabia's approach to AI ethics weaves Islamic privacy norms into cyber governance, stressing human dignity and personal boundaries taught by Islam. The Saudi data law, as mentioned, reflects Quranic values (e.g., the idea that one should not pry into others' affairs). Furthermore, Saudi Arabia's Vision 2030 and AI initiatives are often couched in terms of promoting welfare (maṣlaḥa) and preventing harm (darar), classical Islamic governance aims. At the 2020 Global AI Summit in Riyadh, there were panels on "AI Ethics from an Islamic perspective," and SDAIA's public communications frequently pair "trustworthy AI" with "Islamic values". Additionally, because Saudi's judicial and advisory bodies include religious scholars (ulema), any ethically contentious AI application could be reviewed; while we haven't seen major fatwas on AI yet, the tacit approval by the religious establishment is implied when the government advances initiatives (no senior scholars have objected to AI so far). The net effect is that Islamic ethics is part of the narrative and implementation: whether it's AI for Hajj management (framed as facilitating a pillar of Islam with technology in a way that embodies compassion and care for life), or AI surveillance for security (framed as preserving public order, a Shariah objective, albeit raising privacy questions). Every new tech policy in Saudi must pass a sort of cultural/religious filter – sometimes this is explicit (e.g., calling something "halal AI" informally), other times implicit.

Strengths and Critical Gaps: Saudi Arabia's strengths lie in top-down coherence and cultural legitimacy. With strong state control, when the leadership commits to ethical AI, it can swiftly issue guidelines and implement them in government projects. The alignment with Islamic values gives these guidelines societal resonance – they are not seen as foreign impositions but as continuous with local norms. The country's significant investment in AI (financial and institutional) means ethics can be embedded from the start (e.g., requiring all government AI tenders to follow SDAIA principles). Another strength is the existence of a legal framework that can support ethics (data protection law, cybercrime law, etc.) combined with the moral framework of Shariah. However, critical gaps and questions do exist. One is the breadth of implementation: outside government, how much will private sector actors comply with the voluntary ethics principles? If a foreign tech company operates in Saudi, will it be held to these standards? Currently, enforcement beyond moral suasion is unclear. Another gap is potential tension between ethics and authority. For example, the government's heavy investments in AI for smart surveillance (e.g., NEOM's city-wide monitoring systems) are officially to improve security and services. But there is an ethical balance to strike, given Islamic values

on privacy and moderation. Here arises a pointed question: Is the reference to Shariah in Saudi's AI ethics a deep ethical commitment or could it become a tool for political control? In other words, there is a risk that authorities might cherry-pick Islamic concepts to justify invasive practices – for instance, citing "preventing harm" (darʾ al-mafāsid) to override privacy concerns wholesale. This could lead to a situation where ethical rhetoric masks authoritarian surveillance ("ethics-washing" in a different sense). The maslaḥa (public interest) vs. individual rights debate is real: a government might argue extensive AI surveillance is Islamically justified for public security (a collective benefit), while critics might invoke Islamic protections of personal privacy and due process to oppose unchecked monitoring. To maintain credibility, Saudi's framework will need transparent engagement with such debates, perhaps by involving a range of religious scholars and ethicists in oversight (to avoid a single narrative). In summary, Saudi Arabia demonstrates how a country can leverage religious principles as a source of strength and legitimacy in AI ethics – the ethics aren't an external add-on but rooted in national identity. The challenge ahead is ensuring these principles are applied consistently and not selectively, and that ethics oversight keeps pace with ambitious AI deployment, especially in areas that test the balance between collective good and individual rights.

- **Malaysia**

National Strategy & Key Policies: Malaysia is an ethnically diverse, Muslim-majority constitutional monarchy in Southeast Asia that is increasingly active in AI policy. The federal government has expressed ambitions to make Malaysia a regional AI hub while ensuring "Malaysian values" are upheld. In late 2024, Malaysia took a significant step by launching the National AI Office (NAIO), indicating a move toward a more centralized and strategic approach to AI governance. On December 12, 2024, Prime Minister Anwar Ibrahim formally announced the establishment of NAIO, calling it "another historical moment in our digital transformation" and positioning it under the Ministry of Communications and Digital. The NAIO serves as the central authority for AI policy and regulation in Malaysia, tasked with strategic planning, coordinating AI R&D, and providing regulatory oversight of AI development. The motivation for NAIO was to unify various AI initiatives that had been spread across different agencies (like the Ministry of Science, Technology and Innovation, MOSTI) under one umbrella.

At its launch, NAIO outlined seven key deliverables for its first year, notably including the development of a national AI ethics code and an AI regulatory framework, as well as a five-year action plan through 2030. The explicit mention of an AI ethics code from the outset shows that Malaysia recognizes the importance of guiding AI use with clear ethical standards from the beginning. As of this writing (late 2025), Malaysia is expected to roll out these guidelines, which will likely draw from global best practices (Malaysia often references OECD or IEEE AI guidelines) but be tailored to Malaysian cultural norms. In terms of broader policy, Malaysia had earlier produced a National AI Roadmap (2021–2025) and the Malaysia Digital Economy Blueprint (MyDIGITAL, 2021). These documents set the stage by highlighting the need for "Responsible AI" and including principles like trust and security in digital strategy. While the 2021 AI Roadmap did not explicitly mention Islam, it framed ethics in multicultural terms, reflecting Malaysia's multireligious society.

Explicit/Implicit Ethical Commitments: Malaysia's AI governance is still in formative stages, but some commitments are evident. The forthcoming AI Ethics Code under NAIO is the key explicit commitment. Based on Malaysia's statements, we anticipate it will incorporate elements like fairness, transparency, accountability, and perhaps "Malaysian values" of harmony and respect. Even before the code, Malaysia's officials have signaled an intention to ensure AI is inclusive and human-centric, echoing both Islamic and broader values. For example, when partnering with tech giants (e.g., Google, Microsoft) for AI initiatives, Malaysia has emphasized AI that benefits society and "leaves no one behind," which implies an ethical stance (avoid exacerbating inequality). Malaysia's Communications and Multimedia Ministry already enforces

content standards (e.g., not allowing content that offends religious sentiments), which, while not AI-specific, shows a precedent of intertwining ethics/culture with tech regulation. Implicitly, Malaysia references frameworks like the OECD AI Principles and has indicated alignment with international standards. It also often uses the term "Asian values" alongside Islamic values, suggesting an approach that blends religious and regional cultural ethics. A notable initiative is that NAIO is part of (and reports to) the MyDIGITAL blueprint oversight, meaning ethical AI is being integrated into the country's broader digital transformation agenda, not treated as an afterthought.

Linkage to Islamic Values: Malaysia is officially an Islamic federation (Islam is the religion of the Federation per the constitution) but practices a moderate form of governance with freedom of religion and a significant non-Muslim minority. Thus, its approach to AI ethics tends to be inclusive and multicultural, even as it draws on Islamic principles. Culturally, Islamic values influence laws on morality (censorship of obscene content, etc.), finance (Malaysia is a pioneer in Islamic banking), and education. For instance, Malaysia's concern for "cyber wellness" and fighting hate speech has at times invoked both "Asian values" and Islamic values of maintaining public morality and harmony. In the AI context, Malaysia's emphasis on human-centric development resonates with Islamic ideas of human dignity and communal harmony (e.g., the concept of ummah or social unity). It is likely that as NAIO drafts the AI ethics code, it will consult Islamic scholars or at least consider perspectives of Shariah research institutions. Indeed, Malaysia has bodies like JAKIM (the Islamic Development Department) and various Islamic universities; these could contribute thoughts on AI usage in areas like finance (ensuring AI doesn't facilitate usury or fraud), the halal industry (AI in food or products complying with halal standards), or general matters of public morality. One unique domain is Islamic finance and fintech: Malaysia, being a leader here, is already integrating AI in these sectors (e.g., AI chatbots for Islamic banks, AI to optimize zakat distribution). This has prompted discussions on how AI can comply with Shariah finance rules – for example, ensuring an AI-driven loan approval system does not inadvertently violate the prohibition of riba (usury), or that automated investment advisors adhere to Islamic investment filters. Malaysia's Central Bank (Bank Negara) has Shariah advisory boards; these are likely to weigh in on AI applications in banking to ensure they meet Islamic ethics. Another example: if AI is used for surveillance or law enforcement, considerations from Islamic jurisprudence (such as maintaining privacy, presumption of innocence, avoiding entrapment) might be raised by civil society or religious voices. However, given Malaysia's pluralism, any ethical framing in official policy tends to stress universal values common to all religions (justice, benevolence, etc.), rather than overtly Islamic terminology, so as not to alienate non-Muslim citizens. In summary, Islamic values are an undercurrent in Malaysia's approach – present via the moral fabric and existing regulatory ethos – but balanced within a multicultural narrative.

Strengths and Critical Gaps: Malaysia's emerging approach has the strength of deliberation and balance. By setting up NAIO and planning an ethics code from the start, Malaysia is proactively embedding ethics into AI governance. The country can draw on both international models and its own cultural values, potentially creating a hybrid that is locally legitimated. The multi-stakeholder environment (government, industry, academia, religious bodies) could yield a well-rounded ethical framework if managed inclusively (akin to the concept of shūrā – consultation – which is an Islamic governance principle encouraging broad input). Malaysia also benefits from observing the experiences of early adopters like UAE and KSA, learning what to emulate (or avoid). For instance, it has noted the need for a dedicated authority (hence NAIO) and for aligning with global standards while infusing local norms. A further strength is Malaysia's relatively free press and civil society, which means there can be public discourse and critique of AI initiatives (this can act as a check against ethics-washing or authoritarian uses of AI). However, there are challenges and open questions. One gap is the delay in concrete guidelines – until the ethics code is published and enforced, there is a policy vacuum that fast-moving tech companies could exploit. Also, as a democracy with a developing economy,

Malaysia might face capacity issues: implementing AI ethics will require skilled personnel, awareness in industry, and inter-agency coordination. Will NAIO have sufficient authority and resources to enforce ethical practices across both government and private sector AI projects? Another challenge is balancing Islamic and secular viewpoints. For example, if Malaysia leans too much into Islamic rhetoric for AI ethics, it risks discomfort among non-Muslims; if it leans too far into generic terms, it might not fully leverage the motivational power of Islamic ethics for its Muslim majority. The likely path is moderate: use universal values that overlap with Islamic ones (which is effective but might not fully tap into, say, the amānah concept that could inspire greater personal accountability). A critical analysis point: As Malaysia crafts its AI Ethics Code, will it meaningfully incorporate Islamic ethical insights or default to repackaging Western frameworks with local branding? This question speaks to depth versus appearance. There is a risk of a nominal "Malaysian values" coating on what is essentially the OECD/EU principles, without mechanisms to address uniquely Islamic concerns or leverage Islamic institutions. Finally, Malaysia must consider enforcement: having codes and committees is good, but without legal teeth (e.g., a data protection law with real penalties, or mandatory ethics impact assessments), the policies could remain advisory. On the positive side, Malaysia is already moving to legislate in related areas (its Personal Data Protection Act is being updated, etc.), and NAIO's existence itself shows political will. In conclusion, Malaysia's case underscores the importance of inclusive governance and foresight. It is on the cusp of defining its AI ethics trajectory; if done well, it can create a model that harmonizes Islamic and universal ethics in a democratic context, providing an example of ethical AI in a plural society. If done poorly or too slowly, it could face issues of public trust or play catch-up to regulate harmful AI uses after they occur.

- **Pakistan**

National Strategy & Key Policies: Pakistan, the world's fifth most populous country, is an Islamic republic that has only recently formulated a comprehensive AI policy. On August 1, 2025, Pakistan's federal cabinet approved the National AI Policy 2025, marking the country's first holistic strategy on artificial intelligence. This policy aims to leverage AI for socio-economic development while explicitly emphasizing ethics and inclusivity. In broad strokes, it envisions transforming Pakistan into a "knowledge-based economy powered by ethical, inclusive, and innovative AI". The policy is built around six strategic pillars: (1) establishing a robust AI ecosystem (e.g., a National AI Fund, centers of excellence), (2) massive human capital development (training 200,000 people per year in AI, scholarships, internships), (3) governance and ethics (including regulatory sandboxes, cybersecurity and transparency frameworks to ensure ethical AI use and data protection), (4) sectoral AI integration (in education, healthcare, agriculture, etc.), (5) digital infrastructure (national computing clouds and data hubs), and (6) international collaboration on AI standards. The inclusion of an entire pillar on governance and ethics is noteworthy – it shows policymakers recognizing from the outset that AI's risks (privacy, security, bias, trust) must be managed alongside its promotion.

The policy text explicitly uses terms like "ethical AI" and calls for building public trust in AI systems. It proposes developing transparency frameworks (e.g., guidelines requiring government AI algorithms to be auditable and explainable) and cybersecurity protocols to safeguard data. It also mentions regulatory sandboxes, indicating an approach of controlled experimentation with oversight – reflecting caution in deployment. Importantly, the policy emphasizes alignment with international standards (such as ISO/IEC AI governance standards and UNESCO's AI Ethics Recommendation), showing intent to stay in harmony with global norms while implementing locally.

Explicit/Implicit Ethical Commitments: The National AI Policy 2025 itself is the explicit commitment to ethical AI, as it repeatedly stresses ethics, trust, and inclusion. For example, it speaks of "ethical, inclusive AI" as a guiding vision. It calls for public trust, which implicitly means AI must be transparent, fair, and

accountable. The policy stops short of laying down detailed ethical guidelines (that is left to subsequent frameworks and committees), but it establishes that ethics is a principle at the same level as innovation or growth. One concrete step announced is the creation of an AI Regulatory Authority or Committee – while not yet in place, the policy indicates Pakistan will set up an inter-agency committee including tech experts, ethicists, and Shariah scholars to draft AI ethics guidelines and oversee their implementation. Another explicit measure is to integrate with the forthcoming Personal Data Protection Bill to ensure data privacy in AI contexts. Implicitly, the policy leans on existing legal and ethical frameworks: for instance, Pakistan has an overarching digital governance framework and cyber laws that will interact with AI policy (e.g., the Prevention of Electronic Crimes Act might cover malicious AI use). Also, by referencing UNESCO and OECD guidelines, Pakistan's policy implies it will incorporate typical ethical principles (without necessarily naming them all in the policy document). It's worth noting that in early discussions, officials have mentioned the need for AI to respect human rights and accountability, showing awareness of global discourse.

Linkage to Islamic Values: Pakistan's constitution declares Islam as the state religion and that no law shall be repugnant to Islam. However, the National AI Policy, being a technical policy, is phrased largely in secular development language. It does not explicitly quote the Quran or Hadith, nor does it label any initiative as "Islamic AI." Nonetheless, the values it promotes – inclusion, protection, responsibility – echo Islamic principles. For example, the emphasis on inclusion (training marginalized groups, ensuring AI benefits the underprivileged) aligns with the Islamic emphasis on social justice and caring for the mustaḍʿafīn (vulnerable). The focus on ethics and data protection can be linked to the Islamic concept of amānah (trust): treating people's data and rights as a trust that must not be abused. Indeed, one could interpret the policy's repeated call for trustworthy AI as resonating with amānah – even if not explicitly stated, culturally it carries weight.

In Pakistan's cultural context, the implementation of AI policy will inevitably consider religious sensitivities. For instance, AI applications must respect local norms: no AI system should, say, generate blasphemous content or violate cultural decency laws (Pakistan has strict blasphemy and decency laws). The policy doesn't mention this, but it "leaves room" by using the term "ethical AI" generally – subsequent guidelines could clarify such points. Already, some Pakistani scholars and commentators have begun discussing AI from an Islamic lens, e.g. how AI could assist Islamic finance (similar to Malaysia's case) or whether using AI in certain artistic ways is permissible under Islamic views on depiction of living beings. Additionally, because Pakistan's legal system includes the Federal Shariat Court (which can strike down laws against Islam) and the Council of Islamic Ideology (advisory on Islamic conformity), any AI-related laws or rules may be reviewed for consistency with Islamic injunctions. For example, if the government were to ban or allow a certain AI (like facial recognition in public), someone might petition whether that infringes on Islamic privacy rights or, conversely, whether banning it hampers the state's duty to prevent crime (an Islamic duty). So far, nothing of that sort has occurred with AI, but one can anticipate these discussions. The AI Policy itself acknowledges alignment with local values implicitly: it says align with international standards "without losing local values", and it's understood that local values in Pakistan include Islamic values among others.

Strengths and Critical Gaps: Pakistan's AI policy shows strength in vision and intent. It explicitly ties AI advancement with ethical and inclusive principles, which is commendable especially for a developing country trying to catch up on technology. The document was hailed for its bold vision and comprehensive scope (covering infrastructure, skills, ethics, etc.). It also signals political commitment at the highest level (cabinet approval) to responsible AI. Another strength is that it builds on Pakistan's existing pool of Islamic scholarship and activism that can be tapped for guidance. For instance, Pakistan has a vibrant community of Islamic finance experts, tech entrepreneurs mindful of local norms, and civil society that can rally around causes like data privacy or algorithmic bias affecting disadvantaged groups. The policy's emphasis on

consultation (it mentions multi-stakeholder involvement) resonates with the Islamic principle of shūra (consultation), and indeed Pakistan could leverage forums involving religious leaders, tech experts, and the public to shape AI rules (one recommendation we make later is formalizing such an Ethics Board).

However, the gaps and risks are significant. A major critique, as noted by analysts (Mirza & Baig, 2025), is that while the policy's intent is laudable, it lacks detailed mechanisms for governance and enforcement. It proclaims principles but doesn't establish the concrete "how": e.g., which body will ensure AI systems are audited for bias? who will enforce data protection in AI beyond the pending law? The policy suggests ideas (like a committee, like legal frameworks) but those need to materialize. Currently, Pakistan has no dedicated AI regulatory authority (NAIO equivalent) and its Personal Data Protection Bill has lingered for years without being passed. This means the beautiful principles could ring hollow if not backed by legislation and institutions. The risk of "fine words not translating into practice" is acknowledged even within policy discussions. This is essentially the risk of ethics-washing: that Pakistan might check the box of having an ethical AI policy, but on the ground, nothing changes due to weak governance (the critique "Bold Vision, Fragile Foundations" captured this concern). Specifically, without passing the data protection law, privacy principles have no legal backbone. Without an AI oversight body, accountability is vague. Without mandating ethics impact assessments, developers might ignore the guidelines.

Another gap is capacity and coordination. Pakistan will need to significantly upskill bureaucrats, train regulators, and raise awareness in industry to implement the policy – a tall order given limited resources and other pressing issues. Yet another challenge is political stability: changing governments or priorities could stall the policy's implementation (Pakistan has had frequent changes in leadership). On the ethical front, one potential tension is between rapid development and cautious ethics: Pakistan is eager for economic development through AI and might be tempted to prioritize growth over strict ethics (the global "ethics vs innovation" dilemma, seen through an Islamic lens as well – Islam encourages progress but not via unethical means). Ensuring long-term ethical compliance may require resisting shortcuts.

On the positive side, opportunities exist if gaps are addressed. For example, Pakistan's rich Islamic intellectual tradition and active media could promote an Islamically grounded AI literacy: framing AI's benefits and harms in terms people relate to (e.g., sermons about using AI responsibly as part of one's duty). The policy hints at training scholars in AI for informed guidance – an innovative idea that could yield local "AI ethicists" fluent in both tech and theology.

In conclusion, Pakistan's case highlights the classic challenge of moving from principles to practice. The ethical aspirations are clearly present and rooted in both universal and Islamic values, but the execution machinery is still nascent. The risk of fragmented enforcement or policy paralysis is real if institutional structures aren't firmed up. The upcoming years will test whether Pakistan can enact the laws (like the Data Protection Act), create the bodies (AI Ethics Committee/Board), and allocate the resources to truly implement its ethical vision. If it succeeds, it will not only safeguard against AI harms but could position itself as a leader in the Muslim world for ethical tech governance. If it fails, the lofty words of the policy could become another unfulfilled promise, and AI development might proceed in a Wild West fashion or under external (foreign tech) influences with little local ethical input.

Discussion

The above analysis of Islamic ethical principles and the four national case studies reveals several overarching themes and challenges in integrating Islamic ethics with AI and digital transformation. In this discussion, we synthesize our findings, compare the approaches of the case study countries, and identify opportunities as well as obstacles for leveraging Islamic ethics in AI governance.

Convergence of Islamic Ethics and Global AI Principles: A striking observation is the high degree of alignment between Islamic ethical values and the universally espoused principles of AI ethics. Concepts like justice, fairness, doing no harm, accountability, and privacy are cornerstones in both domains. This convergence means that adopting an Islamic ethics perspective does not entail a completely separate or conflicting framework, but rather a culturally contextualized reinforcement of global best practices. For Muslim-majority societies, framing AI ethics in Islamic terms can enhance public buy-in: people may adhere to guidelines more earnestly if they see them as fulfilling religious obligations (e.g., avoiding algorithmic bias because it's unjust in Islam, not only because it's illegal or unfair by secular standards). It also allows policymakers to articulate AI regulations in a language that resonates locally – as seen in the UAE and KSA, where official narratives deliberately combine modernization with tradition (e.g., referencing both the SDGs and Shariah values). On the global stage, the inclusion of Islamic ethics contributes to ethical pluralism. In international forums like UNESCO or the Global Partnership on AI, Islamic viewpoints offer complementary perspectives, such as emphasizing spiritual well-being or collective welfare, which can enrich the moral discourse beyond a Western secular echo chamber. Crucially, none of the Islamic principles advocated (justice, mercy, trustworthiness, etc.) contradict global norms; rather, they often set a higher bar or add an additional rationale (for instance, treating privacy as inviolable because it's tied to dignity and honor, which might push for stronger privacy protections). In summary, there is a robust common ground between Islamic ethics and mainstream AI ethics, suggesting partnerships and mutual learning are very feasible.

Contextual Implementation – "One Size Does Not Fit All": While high-level principles coincide, our case studies demonstrate that implementation varies greatly with context. Governance style, legal systems, and social norms influence how AI ethics are adopted:

Authoritarian vs. Democratic Contexts: In more centralized states like the UAE and Saudi Arabia, we see top-down, government-driven approaches. Strong state control allows these governments to issue ethics guidelines and expect compliance in public-sector projects (and even nudge private sector adoption). These states project an image of ethical leadership as part of their global branding – essentially using ethical commitments to bolster legitimacy at home and attractiveness abroad. They often integrate Islamic ethics implicitly, embedding values in principles without overt religious language, to maintain international appeal and avoid sectarian specifics. This approach has been effective in achieving quick wins (e.g., published guidelines, showcase initiatives like AI ethics councils), but it relies heavily on political will and central oversight. In contrast, democratic and developing contexts like Malaysia and Pakistan are taking more consultative, iterative approaches. Malaysia's NAIO suggests a structured process involving academia, industry, and possibly civil society in drafting its codes. Being a democracy, Malaysia will need stakeholder engagement and transparency in rule-setting – akin to a shūrā-like model where various voices (including possibly opposition and public opinion) are considered. Pakistan, currently, has articulated ethical intentions in its policy but is still formulating mechanisms and battling capacity issues. The critique of Pakistan's policy warns of ethics-washing risk – grand principles without enforcement. To avoid this, Pakistan might need to create new institutions (an AI Ethics Council) that include not just officials but also independent experts and religious scholars, to flesh out standards and monitor compliance. Otherwise, lofty ideals in the policy may not prevent unethical practices (like biased AI in policing or unchecked facial recognition) from emerging.

These differences imply that any recommendation to incorporate Islamic ethics in AI must be tailored to governance context. Factors such as whether a country is authoritarian or democratic, how closely its legal system ties to Shariah, and whether its society is homogeneous or diverse will shape the strategy. For example, what works in the UAE (with a 90% expatriate population and a controlled political environment) might differ from what works in Pakistan (with active courts, political opposition, and a free press that will question AI failures). Yet, a common ethical vocabulary – rooted in shared values – can still unify these efforts, enabling

collaboration. Contextualization also means acknowledging limitations: a developing country may lack AI experts who are also versed in Islamic theology; hence, capacity-building (training ethicists in both domains) becomes part of the solution.

Tensions and Challenges: Despite general alignment, integrating Islamic ethics with AI governance is not without tensions. Several challenges emerged:

Interpretation and "Ethics Capture": Within Islamic thought, there is a spectrum of views on ethical priorities. Some lean utilitarian (emphasizing public interest/maslaḥa), while others are more deontological (strict on inviolable principles). This can lead to different stances on the same AI issue. For instance, a government might justify extensive AI-powered surveillance by citing maslaḥa – the security of society and prevention of harm – whereas civil libertarians or religious scholars might invoke Islamic individual rights to privacy and property to oppose such surveillance. Different scholars might disagree on what Islam dictates in such scenarios, especially when weighing state authority vs. personal freedom. There's a risk that authorities could cherry-pick Islamic concepts to legitimize what they intended to do anyway (e.g., overriding privacy by overemphasizing "preventing harm" or "commanding right and forbidding wrong" without balance). The antidote is a balanced interpretation via ijtihād (independent reasoning) that involves diverse scholars, including those with tech knowledge, to avoid any single narrative dominating. Essentially, Muslim-majority societies need to guard against "ethics capture" by political interests masquerading as religious justification.

Ethics vs. Innovation Dilemma: This is a universal debate, but in Islamic terms it becomes a question of preventing harm vs. promoting benefit. All our case studies desire AI-driven innovation and economic growth, yet they also profess ethical limits. Islam certainly encourages scientific advancement and the pursuit of benefits for humanity, but not at the cost of core moral breaches. The challenge is finding the right balance – for instance, allowing experimentation with AI (through sandboxes) but imposing boundaries to avoid known harms. A country like Pakistan, eager to not miss out on AI's benefits, might be tempted to cut corners on ethics for the sake of progress (e.g., deploying a technology widely before privacy safeguards are in place). Islamic ethics would caution against this by reminding that ends do not justify means, and that true success (with God's blessing, barakah) comes from righteous conduct, not just technical achievement. Policymakers must internalize that ethical compliance is part of sustainable progress, not an impediment to it. If framed correctly, Islamic ethics can actually reinforce long-term innovation by promoting trust and stability – a morally grounded market is more sustainable.

Global Compatibility and Cultural Relativism: As countries infuse local (Islamic) values into AI regulation, there's a question of whether this will fragment global AI norms. For instance, if an Islamic country like Saudi Arabia or Iran (outside our study but relevant) were to ban certain AI applications on religious grounds (say, generative AI that creates human images, seen by some as a form of creating idols), how does that interface with global AI systems and standards? Will multinational companies or open-source AI platforms need to accommodate such restrictions? There is a legitimate concern about interoperability and whether strict local rules isolate a country technologically. Our study suggests, however, that the Islamic ethics advocated so far are largely complementary to international norms. They might impose a higher standard (e.g., absolutely no use of biometric surveillance beyond a point, where others might allow some), but they do not introduce alien principles. In essence, Islamic ethics can be seen as a regional "accent" on a universal grammar of ethics, rather than a wholly different language. The focus should remain on ethics that enhance universal principles (justice, welfare, accountability) without creating unique rules that completely diverge. For globally integrated technologies, one strategy is flexible implementation: uphold core ethical requirements and allow optional stricter measures for local context. For example, a global AI might allow a

setting that disables certain features to comply with local ethical/legal norms (much like how content filters adapt to local laws). Another aspect is contributing Islamic ethical insights to global frameworks, so they become part of the international consensus rather than a point of conflict. The UNESCO Recommendation on AI Ethics (2021) actually came after consultations that included many cultures; it explicitly mentions respect for cultural diversity and the importance of ethical impact assessments – concepts very much in line with Islamic perspectives on deliberation and social impact. Thus, far from fragmenting, Islamic ethics can enrich the universality of AI norms if approached through dialogue and shared goals.

Enforcement and "Ethics-Washing": A recurring theme, especially in Malaysia and Pakistan, is the gap between policy and implementation. Even the UAE and KSA, for all their authority, primarily have guidelines rather than strict laws on AI ethics. Without enforcement mechanisms, ethical principles risk remaining performative. This problem is not unique to Islamic contexts – it's common worldwide that AI ethics boards and principles often lack teeth. But in Islamic contexts, it's particularly dissonant if lofty values (invoking God and morality) are proclaimed and then ignored, as it could breed cynicism about both governance and religious sincerity. We saw Pakistan being cautioned that without concrete structures (laws, authorities) the fine words may not translate to practice. To address this, countries will need to move from soft governance to hard governance over time: e.g., make certain ethical guidelines mandatory via legislation, set up independent oversight bodies, and impose penalties for violations (as data protection laws do). This institutionalization is further discussed below as a recommendation.

The Need to Institutionalize Ethical AI Governance: Across the board, a key to success will be creating institutions and processes that embed Islamic ethics into the AI life cycle. The case studies show some emerging institutions – like Malaysia's NAIO, Saudi's SDAIA and advisory boards, UAE's AI Council, Pakistan's planned committees. These are promising but need to be empowered and specialized. We recommend that these bodies establish dedicated Ethics Committees involving a mix of expertise: ethicists (including Islamic scholars), technologists, legal experts, and representatives of the public. For example, Malaysia's NAIO could have an ethics advisory council that includes an Islamic scholar from IIUM (International Islamic University Malaysia) alongside AI scientists and legal experts. That ensures that as they develop codes or approve AI projects, both technical feasibility and ethical compliance (in both secular and religious sense) are evaluated. Similarly, Pakistan might consider an AI Ethics Board under its Ministry of IT that includes member(s) from the Council of Islamic Ideology, to evaluate whether certain AI use-cases pose moral concerns under Islam. Such interdisciplinary oversight would operationalize the high-level ethics commitments by giving them a constant watchdog and guide.

Furthermore, training and education emerged as crucial. Ethical AI capacity can be boosted by training not only engineers in ethics but also religious scholars in technology. This cross-fertilization would produce professionals who can interpret Islamic principles in AI contexts accurately. Pakistan's policy even hints at training scholars to issue informed fatwas on tech – an innovative idea that could preempt knee-jerk rejections or approvals with more nuanced guidance.

Finally, accountability frameworks must be clear: requiring government AI systems to undergo ethics approval and audits, and encouraging or mandating the private sector to conduct ethics self-assessments or certification. The UAE's idea of an "AI Seal" and self-assessment toolkit is one model, albeit voluntary. Over time, these may evolve into required certifications for certain high-impact AI systems (similar to how halal certification works for food – a parallel Islamic model where ethical/religious criteria are certified by authorities).

In sum, the discussion underscores that Islamic ethical perspectives can indeed enrich AI governance, but realizing that potential needs deliberate effort. There is broad harmony with global ethics (a boon for

collaboration), and unique strengths like spiritual accountability and community-centric thinking that Islamic ethics brings. The challenges – interpretative differences, ensuring genuine implementation, balancing innovation – are significant but not insurmountable. They mirror classic governance issues, albeit with a cultural-religious dimension. If addressed through inclusive, well-designed institutions and sincere political will, Islamic ethics could move from being a rhetorical flourish in strategy documents to a practical, lived reality in how AI is developed and deployed across these nations.

Conclusion & Recommendations

In the 21st century, as artificial intelligence reshapes economies and daily life, grounding technological advancement in robust ethical frameworks is essential. This research demonstrates that Islamic ethics provides not merely a compatible, but a compelling and holistic framework for AI governance. Through literature analysis and case studies of the UAE, Saudi Arabia, Malaysia, and Pakistan, we found that Islamic ethical principles offer a congruent and culturally resonant foundation for addressing AI's challenges – from ensuring justice and fairness to safeguarding privacy and promoting the common good. Far from being a constraint, Islamic ethical perspectives can serve as a catalyst for responsible innovation, framing ethical AI as not just a legal or social requirement, but as a spiritual duty (amānah) and form of worship. This adds a powerful motivational lever for ethical compliance that goes beyond secular incentives.

Key Conclusions: First, values such as justice (ʿadl), public welfare (maṣlaḥa), compassion (raḥmah), and accountability (mas'ūliyyah) in Islamic ethics closely parallel the principles in global AI ethics frameworks (fairness, beneficence, privacy, accountability, etc.). This alignment means Muslim-majority countries can adopt global best practices in AI ethics without cultural dissonance – by articulating them through familiar ethical language and concepts. Indeed, our cases show that when ethical AI guidelines are mapped to Quranic or Hadith-based values (as done in the UAE and KSA), it enhances their acceptance and legitimacy. Second, the case studies demonstrate proactive efforts to integrate ethics into AI strategies: the UAE and KSA have issued national AI ethics guidelines that harmonize with Islamic teachings on justice and human dignity; Malaysia is establishing governance structures (NAIO) explicitly tasked with ethical AI regulation; and Pakistan's policy aspires to "ethical, inclusive AI" as a core vision. However, implementation and enforcement mechanisms need strengthening. For instance, Pakistan must operationalize its ethical principles via supporting laws (like the data protection act) and create an empowered oversight body – otherwise the principles may not prevent real harms. Third, Islamic ethics adds unique depth to certain issues: it elevates privacy to a near-sacred status (a God-given right), emphasizes intention and benevolence in AI use (encouraging developers to have righteous aims, not just risk mitigation), and frames accountability not just in legal terms but as a spiritual duty to God. These dimensions can reinforce ethical behavior even when legal systems are still catching up. For example, an AI engineer who internalizes that misusing AI breaches amānah may think twice even if regulation is lax. Fourth, there is a growing recognition globally of the need for inclusive AI governance. Incorporating Islamic ethics is part of a broader move toward multicultural approaches to AI norms. This pluralism can enhance global consensus by ensuring that ethical norms for AI have buy-in from different civilizations and are not seen as a one-way cultural imposition. In other words, an AI ethics dialogue that welcomes Islamic perspectives (alongside others) will likely craft principles that are more universally respected and effective.

Building on these conclusions, we present our recommendations as the logical outcome of this research – essentially outlining an "Islamic AI Governance Framework." The recommendations aim to bridge gaps identified and bolster the infusion of Islamic ethical principles into practice. The capstone proposal is the creation of an Islamic AI Ethics Charter, supported by national-level institutions and policies to enforce ethical AI. By embracing this ethically-grounded approach, Muslim-majority nations can transition from

being mere consumers of imported AI technology to leaders in a global movement for responsible innovation, truly fulfilling their role as khalāʾif al-arḍ (stewards of the earth).

Develop an "Islamic AI Ethics Charter": We recommend that leading Muslim-majority countries (perhaps under the auspices of the Organisation of Islamic Cooperation, OIC) collaboratively draft a high-level Islamic AI Ethics Charter or guidelines. This charter would distill foundational principles – many of which have been identified in this paper (justice, benevolence, non-maleficence, accountability, privacy, transparency, etc., all rooted in Islamic values) – and provide a reference framework for national policies. It should be positioned not in opposition to global frameworks but as a complementary document that highlights areas Islamic ethics can augment (e.g., spiritual accountability) and ensures terminology is culturally resonant. An OIC-led charter could help align countries and signal to the global community the collective commitment of the Islamic world to ethical AI. It would also mitigate the risk of divergent interpretations by providing a common baseline (for example, agreeing that mass surveillance without oversight is generally inconsistent with Islamic principles on privacy and justice, thereby encouraging all member states to put checks in place).

Institutionalize Ethics Oversight Bodies: Each country should establish (or strengthen) a dedicated mechanism for AI ethics oversight, integrating Islamic ethical expertise. This means moving beyond ad-hoc committees to permanent bodies with clear mandates. For example:

The UAE and Saudi Arabia can enhance the mandates of their existing councils or committees by including external ethics advisors (such as theologians, sociologists, and independent AI ethics experts) to periodically review AI deployments for compliance with stated ethical guidelines. Currently, much of the oversight is internal or ministerial; adding independent and religious voices can increase credibility and thoroughness. These bodies could publish annual "AI ethics compliance reports" evaluating major projects.

Malaysia's NAIO should, early in its operation, form a multi-stakeholder Ethics Committee that drafts the AI Ethics Code and then monitors its implementation. This committee should include representatives from Islamic advisory councils (e.g., JAKIM or university scholars) as well as industry, civil society, and technical experts. This will ensure that the code and its enforcement reflect both global best practices and local values. Given Malaysia's diversity, the committee should also include non-Muslim ethicists to keep the code inclusive.

Pakistan should expeditiously form an AI Ethics Board (perhaps under the Ministry of IT & Telecom or as an independent commission). This board would develop detailed ethical guidelines (translating the policy's principles into actionable standards, akin to how UAE's principles were detailed) and crucially, have the power to oversee projects. It could be modeled somewhat like Pakistan's Council of Islamic Ideology but focused on technology – including technologists, Shariah scholars, legal experts, and citizen representatives. The board should also coordinate with Pakistan's planned Data Protection Authority once the privacy law is passed, to ensure AI-specific issues (like algorithmic bias or automated decision impacts) are covered. We also advise Pakistan to fast-track the Personal Data Protection Act (pending legislation) and related cyber laws, as these provide the legal backbone for many ethical issues (privacy, consent, cybersecurity).

In all cases, it's important that these bodies are not purely advisory. They should have defined powers – for instance, the ability to halt or demand modification of AI systems that pose significant ethical risks. This could be similar to how an ethics review board can stop an unsafe clinical trial, or how some countries' ICT regulators can block apps that violate laws. Having some "teeth" will prevent ethics guidelines from being easily ignored.

Integrate Ethics into AI Project Lifecycles (Accountability Mechanisms): To operationalize ethics, governments should require that major AI initiatives undergo an Ethical Impact Assessment (EIA) during planning. This is akin to environmental impact assessments for new projects. An EIA would evaluate potential biases, privacy implications, societal impact, and alignment with Islamic ethical values before deployment. For instance, if a city wants to roll out AI surveillance cameras, an EIA would consider privacy and suggest mitigations (like independent audits, data anonymization, strict access controls in line with not violating ḥarīm/private spaces unjustly). Additionally, accountability mechanisms such as:

- **Algorithmic audit requirements:** Government agencies (and perhaps companies, for high-impact systems) should periodically audit AI algorithms for fairness and transparency. The results could be reported to the oversight body mentioned above.
- **Ethics review in procurement:** When government agencies procure AI systems, ethical compliance should be a criterion. Vendors might need to show they have considered bias, privacy, etc., and if possible that their design aligns with local values (for example, not using forbidden data sources). This incentivizes private developers to incorporate ethics from the design phase.
- **Public transparency and complaints channels:** Build trust by allowing some transparency (without jeopardizing security) about how AI is used and how decisions are made. Also establish channels for individuals or whistleblowers to report AI-related grievances (e.g., if someone feels an AI decision was discriminatory or a system is being used in a way that breaches ethical guidelines). These can be directed to the ethics oversight bodies or human rights commissions. Islamic governance historically included the institution of ḥisbah (ombudsmen monitoring market and public morality) – a modern analog could be a tech-ombudsman to hear and address ethical concerns of AI users or those affected.

Capacity Building and Research (Empirical Grounding): We urge investment in interdisciplinary research and training at the intersection of AI, ethics, and Islamic studies. This could include:

- Funding university programs or think-tanks on "AI and Society in Islamic contexts" to produce empirical studies (e.g., surveys on public perception of AI ethics in Muslim communities, case studies on outcomes of implemented policies, etc.). Empirical grounding, as noted by reviewers of our work, is currently a gap – so encouraging local scholars to gather data will inform more nuanced policies.
- Developing curricula for engineers that include modules on Islamic ethics, and conversely, training for religious scholars on the basics of AI technology and its ethical challenges. This will create a new cadre of professionals who can bridge the two fields (we might call them "Islamic AI ethicists"). These individuals could staff the ethics committees, advise industry, or serve in academia to further knowledge.
- Encouraging knowledge exchange among Muslim-majority countries: e.g., an annual "Islamic AI Ethics Conference" under OIC or Islamic Educational, Scientific and Cultural Organization (ISESCO) where policymakers, scholars, and practitioners share experiences, what worked or didn't (UAE could share how its toolkit is used, Saudi could share how it engages ulema, Malaysia and Pakistan could discuss balancing pluralism, etc.). This will refine the collective approach and avoid each country reinventing the wheel.
- Partnering with global initiatives: Ensuring that Islamic perspectives are represented in global AI ethics forums (e.g., the GPAI – Global Partnership on AI – or IEEE's Ethically Aligned Design initiatives). This both contributes Islamic insights (like amānah, stewardship ideas) and allows local frameworks to get feedback and recognition internationally.

Mitigate Ethics-Washing – from Principles to Practice: As a cross-cutting recommendation, we stress moving beyond paper principles to concrete enforcement, which has been addressed above. One specific point: nations should consider tracking and publicizing progress on ethical AI commitments. For instance, a government could publish an annual report or scorecard: how many AI projects underwent ethics review, how many complaints received and resolved, what training programs for civil servants on AI ethics took place, etc. This self-assessment would help ensure accountability for implementing the ethics frameworks. It also signals seriousness to citizens (and international partners) that ethics isn't just talk. Civil society and media have a role here too – encouraging independent watchdog groups or journalists to investigate and spotlight discrepancies (e.g., if a country claims to value privacy but is quietly exporting spyware, that should be called out).

Finally, we return to the big picture: the goal is to embed a trusteeship mindset in AI governance. This means all stakeholders – policymakers, developers, users – see themselves as trustees (amīn) of AI, answerable to society and the Divine for how this technology is used. If the recommendations above are implemented, countries will not only ensure their AI adoption is ethically and socially responsible, but they may also become pioneers in what we term a cross-cultural, faith-informed AI governance movement. By doing so, they harness the best of both worlds: advanced technology steered by timeless moral principles. Inshallah (God willing), this will contribute to an AI future that benefits humanity at large and honors the higher purpose of knowledge and innovation.

References

Ali, F., Bouzoubaa, K., Gelli, F., Khan, S., & Hamzi, B. (2025). Islamic ethics and AI: An evaluation of existing approaches to AI using trusteeship ethics. Philosophy & Technology, 38(3). https://doi.org/10.1007/s13347-025-XXXX-X Rahman, M. A. (2002).

CommandEleven. (2025, August 11). Fragile foundations: Gaps in Pakistan's AI policy. https://commandeleven.com

Digital Policy Alert. (2020, June 17). Saudi Arabia: Adopted national strategy for data & AI. https://digitalpolicyalert.org

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

Gorian, E., & Osman, N. D. (2024). Digital ethics of artificial intelligence in Saudi Arabia and United Arab Emirates. Malaysian Journal of Syariah and Law, 12(3), 583–597.

Mirza, A. A., & Baig, M. A. (2025). National AI policy: Bold vision, fragile foundations. CommandEleven Strategic Insights. https://commandeleven.com

MyDigital Malaysia. (2024). Malaysia launches national AI office (NAIO). https://www.mydigital.gov.my/

Reuters News Agency. (2024, December 12). Malaysia launches a national AI office for policy, regulation. Reuters World News.

Saudi Data & AI Authority. (2023). AI ethics principles (Version 2.0). https://www.sdaia.gov.sa

The Legal Wire. (2025, August 9). Pakistan's national AI policy approved. https://thelegalwire.ai/

Thomson Reuters. (2023). How is AI regulated in the UAE? What lawyers need to know. Legal Insight – UAE. https://www.thomsonreuters.com/

Thomson Reuters. (2024). EU's Artificial Intelligence Act passed as first comprehensive AI law. Legal Insight MENAhttps://www.thomsonreuters.com/

Article Information and Declarations

Declarations

Authors' Contribution:

- [a] Conceptualization, and intellectual revisions
- [b] Data collection, interpretation, and drafting of manuscript
- The authors agree to take responsibility for every facet of the work, making sure that any concerns about its integrity or veracity are thoroughly examined and addressed

Conflict of Interest: NIL

Funding Sources: NIL

Correspondence:

Abdur Rahim

abdurrahim724336@gmail.com

How to Cite:

Islamic Ethical Perspectives on AI and Digital Transformation in the 21st Century. (2025). Wah Academia Journal of Global Religions, 1(3), 31-54. https://doi.org/10.63954/4sz9d580