PENERBIT PRESS
UNIVERSITI TEKNOLOGI MARA

Available online at
https://journal.uitm.edu.my/ojs/index.php/

e-ISSN: 3093-8368

Journal of Halal Science and Management Research 2(2) 2026, 105-133

Journal of Halal Science and Management Research

Halal Knowledge Integrity (HKI): A Proposed Governance Framework for AI-Mediated Islamic Knowledge

Saidatul Akmar Ismail[1*,2]

[1] *Faculty of Information Science, Universiti Teknologi MARA, 40450 Shah Alam, Selangor,Malaysia*
[2] *Office of the Vice-Chancellor, Universiti Teknologi MARA, 40450 Shah Alam, Selangor,Malaysia*

ARTICLE INFO

Article history:
Received 23 July 2026
Revised 6 August 2026
Accepted 15 August 2026
Published 30 September 2026

Keywords:
AI Governance
Governance Theory
Halal Knowledge Integrity
Information Integrity
Islamic Epistemology

DOI:
10.24191/jhsmr.v2i2.12815

ABSTRACT

The proliferation of Generative Artificial Intelligence (GenAI) poses a growing threat to knowledge integrity in domains where information carries spiritual, legal, and communal authority. In Islamic knowledge contexts, AI-generated religious content may circulate without adequate verification, attribution, or epistemic accountability, yet existing international AI governance instruments do not operationalise the source-authority, provenance, and scholarly-verification requirements of Islamic knowledge. This conceptual study introduces Halal Knowledge Integrity (HKI), a proposed domain-specific governance framework designed to safeguard the authenticity, verifiability, contextual validity, traceability, ethical production, and institutional accountability of AI-mediated Islamic knowledge. A structured, concept-centric review across four domains, namely AI governance, knowledge integrity, Islamic epistemology, and halal assurance, was conducted to develop HKI through cross-domain gap analysis and iterative framework construction. The key output is a framework comprising ten governance principles and an eight-stage lifecycle ecosystem. The principal contribution is a theoretically grounded and implementation-oriented governance architecture that extends halal assurance logic from physical products to AI-mediated knowledge, advancing the field of halal science and management into the digital knowledge domain. Five propositions are advanced to guide future empirical validation of the framework.

1. INTRODUCTION

The advent of large language models and generative artificial intelligence systems marks one of the most consequential transformations in the history of human knowledge production. Within the span of fewer than five years, AI systems have acquired the capacity to produce text, audio, imagery, and multimedia content indistinguishable from that created by human experts, including scholars, jurists, and religious authorities. The implications of this development for knowledge systems that derive legitimacy

[1*]Corresponding author. *E-mail address*: saidatulakmar@uitm.edu.my

from centuries of verified scholarship, traceable transmission chains, and community consensus are profound and, in many respects, alarming.

The global Islamic knowledge ecosystem is particularly exposed to these risks. Islamic jurisprudence, theological commentary, and religious guidance have historically relied upon rigorous epistemological mechanisms, including isnad (chain of transmission), tabayyun (verification), ijtihad (scholarly reasoning), and ijma (scholarly consensus), to ensure the authenticity, accuracy, and contextual appropriateness of religious knowledge. These mechanisms took centuries to develop and represent a sophisticated institutional architecture for knowledge governance. Generative AI systems, operating without access to these verification chains and frequently without awareness of the distinctions between different legal schools, cultural contexts, or levels of scholarly authority, pose a structural threat to this architecture.

For the purposes of this study, Halal Knowledge Integrity (HKI) refers to an integrated, domain-specific governance framework designed to ensure that AI-mediated Islamic knowledge is authentic in sourcing, verifiable in provenance, contextually valid in application, traceable in generation, ethically produced, institutionally accountable in dissemination, and aligned with the epistemological standards and objectives of Shariah. This definition establishes HKI at the outset as the focal construct of the study.

The scale of the governance challenge is substantial because religious guidance is increasingly encountered through digital platforms, conversational AI systems, and mobile applications. Generative models can produce fluent but unsupported or fabricated content, while Islamic knowledge additionally requires attention to source authority, juristic context, and legitimate scholarly interpretation (Ji et al., 2023; Elmahjub, 2023). In this setting, misinformation can extend beyond individual confusion to weaken trust in recognised knowledge institutions and blur distinctions between machine-generated synthesis and qualified religious authority.

The wider AI governance landscape has recognised the challenge of AI-generated misinformation but has responded primarily at the level of general principles. The OECD AI Principles (OECD, 2019; updated 2024) advance values of transparency, accountability, and human oversight, but do not address domain-specific knowledge integrity requirements. The NIST AI Risk Management Framework (NIST, 2023) provides a systematic approach to AI risk management but operates within a predominantly technical and organisational paradigm that does not engage with epistemic traditions. The UNESCO Recommendation on the Ethics of AI (UNESCO, 2021) addresses cultural diversity and value pluralism in principle but offers no operationalisation for Islamic knowledge contexts. The EU AI Act (European Commission, 2024) categorises AI risks by application domain but does not specifically recognise religious knowledge systems as a domain requiring specialised governance. ISO/IEC 42001:2023, the international AI management system standard, provides an institutional quality assurance architecture but similarly lacks domain-specificity.

Within the Islamic AI-ethics literature, scholars have begun to develop pluralist and trusteeship-based approaches that bring Islamic moral and epistemic resources into contemporary AI governance debates (Elmahjub, 2023; Ali et al., 2025). This emerging body of work establishes an important normative foundation, but it does not yet provide an end-to-end institutional assurance architecture for AI-mediated Islamic knowledge. Malaysia offers a useful conceptual setting for examining such an architecture because it combines mature halal-governance institutions with a developing national AI-governance ecosystem. The present study treats Malaysia as a prospective pilot setting only; it does not claim existing HKI adoption or validation.

This paper introduces Halal Knowledge Integrity (HKI) as a proposed domain-specific governance framework with paradigm-oriented potential, responding to the intersection of AI capability, Islamic knowledge vulnerability, and institutional governance need. HKI is defined as a framework designed to ensure that AI-mediated Islamic knowledge remains authentic, verifiable, contextually valid, traceable, ethically generated, institutionally accountable, and aligned with Shariah epistemology. The paper develops HKI from an initial conceptual proposition into a theoretically grounded governance framework by articulating its intellectual foundations, defining its governance principles, specifying its ecosystem architecture, and illustrating its prospective applicability through the Malaysian institutional context and comparative analysis of selected international institutional contexts.

The structure of this paper is as follows. Section 2 presents a structured, concept-centric review across four domains: AI governance, knowledge integrity, Islamic epistemology, and halal assurance. Section 3 outlines the research methodology. Section 4 presents the theoretical positioning and development of HKI as a proposed domain-specific governance framework. Section 5 identifies AI risks specific to Islamic knowledge. Section 6 presents the HKI Governance Ecosystem. Section 7 details the ten HKI governance principles. Section 8 addresses operationalisation strategies. Section 9 positions Malaysia as a Conceptual Governance Illustration and Prospective Pilot Setting. Section 10 presents international implications and comparative institutional contexts. Section 11 discusses HKI in relation to existing frameworks. Section 12 concludes with theoretical, policy, and research contributions.

2. LITERATURE REVIEW

2.1 Artificial Intelligence Governance

The governance of artificial intelligence has emerged as one of the defining policy challenges of the twenty-first century. As AI systems have transitioned from narrow tools into general-purpose and generative technologies with broad societal impact, governance scholarship has increasingly moved from high-level ethical principles toward lifecycle, organisational, multi-level, and adaptive governance arrangements (Batool et al., 2025; Janssen, 2025; Papagiannidis et al., 2025; Taeihagh, 2025; Ulnicane, 2025). The international community has likewise mobilised to establish governance frameworks that balance innovation with accountability.

The OECD AI Principles, first adopted in 2019 and updated in 2024, represent the most widely endorsed international AI governance standard, with over 40 countries formally adopting them. The principles advance five value-based commitments: inclusive growth and sustainable development; human-centred values and fairness; transparency and explainability; robustness, security, and safety; and accountability (OECD, 2024). While these principles have been influential in shaping national AI strategies, critics have noted their limitation as voluntary, non-binding guidelines that lack enforcement mechanisms and do not address the specific governance requirements of domain-specific AI applications (Taeihagh, 2021; Corrêa et al., 2023).

The NIST AI Risk Management Framework (NIST, 2023) advances a structured, lifecycle-oriented approach to AI risk identification and mitigation. Organised around the functions of Govern, Map, Measure, and Manage, the framework provides practical guidance for organisations seeking to embed responsible AI practices into their operations. NIST's subsequent Generative AI Profile extends this logic to risks distinctive to generative systems, including confabulation, information integrity, and content provenance (Autio et al., 2024). These instruments strengthen organisational accountability and cross-functional

governance but do not engage with domain-specific epistemological requirements arising in religious, legal, or scholarly knowledge contexts.

The UNESCO Recommendation on the Ethics of AI (2021) represents the first global normative instrument on AI ethics and was adopted by UNESCO's 193 Member States. The Recommendation emphasises human rights, cultural diversity, environmental sustainability, and multi-stakeholder governance. It recognises the importance of cultural diversity and plural values but does not specify an operational assurance model for source authority, jurisprudential context, and scholarly verification in Islamic knowledge settings (UNESCO, 2021).

The EU AI Act (2024), the world's first binding comprehensive AI governance legislation, adopts a risk-based regulatory approach, categorising AI systems by their level of risk across four tiers: unacceptable risk (prohibited), high risk (regulated), limited risk (transparency obligations), and minimal risk (voluntary codes). While the Act introduces important provisions for AI systems used in critical infrastructure, education, and democratic processes, it does not specifically address the governance of AI-generated religious or knowledge-authority content (European Commission, 2024; Veale & Borgesius, 2021).

ISO/IEC 42001:2023 provides an internationally recognised management system standard for AI, analogous to ISO 9001 for quality management. It specifies requirements for establishing, implementing, maintaining, and improving AI management systems within organisations. The standard is notable for its institutional scope and emphasis on continual improvement, but its generic design makes it insufficient as a standalone framework for managing domain-specific AI risks in Islamic knowledge contexts (ISO/IEC, 2023).

A critical review of the AI-governance literature reveals a consistent gap: contemporary governance frameworks increasingly address organisational accountability, lifecycle controls, adaptive governance, and generative-AI risks, yet remain designed for broad applications rather than domains in which knowledge carries epistemic, legal, or spiritual authority (Batool et al., 2025; Janssen, 2025; Papagiannidis et al., 2025; Taeihagh, 2025). This gap justifies the development and testing of a specialised governance framework for AI-mediated Islamic knowledge.

2.2 Knowledge Integrity

Knowledge integrity, as a governance concept, encompasses the accuracy, completeness, authenticity, traceability, and contextual appropriateness of knowledge across its lifecycle from generation to dissemination. The concept bridges several adjacent fields including information integrity, research integrity, knowledge governance, and digital trust.

Information integrity has been conceptualised as a property of information systems that helps ensure information remains accurate, consistent, and trustworthy throughout its lifecycle (Laudon & Laudon, 2022). In the context of generative AI, information integrity is threatened by hallucination: the production of fluent content that is unsupported, unfaithful to source material, or factually incorrect. Recent LLM research distinguishes factuality and faithfulness failures, demonstrates methods for detecting uncertainty-based hallucinations, and shows that known facts may still be generated unreliably under particular model and prompting conditions (Farquhar et al., 2024; Huang et al., 2025; Jiang et al., 2024; Wang et al., 2024), extending earlier work on hallucination and factual consistency (Ji et al., 2023; Maynez et al., 2020). The prevalence of hallucination varies by model, task, prompting conditions, and domain; accordingly, this

study does not assign a single cross-domain error rate and instead treats hallucination as a governance risk requiring verification controls.

Research integrity frameworks, such as those advanced by the Committee on Publication Ethics (COPE, 2022) and the Singapore Statement on Research Integrity emphasise honesty, rigour, transparency, and accountability in the production of scholarly knowledge (McNutt et al., 2018). These frameworks, while applicable to academic knowledge production, do not address the challenges of AI-mediated knowledge generation or the specific requirements of Islamic scholarly traditions.

Knowledge governance, a concept developed primarily within organisational and institutional economics, addresses the institutional arrangements that enable the creation, sharing, and protection of knowledge within and across organisations (Grandori, 2001; Foss, 2007). Digital platforms and algorithmic systems have intensified longstanding governance questions concerning who produces, curates, distributes, and controls knowledge (Zuboff, 2019). This literature provides a useful institutional lens for understanding HKI but does not operationalise the epistemological requirements of Islamic scholarly authority.

Digital trust, increasingly recognised as a foundational requirement for the digital economy, encompasses the confidence of users and institutions in digital systems and the information they produce (OECD, 2022; WEF, 2023). Contemporary AI research highlights accountability, explainability, meaningful human evaluation, and traceable provenance as complementary conditions for justified trust (Kim et al., 2024; Novelli et al., 2024; Pan et al., 2023; Schneider, 2024), building on earlier trustworthy-AI literature (Thiebes et al., 2021; Floridi et al., 2020). However, digital trust frameworks do not address the specific trust requirements of religious communities, where trust is embedded in institutional and theological authority structures.

A critical review of the knowledge integrity literature reveals that existing frameworks address important dimensions of knowledge quality, reliability, and governance but do not fully address the requirements of Islamic knowledge contexts. Specifically, none of the reviewed frameworks incorporates the epistemological principles of Islamic scholarship, the institutional mechanisms of Islamic knowledge certification, or the community governance structures that have historically ensured the integrity of Islamic knowledge.

2.3 Islamic Epistemology

Islamic epistemology constitutes one of the most sophisticated and institutionally embedded knowledge governance systems in human history. Developed over fourteen centuries of scholarly activity, it has produced a rich array of mechanisms for verifying, transmitting, contextualising, and applying knowledge in ways that are responsive to community needs while maintaining fidelity to authoritative sources.

Amanah (trustworthiness, fidelity) represents the foundational ethical principle governing knowledge production and transmission in Islam. Knowledge in the Islamic tradition is not merely information; it is a trust (amanah) entrusted to scholars and communities, to be preserved, transmitted, and applied with integrity. The Quranic injunction to "convey as you have received" (Al-Tirmidhi, Hadith 2519) establishes an explicit governance principle for knowledge transmission that has direct relevance to AI-mediated knowledge systems.

Tabayyun (verification, due diligence) is established in the Quran (Al-Hujurat, 49:6) as an obligation upon believers to verify information before accepting or acting upon it. This principle has been

operationalised in Islamic scholarship through the science of hadith criticism ('ilm al-hadith), which developed sophisticated methodologies for evaluating the reliability of sources, the coherence of transmission chains, and the contextual appropriateness of reported information (Brown, 2009). Tabayyun is directly applicable as a governance principle for AI-generated Islamic content.

Isnad (chain of transmission) represents the distinctive Islamic mechanism for knowledge provenance and traceability. Every hadith (prophetic tradition) in the canonical collections is accompanied by an isnad documenting the complete chain of transmitters from the reporter to the Prophet. The science of rijal (biographical criticism) developed alongside isnad to evaluate the reliability, memory, and character of each transmitter. This system represents perhaps the world's most sophisticated pre-modern knowledge provenance architecture and provides a compelling model for AI-mediated knowledge attribution systems (Lucas, 2004).

Ijtihad (independent juristic reasoning) and its conditions, qualifications, and limits represent the Islamic framework for knowledge innovation and contextualisation. The conditions for qualified ijtihad, including mastery of Arabic linguistics, Quranic sciences, hadith sciences, jurisprudential principles, and awareness of contemporary circumstances, constitute a sophisticated competency framework for knowledge authority that has direct governance implications for AI systems that generate Islamic content (Kamali, 2008).

Maqasid al-Shariah (the objectives of Islamic law) provides the teleological governance framework within which Islamic knowledge is evaluated. The five essential objectives, protecting religion (din), life (nafs), intellect ('aql), lineage (nasl), and property (mal), offer a values-based governance criteria against which the appropriateness of AI-generated Islamic knowledge can be assessed (Auda, 2008). The Maqasid framework has been extended by contemporary scholars to include the protection of human dignity and the environment, providing a dynamic governance architecture adaptable to new challenges (Chapra, 2008).

Hikmah (wisdom, practical discernment) represents the capacity to apply knowledge appropriately within specific contexts, recognising the limitations of general principles and the importance of contextual judgment. Hikmah is the quality that distinguishes a mufti (jurisconsult) from a mere jurisprudential calculator and has no adequate analogue in current AI systems (Al-Ghazali, trans. Winter, 1995). The absence of hikmah in AI systems represents a fundamental epistemic limitation that governance frameworks must address through mandatory human oversight provisions.

2.4 Halal Assurance

The halal industry represents one of the most developed institutional architectures for Islamic quality assurance and provides a valuable model for the extension of assurance principles from products to knowledge.

Malaysia has developed an institutionalised halal-certification architecture through JAKIM's Manual Prosedur Pensijilan Halal Malaysia (Domestik) 2020 and Sistem Pengurusan Halal Malaysia 2020. These instruments formalise requirements for certification procedures, internal halal control, documentation, traceability, monitoring, and corrective action across relevant operations (JAKIM, 2020a, 2020b). Their importance for HKI lies in the assurance logic they embody rather than in any claim that existing halal certification already governs AI-mediated knowledge.

Halal certification and assurance practices share structural features with the knowledge-governance requirements proposed in HKI: controlled inputs, documented processes, conformity checks, traceability, and institutional oversight. HKI adapts this assurance logic conceptually from product and process governance to knowledge provenance and scholarly verification (Riaz & Chaudry, 2004; JAKIM, 2020a, 2020b).

The digitalisation of halal-related services and the growing use of AI across Islamic domains create a conceptual need to examine whether assurance principles can extend beyond physical products and processes to information and knowledge outputs. The present study treats this extension as a theoretical proposition rather than as an already established field of halal certification or digital assurance.

The extension of halal assurance to AI-mediated knowledge, however, requires more than an analogical transfer of product certification principles. It requires a domain-specific governance framework that integrates Islamic epistemological principles with contemporary AI governance standards, institutional knowledge-governance frameworks, and halal-assurance mechanisms into a coherent and potentially operationalisable architecture. This is the contribution that HKI seeks to make.

2.5 Critical Synthesis: Why a Distinct HKI Framework Is Needed

Across the reviewed traditions as shown in Table 1, the central gap is not the absence of ethical principles but the absence of an integrated mechanism that links epistemic authority, verifiable provenance, contextual validity, institutional responsibility, and lifecycle assurance for AI-mediated Islamic knowledge. General AI frameworks privilege cross-sector principles and risk controls; knowledge-integrity approaches emphasise reliability and trust; Islamic epistemology supplies source-authority and verification norms; and halal assurance contributes institutionalised lifecycle control. None of these traditions alone specifies how an AI-generated Islamic output should move from authoritative sources through machine processing, scholarly verification, institutional endorsement, public disclosure, monitoring, and corrective learning. HKI is theoretically necessary only to the extent that it closes this integration gap. It should therefore be understood as a domain-specific synthesis and extension of knowledge-governance and assurance theory, rather than as a claim that existing frameworks are wholly inadequate.

Table 1. Critical cross-framework synthesis and the theoretical necessity of HKI

Framework / Tradition	Primary Strength	Critical Limitation	Unresolved Question	HKI Response
OECD / UNESCO	Normative principles, rights, transparency, accountability	Cross-sector and high-level; limited operational treatment of religious epistemic authority	Who determines authoritative Islamic sources and contextual validity?	Adds source hierarchy, scholarly verification, institutional authority.
NIST AI RMF	Lifecycle risk management	No domain-specific epistemic validation mechanism	How are jurisprudential and source-authority risks governed?	Adds verification checkpoints and provenance controls.
ISO/IEC 42001	Management-system accountability and continual improvement	Generic knowledge-quality criteria	What constitutes acceptable Islamic knowledge integrity?	Adds Shariah-grounded criteria and scholar-led assurance.
EU AI Act	Binding risk-based regulation	Religious knowledge not a specialised risk class; EU-bound	When should religious guidance trigger heightened review?	Proposes risk-sensitive scholarly and institutional review.
Islamic Epistemology	Isnad, tabayyun, ijtihad, maqasid, amanah	Not designed as an AI management system	How can classical norms become auditable digital controls?	Translates epistemic norms into AI-governance controls.

Halal Assurance	Certification, lifecycle control, traceability	Primarily product/service assurance	Can assurance logic govern knowledge provenance?	Adapts assurance to source, process, output and institutional integrity.

3. METHODOLOGY

This study employs a conceptual research methodology appropriate to the development of new theoretical constructs and governance frameworks in contexts where the primary objective is theory synthesis and framework construction rather than empirical testing. Conceptual research is particularly suitable when a phenomenon is fragmented across multiple literatures and the contribution lies in clarifying constructs, integrating theoretical perspectives, and establishing propositions for subsequent empirical investigation (Jaakkola, 2020; MacInnis, 2011). Because the goal is framework construction rather than systematic evidence synthesis, this study does not adopt a PRISMA systematic review protocol. Instead, it follows a purposive, concept-centric review guided by explicit domain scope, eligibility criteria, and analytical coding, as detailed in Sections 3.1 to 3.3. The framework development procedures and their outputs are presented in Sections 4 to 8, which constitute the principal conceptual contribution of this paper.

The study follows a three-phase methodology. In the first phase, a structured conceptual review was undertaken across four domains: AI governance, knowledge integrity, Islamic epistemology, and halal assurance. Searches were conducted iteratively across Scopus, Web of Science, Google Scholar, official standards and regulatory repositories, and relevant Islamic-studies sources. Search terms combined concepts such as artificial intelligence governance, AI ethics, knowledge integrity, information integrity, digital trust, Islamic epistemology, Islamic AI, halal assurance, halal certification, and knowledge governance. The review was purposive and concept-centric rather than a registered systematic review; because a complete record-level search and screening log was not maintained, the study does not report PRISMA counts or claim exhaustive coverage. Recent literature was prioritised for rapidly evolving AI-governance issues, while seminal earlier works were retained where theoretically necessary.

In the second phase, a conceptual synthesis was conducted to identify the gaps and overlaps between the four literature domains and to map the conceptual terrain within which HKI is positioned. This synthesis employed the concept-centric approach to literature review recommended by Webster and Watson (2002), organising the reviewed literature around key concepts rather than individual studies, and identifying the theoretical void that HKI addresses.

In the third phase, the HKI framework was developed through an iterative conceptual construction process, drawing on established methods for framework development in information systems, governance, and management research (Gregor, 2006; Jaakkola, 2020). The framework was assessed conceptually for internal consistency, coherence with the reviewed literature, and alignment with established governance theory. The study also draws on institutional analysis of Malaysia's halal and AI governance infrastructure and comparative analysis of OIC member state AI governance initiatives.

This study acknowledges the limitations inherent in conceptual research, including the absence of empirical validation and the dependence on the researcher's interpretive choices. To address these limitations, five research propositions are advanced to provide a structured programme for future empirical investigation, and the comparative institutional-context analysis provides a preliminary cross-contextual plausibility assessment rather than empirical validation.

3.1 Review Scope, Eligibility and Search Logic

The review was structured around four concept domains: AI governance, knowledge integrity, Islamic epistemology, and halal assurance. The following databases were searched: Scopus, Web of Science, Google Scholar, and the official repositories of ISO, NIST, OECD, UNESCO, and the EU. For Islamic epistemology and halal assurance, supplementary searches were conducted in the JAKIM official publications portal, the SMIIC standards library, and the Islamic Fiqh Academy database. Search terms included combinations of "artificial intelligence governance," "AI ethics," "knowledge integrity," "information integrity," "digital trust," "Islamic epistemology," "Islamic artificial intelligence," "halal assurance," "halal certification," and "knowledge governance," connected using Boolean AND/OR operators and refined through iterative citation chaining. Materials were eligible for inclusion when they were peer-reviewed articles, scholarly books or book chapters, recognised international standards, legislation, or authoritative institutional policy documents making a substantive contribution to at least one of the four domains. Materials were excluded when they lacked identifiable scholarly or institutional provenance, addressed AI without a governance or knowledge-integrity dimension, or treated halal solely as a consumer-market label without assurance or governance relevance. Recent literature published predominantly within the five-year period preceding submission was prioritised for rapidly evolving areas, particularly generative AI, AI governance, information integrity, digital trust, and responsible AI. Earlier sources were retained selectively where they constitute seminal governance theory, foundational conceptual methodology, authoritative Islamic epistemological scholarship, classical Islamic texts, or still-operative institutional standards and normative instruments.

3.2 Source Selection and Review Transparency

Sources were selected for their direct contribution to one or more of the four conceptual domains and for their authority, methodological or normative relevance, and usefulness to framework construction. Peer-reviewed scholarship, recognised standards, legislation, and authoritative institutional documents were prioritised. Because the review process did not preserve a complete record-level screening log, no numerical flow diagram or retrospective inclusion/exclusion counts are reported. This limitation is stated explicitly to avoid presenting reconstructed search figures as if they were auditable systematic-review data. Future review-based extensions should preregister the protocol, document database-specific search strings and dates, and maintain a record-level screening log.

3.3 Quality Appraisal and Analytical Coding

Source quality was assessed according to source authority, methodological or normative relevance, recency for rapidly evolving AI-governance materials, and conceptual contribution. Peer-reviewed research and official standards or regulatory instruments were given greater evidentiary weight than commentary. Coding was concept-centric rather than author-centric. Extracted concepts were organised into four analytical families: (i) source and provenance integrity; (ii) process, transparency, and verification; (iii) authority, accountability, and human oversight; and (iv) lifecycle assurance, monitoring, and continuous improvement. Islamic epistemic concepts were then mapped against functionally comparable governance mechanisms in contemporary AI and assurance frameworks.

3.4 Framework Development Procedure

HKI was constructed iteratively in four steps. First, cross-domain gaps were identified by comparing what existing AI-governance instruments protect against with what Islamic knowledge systems require for legitimate authority and transmission. Second, convergent concepts were mapped across Islamic epistemology, knowledge governance, AI governance, and halal assurance. Third, these concepts were translated into prescriptive governance principles and lifecycle stages. Fourth, the architecture was checked for internal coherence, non-redundancy, traceability to its conceptual foundations, and consistency between principles, actors, stages, and proposed outcomes. This procedure produces an analytically derived framework; it does not constitute stakeholder or empirical validation.

3.5 Validation Status and Planned Validation Strategy

HKI has not yet undergone empirical or stakeholder validation. Accordingly, claims in this paper concern theoretical coherence and prospective applicability, not demonstrated effectiveness. A staged validation programme is proposed: Phase 1, a multidisciplinary Delphi panel involving Islamic scholars, halal-assurance specialists, AI-governance experts, regulators, and information-science researchers to assess relevance, clarity, completeness, and jurisdictional sensitivity; Phase 2, formal content-validity assessment of the principles and stages; Phase 3, interviews or focus groups with implementing institutions and affected users to test acceptability and role feasibility; Phase 4, pilot case studies using selected AI-mediated Islamic knowledge services to assess workflow feasibility, verification burden, traceability, error reduction, and governance responsiveness; and Phase 5, comparative testing across jurisdictions to evaluate transferability. This sequence is a future research protocol and should not be interpreted as validation already completed.

4. RESULT AND DISCUSSION

4.1 Framework Development: The Theoretical Positioning of HKI

The Insufficiency of Existing Frameworks

The case for HKI as a proposed domain-specific governance framework with paradigm-oriented potential rests on a threefold argument: general AI ethics frameworks do not fully operationalise Islamic epistemic requirements; halal assurance frameworks were developed primarily for products and processes; and an integrated governance architecture is therefore needed for AI-mediated Islamic knowledge.

Most influential contemporary AI-governance frameworks have emerged from international regulatory, technical, organisational, and rights-based traditions and are designed primarily for cross-sector applicability. Recent research has further emphasised multi-level governance, lifecycle responsibility, adaptive organisational mechanisms, accountability, and the distinctive governance challenges of generative AI (Batool et al., 2025; Janssen, 2025; Leon, 2026; Novelli et al., 2024; Taeihagh, 2025; Ulnicane, 2025). Their principles of fairness, transparency, accountability, safety, and human oversight are important, but they do not by themselves operationalise the source-authority, jurisprudential, contextual, and scholarly-verification requirements of Islamic knowledge. The limitation is therefore one of domain specificity and institutional translation rather than a wholesale inadequacy of existing AI-governance traditions.

Halal assurance frameworks, by contrast, are institutionally rich and practically operationalised, but they were developed for the governance of physical products and have not been systematically extended to knowledge. The extension of halal assurance to knowledge is not merely a matter of analogy; it requires a fundamental reconceptualisation of what it means for knowledge to be halal, addressing questions of source authority, verification methodology, contextual appropriateness, institutional accountability, and public dissemination that do not arise in product certification contexts.

A domain-specific governance framework for AI-mediated Islamic knowledge should therefore integrate insights from these traditions while addressing their respective limitations in this specialised context. Such a framework should be epistemologically grounded in Islamic scholarship, institutionally informed by halal-assurance practice, and aligned with contemporary responsible-AI principles. HKI is proposed as one such integrative framework.

Formal Definition of Halal Knowledge Integrity

This study advances the following formal definition of Halal Knowledge Integrity as a proposed domain-specific governance framework with paradigm-oriented potential:

> *"Halal Knowledge Integrity (HKI) is a proposed domain-specific governance framework designed to ensure that AI-mediated Islamic knowledge remains authentic in its sourcing, verifiable in its provenance, contextually valid in its application, traceable in its generation, ethically produced in its process, institutionally accountable in its dissemination, and continuously aligned with the epistemological standards and teleological objectives of Shariah."*

This definition reflects six foundational governance commitments: authenticity, verifiability, contextual validity, traceability, ethical production, and institutional accountability. These commitments are not merely descriptive; they are prescriptive, establishing the standards against which AI systems operating in Islamic knowledge contexts should be evaluated.

HKI is positioned more cautiously in this study as a proposed domain-specific governance framework with paradigm-oriented ambitions. Its novelty lies not in replacing established governance theory, but in integrating knowledge governance, lifecycle assurance, and AI-risk management with Islamic epistemic requirements that are not jointly operationalised in existing frameworks. The term 'paradigm' is therefore treated as a prospective theoretical claim: HKI could develop into a broader governance paradigm if future research demonstrates distinctive explanatory value, stakeholder legitimacy, implementation feasibility, and cross-jurisdictional robustness. At its present stage, HKI should be read as an analytically derived framework and research programme rather than an empirically established paradigm.

HKI and Governance Theory

HKI is positioned within the broader tradition of governance theory as applied to knowledge domains. Knowledge governance, following Grandori (2001) and Foss (2007), addresses the institutional arrangements, mechanisms, and norms that shape the creation, sharing, and use of knowledge. HKI extends knowledge governance theory in three ways.

First, it introduces Islamic epistemological principles as legitimate governance resources, demonstrating that the sophisticated knowledge verification mechanisms developed within Islamic

scholarship (isnad, tabayyun, ijtihad) are not merely historical artefacts but are practically applicable to contemporary AI governance challenges.

Second, it integrates domain-specific governance with international AI governance standards, providing a bridge between the universal principles of frameworks such as the OECD AI Principles and the specific requirements of Islamic knowledge governance. This integration is significant because it demonstrates how universal governance principles can be contextualised and enriched through engagement with diverse epistemic traditions.

Third, HKI introduces an institutional accountability architecture that specifies the roles and responsibilities of different institutional actors, including AI developers, Islamic scholars, halal certification bodies, regulatory authorities, and civil society, in ensuring the integrity of AI-mediated Islamic knowledge. This multi-stakeholder governance architecture distinguishes HKI from both AI ethics frameworks, which tend to focus on developers, and halal assurance frameworks, which tend to focus on certification bodies.

4.2 AI Risks in Islamic Knowledge Contexts

Before presenting the HKI framework, it is necessary to map the specific AI risk landscape in Islamic knowledge contexts. This mapping draws on the literature reviewed in Section 2 and on institutional analysis of documented cases of AI-generated Islamic content.

Hallucination and fabrication represent an immediate and well-documented generative-AI risk. Large language models can produce plausible but unsupported statements, invented citations, and unfaithful summaries; recent studies further demonstrate that hallucination is heterogeneous, can persist even with known facts, and requires explicit factuality and uncertainty controls (Farquhar et al., 2024; Huang et al., 2025; Jiang et al., 2024; Wang et al., 2024), alongside the earlier evidence base (Ji et al., 2023; Maynez et al., 2020). In Islamic knowledge contexts, such failures may manifest as misattributed hadith, conflated juristic positions, or unsupported religious claims; these domain-specific manifestations are treated here as risk scenarios requiring empirical investigation rather than as prevalence estimates already established by this study.

Source integrity risks arise from the inability of AI systems to distinguish between authoritative Islamic sources (canonical hadith collections, established works of jurisprudence, recognised scholarly commentary) and low-quality, polemical, or unreliable sources. Without robust source governance, AI systems trained on undifferentiated Islamic content will reflect the errors, biases, and interpretive distortions present in their training data.

Authority and attribution risks reflect the structural inability of AI systems to replicate the institutional authority structures of Islamic scholarship. When an AI system generates a fatwa-like ruling without indicating that it is not a qualified jurist, it potentially misleads users into treating its output as authoritative religious guidance. This risk is compounded by the opacity of AI systems, which may not disclose the sources or reasoning processes underlying their outputs.

Contextual validity risks arise from the context-dependence of Islamic jurisprudential rulings. Many rulings are explicitly conditioned on specific circumstances (time, place, community, custom), and applying a ruling generated for one context to a different context without appropriate qualifications constitutes a jurisprudential error that AI systems are ill-equipped to detect. The collapse of contextual specificity in AI-generated Islamic content represents a systemic governance challenge.

Institutional delegitimisation risks arise at the systemic level, as the proliferation of AI-generated Islamic content without governance undermines the authority and relevance of established Islamic institutions, including religious councils, muftis, and scholarly bodies. If users can obtain instant, authoritative-sounding Islamic guidance from AI systems, the incentive to seek guidance from institutional authorities is reduced, potentially eroding the community governance structures that have historically ensured knowledge integrity.

Table 2 below summarizes the five major risk categories, their specific manifestations in Islamic knowledge contexts, and their governance implications.

Table 2. AI Risk Categories in Islamic Knowledge Contexts

Risk Category	Manifestation	Governance Implication
Hallucination & Fabrication	Factually incorrect hadith, conflated rulings, fabricated scholarly positions	Mandatory source verification, output auditing, institutional review before public dissemination
Source Integrity	Undifferentiated training on authoritative and unreliable sources	Curated Islamic knowledge repositories, source classification systems, authority weighting
Authority & Attribution	AI outputs presented without scholarly authority disclosure	Mandatory provenance labelling, disclaimer requirements, institutional accountability chains
Contextual Validity	Rulings applied across incompatible legal, cultural, or temporal contexts	Contextual metadata requirements, jurisdictional compliance checks, scholarly review for context-sensitive rulings
Institutional Delegitimisation	Erosion of authority of religious councils and scholarly institutions	Institutional oversight requirements, AI-human complementarity mandates, community governance integration

4.3 The HKI Governance Ecosystem

The HKI Governance Ecosystem presents a proposed lifecycle architecture for AI-mediated Islamic knowledge, encompassing eight sequential governance stages from knowledge source to continuous learning feedback. The ecosystem is designed as a closed-loop system in which Islamic scholarly authorities, AI developers, halal certification bodies, regulatory agencies, and end users each occupy defined roles connected by formal accountability flows. Specifically, Islamic scholarly authorities govern Stage 1 (source curation) and Stage 3 (human verification) and participate in Stage 4 (institutional review); AI development teams govern Stage 2 (processing) and contribute to Stage 8 (continuous learning); an HKI Governance Committee bridges Stages 5 (compliance) and 7 (monitoring), providing the connective accountability structure; and the public interacts at Stage 6 (dissemination) and feeds back into Stage 7. Figure 1 illustrates this complete ecosystem, showing directional flows between stages and the institutional actors responsible at each point. Table 3 summarizes responsible and participating actors at each stage of the HKI governance ecosystem.

Fig. 1. The HKI Governance Ecosystem: Halal Knowledge Integrity Governance Architecture for AI-Mediated Islamic Knowledge

Table 3. Responsible and participating actors at each stage of the HKI governance ecosystem

HKI Governance Stage	Responsible / Participating Actor
Stage 1: Knowledge Sources	Islamic Scholars / Knowledge Custodians
Stage 2: AI Processing	AI Developers / Technical Team
Stage 3: Human Verification	Qualified Islamic Scholars
Stage 4: Institutional Review	Recognised Islamic Authority
Stage 5: Governance & Compliance	HKI Governance Committee
Stage 6: Public Dissemination	Authorised Institution / Platform
Stage 7: Monitoring & Feedback	HKI Committee + Users / Community
Stage 8: Continuous Learning	AI Team + Scholars + HKI Committee

Stage 1: Knowledge Sources. The ecosystem begins with the curation and classification of Islamic knowledge sources. This stage involves the establishment of authoritative source repositories, the classification of sources by type (Quran, hadith, classical jurisprudence, contemporary scholarly opinion), authority level (canonical, established, qualified, contested), and contextual domain (worship, transactions, family law, bioethics, etc.). Source governance at this stage is analogous to supplier qualification in halal assurance systems.

Stage 2: AI Processing. At the processing stage, AI systems ingest curated knowledge sources and generate outputs. Governance at this stage encompasses the design and training of AI systems to respect source hierarchies, the implementation of retrieval-augmented generation (RAG) architectures that ground AI outputs in curated sources, and the deployment of content filtering systems trained to detect common error patterns in Islamic content.

Stage 3: Human Verification. All AI-generated Islamic knowledge outputs must pass through a structured human verification process before dissemination. This stage involves qualified Islamic scholars (or trained paraprofessional reviewers for lower-stakes content) assessing the accuracy, appropriateness, and contextual validity of AI outputs. The verification process should be documented and traceable, creating an accountability record.

Stage 4: Institutional Review. For high-stakes content (fatwas, jurisprudential rulings, theological positions), institutional review by a recognised Islamic scholarly body is required. This stage mirrors the institutional certification function in halal assurance and provides the authoritative endorsement that distinguishes certified AI-Islamic knowledge outputs from unverified content.

Stage 5: Governance and Compliance. A dedicated HKI governance function, analogous to a halal assurance committee, oversees the overall integrity of the AI-Islamic knowledge system. This function is responsible for setting governance standards, monitoring compliance, adjudicating disputes, and reporting to regulatory and community stakeholders.

Stage 6: Public Dissemination. Verified and governed Islamic knowledge content is disseminated through designated channels with appropriate provenance labelling, authority disclosures, contextual qualifications, and user guidance. Dissemination channels must be registered with the HKI governance function and must comply with dissemination standards.

Stage 7: Monitoring and Feedback. Post-dissemination monitoring encompasses user feedback collection, error reporting systems, community governance mechanisms for flagging problematic content,

and regular audits of disseminated content against HKI standards. Monitoring data feeds into the continuous learning stage.

Stage 8: Continuous Learning. The ecosystem incorporates structured feedback loops through which monitoring data informs updates to source repositories, AI system fine-tuning, verification protocols, and governance standards. This stage ensures that the HKI ecosystem is adaptive and responsive to emerging challenges and community needs.

The eight-stage ecosystem provides an end-to-end, implementation-oriented architecture integrating institutional responsibilities, technical controls, and community-governance mechanisms. Its sequencing and role allocations are theoretically derived; empirical testing is required to assess implementation feasibility, stakeholder acceptance, verification cost and speed, provenance reliability, and enforcement capacity.

4.4 HKI Governance Principles

The HKI Governance Ecosystem is anchored by ten foundational principles, each linked to a related Islamic epistemic or ethical foundation and a corresponding contemporary AI-governance concept. Table 4 summarizes these relationships and their operational implications. The mappings are interpretive and functional rather than claims of one-to-one equivalence between classical Islamic concepts and modern AI-governance terminology.

Table 4. HKI Governance Principles

#	Principle	Related Islamic Epistemic/Ethical Foundation	AI Governance Equivalent	Operational Implication
P1	Authenticity	Isnad (chain of transmission)	Source attribution and provenance	All AI-generated Islamic content must be traceable to verified primary sources
P2	Tabayyun (Verification)	Quranic obligation to verify (49:6)	Factual accuracy standards	Mandatory multi-layer verification before dissemination
P3	Transparency	Amanah and bayān (clarity/disclosure)	Algorithmic transparency	Disclosure of AI involvement, training sources, and limitations
P4	Explainability	Ijtihad and reasoned justification	Explainable AI (XAI)	AI outputs must include reasoning traces accessible to scholars
P5	Accountability	Amanah (trusteeship)	Organisational accountability	Named institutional accountability for each published AI-Islamic output
P6	Traceability	Silsilah (scholarly lineage)	Audit trails and logging	Complete generation logs maintained for retrospective review
P7	Provenance	Matn (text authenticity)	Data provenance	Source classification and authority-weighting documented for each output
P8	Human Oversight	Qualified scholarly authority and consultative review	Human-in-the-loop	All substantive rulings require human scholarly endorsement
P9	Shariah Compliance	Maqasid al-Shariah (objectives)	Values alignment	Content evaluation against five Maqasid and contemporary scholarly consensus
P10	Continuous Improvement	Islah (reform and renewal)	ISO 42001 continual improvement	Annual governance review, stakeholder feedback integration, standard updates

Principle 1 (Authenticity) requires that all AI-generated Islamic content be traceable to verified primary sources within the recognised canon of Islamic scholarship. Drawing on the isnad principle, this governance requirement is not merely bibliographic citation but epistemological certification: the content of an AI output must be demonstrably grounded in authoritative sources, not reconstructed or interpolated.

Principle 2 (Tabayyun/Verification) operationalises the Quranic injunction of Al-Hujurat 49:6 as a mandatory governance checkpoint. In the HKI ecosystem, tabayyun is implemented as a multi-layer verification process encompassing automated fact-checking, scholarly review, and institutional endorsement for content above defined risk thresholds.

Principle 3 (Transparency) requires full disclosure of the AI nature of knowledge outputs, the training sources and methods employed, the limitations of the AI system, and the verification processes applied. Transparency is associated here with the Islamic ethical commitments to amanah and bayān, understood functionally as duties of trustworthy disclosure and clarity. This mapping is interpretive rather than a claim of direct conceptual equivalence, and it corresponds to contemporary AI-governance expectations concerning transparency and disclosure.

Principle 4 (Explainability) requires that AI systems be capable of providing reasoning traces that can be examined by qualified Islamic scholars. The HKI treatment of explainability draws analogically on the Islamic tradition of reasoned justification associated with ijtihad, in which conclusions are expected to be supported by identifiable sources and reasoning. The analogy does not imply that machine-generated explanation constitutes ijtihad. It does, however, identify a functionally comparable governance requirement: AI systems operating in Islamic knowledge contexts should be capable of providing reasoning traces that qualified scholars can examine and critique. Recent generative-AI explainability research similarly emphasises the need to specify when, what, and how explanations should be provided and to support verifiability rather than treating explanation as a purely technical display (Jang et al., 2024; Schneider, 2024).

Principle 5 (Accountability) requires that named institutional actors bear responsibility for each published AI-Islamic knowledge output. Drawing on the Islamic concept of amanah, accountability is not merely procedural but moral: institutions that publish AI-generated Islamic content are trustees of the community's epistemic heritage and bear responsibility for its integrity.

Principles 6 and 7 (Traceability and Provenance) address the documentation requirements for AI-generated Islamic content, including generation logs, source classifications, authority weightings, and contextual metadata. These principles are derived from the twin Islamic sciences of isnad (chain of transmission) and matn (text authenticity) and correspond to AI governance requirements for audit trails and data provenance.

Principle 8 (Human Oversight) establishes the non-negotiable requirement for human scholarly involvement in substantive Islamic knowledge decisions. This principle reflects the Islamic emphasis on qualified scholarly authority, consultation, and accountable human judgment. It does not equate ordinary human-in-the-loop review with ijma, which has a distinct technical meaning within Islamic jurisprudence. The governance requirement it articulates is that substantive Islamic knowledge decisions must involve qualified human scholars whose judgment cannot be fully delegated to AI systems.

Principle 9 (Shariah Compliance) requires systematic evaluation of AI-generated Islamic knowledge content against the five Maqasid al-Shariah and against contemporary scholarly consensus. This principle provides a values-based governance mechanism that goes beyond factual accuracy to assess the teleological appropriateness of AI-Islamic knowledge outputs.

Principle 10 (Continuous Improvement) institutionalises the Quranic concept of islah (reform, renewal, improvement) as a governance commitment. In the HKI framework, continuous improvement encompasses

annual governance reviews, systematic stakeholder feedback integration, regular updates to source repositories, and ongoing refinement of verification protocols in response to emerging challenges.

4.5 Operationalisation of The HKI Framework

The translation of HKI from a proposed governance framework to operational practice requires institutional design, role specification, technical infrastructure, capacity building, and empirical validation. The operational requirements below are design propositions rather than evidence of current feasibility. Pilot implementations should assess at minimum: source-verification accuracy; frequency and severity of detected errors; inter-rater agreement among scholarly reviewers; time and cost of verification; completeness of provenance records; stakeholder acceptance; institutional role clarity; user trust; and the effectiveness of monitoring and correction mechanisms.

At the institutional level, operationalising HKI requires the establishment of an HKI Governance Committee within relevant institutions (universities, Islamic councils, media organisations, AI developers) comprising qualified Islamic scholars, AI governance experts, halal assurance specialists, and community representatives. This committee is responsible for setting institutional HKI standards, overseeing verification processes, and reporting on governance performance.

At the technical level, operationalising HKI requires investment in curated Islamic knowledge repositories with documented source classifications and authority weightings; retrieval-augmented generation (RAG) architectures that ground AI outputs in authoritative sources; output monitoring systems capable of detecting common error patterns in Islamic content; and provenance labelling systems that document the sourcing and verification history of AI-Islamic knowledge outputs. Contemporary RAG research supports the use of external retrieval to improve grounding while also cautioning that retrieval pipelines do not eliminate hallucination or traceability problems (Gao et al., 2024; Huang et al., 2025). Provenance should therefore be treated as an auditable lifecycle record rather than merely a citation layer (Carey-Wilson et al., 2024; Pan et al., 2023).

At the capacity level, operationalising HKI requires the development of HKI literacy programmes for Islamic scholars (equipping them to evaluate AI-generated content), for AI developers (equipping them to understand Islamic epistemological requirements), and for the Muslim public (equipping them to be critical consumers of AI-generated Islamic content). This multi-stakeholder capacity development is essential for the sustainable governance of AI-mediated Islamic knowledge.

At the regulatory level, operationalising HKI at scale may eventually require consideration of dedicated standards or guidance analogous to assurance standards in established halal-governance systems. In Malaysia, one possible institutional pathway for future exploration could involve consultation with JAKIM and other relevant religious, digital-governance, standards, academic, and regulatory stakeholders. At the OIC level, organisations such as SMIIC could provide a platform for future comparative discussion, subject to formal institutional consultation.

Research Propositions for Future Empirical Investigation: The following five propositions are advanced to provide a structured programme for empirical research on HKI.

> *P1: Source Integrity (adherence to HKI Principle 1) positively influences the perceived reliability of AI-generated Islamic knowledge outputs.*

P2: Greater Process Integrity, operationalised through adherence to HKI Principles 2-4, is associated with a lower frequency of factual and jurisprudential errors in AI-generated Islamic content.

P3: Greater Output Integrity, operationalised through adherence to HKI Principles 5-7, is positively associated with institutional trust in AI systems used in Islamic knowledge contexts.

P4: Greater Human Oversight, operationalised through HKI Principle 8, is positively associated with perceived governance effectiveness in AI-mediated Islamic knowledge systems.

P5: Perceived adherence to HKI governance requirements is positively associated with Muslim users' confidence in AI-mediated religious knowledge services.

4.6 Malaysia as a Conceptual Governance Illustration and Prospective Pilot Setting

Malaysia is presented here as a conceptual governance illustration and prospective pilot setting for HKI because it combines an established halal-governance infrastructure, an evolving national AI-governance environment, a substantial Muslim-majority population, and active participation in ASEAN and the OIC.

Importantly, this section does not claim that HKI has been adopted, implemented, endorsed, or validated by Malaysian authorities or institutions. The Malaysian case is used analytically to examine institutional fit: it asks whether existing halal-assurance, Islamic-governance, digital-regulation, and AI-policy capacities could provide a plausible environment for future HKI piloting. References to possible institutional roles are therefore prospective and require formal stakeholder consultation and regulatory review.

Malaysia has a mature halal-governance infrastructure centred on JAKIM certification and a national industry-development agenda. The Halal Industry Master Plan 2030 (HDC, 2023) explicitly seeks to leverage Malaysia's halal infrastructure and strengthen the country's international leadership in the halal economy. For HKI, the relevance of this infrastructure is institutional: it provides established experience with documented assurance, certification, governance, and continuous improvement that could inform a future knowledge-integrity pilot.

Malaysia's developing AI-governance ecosystem provides a second institutional foundation for prospective HKI testing. The National Guidelines on AI Governance and Ethics provide voluntary national principles for responsible and ethical AI, while the National AI Office currently coordinates national AI development and promotes responsible, transparent, and ethical deployment. These instruments do not endorse HKI, but they provide a contemporary governance environment within which a domain-specific knowledge-integrity framework could be examined (MOSTI, 2024; Malaysia National AI Office, n.d.).

The broader Malaysian policy environment also emphasises values-based and responsible digital development.

Malaysia's Digital Economy initiatives, including Malaysia Digital and the Jalinan Digital Negara (JENDELA) broadband expansion programme, are creating the digital infrastructure through which AI-mediated services, including Islamic knowledge services, will reach the Malaysian population. The combination of expanding digital access and widespread mobile connectivity creates both an opportunity and a governance need for examining HKI in the Malaysian context.

As ASEAN Chair in 2025 and an active member of the OIC, Malaysia may provide a conducive institutional setting for piloting and further developing HKI, subject to stakeholder validation, jurisdictional alignment, institutional capacity, and regulatory assessment. ASEAN's Guide on AI Governance and Ethics (ASEAN, 2024) provides a regional framework within which a future HKI module could be explored pending appropriate consultation with regional partners, including Indonesia, Brunei, and Muslim-minority communities in Thailand, the Philippines, and Singapore.

The institutional actors potentially relevant to Malaysia's prospective HKI institutional ecosystem include JAKIM (as the primary Islamic governance authority), the Malaysian Institute of Islamic Understanding (IKIM) (as a think-tank for Islamic governance policy), the Halal Development Corporation (as the industry development agency), the National AI Office (as the AI governance policy coordinator), Universiti Teknologi MARA and other research universities (as research and capacity development actors), and the Malaysian Communications and Multimedia Commission (as the digital content regulator). A candid assessment of Malaysia's readiness must also acknowledge substantial governance challenges. First, there is currently no institutional mechanism within JAKIM or the National AI Office that formally governs AI-mediated Islamic knowledge, meaning that HKI adoption would require new policy instruments rather than adaptation of existing ones. Second, the supply of Islamic scholars with sufficient AI and digital-governance literacy to serve as credible HKI verifiers is limited; building this capacity would require sustained investment in multidisciplinary training programmes. Third, jurisdictional complexity within Malaysia's federal religious governance structure means that the authority of a centralised HKI body to govern AI-generated Islamic content across states with their own Shariah authorities would need careful constitutional and legislative resolution. Fourth, the commercial AI ecosystem in Malaysia is developing rapidly and largely without domain-specific Islamic-knowledge governance requirements, meaning that any HKI mandate would face commercial compliance burdens. These challenges are not insurmountable, but they underscore that Malaysia's institutional strengths represent a necessary rather than sufficient condition for HKI implementation, and that the pathway from conceptual illustration to operational pilot requires deliberate policy design, stakeholder engagement, and adequate resourcing.

4.7 International Implications and Comparative Institutional Contexts

The HKI framework, while developed with particular attention to the Malaysian context, is proposed as a domain-specific governance framework with potential relevance across Muslim-majority jurisdictions. This section as shown in Table 5 offers an illustrative comparison of institutional contexts and identifies questions for future cross-jurisdictional testing; it does not constitute empirical benchmarking or evidence of readiness.

Indonesia represents an important comparative context for future HKI research because of its large Muslim population, plural Islamic institutional landscape, developing AI-policy environment, and established halal-governance infrastructure (Stranas KA, 2020). The Indonesian institutional context includes national Islamic scholarly bodies and halal-governance institutions that provide useful points of comparison with Malaysia. The greater diversity of Islamic legal opinion and the federalised religious authority structure in Indonesia suggest that any HKI-type framework would need to accommodate substantive institutional pluralism, making Indonesia a valuable test case for framework transferability.

Brunei Darussalam offers a useful comparative context because of its relatively centralised Islamic governance arrangements and its participation in ASEAN and the OIC. Future research could examine

whether such institutional characteristics facilitate or constrain HKI-type governance, and how a smaller jurisdiction with concentrated religious authority might approach knowledge-integrity governance differently from a larger, more pluralist one.

Saudi Arabia offers another important comparative context because of its globally influential Islamic scholarly institutions and its significant investment in AI governance and digital transformation. Future comparative research could examine how HKI principles align with Saudi AI-governance initiatives and established scholarly-authority structures. Such comparison would provide evidence concerning the cross-jurisdictional applicability of HKI rather than presuming institutional endorsement.

The United Arab Emirates provides a useful comparative setting because of its substantial investment in AI strategy, institutional development, and AI-focused higher education. Future HKI research could examine how a domain-specific Islamic knowledge-governance framework might interact with such innovation-oriented governance environments, particularly where AI development objectives may create tensions with verification-intensive governance requirements.

The Organisation of Islamic Cooperation (OIC) and related Islamic standardisation and research bodies may provide future platforms for comparative dialogue on AI-mediated Islamic knowledge governance. Any movement toward regional or cross-jurisdictional standardisation would, however, require extensive scholarly, regulatory, and institutional consultation, and should be preceded by national-level validation studies in interested member states.

Table 5. Illustrative Institutional Contexts for Prospective HKI Research Across Selected OIC Jurisdictions. Note: "Islamic Auth." describes the primary national Islamic-governance authority and its jurisdictional scope (High = centralised national body; Moderate = pluralist or advisory structure). "AI Governance" indicates the existence of a national AI strategy or policy office. "Halal Infra." reflects halal-certification capacity. "Potential Institutional Relevance" is an analytical inference derived from the combination of these factors; it does not constitute validated readiness assessment and is subject to ongoing policy change.

Country	Islamic Authority	AI Governance	Halal Infra.	Potential Institutional Relevance	Illustrative Research Role
Malaysia	JAKIM (High)	National AI Office	Highly developed	Potential pilot context	Prospective pilot
Indonesia	MUI (High)	Stranas KA	Developed	Comparative context	Comparative case
Brunei	MUIB (High)	ICT Authority	Developing	Comparative context	Comparative case
Saudi Arabia	Senior Scholars (Global)	SDAIA	Very High	Comparative context	Comparative case
UAE	GCAAI (Moderate)	Ministry of AI	Developing	Comparative context	Comparative case
OIC (Platform)	SMIIC / IFA	COMSTECH	Varied	Varied	Cross-jurisdictional comparison

4.8 Discussion

This section situates HKI within the landscape of existing governance frameworks, identifying the distinctive theoretical and practical contributions that HKI makes to each. Table 6 presents a structured comparison.

Table 6. HKI Compared to Existing Governance Frameworks

Framework	Scope	Limitation	HKI Contribution

OECD AI Principles	General AI governance	No Islamic epistemic context; voluntary; non-enforceable	Islamic governance principles as operationalisation mechanism; institutional enforcement architecture
UNESCO AI Ethics	Education and culture	Value pluralism stated but not operationalised for Islamic knowledge	Concrete epistemic assurance mechanisms rooted in Islamic scholarship
NIST AI RMF	Organisational AI risk	Technical/organisational focus; no domain-specific knowledge integrity	Knowledge verification protocols; scholarly authority integration; isnad-based traceability
ISO/IEC 42001	AI management systems	Generic; no accommodation of religious or epistemic authority	Shariah-grounded quality assurance layer; HKI Governance Committee model
EU AI Act	Risk-based AI regulation	Binding only in EU; religious knowledge not a recognised risk category	OIC-adaptable governance model; recognition of Islamic knowledge as high-risk AI domain
Halal Assurance (JAKIM/MHMS)	Physical product certification	Product-only scope; no AI or knowledge domain application	Extension to AI-mediated knowledge; governance ecosystem model; knowledge provenance standards

The comparative analysis suggests that HKI occupies a distinctive position within the reviewed governance landscape by integrating elements that are otherwise dispersed across existing AI-governance, knowledge-governance, Islamic epistemological, and halal-assurance approaches. It is more domain-specific than international AI governance frameworks, more epistemologically grounded than Islamic AI ethics literature, more explicitly oriented toward AI-mediated knowledge governance than conventional halal-assurance systems, while offering a more operationally structured model than purely philosophical treatments of Islamic knowledge governance.

The most important theoretical implication of HKI is its proposition that Islamic epistemological principles can function as governance resources for contemporary AI-mediated knowledge. Much influential AI-governance work has emerged from regulatory, technical, organisational, and rights-based traditions designed for broad cross-sector use. HKI contributes to the diversification of this knowledge base by examining how Islamic verification, provenance, responsibility, and contextual-reasoning traditions may complement those established approaches in a domain-specific setting.

A critical perspective on HKI must acknowledge its limitations. As a conceptual framework, HKI awaits empirical validation. The operationalisation challenges identified in Section 8, including the scarcity of AI-literate Islamic scholars, the technical complexity of implementing isnad-based provenance systems, and the political economy of Islamic institutional reform, are real and significant. Future empirical research should test the five propositions advanced in Section 8 and engage directly with practitioners in the Islamic knowledge, AI development, and halal assurance sectors to refine and contextualise the framework.

The international dimensions of HKI also raise important questions about governance harmonisation and cultural autonomy. Muslim communities in different jurisdictions have different traditions of Islamic legal interpretation, different institutional governance arrangements, and different relationships between religious and state authority. A future research agenda for HKI must include comparative empirical studies of HKI applicability across diverse Islamic governance contexts to ensure that the framework does not inadvertently impose a particular national or legal school perspective on the diversity of the global Muslim community.

From Conceptual Coherence to Validated Governance

The appropriate maturity claim for HKI at this stage is conceptual coherence, not validated governance effectiveness. A Q1-level research programme should distinguish four evidentiary thresholds: conceptual validity (clear constructs and theoretically defensible relationships); content validity (expert agreement that principles and stages are relevant and sufficiently comprehensive); implementation validity (evidence that institutions can execute the workflow with acceptable burden and role clarity); and outcome validity (evidence that HKI reduces errors, improves provenance and accountability, and strengthens justified user trust). The present study addresses primarily the first threshold. The remaining thresholds constitute explicit tests for subsequent research and guard against overstating the framework's maturity.

5. CONCLUSION

This paper has introduced Halal Knowledge Integrity (HKI) as a proposed, domain-specific governance framework for AI-mediated Islamic knowledge. Its central contribution is an integration of Islamic epistemic safeguards with contemporary knowledge-governance, AI-risk, and lifecycle-assurance mechanisms. The framework specifies ten principles and an eight-stage governance ecosystem, but it is not presented as empirically validated. Malaysia functions only as a conceptual institutional illustration and prospective pilot setting. The practical value of HKI will depend on whether future expert review and real-world pilots demonstrate that its source controls, scholarly verification, institutional review, provenance labelling, monitoring, and corrective-learning mechanisms are feasible, legitimate, scalable, and capable of improving knowledge integrity. From the perspective of halal science and management, HKI represents a principled extension of the halal paradigm from its established domain of physical products and services into the emergent domain of AI-mediated religious knowledge. The halal assurance literature has shown that integrity, traceability, and institutional accountability can be operationalised within a certification architecture that combines Islamic normative standards with management-system controls. HKI applies this same logic to a new object of governance: knowledge itself. In doing so, it addresses a gap that halal science and management research has not yet systematically explored, namely, the question of what it means for knowledge to be halal in the age of generative AI, and what institutional architecture is required to ensure it. This contribution situates HKI firmly within the scope of the Journal of Halal Science and Management Research and invites the halal science community to engage with the knowledge-integrity dimension of the halal economy.

5.1 Theoretical Contributions

The primary theoretical contribution of this paper is the introduction of HKI as a proposed domain-specific governance framework with paradigm-oriented potential, integrating Islamic epistemological principles with contemporary AI-governance standards and knowledge-governance theory. HKI extends the discussion in knowledge governance by proposing that effective governance of high-authority knowledge may require domain-specific epistemic controls in addition to general AI-governance safeguards. The paper also contributes to a more plural AI-governance discourse by examining the relevance of Islamic epistemic traditions to contemporary governance challenges.

5.2 Governance Contributions

The governance contribution of this paper is the development of a comprehensive, implementation-oriented governance architecture for AI-mediated Islamic knowledge, encompassing an eight-stage

governance ecosystem, ten governance principles with institutional operationalisation guidance, and a research propositions programme for future empirical development. This architecture offers a proposed governance roadmap for future institutional testing by Islamic institutions, AI developers, halal-governance bodies, and regulatory authorities.

5.3 Policy Contributions

The policy contribution of this paper lies in identifying how existing halal-assurance, Islamic-governance, and AI-governance infrastructures might support future experimentation with domain-specific knowledge-integrity controls. Malaysia is used as a prospective pilot setting rather than as evidence of implementation readiness. The framework therefore offers policymakers a conceptual basis for considering how HKI might be piloted, evaluated, refined, and potentially adapted to other jurisdictions.

5.4 Societal Contributions

The societal contribution of this paper is its contribution to the protection of the epistemic integrity of the Islamic knowledge tradition in an era of unprecedented AI capability. By providing a governance framework that operationalises the sophisticated epistemic mechanisms of Islamic scholarship, HKI is intended to contribute to efforts to safeguard AI-mediated Islamic knowledge so that it may serve the community's spiritual, legal, and intellectual needs without compromising the integrity of its knowledge heritage. This contribution extends beyond the Muslim community to the broader project of governing AI in ways that respect the diversity of human epistemic traditions.

5.5 Implications for Muslim-Majority Nations

For Muslim-majority nations, HKI provides a conceptual basis for examining how Islamic governance principles may contribute constructively to global AI-governance debates. The Islamic tradition's sophisticated mechanisms for knowledge verification, provenance documentation, and scholarly authority represent genuine governance innovations that the international AI governance community has not yet engaged with seriously. Future comparative research in jurisdictions such as Malaysia, Indonesia, Brunei Darussalam, Saudi Arabia, and the United Arab Emirates could assess whether different combinations of Islamic governance, halal assurance, and AI-policy capacity affect the feasibility and legitimacy of HKI implementation.

5.6 Future Research Agenda

The next research priority is staged validation. First, a multidisciplinary Delphi study should assess the relevance, clarity, completeness, jurisdictional sensitivity, and stakeholder legitimacy of the ten principles and eight stages. Second, content-validity indices and inter-expert agreement should refine the framework. Third, interviews and focus groups should test role feasibility among Islamic scholars, halal authorities, AI developers, regulators, information professionals, and users. Fourth, bounded pilot case studies should evaluate operational metrics such as verification accuracy, error reduction, provenance completeness, review time, cost, escalation rates, and user trust. Fifth, comparative studies across Malaysia and other OIC contexts should examine transferability and the effects of different legal schools and institutional structures. Only after these stages should stronger claims be made that HKI constitutes an empirically established governance framework or broader paradigm.

Halal Knowledge Integrity begins as a governance proposition whose ultimate value must be demonstrated through empirical validation, institutional testing, and measurable contribution to the integrity of AI-mediated Islamic knowledge.

6. ACKNOWLEDGEMENTS/FUNDING

No external funding was received for this conceptual study. The author acknowledges Universiti Teknologi MARA (UiTM) for the institutional environment supporting scholarly work in information science, governance, and halal knowledge research.

7. CONFLICT OF INTEREST STATEMENT

The author declares no commercial, financial, or personal conflict of interest in relation to this study.

8. AUTHOR'S CONTRIBUTION

Saidatul Akmar Ismail conceptualised the study, conducted the conceptual review and synthesis, developed the HKI framework, and wrote and revised the manuscript.

9. DECLARATION OF GENERATIVE AI IN THE WRITING PROCESS

During the preparation of this work, the author used Grammarly to improve the grammar, clarity, and readability of the manuscript. After using this service, the author reviewed and edited the content as needed and take full responsibility for the content of the publication

10. DATA AVAILABILITY/SUPPLEMENTARY MATERIALS

No empirical dataset was generated or analysed for this conceptual study. The sources underpinning the analysis are identified in the reference list.

11. ETHICS STATEMENT

Ethics approval was not required because this study is conceptual and did not involve human participants, personal data, or animal subjects.

12. REFERENCES

Al-Ghazali, A. H. (1995). The remembrance of death and the afterlife: Kitab dhikr al-mawt wa-ma ba'dahu. Book XL of the Revival of the Religious Sciences (T. J. Winter, Trans.). Islamic Texts Society.

Al-Tirmidhi, M. I. (n.d.). Jami' al-Tirmidhi (Hadith 2519). Compiled by Abu 'Isa Muhammad ibn 'Isa al-Tirmidhi. [Various editions available; see Darussalam edition, 2007].

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

ASEAN. (2024). ASEAN guide on AI governance and ethics. Association of Southeast Asian Nations.

https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf

Auda, J. (2008). Maqasid al-Shariah as philosophy of Islamic law: A systems approach. International Institute of Islamic Thought.

Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1

Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. AI and Ethics, 5, 3265-3279. https://doi.org/10.1007/s43681-024-00653-w

Brown, J. (2009). Hadith: Muhammad's legacy in the medieval and modern world. Oneworld Publications.

Carey-Wilson, T., Simperl, E., Das, A., & Reklos, I. (2024). Understanding data governance in AI: A lifecycle perspective. Open Data Institute. https://doi.org/10.61557/KSQF4883

Chapra, M. U. (2008). The Islamic vision of development in the light of Maqasid al-Shariah. International Institute of Islamic Thought.

Committee on Publication Ethics (COPE). (2022). COPE code of conduct and best practice guidelines for journal editors. COPE.

Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E., & de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. Patterns, 4(10), Article 100857. https://doi.org/10.1016/j.patter.2023.100857

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/s13347-023-00668-x

European Commission. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. Nature, 630, 625-630. https://doi.org/10.1038/s41586-024-07421-0

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: Seven essential factors. Science and Engineering Ethics, 26(3), 1771-1796. https://doi.org/10.1007/s11948-020-00213-5

Foss, N. J. (2007). The emerging knowledge governance approach: Challenges and characteristics. Organization, 14(1), 29-52. https://doi.org/10.1177/1350508407071087

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. arXiv. https://doi.org/10.48550/arXiv.2312.10997

Grandori, A. (2001). Neither hierarchy nor identity: Knowledge-governance mechanisms and the theory of the firm. Journal of Management and Governance, 5(3-4), 381-399. https://doi.org/10.1023/A:1013442830067

Gregor, S. (2006). The nature of theory in information systems. MIS Quarterly, 30(3), 611-642. https://doi.org/10.2307/25148742

Halal Development Corporation (HDC). (2023). Halal Industry Master Plan 2030. Halal Development Corporation. https://hdcglobal.com/halal-industry-master-plan-2030/

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. ACM Transactions on Information Systems, 43(2), Article 42. https://doi.org/10.1145/3703155

ISO/IEC 42001:2023. (2023). Information technology - Artificial intelligence - Management system. International Organization for Standardization. https://www.iso.org/standard/81230.html

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. AMS Review, 10(1-2), 18-26. https://doi.org/10.1007/s13162-020-00161-0

JAKIM. (2020a). Manual Prosedur Pensijilan Halal Malaysia (Domestik) 2020. Jabatan Kemajuan Islam Malaysia. https://www.halal.gov.my/?data=bW9kdWxlcy9jb2xsYXBzaWJsZV9jb250ZW50Ozs7Ow%3D%3D&utama=panduan

JAKIM. (2020b). Sistem Pengurusan Halal Malaysia 2020. Jabatan Kemajuan Islam Malaysia. https://www.islam.gov.my/uploads/LAPORAN-TAHUNAN-JAKIM-2020-latest.pdf

Jang, S., Lee, H., Kim, Y., Lee, D., Shin, J., & Nam, J. (2024). When, what, and how should generative artificial intelligence explain to users? Telematics and Informatics, 93, Article 102175. https://doi.org/10.1016/j.tele.2024.102175

Janssen, M. (2025). Responsible governance of generative AI: Conceptualizing GenAI as complex adaptive systems. Policy and Society, 44(1), 38-51. https://doi.org/10.1093/polsoc/puae040

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), Article 248. https://doi.org/10.1145/3571730

Jiang, C., Qi, B., Hong, X., Fu, D., Cheng, Y., Meng, F., Yu, M., Zhou, B., & Zhou, J. (2024). On large language models' hallucination with regard to known facts. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Vol. 1, pp. 1041-1053). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.naacl-long.60

Kamali, M. H. (2008). Shariah law: An introduction. Oneworld Publications.

Kim, J., Maathuis, H., & Sent, D. (2024). Human-centered evaluation of explainable AI applications: A systematic review. Frontiers in Artificial Intelligence, 7, Article 1456486. https://doi.org/10.3389/frai.2024.1456486

Laudon, K. C., & Laudon, J. P. (2022). Management information systems: Managing the digital firm (17th ed.). Pearson.

Leon, M. (2026). Lifecycle-based governance to build reliable ethical AI systems. Systems Research and Behavioral Science, 43(3), 1116-1132. https://doi.org/10.1002/sres.70014

Lucas, S. W. (2004). Principles of hadith literature. In J. D. McAuliffe (Ed.), Encyclopedia of the Quran (Vol. 4, pp. 376-397). Brill.

MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. Journal of Marketing, 75(4), 136-154. https://doi.org/10.1509/jmkg.75.4.136

Malaysia National AI Office. (n.d.). National AI governance. National AI Office Malaysia. Retrieved 2026, from https://ai.gov.my/

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 1906-1919. https://doi.org/10.18653/v1/2020.acl-main.173

McNutt, M. K., Bradford, M., Drazen, J. M., Hanson, B., Howard, B., Jamieson, K. H., ... & Verma, I. M. (2018).

Transparency in authors' contributions and responsibilities to promote integrity in scientific publication. Proceedings of the National Academy of Sciences, 115(11), 2557-2560. https://doi.org/10.1073/pnas.1715374115

Ministry of Science, Technology and Innovation (MOSTI). (2024). National Guidelines on AI Governance and Ethics. Government of Malaysia. https://mastic.mosti.gov.my/storage/2024/09/THE-NATIONAL-GUIDELINES-ON-AI-GOVERNANCE-ETHICS.pdf

NIST. (2023). Artificial intelligence risk management framework (AI RMF 1.0). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. AI & Society, 39, 1871-1882. https://doi.org/10.1007/s00146-023-01635-y

OECD. (2019). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449; amended 2024). OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

OECD. (2022). Building trust in AI: OECD principles in practice. OECD Publishing.

OECD. (2024). Recommendation of the Council on Artificial Intelligence (updated 3 May 2024). OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Pan, B., Stakhanova, N., & Ray, S. (2023). Data provenance in security and privacy. ACM Computing Surveys, 55(14s), Article 323. https://doi.org/10.1145/3593294

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. Journal of Strategic Information Systems, 34(2), Article 101885. https://doi.org/10.1016/j.jsis.2024.101885

Riaz, M. N., & Chaudry, M. M. (2004). Halal food production. CRC Press.

Schneider, J. (2024). Explainable generative AI (GenXAI): A survey, conceptualization, and research agenda. Artificial Intelligence Review, 57, Article 289. https://doi.org/10.1007/s10462-024-10916-x

Stranas KA [Indonesian National AI Strategy Team]. (2020). National strategy for artificial intelligence: Indonesia. Ministry of Research and Technology/National Research and Innovation Agency of the Republic of Indonesia.

Taeihagh, A. (2021). Governance of artificial intelligence. Policy and Society, 40(2), 137-157. https://doi.org/10.1080/14494035.2021.1928377

Taeihagh, A. (2025). Governance of generative AI. Policy and Society, 44(1), 1-22. https://doi.org/10.1093/polsoc/puaf001

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. Electronic Markets, 31(2), 447-464. https://doi.org/10.1007/s12525-020-00441-4

Ulnicane, I. (2025). Governance fix? Power and politics in controversies about governing generative AI. Policy and Society, 44(1), 70-84. https://doi.org/10.1093/polsoc/puae022

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. United Nations Educational, Scientific and Cultural Organization. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

Veale, M., & Borgesius, F. Z. (2021). Demystifying the draft EU Artificial Intelligence Act. Computer Law Review International, 22(4), 97-112. https://doi.org/10.9785/cri-2021-220402

Wang, Y., Wang, M., Manzoor, M. A., Liu, F., Georgiev, G. N., Das, R. J., & Nakov, P. (2024). Factuality of large language models: A survey. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language

Processing (pp. 19519-19529). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.1088

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii-xxiii. https://doi.org/10.2307/4132319

World Economic Forum (WEF). (2023). Digital trust: Advancing consumer rights in the data economy. World Economic Forum. https://www.weforum.org/publications/

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

© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) license (http://creativecommons.org/licenses/by/4.0/).