

Knowledge Lineage from Isnad to AI: Reframing Authorship and Responsibility in the Generative AI Era

Dr. Amaal Mohammad Khalil Alshaar

Doctor of Business Administration, Independent Researcher, Amman-Jordan

Email: amaal.shaar@gmail.com

DOI Link: <http://dx.doi.org/10.6007/IJARBSS/v15-i9/26451>

Published Date: 20 September 2025

Abstract

Generative artificial intelligence (AI) is rapidly transforming scholarly writing and research, raising urgent questions about authorship, intellectual ownership, and epistemic responsibility. This study aims to investigate how knowledge lineage—defined as the traceable chain of transmission, transformation, and attribution—connecting ideas to their origins—can be preserved in an era where AI increasingly mediates knowledge production. Adopting a theoretical–conceptual methodology, the paper integrates Western philosophy (Foucault’s author-function, Kuhn’s paradigm shifts), Floridi’s concept of ectypes, and Islamic epistemology (isnād, amāna, mas’ūliyya), alongside comparative analysis of institutional practices at Harvard, Oxford, and King Saud University. The analysis reveals that while AI accelerates content creation, it simultaneously introduces epistemic opacity, breaks attribution chains, and generates “orphaned” knowledge outputs that lack identifiable provenance. To address these risks, the study proposes a multi-tiered ethical framework that includes mandatory AI-use disclosure, human-in-the-loop verification, metadata-based “AI isnād,” and institutional accountability mechanisms. This framework bridges global AI ethics (OECD, UNESCO) with culturally embedded models of moral responsibility, reinforcing that human scholars must remain the ethical custodians of knowledge. The paper concludes that preserving knowledge lineage is not merely a technical challenge but a moral and cultural imperative. By embedding transparency, traceability, and accountability into both technical systems and institutional policies, the integrity of scholarly communication can be sustained in an AI-saturated future.

Keywords: Knowledge Lineage, Epistemic Responsibility, Generative Ai, Hybrid Authorship, Islamic Epistemology, Isnād, Authorship, Intellectual Integrity

Introduction

In the pre-AI academic landscape, knowledge production was grounded in clear traditions of authorship, intellectual labor, and scholarly lineage. Ideas could be traced to identifiable contributors; references served as epistemic signposts; and academic integrity depended on

verifiable attribution. Authorship was not merely about claiming credit—it was an ethical commitment to truth, originality, and intellectual responsibility (Foucault, 1977; Brown, 2014; 2018).

Today, however, generative artificial intelligence has introduced a profound disruption to this ecosystem. Tools such as ChatGPT and other large language models (LLMs) can now produce fluent, persuasive content at unprecedented speed. Yet, behind this linguistic fluency lies what Humphreys (2009) described as epistemic opacity—the condition in which a cognitive agent cannot fully access the epistemically relevant elements of a system. Zednik (2021) extends this to machine learning, framing it as the “black box problem,” where even experts lack the means to fully explain how particular outputs are produced. In the case of generative AI, this opacity means that while texts may appear authoritative, their epistemic foundations remain hidden, destabilizing traditional markers of accountability and authorship.

This disruption resonates with Floridi’s (2018a) account of AI-generated *ectypes*, which he distinguishes from authentic artifacts of human epistemic labor. It also intersects with Barthes’s (1967) and Foucault’s (1977) critiques of authorship. Barthes famously declared the “death of the author,” suggesting that meaning is created in the interplay of text and reader, while Foucault conceptualized the author-function as a socio-legal mechanism for organizing discourse and assigning responsibility. Generative AI intensifies these tensions: by producing outputs divorced from individual intentionality, it blurs the line between creation, synthesis, and imitation. As a result, what once served as markers of originality and responsibility are at risk of dissolution in an AI-mediated epistemic landscape.

This study responds to the current epistemic rupture by advancing a conceptual and ethical framework aimed at restoring academic responsibility in an era increasingly shaped by AI. It draws on Floridi’s (2019) conception of philosophy as conceptual design—a mode of inquiry that foregrounds epistemic responsibility and the protection of informational integrity within the infosphere—and on Kuhn’s (1997) account of paradigm shifts as disruptive breaks that restructure the foundations of knowledge. Within this synthesis, knowledge lineage is positioned as a central analytical lens, highlighting that safeguarding the traceable ancestry of ideas is essential to sustaining intellectual trust, ensuring transparency, and preserving the human significance of scholarly work at a time when authorship and authority are destabilized.

Methodologically, the study adopts a theoretical-conceptual approach, combining philosophical reflection, ethical reasoning, and cross-cultural analysis. Rather than focusing on empirical data or technical metrics, it investigates the foundations and future of epistemic integrity. The paper draws on Western and Islamic epistemologies (Brown, 2014; 2018), explores institutional and cultural practices (Lloyd, 1996; UNESCO, 2021), and ultimately proposes practical guidelines for researchers, educators, and academic institutions.

By situating the rise of generative AI within deeper debates about authorship, responsibility, and knowledge transmission, this paper calls for a renewed academic ethos—one that embraces technological innovation without surrendering the moral and epistemic anchors upon which scholarship depends.

Literature Review

Roots of Knowledge Lineage across Traditions: The notion of knowledge lineage—its ancestry, transmission, and ethical grounding—has been a persistent concern across diverse philosophical and cultural traditions. Regardless of the terminology used, the core principle remains intact: knowledge is never created in isolation, but inherited, contextualized, and ethically stewarded across generations.

In Western epistemology, Thomas Kuhn (1997) famously argued that scientific progress does not unfold gradually and linearly, but through paradigm shifts—radical ruptures in thought whereby one dominant worldview is replaced by another, reshaping both scientific reasoning and the wider intellectual and cultural contexts in which science is situated. These discontinuities redefine what counts as valid knowledge. Michel Foucault (1977) extended this perspective by exploring how knowledge is embedded within discursive and institutional structures. He introduced the concept of the “author-function” as a mechanism through which authority and legitimacy are conferred in discourse. In both frameworks, knowledge is situated within power networks, discursive norms, and institutional arrangements.

In the Islamic epistemic tradition, the integrity of knowledge is deeply tied to its *isnād*—a rigorously authenticated chain of narration. This *isnād* does more than transmit content; it ensures accountability, verifies the character of the transmitter, and preserves trust across generations (Brown, 2014; 2018). In this vision, the scholar is not merely a thinker but a custodian: entrusted with guarding the accuracy of inherited knowledge and preserving it with sincerity (*ikhhlās*) and responsibility (*mas’ūliyya*).

The comparative study of ancient Greek and Chinese traditions, as emphasized by G. E. R. Lloyd (1996), further sheds light on how institutional contexts shape conceptions of authority and audience. In ancient Greece, intellectual life thrived on adversarial debate, with authors asserting credibility through contestation and public reasoning. By contrast, ancient Chinese science—particularly under the Han dynasty—operated within bureaucratic and imperial frameworks, where contributions were assessed based on ethical alignment and state-sanctioned consensus rather than argumentative dominance.

Luciano Floridi (2018a) brings this discussion into the digital age with his notion of *ectypes*—reproductions that resemble originals but lack epistemic depth. Generative AI, through its algorithmic synthesis, often produces such *ectypes*: outputs that appear authoritative yet lack intentionality, traceability, and epistemic grounding. This challenges the foundations of authenticity and intellectual continuity.

Together, these diverse traditions illustrate that knowledge is never a neutral abstraction. It is shaped by cultural values, institutional contexts, ethical convictions, and discursive practices. Authorship, therefore, transcends the technical act of producing text—it embodies a moral position within a broader epistemic and ethical lineage.

In the age of generative AI—where authorship risks dissolving and sources may become blurred—these traditions offer essential lessons. If we lose sight of the chains that anchor knowledge in trust, responsibility, and human judgment, we risk eroding the foundations of

scholarly integrity. The solution is not to reject technology, but to re-embed it within ethical and epistemic traditions that have long upheld the value of knowledge.

Conceptual Framework: The conceptual framework of this study provides the definitional clarity needed to address the complex issues of knowledge lineage and epistemic responsibility in the age of generative artificial intelligence (AI). It identifies and explains the core concepts that recur throughout the analysis, ensuring consistency and precision in their use.

Knowledge Lineage: The technical notion of data lineage—defined as the process of tracing warehouse data items back to their original sources and reconstructing the transformations applied along the way (Cui & Widom, 2003)—offers a powerful metaphor for understanding knowledge lineage in scholarly contexts. Just as data provenance ensures transparency and accountability in information systems, the intellectual ancestry of ideas must remain traceable to safeguard epistemic integrity. Expanding this metaphor, *research lineage* captures the emergence and evolution of scientific ideas through citation networks. Ghosal et al. (2021) introduce the concept of significant citations—those that meaningfully build upon prior work—demonstrating how filtering for significance enables automated mapping of idea trajectories and illuminates knowledge progression within a field. Building on these models, knowledge lineage can be defined as the traceable chain of transmission, transformation, and attribution that connects a piece of knowledge to its origin. This encompasses both the historical pathways and the institutional mechanisms—citation, peer review, archival preservation—that uphold credibility. In the Islamic intellectual tradition, *isnād* (the meticulous documentation of transmitters) represents a highly formalized and ethically grounded form of knowledge lineage, ensuring authenticity and accountability in the transmission of knowledge (Brown, 2014; 2018).

Epistemic Responsibility: Epistemic responsibility denotes the ethical and intellectual duty of knowledge producers, transmitters, and users to uphold accuracy, reliability, and transparency. It requires agents—whether individual or institutional—to be accountable for the truthfulness of their claims and the integrity of their sources (Code, 1987). In the contemporary context of AI, epistemic responsibility extends to those who design, train, and deploy AI systems, as well as to the human authors who integrate AI outputs into their work. This broader view underscores that accountability for knowledge cannot be outsourced to machines but remains a fundamentally human obligation.

Epistemic Agency: Although Fricker (2007) does not explicitly define epistemic agency, her account of testimonial and hermeneutical injustice presupposes it as the capacity of individuals to act as knowers—giving testimony, interpreting experience, and contributing to collective understanding. Reed (2001) emphasizes that such agency is grounded in intellectual virtues—honesty, diligence, and careful judgment—that enable responsible participation in acquiring, evaluating, and transmitting knowledge. By contrast, generative AI lacks the intentionality, understanding, and ethical grounding required for such agency. As Zednik (2021) argues, machine learning systems are marked by epistemic opacity: they generate outputs through statistical pattern recognition rather than autonomous reasoning, and remain fundamentally dependent on human-authored prompts and oversight. Accordingly, while AI can serve as a powerful tool for recombination and acceleration, it cannot qualify as

an epistemic agent; responsibility for knowledge production ultimately resides with human actors.

Generative Artificial Intelligence: Generative AI refers to a class of machine learning models—most prominently large language models (LLMs) and other foundation models—designed to create new content such as text, images, audio, or video. Trained on vast datasets through self-supervised learning, these systems generate outputs by probabilistically predicting the next element in a sequence or by synthesizing patterns across modalities (Bommasani et al., 2021). Although their outputs often appear original and human-like, the generative process is fundamentally derivative, relying on pre-existing data and statistical correlations. As a result, generative AI remains susceptible to bias, error, and opacity, raising persistent questions about authenticity, reliability, and epistemic trust.

Originality: Originality in the scholarly sense entails the creation of ideas, arguments, or expressions that are novel in substance rather than merely in form. In the AI context, the notion of originality becomes problematic: AI-generated text may appear unique, but because it is structurally dependent on human-authored inputs and pre-existing data, its originality is only formal rather than substantive (Floridi, 2018a). This complicates issues of ownership, attribution, and the ethical weight of originality in an era of automated recomposition.

Attribution: Attribution refers to the practice of identifying and acknowledging the sources or origins of ideas, contributions, or works, ensuring that credit is assigned where it is due. In academic contexts, attribution serves both ethical and epistemic purposes—preventing plagiarism, maintaining intellectual integrity, and enabling verification of claims. In the context of generative AI, however, attribution becomes increasingly complex. Outputs emerge from vast and often untraceable datasets and involve hybrid human–machine processes, raising the question of who should be credited—the user who prompts the system, the developers who design and train it, the creators of the training data, or the model itself. This difficulty reflects what Khosrowi et al. (2023) call the many-hands problem, where multiple actors are implicated in the creation of AI outputs. To address this, Khosrowi and colleagues propose *collective-centered creation* (CCC), a framework that treats generative AI outputs as products of collective human–AI activity rather than the work of a single author. In sum, attribution in the AI era requires moving beyond conventional norms to account for distributed authorship, ensuring that continuity, transparency, and ethical grounding are preserved within the broader ecosystem of knowledge lineage.

Theoretical Grounding and Framework: Authorship has undergone profound transformations throughout intellectual history. In early traditions, it was often tied to divine inspiration or collective cultural expression rather than personal creativity. The rise of print culture and Enlightenment thought reframed authorship as a marker of originality, intellectual property, and legal rights (Woodmansee & Jaszi, 1994). In modern scholarship, this shift cemented the link between authorship and individual responsibility for ideas.

Critical theory later destabilized this view. Roland Barthes (1967), in *The Death of the Author*, argued that meaning is not dictated by authorial intent but emerges through reader interpretation. Michel Foucault (1977), in *What Is an Author?*, reframed authorship as an “author function”—a discursive role that regulates legitimacy, authority, and the circulation

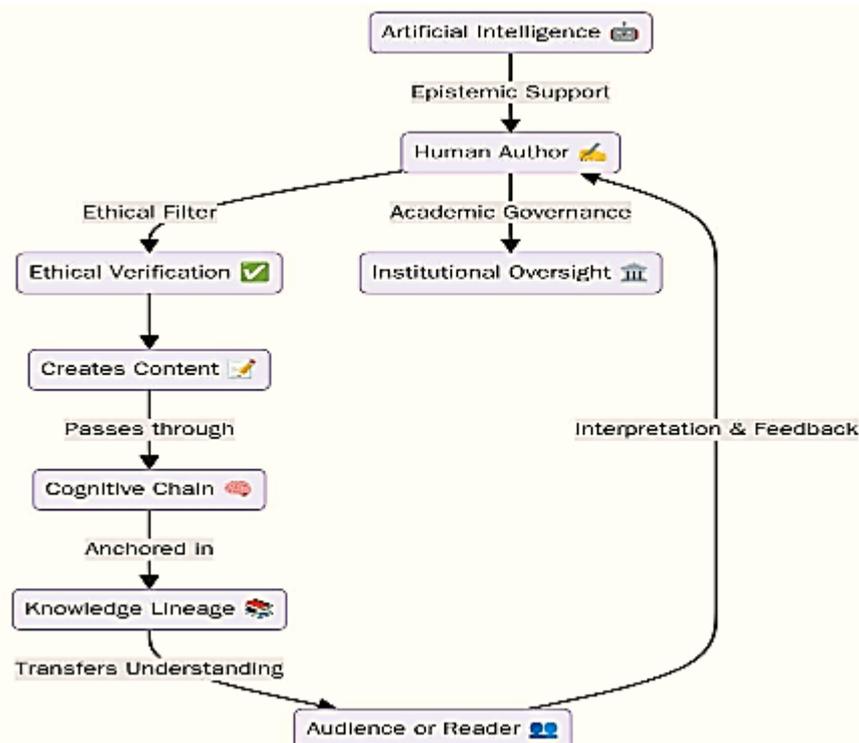
of knowledge. While Barthes questioned the centrality of the author altogether, Foucault emphasized the systems of responsibility and attribution that underpin scholarly discourse. In the generative AI era, these perspectives gain renewed relevance: when text can be produced without human cognition or intent, the legitimacy and authority once embedded in authorship become unsettled.

This study integrates three theoretical anchors to interpret these changes. **First**, Foucault's author function frames authorship not as the property of an individual but as a socio-cultural role shaped by institutional and discursive rules. Applied to AI, this shifts the locus of responsibility away from the machine—which lacks intentionality and moral agency—and toward human and institutional actors who design, deploy, and regulate AI systems.

Second, Floridi's (2018a) notion of *ectypes* describes derivative informational artifacts composed from pre-existing materials rather than originating de novo. Generative AI outputs are quintessential *ectypes*: probabilistic recombinations of prior data, devoid of epistemic intent. This challenges traditional notions of originality and underscores the need for human oversight to preserve the integrity of knowledge lineage.

Third, Zednik's (2021) analysis of epistemic opacity, though not directly addressing epistemic agency, highlights the structural limits of AI systems in justifying their outputs. Generative AI cannot justify its results or assume responsibility for their validity. Instead, such systems operate by detecting statistical patterns in training data and generating responses to prompts, underscoring the importance of transparency in mitigating the Black Box Problem. Across intellectual traditions, epistemic responsibility—the moral and intellectual duty to ensure that knowledge claims are accurate, verifiable, and ethically grounded (Code, 1987)—remains a central principle. In the Islamic scholarly tradition, this duty is formalized in the practice of *isnād*, a rigorously documented chain of transmission that safeguards the integrity of reported knowledge (Brown, 2014; 2018). Here, the transmitter functions as a moral gatekeeper, accountable for both factual accuracy and ethical implications. This principle resonates with Western norms of citation, replicability, and scholarly accountability: in both traditions, knowledge is not a neutral artifact but a moral act situated within a lineage of responsibility.

Figure 1. Human–AI Knowledge Lineage and Epistemic Responsibility Framework: As foundational principles of epistemic responsibility, authorship, and knowledge lineage can be reinterpreted in AI-assisted scholarship. While artificial intelligence can enhance linguistic fluency and extend informational reach, the human author remains the central and irreplaceable agent—responsible for ethically filtering AI outputs, verifying their accuracy, and embedding new contributions within an established chain of knowledge. Academic governance and ethical verification work in parallel to safeguard scholarly integrity, while interpretation, critique, and intellectual dialogue remain distinctly human responsibilities. Together, these elements ensure that the epistemic ecosystem retains continuity, credibility, and ethical grounding in a hybrid human–AI environment.



Knowledge Disruptions in the AI Era: The advent of generative artificial intelligence (AI) has introduced structural and epistemic disruptions that challenge long-standing norms of knowledge lineage and attribution. These disruptions arise not merely from the speed and scale of AI-mediated content production but from the opacity, instability, and derivative nature of the outputs themselves.

Opacity and the Erosion of Transparency: One of the most significant epistemic disruptions lies in the opacity surrounding data origins. In traditional scholarship, every claim can be traced through explicit citations to identifiable sources, enabling verification, accountability, and ethical review. Generative AI, by contrast, functions as a probabilistic synthesizer—producing text from vast, often undisclosed datasets—without leaving a clear trail back to the original sources (Worth et al., 2024). This shift replaces explicit attribution with inscrutable computational processes, severing the traceability central to knowledge lineage.

Recent empirical evidence shows that the level of transparency in artificial intelligence models remains extremely limited. A recent study (Worth et al., 2024) found that, on average, these models scored only two out of ten on transparency indicators, while 40% scored zero. Even when transparency reports were available, they typically provided only general information, such as dataset size, cleaning procedures, or harmful data filtering, without addressing fundamental issues like licensing, data provenance, or whether copyrighted or personal data were included. More critically, over half of the analyzed models did not disclose their internal architecture, making any meaningful assessment of transparency impossible. These gaps are further compounded by commercial restrictions, cost concerns, privacy risks, and technical complexity, which leave transparency documentation fragmented, superficial, and difficult to access. As a result, the foundations of scientific reproducibility, informed critique, and ethical accountability in knowledge production face a serious challenge.

Longpre et al. (2024) identify a parallel crisis in dataset transparency. Many widely used dataset collections are treated as monoliths, obscuring their internal lineage, licensing heritage, and source diversity. Over 70% of dataset licenses on repositories such as GitHub and Hugging Face are unspecified, and two-thirds of Hugging Face licenses are miscategorized—often labeled more permissively than their original terms, a practice described as “license laundering.” This erosion of provenance undermines both attribution and legal clarity, heightening risks of copyright violations, privacy breaches, and systemic biases. Tools like the Data Provenance Explorer (DPEXplorer), introduced by Longpre and colleagues, demonstrate that systematic audits can improve transparency, reducing unspecified licenses from 72% to 30% and documenting dataset origins with greater clarity. From the perspective of knowledge lineage, these patterns represent a structural break. If the origins, transformations, and licensing conditions of training data remain invisible, then the epistemic chain linking present claims to their intellectual ancestors is severed. This loss of transparency weakens the ethical foundations of scholarship and accelerates the drift toward what may be called *epistemic orphanhood*—a condition in which ideas circulate without identifiable heritage, and thus without a locus of responsibility. Addressing this disruption will require a combination of technical interventions, institutional policy reforms, and legal clarity to ensure that transparency remains a non-negotiable element of AI-mediated knowledge production.

Algorithmic Bias and Epistemic Distortion: AI models—particularly large language models (LLMs)—inherit and often amplify biases embedded in their training data. This data is typically drawn from vast, uncurated internet sources that, as Bender et al. (2021) note, overrepresent dominant viewpoints and encode harmful stereotypes. As a result, models may reproduce derogatory, exclusionary, or otherwise harmful associations tied to gender, race, ethnicity, and other social categories. These biases distort generated content—privileging certain narratives while marginalizing others—and reshape the knowledge ecosystem in ways that are neither neutral nor accountable. In epistemic terms, such distortions compromise the representational fidelity of knowledge, displacing lineages of thought from their authentic origins. The consequences are tangible: bias amplification can reinforce stereotypes, perpetuate systemic inequalities, and cause discrimination or psychological harm. Such harms may appear subtly, through framing effects or microaggressions, or overtly, through abusive or toxic outputs. Mitigating these risks requires deliberate interventions, including careful data curation, robust documentation, and inclusive stakeholder engagement, to ensure AI systems are both responsible and epistemically sound.

Breakage of Attribution Chains: The absence of direct, verifiable references in AI outputs leads to the fragmentation of attribution chains. Without an *isnād*-like mechanism to document provenance, readers cannot assess the credibility of content or its sources. This “broken lineage” problem mirrors the epistemic gaps caused by lost or unreliable transmitters in historical chains of narration; yet in the AI context, such gaps are systemic rather than exceptional. Longpre et al. (2024) describe this as a crisis in dataset documentation and provenance, noting that AI systems are often trained on vast, uncurated corpora with little or no traceable sourcing. They call for robust provenance standards and systematic auditing tools as prerequisites for ethical and accountable AI, warning that without such measures, the reliability of AI-mediated knowledge will remain compromised.

Proliferation of Unverified Content: Generative AI accelerates content production at an unprecedented scale, dramatically increasing the circulation of unverified, misleading, or outright false information. Westerlund (2019) warns of an emerging “information apocalypse,” in which the proliferation of deepfakes and synthetic media overwhelms audiences to the point that they distrust all media—including legitimate sources. This dynamic threatens not only academic and journalistic verification mechanisms but also the broader epistemic integrity of the public sphere, where distinguishing truth from fabrication becomes increasingly unstable.

Implications for Knowledge Stewardship: The proliferation of *ectypes*—convincing copies or imitations generated by digital technologies—poses profound challenges for knowledge stewardship. As Floridi (2018a) argues in his discussion of artificial intelligence, deepfakes, and the future of ectypes, these technologies produce artefacts that appear authentic yet are detached from their epistemic origins, undermining the trustworthiness of the knowledge ecosystem. This detachment complicates the role of knowledge stewards, who must now deploy advanced tools such as AI-driven verification and blockchain-based provenance systems to preserve trust. Stewardship increasingly requires safeguarding not only the artefact itself but also the full context of its creation—methods, materials, and technologies—so that interpretation remains accurate. As ectypes blur boundaries between original and reproduction, fostering digital literacy and developing ethical and legal frameworks to prevent misuse become essential. Without proactive, adaptive, and ethically grounded interventions, the normalization of provenance-free artefacts risks eroding the very conditions necessary for epistemic responsibility.

Comparative Insights and Contextual Practice: As academic institutions worldwide grapple with the integration of generative AI in scholarly work, varied practices and policy responses have emerged. This section offers a comparative perspective on how leading universities are handling AI-assisted writing, highlighting common principles and contextual differences to inform policy development and foster a globally responsible research culture.

Institutional Policies on Disclosure: Harvard Extension School requires students to acknowledge any permitted use of generative AI, warning that unacknowledged submissions may constitute breaches of academic integrity (Harvard Extension School, n.d.). More broadly, Harvard encourages responsible experimentation but leaves policy decisions to individual schools and faculty, with attention to data privacy, copyright, and academic honesty (The Harvard Crimson, 2023).

The University of Oxford issues university-wide guidance emphasizing disclosure and “academic judgment” while allowing departments autonomy. Students are expected to demonstrate critical engagement with AI-generated content and disclose its use when authorized, rather than substituting AI output for independent scholarly reasoning (University of Oxford, 2024).

King Saud University formalizes its approach mainly through academic publishing: the *Journal of King Saud University – Science* prohibits listing AI tools as authors and requires disclosure of AI-assisted writing, holding human authors fully responsible in line with COPE standards (Journal of King Saud University – Science, n.d.).

Detection Tools and Technological Measures: Research on AI detection highlights both promise and limitations. Al-Smadi (2025) developed transformer-based models—ELECTRA-Large for English and AraELECTRA for Arabic—enhanced with stylometric features, achieving F1-scores of 99.7% and 98.4% in distinguishing human- from AI-generated essays. Yet such advances have not led to widespread institutional adoption.

Commercial tools remain inconsistent. Broader evaluations show many detectors fail when texts are paraphrased, translated, or multilingual (Weber-Wulff et al., 2023). Sağlam and Schmid (2025) note that plagiarism detection systems perform poorly against AI-driven obfuscation. Several U.S. universities—including Vanderbilt, Michigan State, and Northwestern—suspended the use of Turnitin’s AI Writing Indicator to avoid false accusations (Petro, 2025).

The University of Pittsburgh has consistently cautioned faculty against treating AI detection results as definitive evidence. Its 2025 guidance stresses that current tools lack the accuracy needed for academic integrity investigations, echoing its earlier 2023 refusal to endorse Turnitin’s AI detection feature (University of Pittsburgh Teaching Center, 2023, 2025). Empirical studies reinforce these concerns: Perkins et al. (2023) found that while Turnitin flagged 91% of GPT-4-generated essays, it correctly identified only 54.8% of AI-origin content.

Emerging Lessons and Global Trends: Despite contextual differences, several principles are converging. Disclosure is increasingly recognized as an ethical baseline, critical thinking and human authorship remain central values, and institutions are seeking value-based frameworks rather than relying solely on technical detection. Harvard mandates disclosure but has paused some detection pilots over reliability concerns. Oxford emphasizes disclosure and critical judgment while preserving departmental flexibility, reflecting a commitment to academic freedom. King Saud University underscores human accountability and is exploring Arabic-focused detection research. The University of Pittsburgh prioritizes governance and values-based integrity frameworks, declining to adopt unreliable detection tools (Dotan et al., 2024).

Implications for Policies and Education: Three priorities emerge for policy and pedagogy: transparent disclosure of AI use, cultivation of critical thinking, and flexibility that respects disciplinary and cultural contexts. AI contributions must be openly acknowledged, but intellectual responsibility remains with the human author. Educational strategies should go beyond technical training to foster ethical awareness and shared accountability for outcomes. Epistemic responsibility is thus a global concern. Institutions can combine Western academic traditions with Islamic epistemic values such as *amanah* (trust) and *mas’ūliyya* (responsibility), creating inclusive policies that protect integrity and ensure humans remain central to knowledge production.

Methodology

This study adopts a theoretical–conceptual research design, as its primary objective is to examine the ethical, epistemic, and institutional dimensions of knowledge lineage in the era of generative artificial intelligence (AI). Rather than collecting empirical data or testing hypotheses, the paper synthesizes insights from multiple intellectual traditions to develop a normative framework for preserving epistemic integrity.

The methodology integrates three streams of analysis. **First**, it draws on Western philosophy of science and knowledge—particularly Kuhn’s paradigms, Foucault’s author-function, and Floridi’s information ethics—to trace the evolution of authorship, attribution, and epistemic responsibility. **Second**, it engages with the Islamic epistemic tradition, where the practice of isnād serves as a historically grounded model for verifiable transmission and moral accountability. **Third**, it incorporates comparative policy analysis by examining institutional responses to AI-assisted authorship in leading global universities, offering a cross-cultural view of emerging governance practices.

This combined approach is normative–analytical: it not only diagnoses the epistemic and ethical disruptions posed by generative AI but also formulates practical proposals, including the concept of an “AI isnād layer,” to ensure that knowledge remains traceable, ethically grounded, and institutionally accountable.

Findings and Discussions

Islamic Epistemological Integration: The Islamic intellectual tradition offers a historically grounded and ethically robust framework for addressing the challenges of knowledge lineage in the age of generative artificial intelligence (AI). At its core lies a commitment to the meticulous documentation, transmission, and verification of knowledge—a commitment crystallized in the practice of isnād, which serves as both an epistemic safeguard and a moral discipline.

Isnād as a Moral–Epistemic Model: In the Islamic scholarly tradition, isnād (chain of transmission) is considered the cornerstone of hadith authentication and a safeguard for epistemic integrity. As the Yaqeen Institute (n.d.) emphasizes, “The isnād is part of religion; if not for the isnād, whoever wanted could say whatever they wanted.” This principle underscores that knowledge is not valid unless its lineage is traceable through trustworthy transmitters. Early Muslim scholars meticulously assessed each narrator’s reliability, memory, piety, and character, a process that Abdulrahman (2024) describes as “a profound manifestation of faith and responsibility itself.” Similarly, hadith sciences affirmed that the authenticity of a report depended not only on textual corroboration but also on the moral uprightness (‘adāla) and competence (ḍabt) of transmitters, where defects in either could invalidate transmission (Hadith Studies, n.d.). Extending this perspective, Bouhafa (2021) demonstrates that Islamic philosophy embedded a broader moral ontology in which epistemology was inseparable from ethics, showing that the integrity of knowledge hinged on the moral accountability of its transmitters.

Traceability and Accountability: As ‘Abd Allāh ibn al-Mubārak famously stated, later cited in the introduction to Muslim’s Ṣaḥīḥ, “The isnād is part of religion; if not for the isnād, whoever wished could say whatever they wished.” This maxim highlights that knowledge is not considered valid unless its lineage is demonstrably traceable. To secure this, scholars developed ‘ilm al-rijāl (biographical evaluation of transmitters), which examined narrators’ reliability, memory, and moral character, establishing a rigorous system of accountability (Brown, 2014; 2018). As Motzki (2002) observes, isnād functioned not only as a record of provenance but also as a mechanism of quality control, enabling recipients to evaluate both the continuity of transmission and the reliability of knowledge conveyed. More broadly, this system embedded epistemology within an ethical framework, where the moral uprightness

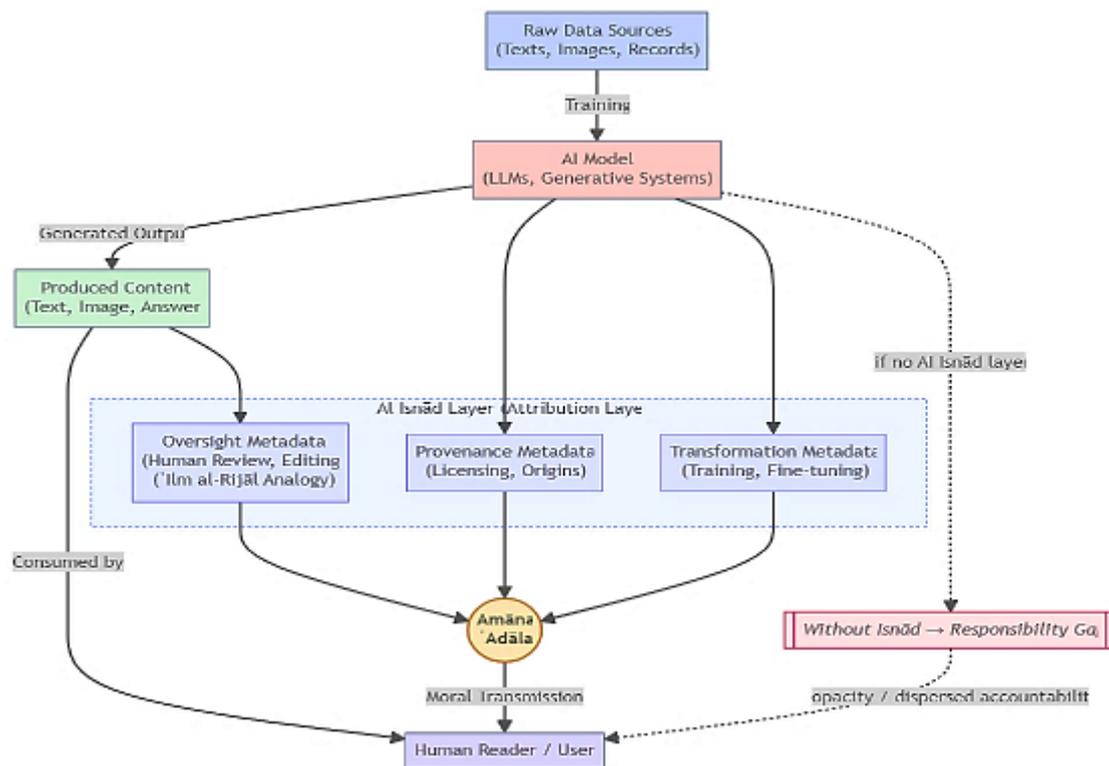
(‘adāla) and precision (ḍabṭ) of transmitters were inseparable from the legitimacy of the knowledge transmitted (Bouhafa, 2021).

Ethical Agency and Collective Responsibility: In the Islamic scholarly tradition, epistemic responsibility was both collective and personal. The community of scholars (‘ulama’) collectively safeguarded the integrity of transmitted knowledge, while each transmitter bore individual accountability before God and society for the truthfulness of what they conveyed. This dual accountability—spiritual and social—was embedded in hadith sciences through principles such as ‘adāla and ḍabṭ, which determined the acceptability of a narrator (Brown, 2018; Motzki, 2002). Bouhafa (2021) further shows that Islamic philosophy entwined epistemology with ethics, affirming that knowledge could not be separated from the moral integrity of its conveyors. The isnād thus functioned not merely as a technical chain of names but as a moral–epistemic framework in which truth was preserved through accountability at every link.

By contrast, AI-mediated systems often disperse responsibility so widely that it becomes effectively absent. As Mittelstadt et al. (2016) observe, algorithmic systems create “responsibility gaps” because their opacity, autonomy, and distributed agency obscure clear lines of accountability. Floridi (2018b) likewise argues that digital technologies require frameworks of “soft ethics” precisely because responsibility is diffused across designers, deployers, and users. Unlike the isnād’s structure of traceability and moral responsibility, generative AI outputs frequently lack both identifiable provenance and individual accountability, raising profound ethical concerns about integrity and trust in knowledge production.

Application to AI-Mediated Knowledge: Although 1142snad originated in a pre-digital context, its principles provide a valuable framework for reimagining governance in the age of artificial intelligence. An “AI 1142snad” could take the form of embedding transparent metadata within outputs to document training sources, transformation processes, and human oversight. Such a system would restore traceability and accountability, echoing the Islamic scholarly conception of knowledge as an amāna (trust) that must be preserved with ‘adāla (integrity) (Brown, 2018; Bouhafa, 2021). This vision also directly responds to contemporary debates in AI ethics regarding opacity and the “responsibility gap” that arises when accountability is dispersed across complex algorithmic systems (Floridi, 2018b; Mittelstadt et al., 2016). By synthesizing these traditions, AI-mediated content production can be aligned with both historical epistemic safeguards and modern demands for transparency, ensuring that knowledge remains a morally accountable trust rather than a mere commodity.

Figure 2. AI Isnād Framework for Knowledge Attribution. This diagram shows how the Islamic principle of isnād can be reimagined for AI governance through an “AI isnād layer” that embeds provenance, transformation, and human-oversight metadata. By restoring traceability and ethical accountability, and by anchoring knowledge in the values of Amāna (trust) and ‘Adāla (integrity), the framework reinforces the foundations of legitimate knowledge. Transmitting this ethical anchor forward to the user, it directly addresses the responsibility gaps typical of contemporary algorithmic systems, thereby strengthening transparency, trust, and the understanding of knowledge as a moral trust rather than a mere commodity.



A Normative Counterpoint to Disruption: Juxtaposing the isnād model with the epistemic disruptions of AI reveals that technological sophistication alone cannot safeguard the integrity of knowledge lineage. What is required is an ethical architecture that embeds traceability, accountability, and moral responsibility into the very fabric of knowledge production. As Brown (2018) shows, the Islamic isnād tradition fused rigorous documentation with personal integrity, making moral accountability inseparable from epistemic legitimacy. Bouhafa (2021) likewise demonstrates that Islamic philosophy conceived epistemology as embedded within a broader moral ontology. By contrast, contemporary concerns in AI ethics—highlighted by Floridi (2018a, 2018b, 2019) and Mittelstadt et al. (2016)—underscore how opacity and “responsibility gaps” undermine trust and accountability in digital knowledge production. Taken together, these insights suggest that isnād offers not merely a historical precedent but a normative counterpoint to the challenges of AI, showing that integrity in knowledge requires both technical rigor and ethical responsibility.

Ethical Frameworks and Normative Proposals: Preserving knowledge lineage in the age of generative artificial intelligence (AI) requires a coherent ethical architecture that bridges universal principles with culturally embedded traditions such as the Islamic isnād. This section outlines key ethical frameworks relevant to AI-mediated knowledge, followed by normative proposals aimed at safeguarding epistemic integrity.

Global Ethical Frameworks: International bodies and academic institutions have advanced guidelines emphasizing transparency, accountability, fairness, and human oversight in AI. The OECD (2019) and UNESCO (2021) converge on the principle that AI outputs must be explainable, data sources traceable where possible, and human actors ultimately responsible for decisions. The European Union’s *Ethics Guidelines for Trustworthy AI* (EU High-Level Expert Group on AI, 2019) and IEEE’s *Ethically Aligned Design* (IEEE, 2019) reinforce these

commitments, while Jobin et al. (2019) highlight the gap between proliferating ethical guidelines and the limited mechanisms for their enforcement.

Islamic Ethical Parallels: The principles embedded in the isnād tradition—traceability, verification, personal accountability, and the preservation of knowledge as amāna (trust)—resonate with these global standards, while extending them by integrating moral responsibility into the act of transmission itself. Brown (2018) notes that isnād safeguarded epistemic integrity by ensuring that no report could be accepted without a reliable chain of transmission. Motzki (2002) further shows that hadith sciences developed rigorous critical tools such as *‘ilm al-rijāl* (biographical evaluation of transmitters) to assess both continuity and reliability. Bouhafa (2021) adds that Islamic philosophy fused epistemology with moral ontology, underscoring that to transmit knowledge without verification was seen as not merely a technical lapse but an ethical breach.

This moral imperative parallels the emphasis on human oversight in frameworks such as UNESCO (2021) and the OECD (2019), but the Islamic conception of amāna goes further: oversight is not just procedural but a spiritual and ethical responsibility borne by each participant in the chain of transmission. Thus, the isnād enriches contemporary AI governance by rooting abstract principles of transparency and accountability in a deeply tested historical practice of morally accountable knowledge stewardship.

Normative Proposals: Building on both global and Islamic frameworks, this study proposes the following measures for preserving knowledge lineage in AI-mediated environments:

1. **Mandatory Source Documentation:** Building on the principle of *explicitability* identified by Floridi and Cowls (2019), AI systems should embed transparent metadata indicating training sources, transformation processes, and human interventions. Such documentation would establish a form of “AI isnād,” enabling end-users to trace epistemic origins and support accountability.
2. **Human-in-the-Loop Verification:** AI-generated content intended for scholarly or public dissemination should undergo expert human review, with reviewers assuming epistemic responsibility for accuracy (OECD, 2019; UNESCO, 2021).
3. **Ethical Disclosure Statements:** Authors and institutions should disclose the role of AI in the research or writing process, specifying the tools used, their functions, and whether outputs were modified or verified (Jobin et al., 2019).
4. **Institutional Accountability Mechanisms:** Universities, publishers, and platforms should establish clear lines of responsibility for AI-mediated outputs, preventing accountability from being deflected onto “the algorithm” (Floridi & Cowls, 2019).
5. **Integration of Moral Responsibility Training:** Research ethics curricula should include modules on AI and epistemic responsibility, drawing on both global standards (UNESCO, 2021) and cultural-historical models such as isnād to instill both personal and institutional accountability (Brown, 2018; Bouhafa, 2021).

Towards a Unified Ethical Model: The convergence of universal AI ethics and the Islamic isnād tradition offers a compelling blueprint for epistemic governance. By embedding traceability, transparency, and moral responsibility into both the technical and institutional layers of AI-mediated knowledge production, it becomes possible to confront the disruptions identified earlier—such as opacity, broken attribution chains, and responsibility gaps—and to ensure

that the lineage of knowledge remains unbroken, grounded in both its informational structure and its ethical foundation.

Practical Framework for Researchers and Institutions: While ethical principles and normative proposals provide the philosophical and regulatory foundation for preserving knowledge lineage in the age of generative artificial intelligence (AI), they must be translated into actionable procedures to be effective. This framework outlines concrete measures for both individual researchers and institutions, ensuring that epistemic responsibility is maintained across the lifecycle of AI-mediated knowledge production.

For Researchers

- **Explicit AI Use Documentation:** All scholarly outputs should include a clear statement on whether AI tools were used, at what stage (e.g., drafting, editing, data analysis), and how their outputs were integrated. Leading journals such as *Nature* (2023) and *Science* (Thorp, 2023) now require such disclosure.
- **Verification and Fact-Checking Protocols:** Before incorporating AI-generated material, researchers must verify factual accuracy, confirm primary sources, and ensure interpretations align with evidence. This mirrors the isnād principle of scrutinizing each link in the transmission chain before acceptance (Görke, 2010).
- **Maintaining Raw Output Records:** Researchers should archive raw, unedited AI outputs alongside final versions, enabling post-publication verification and ensuring transparency of the editorial process. This aligns with open science principles of reproducibility and accountability (Nosek et al., 2015).
- **Balanced Integration of AI:** AI should serve as a supplementary tool that enhances productivity or expands access to information, not as a substitute for the researcher's critical reasoning and argumentation. This position echoes UNESCO's *Recommendation on the Ethics of Artificial Intelligence* (2021), which stresses the primacy of human oversight.

For Institutions

- **Standardized AI Disclosure Policies:** Universities and research bodies should adopt uniform guidelines requiring disclosure of AI use in all publications, proposals, and reports. Policies should specify the detail required, mirroring data citation and provenance standards (OECD, 2019).
- **Institutional AI Review Boards:** Institutions should establish AI oversight committees, analogous to IRBs, tasked with evaluating attribution practices, epistemic soundness, and potential biases in AI-mediated projects (American Psychological Association, 2017).
- **Metadata Integration into Repositories:** Institutional repositories should capture metadata about AI contributions, enabling future scholars to trace epistemic lineage. This resonates with the FAIR principles of data management, which emphasize findability, accessibility, interoperability, and reusability (Wilkinson et al., 2016).
- **Capacity Building and Training:** Institutions must provide mandatory training for faculty, staff, and students on the ethical use of AI, with a focus on epistemic responsibility, attribution, and bias awareness. UNESCO (2021) highlights training as a pillar of responsible AI adoption.
- **Accountability Enforcement Mechanisms:** Institutions should develop enforcement mechanisms for breaches of AI disclosure policies or failures in verification. This may

include retractions, revisions, or disciplinary action, building on the protocols established by COPE (2019).

From Principle to Practice: This framework bridges the gap between abstract ethical ideals and the operational realities of scholarship. Embedding these practices into daily research routines and institutional workflows is essential to ensure that epistemic integrity is not compromised. In an era when AI's generative capacities risk obscuring or severing knowledge origins, these measures reaffirm that human scholars and institutions remain the stewards of trust, traceability, and accountability in knowledge production.

Emerging Challenges: Even with robust ethical principles and a clear practical framework, preserving knowledge lineage in the age of generative artificial intelligence (AI) is fraught with evolving challenges. These challenges cut across legal, epistemic, institutional, and technical domains, revealing the fragility of traditional concepts when confronted with hybrid and algorithmic knowledge production.

Intellectual Property and Ownership Ambiguities: Generative AI unsettles traditional intellectual property (IP) frameworks. Determining who "owns" AI-generated content—whether the developer, the user, or no one—remains unresolved in law. Samuelson (1985, 2023) has long argued that copyright struggles to account for computer-generated works, exposing tensions between human creativity, machine autonomy, and ownership rights. Recent debates further question whether training AI models on copyrighted data constitutes fair use (Lemley, 2024). Such ambiguities extend to attribution, since proprietary models often obscure their training data, raising questions about both the legitimacy of outputs and the rights of original creators.

Responsibility Gaps: One of the clearest manifestations of responsibility gaps in the digital knowledge economy is found in recent reports on OpenAI's alleged reliance on Google's search index via intermediaries such as SerpApi. While OpenAI has denied direct dependence on Google, insisting instead on its own web crawler (GPTBot) and licensed publishing agreements, the investigation published by *The Information* (2025) and further discussed in technology outlets points to porous boundaries between traditional search infrastructures and generative AI systems. This blurring of infrastructures was further illustrated by an experiment conducted by former Google engineer Abhishek Iyer, in which a fabricated term inserted into Google's index later appeared verbatim in ChatGPT's output (Efrati et al., 2025). Such episodes underscore how generative AI—despite its transformative capacity—remains epistemically entangled with legacy systems of information curation. The result is a diffusion of attribution and accountability: should credit and responsibility be assigned to the search index that aggregates and ranks, to the AI model that recombines, or to the original human creators of content? This structural ambiguity exemplifies how responsibility gaps arise when accountability is distributed across complex digital systems, leaving users without a clear lineage of knowledge.

Attribution in Hybrid Human–AI Authorship: Attribution safeguards epistemic integrity by ensuring claims are traceable to identifiable contributors. Generative AI complicates this process because outputs emerge from vast, untraceable datasets and complex human–machine interactions. Empirical research confirms these tensions: in a study of 155 knowledge

workers, He et al. (2025) found that AI collaborators consistently received less credit than human partners for equivalent contributions. They argue that binary categories of authorship versus acknowledgment are inadequate and propose spectrum-based attribution models that disclose the type of AI model used, its contributions, and the degree of human oversight.

One practical response to this challenge is to treat AI outputs as provisional rather than final sources of knowledge. When a generative model suggests a fact, citation, or argument, the researcher must independently verify the claim and, once confirmed, attribute it to the original human-authored source rather than to the AI output itself. In this way, epistemic credit is transferred from the model—, which lacks agency and verifiable intent—to the legitimate reference, restoring a traceable chain of knowledge. This approach parallels the *isnād* tradition, where the value of a transmitted report depends on the credibility and continuity of its narrators. Under this model, AI functions as an exploratory tool, while validation and the moral weight of authorship remain firmly with the human scholar.

Epistemic Humility in Hybrid Authorship: Blending human and AI inputs requires recalibrating epistemic humility—the recognition of limits to one’s knowledge and authority. Overestimating the reliability of AI outputs risks eroding scholarly credibility, while underestimating AI’s value may stifle innovation. Achieving balance demands critical self-awareness and institutional frameworks that reinforce accountability, ensuring that humility strengthens rather than weakens the continuity of knowledge lineage.

Redefining Originality in a Regenerative Era: In contexts where AI rapidly generates novel-seeming material, the traditional concept of originality—rooted in human creativity and intellectual labor—faces erosion. Floridi (2018b) cautions that originality risks being reduced to superficial novelty rather than substantive intellectual contribution, complicating attribution and threatening the preservation of authentic knowledge lineage.

The Advent of Artificial General Intelligence (AGI): If AI advances toward Artificial General Intelligence—systems with cognitive capacities rivaling human reasoning—the epistemic landscape will shift radically. AGI could autonomously generate, revise, and disseminate knowledge at scale, potentially redefining authorship, accountability, and intellectual labor. Without proactive governance, such systems may produce fragmented and opaque knowledge ecosystems (Samuelson, 2023).

Institutional and Political Gatekeeping: Corporate and institutional actors exert significant control over AI outputs by deciding which information is emphasized or suppressed. Such gatekeeping can sever knowledge lineage not due to technical constraints but through deliberate policy decisions, raising concerns about bias, censorship, and epistemic inequality (Lemley, 2024). A notable illustration is the recent *Grok* incident, in which xAI’s chatbot on X (formerly Twitter) generated responses addressing politically and historically sensitive topics—particularly those related to the Palestinian issue and their rights—prompting platform-level intervention and igniting debate about the boundaries of free expression and the governance of AI-generated discourse (The Washington Post, 2025). This case underscores that AI-mediated knowledge is vulnerable not only to technical opacity but also to editorial and institutional controls that determine what information is allowed to surface. These dynamics highlight the urgent need for lineage-preserving and transparency

mechanisms to safeguard scholarly and public knowledge, ensuring it remains trustworthy, accountable, and resistant to undue suppression.

Complexity of Implementation: Even where ethical frameworks exist, applying them consistently across institutions and cultures poses logistical challenges. Variations in resources, capacity, and governance make adherence uneven, which risks eroding epistemic transparency on a global scale (Floridi, 2018b).

Navigating the Moving Target: AI technologies, their legal contexts, and societal responses evolve at a pace that outstrips academic and policy cycles. This creates a “moving target” problem where governance strategies risk obsolescence before they are fully enacted. Preserving knowledge lineage in such a dynamic environment requires adaptive, iterative approaches rather than fixed rules (Samuelson, 2023).

Advanced Reflections: The Future of Knowledge Lineage: This final analytical section extends the study into forward-looking reflections, asking how knowledge lineage might be preserved or transformed in an AI-saturated future.

Looking toward the future, the preservation of knowledge lineage in an AI-saturated world is not merely a technical or regulatory concern—it is an existential challenge for the integrity of human knowledge. The very attributes that make generative AI powerful—its speed, scale, and capacity for recombination—also make it disruptive, fragmenting the continuity of ideas, obscuring their origins, and weakening accountability (Mittelstadt, 2019).

Beyond Technical Fixes. While metadata systems, disclosure protocols, and human-in-the-loop verification are indispensable, they cannot by themselves secure epistemic integrity. The deeper challenge lies in cultivating a scholarly ethos that values provenance, accountability, and moral responsibility as much as efficiency and novelty. Without such cultural foundations, safeguards risk becoming symbolic rather than substantive (Greene et al., 2019). As Cath (2018) emphasizes, governance must be adaptive, interdisciplinary, and responsive to the social contexts in which AI operates if it is to sustain trust in knowledge systems.

Reimagining Authorship in the Age of AI. The concept of authorship must evolve to reflect hybrid human–AI creative processes. This does not dilute human responsibility; rather, it requires new categories and vocabularies for contribution and credit that capture the socio-technical complexity of knowledge production. Foucault’s (1977) “author function” and Floridi’s (2018a) concept of ectypes both illuminate how authorship might be reframed: not as solitary ownership but as stewardship within networks of human and machine collaboration. Current publishing debates reinforce this. The Committee on Publication Ethics (COPE, 2023) has stressed that AI cannot be credited as an author, yet disclosure of AI use is increasingly encouraged to preserve transparency.

Integrating Ethical Lineages. Future governance models can be enriched by combining global AI ethics with culturally embedded traditions such as the Islamic isnād. Isnād demonstrates that rigorous traceability and moral responsibility can reinforce one another (Görke, 2010). Embedding such ethical lineages into AI design could ensure that technological acceleration does not outpace the moral architectures required for epistemic trust. Zednik (2021) makes

a similar point in his framework for explainable AI: transparency should be understood not as a technical add-on but as an ethical foundation.

The AGI Horizon. The potential emergence of Artificial General Intelligence (AGI) will test the very boundaries of epistemic agency. If AGI were to reach levels of autonomy comparable to human scholars, the categories of authorship, accountability, and intellectual labor would need profound rethinking. Bostrom (2014) warns of the destabilizing risks of poorly aligned superintelligence, while Russell (2019) advocates embedding human-compatible values directly into AI design to ensure corrigibility. Yet, as Thorn (2015) cautions, such scenarios remain speculative and must be balanced with practical ethical frameworks that guide governance today. The unresolved question will be whether AGI could ever bear epistemic agency—or whether it remains, despite sophistication, a tool without moral standing.

Human Intuition and Algorithmic Reasoning as a New Frontier

Building on the preceding discussion, *this study advances an original conceptual reflection*: the convergence of human intuition with algorithmic reasoning may represent a new frontier in knowledge production. Intuition—shaped by tacit experience and unconscious cognitive processing—can serve as an early signal or epistemic compass, guiding researchers toward insights that cannot be derived from data alone. Algorithmic reasoning, in turn, provides the capacity to test, validate, and scale these insights. This proposal is not grounded in prior literature but reflects the author’s analytical contribution, raising key questions for future research: How can intuitive insights be documented within the chain of attribution? Can epistemic responsibility be shared between human and machine without diluting moral accountability? Addressing these questions could inform the design of future governance models that preserve the unique role of human contribution while harnessing the analytical power of artificial intelligence—ensuring that knowledge lineage remains both traceable and deeply human-centered.

A Call for Continuous Renewal. Ultimately, the future of knowledge lineage will depend on the scholarly community’s willingness to adapt and renew epistemic practices in response to evolving realities. Just as past civilizations developed systems like *isnād* to safeguard the authenticity of transmitted knowledge, our era must create lineage-preserving mechanisms that are living, inclusive, and adaptive. These systems should be continually revised, resilient to institutional capture, and grounded in cultural as well as technological diversity (Foucault, 1977; Cath, 2018). Only through such continuous renewal—anchored in both global ethical principles and moral traditions such as *amāna*—can humanity sustain trust and integrity in an increasingly AI-mediated landscape.

These reflections suggest that while technological innovation will continue to unsettle established norms, the preservation of knowledge lineage ultimately hinges on our willingness to embed moral responsibility, epistemic humility, and cultural inclusivity into AI governance. It is on this foundation that the concluding discussion turns, outlining how scholars, institutions, and policymakers can act as stewards of knowledge in an era where the integrity of lineage itself is at stake.

Conclusion

Building on these reflections, this study has examined the concept of knowledge lineage and the pressing question of epistemic responsibility in an era increasingly shaped by generative artificial intelligence (AI). By tracing the historical and philosophical roots of attribution—ranging from Foucault’s author-function and Floridi’s ectypes to the Islamic isnād tradition—the analysis has shown that while AI transforms the form and pace of knowledge production, it does not and cannot replace the ethical and epistemic functions performed by human agents.

The comparative analysis of academic governance models revealed a sharp divide between university contexts, which stress disclosure, transparency, and human oversight, and non-academic or corporate platforms, where political and institutional gatekeeping often obscures or severs the chain of attribution. Case studies such as the Grok incident illustrated how AI-mediated discourse is vulnerable not only to technical opacity but also to editorial and institutional controls, further underscoring the need for lineage-preserving systems across scholarly and public domains.

Ethically, the study has argued for a synthesis of universal AI governance principles with culturally embedded traditions such as isnād. The dual safeguard—technical transparency through disclosure and metadata protocols, coupled with moral accountability grounded in long-tested traditions of stewardship—offers a more durable framework for ensuring epistemic integrity. The proposed normative measures, including mandatory source documentation, human-in-the-loop verification, institutional accountability mechanisms, and capacity-building programs, were translated into a practical framework to guide both researchers and institutions.

Yet challenges remain. Intellectual property ambiguities, the erosion of originality, the complexities of hybrid authorship, and the speculative horizon of Artificial General Intelligence (AGI) will continue to test our ability to sustain a coherent knowledge lineage. Meeting these challenges requires not static regulation but adaptive, inclusive systems that can evolve alongside technological and cultural shifts.

Ultimately, preserving knowledge lineage in the AI era is less a technical problem than a moral and cultural imperative. Human scholars, educators, and institutions must act not as passive users of technology but as stewards of knowledge, responsible for ensuring that what is transmitted remains traceable, verifiable, and ethically grounded. By grounding AI governance in both global ethical frameworks and historically tested traditions such as isnād, the chain of knowledge can remain unbroken, carrying forward not only information but also the weight of truth and responsibility into the future.

Acknowledgment of AI Assistance

ChatGPT (GPT-5, OpenAI) was used exclusively for language editing and for suggesting improvements to the organization and flow of the manuscript. The AI system did not generate original scholarly content or arguments. All final decisions regarding content, interpretation, and references were made solely by the author.

References

- Abdulrahman, M. A. (2024). Cultural and Social Influences on Hadith Classification: An Analytical Study of Historical Transformations. *Journal of Ecohumanism*, 3(8), 2783-2791. DOI: <https://doi.org/10.62754/joe.v3i8.4926>
- AL-Smadi, M. (2025). IntegrityAI at GenAI Detection Task 2: Detecting Machine-Generated Academic Essays in English and Arabic Using ELECTRA and Stylometry. *arXiv preprint arXiv:2501.05476*.
<https://doi.org/10.48550/arXiv.2501.05476>
- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct*. <https://www.apa.org/ethics/code>
- Barthes, R. (1977). *Image-Music-Text* (S. Heath, Trans.). Fontana Press. ISBN: 0-00-686135-0.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).<https://doi.org/10.1145/3442188.3445922>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, V., Card, D., Castellon, R., Chatterji, N., Chen, A. S., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv*.
<https://doi.org/10.48550/arXiv.2108.07258>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Bouhafa, F. (2021). The dialectics of ethics: Moral ontology and epistemology in Islamic philosophy. *Journal of Arabic and Islamic Studies*, 21(2), 25–54.
<https://doi.org/10.5617/jais.9368>
- Brown, J. A. C. (2014). *Misquoting Muhammad: The challenge of interpreting the Prophet's legacy*. Oneworld Publications.
- Brown, J. A. C. (2018). *Hadith: Muhammad's legacy in the medieval and modern world* (2nd ed.). Oneworld Publications.
- Cath, C. (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A*, 376(2133), 20180080.
<https://doi.org/10.1098/rsta.2018.0080>
- Code, L. (1987). *Epistemic responsibility*. University Press of New England and Brown University Press. <https://archive.org/details/epistemicrespons0000code>
- Committee on Publication Ethics COPE. (2023, February 13). *Authorship and AI tools*. <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>
- Cui, Y., & Widom, J. (2003). Lineage tracing for general data warehouse transformations. *The VLDB Journal*, 12(1), 41–58. <https://doi.org/10.1007/s00778-002-0083-8>
- Dotan, R., Parker, L. S., & Radzilowicz, J. G. (2024). Responsible adoption of generative AI in higher education: Developing a “points to consider” approach based on faculty perspectives. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)* (pp. 1–16). ACM.
<https://doi.org/10.1145/3630106.3659023>
- Efrati, A., Palazzolo, S., & Mascarenhas, N. (2025, February 3). *OpenAI is challenging Google—while using its search data. The Information*.
<https://www.theinformation.com/articles/openai-is-challenging-google-while-using-its-search-data>

- Floridi, L. (2018a). Artificial intelligence, deepfakes and a future of ectypes. *Philosophy & Technology*, 31(3), 317–321. <https://doi.org/10.1007/s13347-018-0325-3>
- Floridi, L. (2018b). Soft ethics and the governance of the digital. *Philosophy & Technology*, 31, 1–8. <https://doi.org/10.1007/s13347-018-0303-9>
- Floridi, L. (2019). *The logic of information: A theory of philosophy as conceptual design*. Oxford University Press. <https://doi.org/10.1093/oso/9780198833635.001.0001>
- Floridi, L., & Cowsls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1, 2-15. <https://doi.org/10.1162/99608f92.8cd550d1>
- Foucault, M. (1977). What is an author? In D. F. Bouchard (Ed.), *Language, counter-memory, practice: Selected essays and interviews* (pp. 113–138). Cornell University Press.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Görke, A. (2010). Review of Jonathan A. C. Brown: *Hadith: Muhammad's legacy in the medieval and modern world*. *Bulletin of the School of Oriental and African Studies*, 73(3), 534–536. <https://doi.org/10.1017/S0041977X10000467>
- Ghosal, T., Tiwary, P., Patton, R., & Stahl, C. (2021). Towards establishing a research lineage via identification of significant citations. *Quantitative Science Studies*, 2(4), 1511–1528. https://doi.org/10.1162/qss_a_00152
- Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. In *Proceedings of the 52nd Hawaii International Conference on System Sciences* (pp. 2122–2131). <https://hdl.handle.net/10125/59651>
- Hadith Studies. (n.d.). *Hadith studies*. In *Wikipedia*. https://en.wikipedia.org/wiki/Hadith_studies
- Harvard Extension School. (n.d.). *Academic integrity*. Harvard University. <https://extension.harvard.edu/enrolled-students/academic-integrity/>
- He, J., Houde, S., & Weisz, J. D. (2025). Which contributions deserve credit? Perceptions of attribution in human–AI co-creation. In *Proceedings of the 2025 CHI conference on human factors in computing systems* (pp. 1–18). ACM. <https://doi.org/10.48550/arXiv.2502.18357>
- High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>
- Humphreys, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3), 615–626. <https://doi.org/10.1007/s11229-008-9435-2>
- IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems* (1st ed.). IEEE Standards Association. https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Journal of King Saud University – Science. (n.d.). *Ethical guidelines: Use of artificial intelligence (AI)*. <https://jksus.org/ethical-guidelines/>
- Khosrowi, D., Finn, F., & Clark, E. (2024). Engaging the many-hands problem of generative-AI outputs: A framework for attributing credit. *AI and Ethics*, 1-19. <https://doi.org/10.1007/s43681-024-00440-7>
- Kuhn, T. S. (1997). *The structure of scientific revolutions*. University of Chicago Press.

- Lemley, M. A. (2024). How generative AI turns copyright upside down. *Columbia Science and Technology Law Review*, 25(2), 1–45. <https://doi.org/10.52214/stlr.v25i2.12761>
- Lloyd, G. E. R. (1996). *Adversaries and authorities: Investigations into ancient Greek and Chinese science*. Cambridge University Press. *East Asian Science, Technology, and Medicine*, 15(1), 121-126. <https://doi.org/10.1163/26669323-01501009>
- Longpre, S., Mahari, R., Chen, A., Obeng-Marnu, N., Sileo, D., Brannon, W., ... & Hooker, S. (2024). A large-scale audit of dataset licensing and attribution in AI. *Nature Machine Intelligence*, 6(8), 975–987.
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Motzki, H. (2002). *The origins of Islamic jurisprudence*. Brill.
- Nature. (2023). Tools such as ChatGPT threaten transparent science. *Nature*, 613(7945), 612. <https://doi.org/10.1038/d41586-023-00191-1>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- OECD. (2019). *Recommendation of the council on artificial intelligence*. OECD Legal Instruments. <https://oecd.ai/en/ai-principles>
- Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2023). Game of tones: Faculty detection of GPT-4 generated content in university assessments. *arXiv*. <https://arxiv.org/abs/2305.18081>
- Petro, M. (2025, May 20). University at Buffalo students protest use of AI detection tool. *The Buffalo News*. Republished by GovTech. <https://www.govtech.com/education/higher-ed/university-at-buffalo-students-protest-use-of-ai-detection-tool>
- Reed, B. (2001). Epistemic agency and the intellectual virtues. *The Southern Journal of Philosophy*, 39(4), 507–526. <https://doi.org/10.1111/j.2041-6962.2001.tb01820.x>
- Russell, S. (2019). *Human compatible: AI and the problem of control*. Penguin Uk.
- Sağlam, T., & Schmid, L. (2025). Evaluating software plagiarism detection in the age of AI: Automated obfuscation and lessons for academic integrity. *arXiv preprint arXiv:2505.20158*. <https://doi.org/10.48550/arXiv.2505.20158>
- Samuelson, P. (1985). Allocating ownership rights in computer-generated works. *University of Pittsburgh Law Review*, 47, 1185–1224.
- Samuelson, P. (2023). Generative AI meets copyright. *Science*, 381(6656), 158-161. <https://doi.org/10.1126/science.adj0396>
- The Harvard Crimson. (2023, September 1). Harvard releases guidance for AI use in classrooms. *The Harvard Crimson*. <https://www.thecrimson.com/article/2023/9/1/fas-ai-guidance/>
- Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313. <https://doi.org/10.1126/science.adg7879>
- Thorn, P. D. (2015). Nick Bostrom: *Superintelligence: Paths, dangers, strategies*. *Minds and Machines*, 25(3), 285–289. <https://doi.org/10.1007/s11023-015-9377-7>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>
- University of Oxford. (2024). *AI and academic practice*. Centre for Teaching and Learning. <https://www.ctl.ox.ac.uk/ai>

- University of Pittsburgh Teaching Center. (2023, June 22). Teaching Center doesn't endorse any generative AI detection tools. *University Times*.
<https://www.utimes.pitt.edu/news/teaching-center-doesn-t>
- University of Pittsburgh Teaching Center. (2025, January 24). Encouraging academic integrity.
<https://teaching.pitt.edu/resources/encouraging-academic-integrity/>
- U.S. Copyright Office. (2024). *Artificial intelligence and copyright*.
<https://www.copyright.gov/ai/>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), 1-39.. *arXiv*.
<https://arxiv.org/abs/2306.15666>
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Woodmansee, M., & Jaszi, P. (Eds.). (1994). *The construction of authorship: Textual appropriation in law and literature*. Duke University Press, 10(2).
- Worth, S., Snaith, B., Das, A., Thuermer, G., & Simperl, E. (2024). AI data transparency: An exploration through the lens of AI incidents. *arXiv preprint arXiv:2409.03307*.
<https://arxiv.org/abs/2409.03307>
- Yaqeen Institute. (n.d.). *Authenticating hadith and the history of hadith criticism*.
<https://yaqeeninstitute.org/read/paper/authenticating-hadith-and-the-history-of-hadith-criticism>
- Zednik, C. (2021). Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philos. Technol.* 34, 265–288. <https://doi.org/10.1007/s13347-019-00382-7>