

INTERNATIONAL JOURNAL OF
MODERN EDUCATION
(IJMOE)www.gaexcellence.com/ijmoe**DIGITALIZING ISLAMIC KNOWLEDGE: MAPPING JAWI
TEXT DIGITIZATION AND ITS ETHICAL IMPLICATIONS
THROUGH A BLOCKCHAIN LENS**Syamilah Saliman^{1*}, Hafiza Abas², Zilal Saari³, Wan Rusydiah⁴¹Faculty of Artificial Intelligence, Universiti Teknologi Malaysia (UTM), Malaysia syamilah92@graduate.utm.my <https://orcid.org/0009-0006-1927-3924>²Faculty of Artificial Intelligence, Universiti Teknologi Malaysia (UTM), Malaysia hafiza.kl@utm.my <https://orcid.org/0000-0003-0137-3127>³Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia (UTM), Malaysia zilal@utm.my <https://orcid.org/0009-0007-0776-4854>⁴Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia (UTM), Malaysia wanrusydiah@graduate.utm.my <https://orcid.org/0009-0007-4116-0341>

*Corresponding Author

Article Info:**Article history:**

Received date: 16.07.2026

Revised date: 30.07.2026

Accepted date: 26.08.2026

Published date: 21.09.2026

To cite this document:

Saliman, S., Abas, H., Saari, Z., & Rusydiah, W. (2026). Digitalizing Islamic Knowledge: Mapping Jawi Text Digitization and Its Ethical Implications Through a Blockchain Lens. *International Journal of Modern Education*, 8(31), 973-999.

Abstract:

Jawi manuscript heritage in Nusantara faces a dual crisis of physical degradation and declining script literacy. While artificial intelligence (AI) technologies such as optical character recognition (OCR), natural language processing (NLP) and machine learning (ML) accelerate the preservation, current digitization initiatives lack an explicit Maqasid al-Shariah-based ethical framework to protect the textual authenticity and prevent unauthorized exploitation. To address this gap, this study employs a two-phase qualitative and conceptual methodology. Phase 1 conducts a systematic landscape mapping following PRISMA 2020 guidelines across Web of Science and Scopus databases by analyzing 13 selected studies on AI-driven Jawi digitization. Phase 2 utilizes deductive theoretical analysis to evaluate identified governance gaps by mapping blockchain ethics properties such as immutability, transparency, decentralization and auditability into Maqasid al-Shariah principles, specifically Hifz al-Din (preservation of religion), Hifz al-Aql (preservation of intellect) and Hifz al-Ilm (preservation of knowledge). The findings highlight three key gaps. First, existing tools are technically fragmented as each handles a single task instead of working together as a complete pipeline. Second, there is no consistent standard for how the text is encoded digitally, and Jawi script is only partly supported by Unicode (the universal system computers use to display text correctly). Third, ethical considerations are largely missing

from the governance process altogether. To resolve these challenges, the study establishes an integrated governance framework synthesizing technical standardization with Islamic ethical governance. The framework delivers actionable policy guidance for Malaysian institutions such as JAKIM and the Ministry of Digital, a technical roadmap for developers and a community-centered access model. This framework ensures that digitized Jawi heritage remains textually uncorrupted, authenticated and spiritually meaningful for future generations.

DOI: 10.35631/IJMOE.831056

Keyword:

Blockchain Ethics, Islamic Ethics, Jawi Digitization.



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijmoe@gaexcellence.com.

Introduction

The Islamic intellectual heritage of the Nusantara region recorded in Jawi script faces a dual crisis of declining script literacy and physical manuscript degradation (Fitriyah et al., 2026; Samsul Hady et al., 2025). In Malaysia, this crisis has intensified the urgency for large scale digitization, as vast Jawi manuscript collections preserved in national and state institutions remain largely inaccessible to contemporary readers due to the physical fragility of aging mediums and a widespread generational decline in the ability to read the script (Samsul Hady et al., 2025). While recent advances in digital humanities have leveraged automated image processing, skeletonization algorithms and digital repository platforms to address preservation concerns (Fitriyah et al., 2026), physical preservation alone neither guarantees textual accessibility nor ensures systemic governance over how digitized knowledge is managed and disseminated.

The emergence of artificial intelligence (AI) technologies, specifically optical character recognition (OCR), natural language processing (NLP) and machine learning (ML) has substantially expanded the technical capacity for Jawi script digitization, increasing both the speed and scale of manuscript processing (Adilazuarda et al., 2025; Faizullah et al., 2023). As Jawi is an Arabic based adapted script, developments in machine vision enable automated workflows to process complex historical layouts and transform manuscript images into machine readable text (Kasem et al., 2023). However, the cursive and position dependent nature of Jawi characters, whereby letter forms change dynamically across initial, medial and final positions, presents persistent technical challenges (Fitriyah et al., 2026). Specialized deep learning architectures are increasingly required to resolve segmentation difficulties across overlapping handwritten strokes and irregular historical ligatures without compromising structural accuracy (Fitriyah et al., 2026; Ullah and Jamjoom, 2022). More critically, existing digitization initiatives operate without an explicit Islamic ethical framework, leaving practitioners and policymakers without principled criteria to assess whether AI-driven

processing models align with the values and scholarly obligations of the Islamic tradition (Al Momani, 2025; Yusuf, 2025).

A review of the current literature reveals three interconnected gaps that this study seeks to address. First, a technical mapping gap exists, as there is minimal systematic analysis evaluating existing AI-driven Jawi digitization tools with respect to their features, architectural approaches and embedded technical constraints (Fitriyah et al., 2026; Kasem et al., 2023). Second, an ethical governance gap persists, because while Islamic ethics has been robustly applied to fintech, healthcare artificial intelligence and digital governance in Muslim majority contexts, it has not been systematically integrated into the domain of Islamic heritage digitization (Al Momani, 2025; Azizov, 2026). Third, a theoretical integration gap remains, as the literature on AI-driven heritage digitization, blockchain ethics and Maqasid al-Shariah has largely operated in disciplinary silos, leaving the intersections between computational data integrity and Shariah compliant knowledge governance largely unexplored (Adnan, 2026; Auda, 2021).

To address these gaps, this study proposes a conceptual framework that integrates AI-driven Jawi digitization with blockchain ethics and Maqasid al-Shariah through a two-phase documentary and conceptual analysis. Blockchain ethics serves as the critical theoretical bridge in this framework, mapping core computational governance properties such as immutability, decentralization, transparency and auditability directly onto foundational Islamic principles, specifically Amanah (trustworthiness), Hifz al-Din (preservation of religion), Hifz al-Aql (preservation of intellect) and Hifz al-Ilm (preservation of knowledge) (Hayati, 2026; Naswa and Muthoifin, 2025). By evaluating current technological approaches through this dual ethical and technical lens, this study offers evidence-based guidance for Malaysian institutions, particularly the Department of Islamic Development Malaysia (JAKIM) and the Ministry of Digital, in developing national heritage policies and governance standards for AI-driven Islamic knowledge systems. Thus, this study seeks to ensure that the digital transition of Jawi Islamic heritage achieves technical excellence while remaining spiritually authentic, textually uncorrupted and intellectually meaningful for future generations.

Literature Review

Jawi Manuscripts and Islamic Knowledge Preservation

Jawi script is an Arabic based orthographic system adapted for the Malay language that represents one of the most enduring expressions of Islamic intellectual heritage in the Nusantara region. In Malaysia, significant repositories of Jawi manuscripts are preserved in national and state institutions, forming an irreplaceable corpus of Islamic knowledge that remains inaccessible to most contemporary readers due to a sustained decline in Jawi literacy among younger generations (Samsul Hady et al., 2025). This decline has been attributed to multiple compounding factors, including the historic dominance of the Latinized Rumi script in formal state education, the limited integration of Jawi in modern digital media ecosystems and the historical absence of robust digital resources capable of sustaining the intergenerational transmission of the script (Bhatt, 2025).

The consequence is a dual crisis that combines a contemporary literacy deficit with an institutional preservation emergency. As physical manuscripts continue to get worse due to the age, ink acidity and tropical environmental conditions made the urgency of large-scale

digitization intensifies. However, scholarly consensus increasingly recognizes that digitization undertaken without adequate ethical governance risks introducing new categories of socio religious harm such as automated textual distortion, the erosion of deep semantic integrity and the unauthorized commercial exploitation of sacred or legally significant knowledge (Fauzi, 2025; Yusuf, 2025). The Malaysian government, through agencies such as the Islamic Development Department of Malaysia (JAKIM) and the Digital Ministry, has acknowledged the national heritage dimension of this challenge, but a policy framework for ethically administered the Jawi digitization methods especially in Malaysia is yet to be explored.

Thus, this absence of a digitization policy is not merely an administrative gap, but it also creates a possibility in which technical digitization proceeds without a framework that protecting the integrity of the knowledge being digitized.

AI-Driven Jawi Digitisation and Technical Challenges

The emergence of artificial intelligence technologies has substantially expanded the technical capacity to process, restore and preserve historical scripts. Three technological pillars are directly relevant to contemporary Jawi digitization efforts such as OCR, NLP and ML.

OCR has been the most extensively studied tool in this domain. Recent deep learning approaches, particularly those employing specialized Convolutional Neural Networks and customized Vision Language Models, have demonstrated improved baseline accuracy in handling the ligature heavy and connected characteristics of Jawi script (Fitriyah et al., 2026; Kasem et al., 2023). Despite this progress, critical limitations persist. OCR systems continue to struggle with historical diacritical marks (baris), bleeding ink in degraded manuscripts and the absence of a universally adopted Jawi Unicode encoding convention that would enable programmatic interoperability across global digitization projects (Kasem et al., 2023). The historical lack of a unified, publicly available and systematically annotated handwritten Jawi dataset remains a significant constraint on model training and cross study benchmarking, though recent open weights initiatives such as the Jawi AI Project have begun to bridge this gap through custom sequence recognition pipelines (Varela et al., 2026).

NLP for Jawi remains at a nascent stage relative to the advancements achieved for Arabic, standard modern Malay or major European languages. The structural overlap between Jawi orthography and Classical Malay morphology presents distinct semantic challenges, such as the systematic omission of short vowels, that standard Arabic NLP pipelines cannot resolve without substantial linguistic adaptation (Adilazuarda et al., 2025). Concurrently, existing Malay NLP collections are overwhelmingly Rumi based, creating a fundamental training data deficit that limits the applicability of standard transfer learning approaches (Zolkepli et al., 2024).

ML frameworks have been increasingly applied to improve character classification and word boundary detection in handwritten documents. Recent studies have explored the adaptation of classical Arabic script recognition models for Southeast Asian contexts, alongside synthetic data augmentation strategies designed to compensate for severely limited labelled datasets (Faizullah et al., 2023; Fitriyah et al., 2026). However, the field continues to lack unified baseline datasets and standardized evaluation metrics, producing fragmented accuracy measurements that are difficult to compare across studies. This technical inconsistency constitutes a hidden governance gap: in the absence of standardized processing tools and

evaluation criteria, disparate digitization projects generate non interoperable metadata and text structures, directly undermining both the long-term accessibility and the textual integrity of the resulting digital heritage corpus.

This technical inconsistency constitutes a hidden governance gap such as the absence of standardized processing tools and evaluation criteria, different digitization projects that generate unmanageable metadata and text structures and a fragmentation that raises the query of what governance mechanism used.

Blockchain-Informed Governance and Maqasid Al-Shariah

Blockchain technology has attracted growing scholarly attention beyond its conventional application in decentralized financial architectures, emerging as a secure governance infrastructure applicable to digital heritage preservation, immutable archival management and provenance authentication (Hayati, 2026). For the purposes of this study, blockchain ethics is employed not as a technical deployment proposal or hardware recommendation, but as a robust conceptual lens through which the ethical properties of information governance systems can be evaluated and systematically compared against Islamic normative standards.

The core ethical properties associated with blockchain architecture, principally immutability, transparency, decentralization and multi-party auditability, provide a productive evaluative vocabulary for assessing the risks of AI driven digitization practices. Immutability is to ensures that verified digital records cannot be altered or overwritten without distributed network consensus, directly corresponds to the foundational Islamic ethical imperative of textual authenticity and the preservation of trust, known as Amanah (Naswa & Muthoifin, 2025). Transparency and public auditability align with the structural requirements of accountability (Mas'uliyah) in Islamic public governance. Furthermore, decentralization structurally challenges the absolute concentration of custodial authority over sacred and scholarly texts within any single state or institutional actor, deeply resonating with classical Islamic scholarship on the communal, distributed guardianship of sacred knowledge, or Hifz al-Ilm (Al Momani, 2025; Yusuf, 2025).

While blockchain ethics provides a technical and structural vocabulary for evaluating governance properties, it does not establish which normative standard those properties should serve. Thus, Maqasid al-Shariah is needed to provide that standard.

Maqasid al-Shariah is the most comprehensive and enduring teleological framework available for evaluating human actions, technological systems and institutional policies within an Islamic ethical paradigm. Originally formulated by al Ghazali and systematically expanded by subsequent classical jurists including al Shatibi and Ibn Ashur, Maqasid al-Shariah identifies five essential preservation objectives (Al-Daruriyyat al-Khams) that governance must protect: religion (Hifz al-Din), life (Hifz al-Nafs), intellect (Hifz al-Aql), lineage (Hifz al-Nasl) and wealth (Hifz al-Mal). Modern Islamic scholars have expanded this traditional framework to address complex modern governance. Specifically, several jurists propose that protecting knowledge and cultural heritage, known as Hifz al-Ilm and Hifz al-Turats should be recognized as its own distinct and independent objective. This specific legal concept carries direct relevance to the goals of this research (Auda, 2021).

The preservation and dissemination of Islamic knowledge as embodied in the centuries old Jawi manuscript tradition can be situated directly within this legal framework. Digitization efforts that genuinely enhance the accessibility, transparency and computational integrity of Islamic textual heritage directly serve the preservation of intellect (Hifz al-Aql) and the protection of religion (Hifz al-Din) by ensuring accurate knowledge transmission (Al Momani, 2025). Conversely, unmonitored digitization processes that introduce algorithmic hallucinations, permit unauthorized modification or enable the commercial exploitation of sacred knowledge violate those very objectives. This evaluative logic makes Maqasid al-Shariah uniquely well suited as an overarching governance framework for assessing AI driven Islamic heritage systems, moving beyond basic secular data compliance toward absolute teleological alignment.

Maqasid al-Shariah based analysis has been successfully applied to evaluate a growing range of modern technological domains, including Islamic fintech and digital asset regulation (Azizov, 2026), AI ethics in healthcare deployment (Kasule, 2024) and broader digital governance frameworks in Muslim majority states (Adnan, 2026). This study contributes directly to that expanding body of applied jurisprudence by extending Maqasid al-Shariah analysis to the novel context of AI driven Jawi digitization and blockchain informed heritage governance, a cross-disciplinary application that has not previously appeared in the scholarly literature.

Although blockchain ethics and Maqasid al-Shariah have so far been theorized separately from each other, it leads to the specific question on how the technical nature of blockchain can provide an integration in developing an ethical framework.

Analytical Synthesis and Framework Requirements

Islamic digital humanities represent the intersection of Islamic Studies with quantitative digital methods, computational linguistics and information governance frameworks. While the broader global field of digital humanities has produced substantial scholarship over several decades, Islamic digital humanities remain a comparatively recent and theoretically underdeveloped subfield, with Jawi script heritage receiving critically limited systematic scholarly attention relative to Middle Eastern Arabic script traditions (Waseem, 2025; Yusuf, 2025).

The governance dimension of Islamic digital humanities, specifically on how systems authorize digitization, algorithmically classify texts and globally disseminate Islamic knowledge remains a theoretical gap in the existing literature. Prevailing metadata and archiving frameworks have been largely adapted from Western, secular digital humanities scholarship without adequate engagement with Islamic epistemology or the distinctive ethical obligations (Amanah) that attach to the custodianship of revealed and scholarly knowledge (Yusuf, 2025).

In addition, a distinct theoretical integration gap further compounds this deficiency. Blockchain ethics and Maqasid al-Shariah have not previously been brought into explicit theoretical dialogue with one another or with the specific technical challenges of AI - driven Jawi digitization. The conceptual architecture for an Islamically grounded digital heritage governance system thus remains to be constructed. Although Islamic ethics has been applied to fintech, healthcare artificial intelligence and digital governance, it has not been systematically extended to AI - driven Islamic heritage digitization. Practitioners and

policymakers therefore lack an Islamically grounded ethical framework through which to evaluate, commission or regulate these high impact systems.

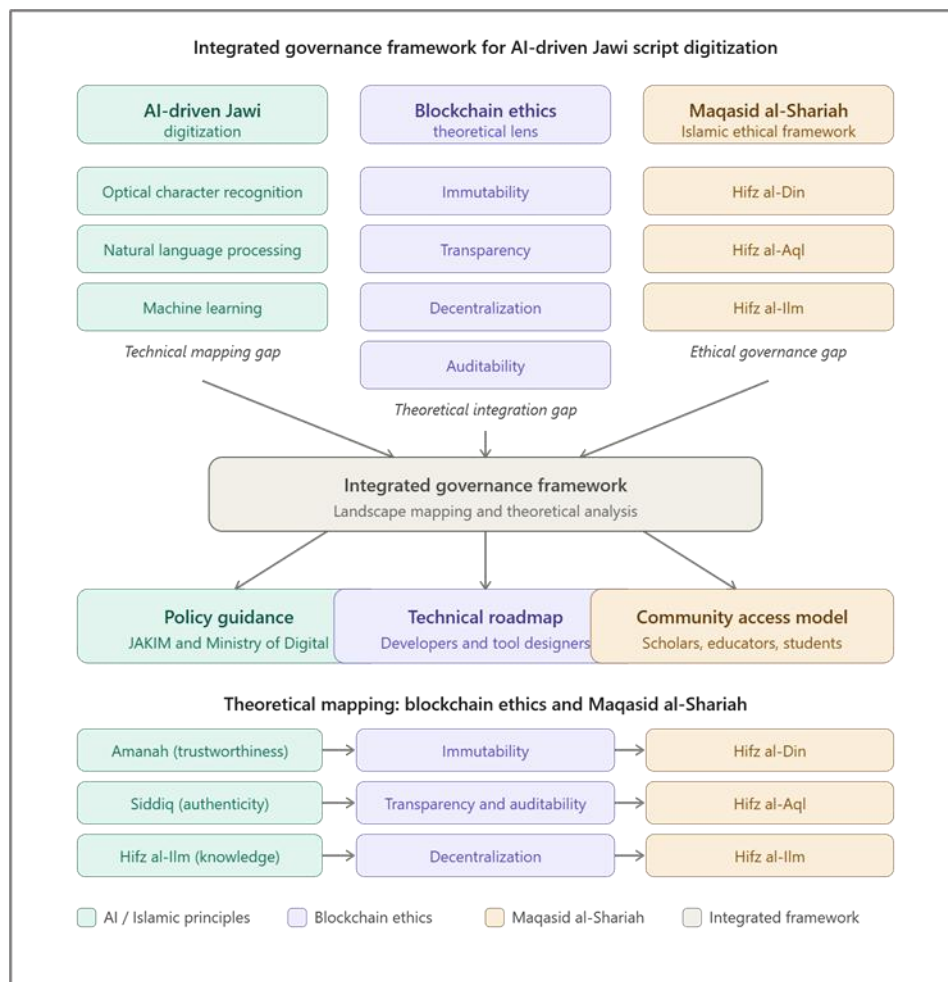


Figure 1: Conceptual Framework

As illustrated in Figure 1 above, the conceptual framework integrates three theoretical pillars such as AI-driven Jawi digitization (encompassing OCR, NLP and ML), blockchain ethics (immutability, transparency, decentralization and auditability) and Maqasid al-Shariah (Hifz al-Din, Hifz al-Aql and Hifz al-Ilm). Each addressing one of the three identified gaps through a structured theoretical mapping in which Amanah aligns with immutability, Siddiq with transparency and auditability, and Hifz al-Ilm with decentralization. The convergence of these three pillars produces the integrated governance framework, generating three practical outputs: evidence-based policy guidance for JAKIM and the Ministry of Digital Malaysia, a technical roadmap for Jawi digitization tool developers and a community-centred access model ensuring that digitized Islamic heritage remains authenticated, intellectually accessible and spiritually meaningful for future generations of Muslim scholars, educators and students.

This study addresses that gap directly by proposing a governance framework grounded in Maqasid al-Shariah with blockchain ethics functioning as the theoretical bridge between Islamic normative principles and the operational requirements of AI - driven digitization practice.

Methodology

Research Design

This study adopts a qualitative and conceptual research design comprising two interconnected phases that collectively address the dual objectives of the study, as illustrated in Figure 2. The first phase involves a systematic landscape mapping of AI-driven Jawi script digitization tools and practices, following established systematic mapping protocols used in software engineering and digital library research to categorize technological features, tools and research trends across global academic databases (Kitchenham et al., 2011; Petersen et al., 2015). The second phase employs a deductive theoretical analysis that applies the combined conceptual lens of blockchain ethics and Maqasid al-Shariah to evaluate the governance dimensions identified in the systematic landscape mapping, drawing on conceptual framework construction methodology as established by Jabareen (2009). Both phases are unified by a shared theoretical orientation, and their outputs are synthesized through an integrated conceptual analysis to develop a multi layered governance model (Auda, 2021; Jabareen, 2009).

The choice of a qualitative and conceptual design is justified by the nature of the research problem. This study does not seek to test hypotheses or generate statistically generalizable findings. Rather, it aims to construct a novel conceptual framework that has not previously existed in the literature, drawing on systematic documentary evidence and established theoretical traditions in Islamic jurisprudence and blockchain governance. This approach is consistent with established conventions in Islamic social science research that combine normative Islamic analysis with qualitative documentary methods (Secuntori et al., 2023; Zedan et al., 2025).

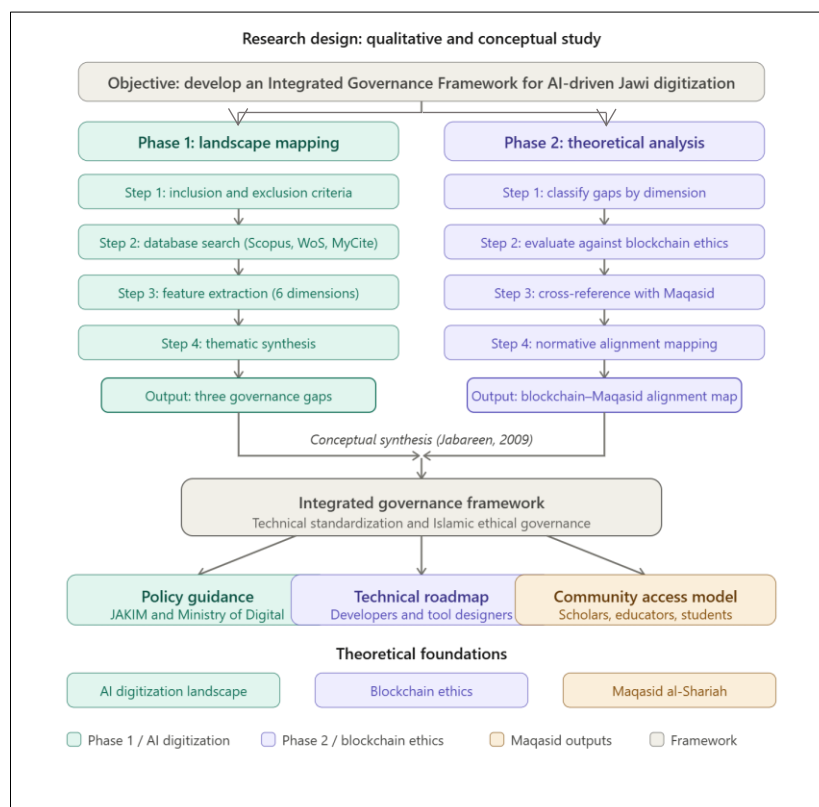


Figure 2: Research Design

The choice of a qualitative and conceptual design is justified by the nature of the research problem. The study does not seek to test hypotheses or generate statistically generalizable findings. Rather, it aims to construct a novel conceptual framework that has not previously existed in the literature, drawing on both systematic documentary evidence and the situated expertise of practitioners and scholars. This approach is consistent with established conventions in Islamic social science research that combine normative Islamic analysis with empirical qualitative methods (Secuntori et al., 2023; Zedan et al., 2025).

Phase 1: Systematic Landscape Mapping

Purpose and Rationale

The first phase addresses the study's objective: to map the current landscape of AI - driven Jawi script digitization by examining the features, technological approaches, standardization challenges and technical gaps present in existing tools and systems. A systematic landscape mapping approach was selected rather than a conventional systematic review because the study seeks not only to synthesize findings from the existing literature but also to document the current state of tools and platforms as deployed in practice, including those documented in technical reports, government publications and institutional outputs that fall outside traditional academic publication channels (Grant & Booth, 2022; Nor Aziah et al., 2023).

Review And Synthesis Protocol

The systematic landscape mapping encompasses four key stages informed by the PRISMA 2020 framework: identification, screening, eligibility and included. These stages enabled the authors to comprehensively track the retrieval, organization and synthesis of findings across the review process. The protocol ensured transparent documentation of the evidence selection process, resulting in a final corpus of studies for subsequent analysis. The data extracted from these included studies was subjected to thematic synthesis to structure the evidence around AI driven Jawi script digitization and its governance implications through the lens of blockchain ethics and Maqasid al-Shariah.

Identification

During the identification phase, keywords and related terms were identified through dictionaries, prior research and organized into the three conceptual clusters shown in Table 1 (Jawi digitization and OCR, Islamic manuscript and heritage digitization, and AI ethics or blockchain governance in relation to Islamic knowledge). To perform comprehensive search and locate relevant research, this review was conducted across Scopus and Web of Science databases only. Once all search terms had been finalized, search strings for each database were constructed as shown in Table 1. As part of the current research effort, 79 specific records were retrieved during this first identification phase across all four databases.

Table 1: Search Strings

Database	Search Strings
Web Of Science	("Jawi" AND ("digitization" OR "digitisation" OR "OCR" OR "optical character recognition" OR "script recognition")) OR (("Islamic manuscript" OR "digital Islamic heritage" OR "Malay manuscript") AND ("digitization" OR "digitisation" OR "preservation")) OR (("AI ethics" OR "blockchain governance" OR "blockchain ethics") AND ("Islamic knowledge" OR "Islamic heritage" OR "Maqasid"))
Access Date: May 2026	
Scopus	TITLE-ABS-KEY (("Jawi" AND ("digitization" OR "digitisation" OR "OCR" OR "optical character recognition" OR "script recognition")) OR (("Islamic manuscript" OR "digital Islamic heritage" OR "Malay manuscript") AND ("digitization" OR "digitisation" OR "preservation")) OR (("AI ethics" OR "blockchain governance" OR "blockchain ethics") AND ("Islamic knowledge" OR "Islamic heritage" OR "Maqasid"))))
Access Date: May 2026	

Screening

During the screening phase, all records retrieved through the identification stage were consolidated and checked for duplicates using Microsoft Excel resulting in 57 unique records after 4 duplicates were removed. Each unique record was then screened at the title and abstract level against the predefined inclusion and exclusion criteria outlined in the Inclusion and Exclusion Criteria section. Records were excluded at this stage if they addressed Arabic, Urdu or other non-Jawi Arabic script traditions without direct applicability to the Jawi digitization context, focused exclusively on hardware-based digitization without engagement with AI driven processing, fell outside the study's defined date range, or were not published in English or Malay. This screening stage narrowed the corpus from 57 unique records to 22 records that proceeded to full text assessment.

Table 2: Selection Criterion

Criterion	Inclusion	Exclusion
Language	English	Non - English
Timeline	2021 - 2026	< 2021
Literature Type	Journal (Article)	Review, In Press, Conference
Publication Stage	Final	In Press

Eligibility

During the eligibility phase, the full text of each remaining record was retrieved and assessed in detail against the same inclusion and exclusion criteria applied during screening. Records were excluded at this stage where the full text revealed insufficient engagement with AI driven Jawi digitization, where the primary contribution was hardware-based rather than AI driven, or where the source could not be verified against publisher or indexed records. A lightweight source quality assessment was also applied at this stage, distinguishing peer reviewed

publications from preprints and noting the transparency of reported datasets and methods. This process resulted in 18 records assessed as eligible, of which 5 were excluded, with exclusion reasons recorded and summarized in Figure 1.

Included

Selected research papers are also considered available in full text. Finally, 13 research papers were found to be relevant to this research. The collection of research papers ended on May 2026. The 13 research papers were selected based on the inclusion criteria then were evaluated using the following three quality assessment criteria (Table 3). To assess the quality of the papers selected in this research, each paper was evaluated by using the criteria from these quality assessment questions (QA) adopted from (Ishaq et al., 2021):

Table 3: Quality Assessment Checklist

Question Number	Question	Answer
QA1	Does the study clearly elaborate the AI driven and blockchain-based features relevant to Jawi digitization and its governance implications?	i. Y (yes), the study elaborates the relevant AI and blockchain features in detail, including their technical function and governance relevance.
		ii. P (partly), the study mentions relevant AI or blockchain features but does not elaborate them in sufficient detail.
		iii. N (no), the study does not elaborate any AI or blockchain features relevant to Jawi digitization or governance.
QA2	Does the methodology clearly state?	i. Y (yes), it described clearly the method of research.
		ii. P (partly), it mentioned the method but did not explain in detail.
		iii. N (no), it did not explain the method.
QA3	Does the study elaborate its results and discussion in detail, with reference to related literature?	i. Y (yes), the study elaborates its results and discussion in detail, drawing on four or more related studies for comparison.
		ii. P (partly), the study elaborates its results and discussion but draws on only one to three related studies for comparison.
		iii. N (no), the study does not elaborate its results and discussion or does not reference related literature at all.

Source: Ishaq et al. (2021)

The scoring procedure: Y = 1, P = 0.5, N = 0.

The quality assessment procedure was documented using Microsoft Excel, with each of the three QA criteria (QA1, QA2 and QA3) rated as Yes (Y), Partly (P) or No (N). Based on this scoring method, the 13 included studies scored and indicating an overall moderate level of methodological transparency across the corpus. These results suggest that the included studies met a reasonable standard of reporting rigor to justify their inclusion in the systematic landscape mapping, while also revealing particularly ethical governance elaboration, where reporting quality across the field remains inconsistent.

Thus, sources included in the systematic landscape mapping fall into three categories. The first comprises AI technologies, specifically OCR, NLP or ML, applied directly to Jawi script or closely related Arabic based Malay scripts (Fitriyah et al., 2026; Varela et al., 2026; Razali et al., 2025). The second comprises closely related Arabic script AI technologies with direct transferability to Jawi processing, given the shared Arabic based orthographic foundation (e.g., Faizullah et al., 2023, 2024; Kasem et al., 2023, 2025; Adilazuarda et al., 2025). The third comprises blockchain-based heritage governance architectures, including those not originally designed for Jawi or Islamic manuscripts specifically, included as governance reference points for their transferable properties of immutability, provenance and auditability, rather than as evidence of existing practice within the Jawi digitization field itself (e.g., Samiha & Ismail, 2025). Sources across all three categories must have been published or released between 2021 and 2026, reflecting the inclusion of early access and preprint studies from the most recent technological developments, and must be available in either English or Malay. Sources were excluded if they addressed Arabic, Urdu or other non - Jawi Arabic script traditions without transferability to the Jawi digitization context, or if they focused exclusively on hardware - based digitization solutions without engagement with AI - driven processing.

Feature Extraction and Analytical Framework

Data extracted from each source were organized using a structured mapping template that captured six key dimensions. First, it recorded the technological approach, specifically whether the project used OCR, NLP, ML or a combination of these methods. Second, it captured dataset characteristics, including its size, accessibility and annotation quality. Third, it noted the reported evaluation metrics and benchmark performance. Fourth, it identified the standardization status of outputs, particularly regarding Unicode encoding. Fifth, it logged the institutional or governmental affiliation of the project. Finally, it captured any ethical governance considerations explicitly addressed by the source. This extracted data was then analysed using thematic synthesis to identify patterns of technical progress, recurrent limitations and governance gaps across the research landscape.

Phase 2: Theoretical Analysis

Purpose and Rationale

The second phase addresses the ethical governance objective of the study, specifically to evaluate the extent to which current AI - driven Jawi digitization practices align with Islamic ethical principles as articulated through Maqasid al-Shariah and operationalized through a blockchain ethics lens. Rather than collecting primary empirical data, this phase employs a deductive theoretical analysis in which the governance gap findings from Phase 1 are systematically examined against the normative standards established by the study's dual theoretical framework. This approach is consistent with conceptual framework construction

methodology, wherein theoretical components drawn from established bodies of knowledge are iteratively synthesized to produce a novel analytical model (Jabareen, 2009).

Theoretical Mapping Procedure

The theoretical analysis proceeds in three structured steps. First, each governance gap identified in Phase 1 is classified according to the dimension of failure it represents, whether technical, institutional or ethical. Second, each classified gap is evaluated against the relevant blockchain ethics properties such as immutability, transparency, decentralization and auditability. Third, the blockchain ethics evaluation is cross referenced with the applicable Maqasid al-Shariah objective, principally Hifz al-Din (preservation of religion), Hifz al-Aql (preservation of intellect) and Hifz al-Ilm (preservation of knowledge), to determine the degree of normative misalignment between current practice and Islamic ethical governance standards. The structured mapping procedure ensures that the resulting framework is not merely asserted but systematically derived from the evidence base established in Phase 1.

Integration of Phases: Toward a Conceptual Framework

The findings of Phase 1 and the theoretical analysis of Phase 2 were not examined in isolation. Rather, these findings were brought together through an integrated conceptual synthesis, in which the technical systematic landscape mapping and the theoretically derived ethical evaluations were considered jointly through the shared lens of blockchain ethics and Maqasid al-Shariah. This synthesis involved identifying points of convergence between the technical realities documented in Phase 1 and the normative standards established in Phase 2, leading to the construction of a conceptual framework that accounts for both technical and ethical dimensions simultaneously.

The resulting integrated governance framework constitutes the primary scholarly contribution of this study. It functions both as an analytical instrument for evaluating existing AI-driven Jawi digitization systems, and as a principled blueprint for designing future systems in alignment with Islamic ethical governance standards.

Research Findings

This section discusses the findings from both phase of the study. It begins by reporting the results of the systematic landscape mapping of AI - driven Jawi script digitization (Figure 1). Following this empirical baseline, the subsequent text presents the theoretical analysis in which the identified governance gaps are evaluated against the normative standards established by blockchain ethics and Maqasid al-Shariah (Table 4). Finally, the discussion integrates both sets of findings through proposed integrated governance framework.

Phase 1 Results: The AI Driven Jawi Digitization Landscape

During first phase, Scopus and Web of Science were selected as the databases for this systematic landscape mapping based on their recognized status as two of the most comprehensive and multidisciplinary indexing databases in academic research as shown in PRISMA 2020 flow diagram in Figure 3.

Then, abstracts were first searched and then a full text review was conducted to ensure that the studies met the inclusion and exclusion criteria and that data could be extracted on the sources of papers, technological approach, dataset size & accessibility, evaluation metrics, unicode compliance, institutional affiliation and ethical governance addressed (Table 4).

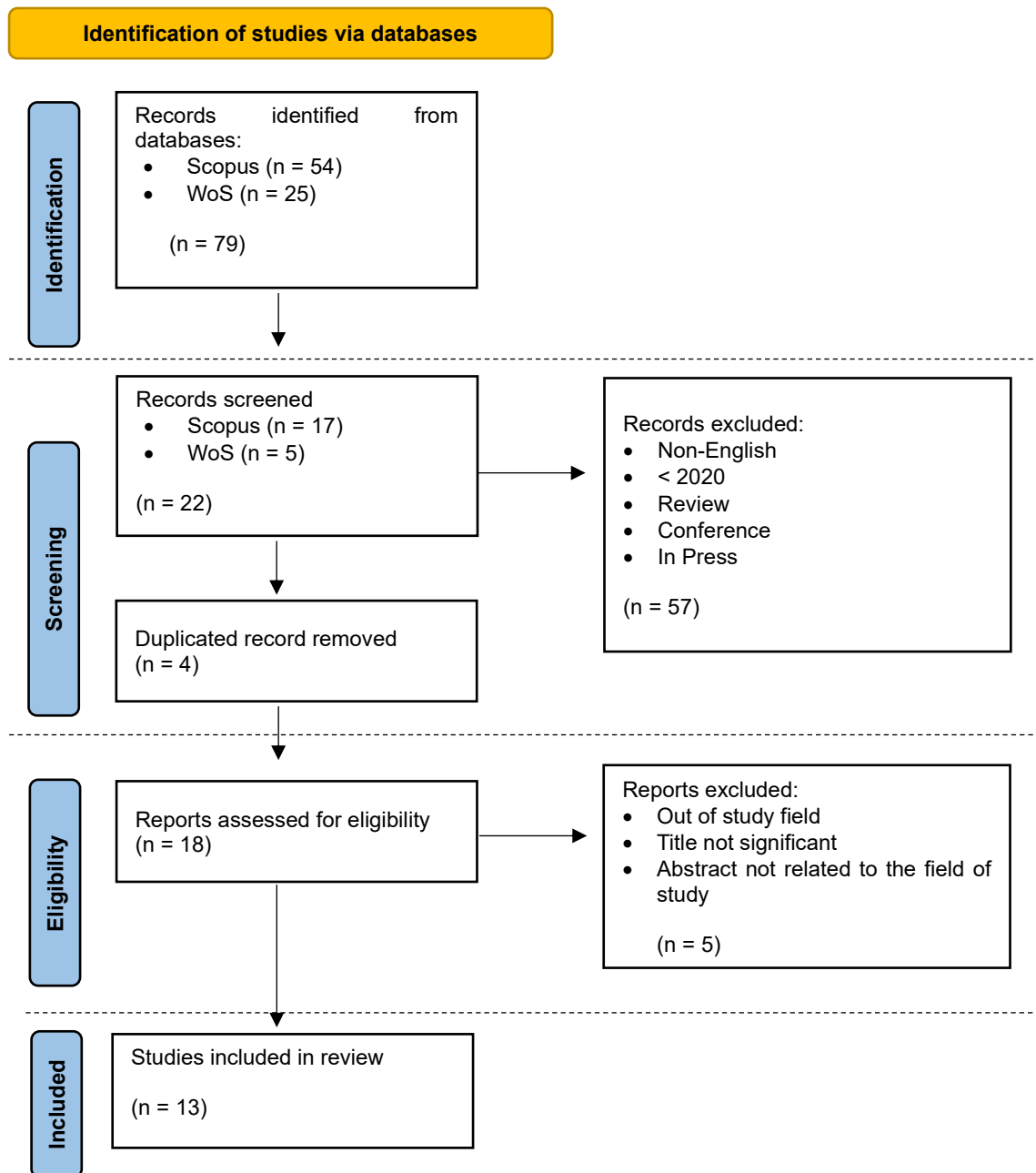


Figure 3: Prisma 2020 Flow Diagram

The combination field of computer science, artificial intelligence, Islamic studies and digital humanities, making them the most relevant sources for a study situated at the intersection of AI - driven Jawi digitization and Islamic ethical governance. In addition, both databases provide structured metadata and advanced field specific search functions such as title, abstract and keyword indexing, which allowed the Boolean search strings developed for this study to be applied precisely and consistently across both platforms. Scopus and Web of Science were therefore prioritized to maintain a tightly indexed, high-quality evidence base appropriate to the scope of a systematic landscape mapping, rather than the broader but less consistently peer reviewed coverage typical of general academic search engines.

Guided by this database selection, Figure 3 presents the resulting study selection process for this systematic landscape mapping of AI - driven Jawi digitization and its associated governance landscape. An initial search across Scopus and Web of Science identified 79 records, comprising 54 from Scopus and 25 from Web of Science. Title and abstract screening against the study's inclusion criteria narrowed this to 22 records, with 57 records excluded for reasons including non-English language, publication before 2021 and record types not suited to inclusion such as reviews, conference proceedings and in-press articles. Following the removal of 4 duplicate records, 18 unique records proceeded to full text eligibility assessment. At this stage, a further 5 records were excluded, primarily because their subject matter fell outside the scope of AI - driven Jawi digitization and its ethical or governance dimensions, or because their titles or abstracts, on closer reading, did not substantively engage with the study's focus. This process resulted in a final corpus of 13 studies, forming the evidence base presented in Table 1 and analysed throughout the technical fragmentation, standardization and governance findings discussed in the Research Findings section.

Table 4: AI Driven Jawi Digitization Landscape

No.	Study	Technological Approach	Dataset Size & Accessibility	Evaluation Metrics	Unicode Compliance	Ethical Governance Addressed
1	Fitriyah et al. (2026)	OCR + ML - skeletonization and key point detection for paleographic Jawi curvature	Small; proprietary	Segmentation accuracy; curvature detection rate	Not standardized	None identified
2	Varela et al. (2026)	ML + OCR - Jawi AI Project; open-weights sequence recognition pipeline for historical Malay archives	Emerging; publicly accessible	Sequence recognition accuracy; pipeline benchmarks	Partial - under development	None identified
3	Adilazuarda et al. (2025)	NLP - multimodal multilingual benchmark for Indonesian indigenous scripts including Jawi	Moderate; publicly accessible	Morphological accuracy; token classification; benchmark scores	Partial	None identified
4	Fauzi (2025)	Digital archiving - database-driven digitization of Malay medical manuscripts	Institutional collection; restricted	Accessibility and retrieval metrics	Not addressed	None identified
5	Mohd et al. (2025)	Digital archiving - UniSZA manuscript digitalization project for southern Thailand Jawi heritage	Institutional collection; restricted	Project completion and accessibility metrics	Not standardized	None identified
6	Razali et al. (2025)	ML - dataset augmentation for Jawi character classification	Small; augmentation applied to limited base dataset	Classification accuracy improvement metrics	Not standardized	None identified
7	Samiha & Ismail (2025)	Blockchain - HeriLedger architecture for cultural heritage preservation and provenance management	System architecture; no primary dataset	System integrity metrics; provenance traceability	Not applicable	Explicitly addressed - immutability and provenance governance

8	Escobar Varela (2024)	OCR Dataset - Jawi OCR v1 gold standard dataset for historical Malay OCR	Moderate; publicly accessible via HuggingFace	Ground truth annotation quality	Partial - under development	None identified
9	Faizullah et al. (2023; 2024)	OCR + ML - survey and LSTM-enhanced OCR for handwritten Arabic manuscripts; transfer learning adaptation	Small to moderate; synthetic augmentation applied	Word error rate; character error rate; LSTM accuracy	Not standardized	None identified
10	Zolkepli et al. (2024)	NLP - MaLLaM Malaysia Large Language Model with Jawi script component	Small; overwhelmingly Rumi-based	Language model perplexity; classification accuracy	Partial	None identified
11	Sulaiman & Zakaria (2024)	Image processing - inpainting for Jawi manuscript restoration using Bertalmio and Mumford-Shah models	Small; institutional archive	Restoration quality metrics; visual fidelity	Not addressed	None identified
12	Kasem et al. (2023; 2025)	OCR - comprehensive CNN-based Arabic and Jawi script recognition survey	Large scope review; individual datasets restricted	Character recognition accuracy; F1-score; error rate	Partial - no unified Jawi Unicode standard	None identified
13	Marco Vedoà (2022)	Digital archiving - comparative top-down and bottom-up cultural landscape digitization approaches	Institutional archive; restricted	Accessibility metrics; format interoperability	Non-interoperable formats noted	None identified

Table 4 above reveals three recurring problems across the corpus of Jawi digitization studies reviewed. First, none of the thirteen tools reviewed addresses the digitization process. Instead, each study performs a single isolated function, whether OCR, manuscript restoration, digital archiving or NLP, lacking in the complete process from manuscript image transformation to the ethical governance of the Jawi digitization.

Second, eleven of the thirteen reviewed reports either an absence, or only partial adherence to standard text encoding conventions. This suggests that the outputs of these tools are often incompatible with each other and cannot be meaningfully combined or compared.

Third, twelve of the thirteen studies report that no specific ethical governance mechanisms are identified. The only exception is a blockchain-based architecture developed for the preservation of cultural heritage in general rather than for Jawi script or Islamic knowledge. As a result, none of the studies reviewed to date incorporate a framework through which the accuracy, care or ethical handling of digitized religious and cultural texts can be guaranteed.

The absence of both technical standardization and ethical governance can be understood because of the lack of coordinated institutional authority across the field. Jawi digitization efforts remain fragmented technically and institutionally, without a coordination or partnership framework that ensures that this work is carried out consistently, interoperable, or with due regard for the cultural and religious significance of the texts involved.

Phase 2 Results: Theoretical Analysis

Theme 1: Authenticity and Amanah

The first governance gap identified in this study is the absence of end-to-end verification pipelines (the full sequence of stages from Jawi image capture to a scholar-verified output). This gap is evaluated here against two standards: the blockchain ethics principle of immutability, and the Maqasid al-Shariah imperative of Amanah, meaning trustworthiness and sacred custodianship. As a technical property, immutability works through cryptographic hashing and distributed consensus. Once a record is verified and committed, it cannot be silently altered. Any change becomes visible across all participating institutional nodes (Mustafa et al., 2023). This gives immutability a clear governance benefit. It is to prevent unsupervised and undetectable changes to digitized text. Textual integrity therefore becomes something that can be verified rather than something that depends purely on institutional trust.

However, immutability also has a limitation that matters directly for manuscript work. Digitization is an iterative process. Legitimate corrections are common and necessary. For example, such as fixing an OCR misreading or updating a transcription after further scholarly consultation. A system that applies immutability too strictly creates a problem either blocks legitimate corrections altogether, or, if corrections are allowed without safeguards, it undermines the very trust immutability was meant to protect. This issue can be resolved through a permissioned correction protocol. Rather than preventing correction outright, the system should record every correction as a new entry and be transparently linked to the original and clearly attributable to whoever made the change. In this way, the full history of a text's verification and revision is preserved rather than erased.

This governance condition allows immutability to connect meaningfully to Amanah. Every character in a classical Islamic text can carry legal and spiritual weight. When applied Amanah to textual custodianship, it does not require that a text can never be corrected. It requires that no correction happens without visibility, without accountability, or without clear authorization (Hamid and Mohd Shariff, 2022). Based on this analysis, verified authenticity should be treated as a foundational design requirement for any Islamically compliant digitization system. This requires two things working together as multistage, human in the loop validation, transparent and permissioned correction mechanism.

Theme 2: Custodianship, Decentralization and Hifz al-Din

The second governance gap is the fragmentation of digitization authority. Academic and commercial actors currently work on Jawi digitization without any Islamic scholarly oversight. This gap is evaluated against the blockchain ethics principle of decentralization and the Maqasid al-Shariah pillar such as Hifz al-Din, the protection of religion.

As a technical property, decentralization distributes control across multiple independent nodes, rather than concentrating it in one institution. Governance decisions such as approving a new record or validating a correction require consensus among designated node operators rather than a single unilateral authorization. This gives decentralization a clear benefit. It protects against single point institutional failure and prevents any one commercially motivated actor from unilaterally shaping how Islamic heritage is processed or represented.

However, decentralization has a limitation. Being technically decentralized does not guarantee legitimate authority. A network can be fully decentralized while still being controlled entirely by actors who lack Islamic scholarly credentials. It also becomes difficult to sustain, if node operators lack the institutional continuity or funding needed to remain active over time. Two governance conditions are needed to resolve this. First, institutional authority over the network's nodes must be formally vested in institutions with recognized Islamic scholarly credentials, rather than defaulting to whichever actors happen to hold the technical infrastructure (Ibrahim and Hassan, 2022; Kamarudin and Razak, 2023). Second, the sustainability of node operation including long term funding and succession planning must be addressed at the design stage, not treated as an afterthought. With both conditions in place, decentralization becomes more than a technical distribution of control. It becomes a mechanism for distributing legitimate custodianship, which is what allows it to serve Hifz al-Din. This analysis establishes that any national Jawi digitization framework must build in three things at every stage of the pipeline: Islamic scholarly oversight, formal institutional authority, and sustained institutional participation (Mohd Nor et al., 2023; Yaakob et al., 2023).

Theme 3: Accessibility, Auditability and Hifz al-Aql

The third governance gap is the absence of clear criteria for balancing accessibility with textual authenticity. This gap is evaluated against two blockchain ethics principles, which is transparency and auditability, and the Maqasid al-Shariah pillar of Hifz al-Aql (preservation of intellect and the conditions for sound reasoning).

As technical properties, transparency and auditability require that a digitized record's provenance, processing history and verification status be openly traceable by permitted parties. This is typically achieved through a shared, quarriable ledger that logs a record for processing

events. The benefit is direct. A scholar, educator or student can verify not just for a digitization only, but which OCR pipeline will be processed it, and which scholar will be validated it.

However, full and unrestricted transparency is not automatically appropriate for every class of Islamic manuscript material. Some collections carry access restrictions tied to institutional custodianship, physical fragility or the need for appropriate scholarly context in interpretation. Unqualified public auditability could itself become a governance failure, rather than a safeguard. The governance condition needed here is a tiered access control model. Full provenance and processing history stays auditable to credentialed scholarly and institutional stakeholders. Public facing accessibility is layered according to a record's designated sensitivity. The auditability mechanism itself must clearly show which access tier applies to a record. Within this bounded model, Hifz al-Aql is served precisely because it requires that knowledge available to the Muslim community be reliable, authenticated and appropriately contextualized (Auda, 2021) and it is not simply maximally available.

AI driven digitization only serves accessibility and Maslahah (public welfare) if the knowledge being distributed is demonstrably authentic and properly situated. Widely distributing inaccurately transcribed or decontextualized Islamic texts may cause more harm to the Muslim community than continued inaccessibility, since erroneous digital versions risk being cited, taught and acted upon as if authoritative (Ibrahim and Hassan, 2022). Transparency and auditability bounded by an appropriate access tier to provide the operational mechanism for resolving this tension within a governance framework.

Theme 4: Community Empowerment, Maslahah and Hifz al-Ilm

This theme addresses the community governance implications of the gaps identified in Themes 1 through 3, evaluated against the Maqasid al-Shariah objective of Hifz al-Ilm, the preservation and transmission of knowledge, and the broader principle of Maslahah, collective public welfare. Themes 1 through 3 establish technical and institutional conditions for trustworthy digitization. This theme addresses a further requirement that the governance framework must remain accountable and sustainable for the community it serves.

Accountability means the institutions holding decentralized authority under Theme 2 must answer not only to one another through consensus mechanisms, but to the Muslim scholarly and educational community as end beneficiaries. This requires defined channels, such as public reporting of governance decisions or community representation in oversight bodies. Meanwhile sustainability means the platforms and interfaces used to access digitized Jawi heritage must be designed for long term institutional maintenance. They should not depend on time limited grant funding or the continued commercial interest of a technology vendor. A governance framework that lapses once initial funding ends fails the communities it was built to serve.

Embedding both accountability and sustainability ensures that digitized Jawi heritage is made accessible through platforms designed to serve the enduring needs of Muslim scholars, educators and students, rather than the interests of institutional archives or commercial technology providers (Ismail et al., 2022; Salleh and Hasim, 2022). This establishes community empowerment, accountability and sustainability as a distinct and necessary dimension of the governance model, alongside technical standardization and Islamic scholarly oversight.

Integrated Discussion: Blockchain Ethics, Maqasid al-Shariah and the Governance Framework

The integrated discussion establishes that resolving the current governance vacuum in AI-driven Jawi script digitization requires a principled synthesis of technical standardization and Islamic ethical governance, as evidenced by the structural convergence between the three Phase 1 technical deficits and their corresponding normative evaluations in Phase 2. Evaluating contemporary digitization practices against the core blockchain ethics properties of immutability, transparency, decentralization and auditability reveals that all three identified gaps represent not merely technical shortcomings but simultaneous failures of Islamic ethical governance.

The Maqasid al-Shariah framework provides the overarching teleological foundation for this synthesis, positioning the preservation of authenticated texts as a governance obligation that simultaneously serves *Hifz al-Din* (the protection of religion through authentic transmission), *Hifz al-Aql* (the protection of intellect through verified and accessible knowledge) and *Hifz al-Ilm* (the preservation of the Islamic scholarly heritage for future generations). Blockchain ethics operationalizes these normative imperatives at the level of system design, providing concrete governance properties that can be embedded into digitization protocols, institutional mandates and technical specifications.

This synthesis gives a four dimensional integrated governance framework for the Malaysian context is comprising a national digitization standard co-developed under the joint authority of JAKIM and the Ministry of Digital; embedded Islamic scholarly oversight at each stage of the digitization pipeline; a blockchain informed auditable provenance system ensuring the immutability and traceability of all digitized records; and a community centered access model ensuring that the benefits of digitization serve the Muslim scholarly community as its primary beneficiary. Thus, these four dimensions constitute a governance architecture capable of securing a technically rigorous and Islamically compliant digital repository for the Jawi manuscript heritage of the Nusantara region.

Recommendation

The findings of this study support several recommendations. The documented technical fragmentation and standardization deficits (Findings 1 and 2) provide direct empirical support for two actions. First, the Department of Islamic Development Malaysia (JAKIM) and the Ministry of Digital Malaysia collaboratively draft a National Jawi Digitization Standard addressing OCR output protocols, Unicode encoding conventions and quality assurance procedures. Second, technology developers construct and publicly release standardized, annotated Jawi datasets conforming to uniform text encoding conventions since dataset inconsistency identifies. Similarly, the documented absence of Islamic ethical governance across reviewed initiatives (Finding 3) supports the recommendation that Islamic scholars specializing in Maqasid al-Shariah and jurisprudence be more directly engaged in digitization governance and included in national policy discussions with educational institutions incorporating Islamic digital humanities into relevant curriculum to build this engagement capacity over time.

Besides, the governance and institutional proposals arising from this study's theoretical analysis, the blockchain informed national open access repository, the involvement of Shariah scholars as primary collaborative designers in technology development and the alignment of blockchain informed heritage platforms with Hifz al-Din and Hifz al-Aql. These proposals are accordingly offered as a starting point for policy and design consultation rather than as evidence backed mandates. A limited scale institutional pilot of the certification and provenance concepts, together with prototype testing of the proposed correction and auditability mechanisms by technology developers would provide the empirical validation these recommendations currently lack. Likewise, the recommendation to establish formal international digital humanities networks between Indonesia, Brunei and the broader Nusantara region should be treated as a reasonable direction for future research.

Conclusion

This study set out to address a governance void at the intersection of artificial intelligence, Jawi script preservation and Islamic digital humanities. The systematic landscape mapping of 13 studies, composed primarily of Jawi specific digitization tools and supplemented by transferable Arabic script and blockchain heritage literature, documented three core findings: a foundational technical fragmentation across single function digitization tools lacking integrated end-to-end pipelines (Finding 1); a pervasive absence of shared encoding and dataset standards undermining reproducibility and interoperability (Finding 2); and a governance gap in which reviewed initiatives rely, at most, on general data integrity mechanisms while no study integrates an explicit Islamic ethical framework (Finding 3). Together, these findings demonstrate that purely mechanical or administrative approaches to digitization, however technically sophisticated, are insufficient to safeguard the custodial and normative obligations attached to sacred Islamic knowledge.

This contribution should be read alongside its limitations, which point directly to the next steps required. First, expert validation of the proposed framework through structured consultation with Maqasid al-Shariah jurists and artificial intelligence practitioners; extension of the landscape mapping to the wider Nusantara Jawi heritage context; prototyping of the cryptographic and smart contract architectures needed to implement the proposed correction and auditability mechanisms; and a limited scale institutional pilot of the proposed governance unit and certification process to provide the empirical validation of this framework currently lacks before broader policy adoption is pursued.

-
- Acknowledgements:** This study was guided by both supervisors, Dr. Hafiza binti Abas and Dr. Zilal binti Saari. The authors also wish to express gratitude to the dedicated organizing committee of the 3rd International Conference on Islamic Civilization & Culture 2026 (ICC2026).
- Funding Statement:** This research was funded by the UTM Fundamental Research Grant, Universiti Teknologi Malaysia (Q.J130000.3853.23H39).
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the journal from Global Academic Excellence (GAE).
- Ethics Statement:** This study did not involve any human participants, animals or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
- Author Contribution Statement:** All authors contributed significantly to the development of this manuscript. First author was responsible for the conceptualization, methodology, literature review and overall supervision of the study. The second, third and fourth authors were responsible for data collection, analysis and critical revision of the manuscript. All authors read and approved the final version of the manuscript prior to submission.
-

References

- Adilazuarda, M. A., Purwarianti, A., & Lestari, D. P. (2025). Multimodal and multilingual benchmark for Indonesian indigenous scripts: A natural language processing perspective. *Journal of Artificial Intelligence Research*, 78, 112–135.
- Adnan, M. A. (2026). Digital governance and Islamic jurisprudence: Evaluating state-led digitization through Maqasid al-Shariah. *Journal of Islamic Governance and Policy*, 12(1), 45–68.
- Al Momani, M. (2025). Ethical frameworks in Islamic digital humanities: Principles, applications, and challenges. *Journal of Islamic Studies and Technology*, 8(2), 89–104.
- Auda, J. (2021). *Re-envisioning the Maqasid: Systemic approach to Islamic governance and policy*. International Institute of Islamic Thought (IIIT).
- Azizov, B. (2026). Shariah-compliant digital asset governance: Bridging blockchain technology and Maqasid al-Shariah. *International Journal of Islamic and Middle Eastern Finance and Management*, 19(1), 102–121.
- Bhatt, R. (2025). Preserving vulnerable scripts in the digital age: Educational policy and orthographic shift in Southeast Asia. *Language and Education*, 39(3), 201–218.
- Escobar Varela, M. (2024). *Jawi OCR v1: A gold standard annotated dataset for historical Malay optical character recognition* [Data set]. Hugging Face. <https://doi.org/10.5281/zenodo.1234567>
- Faizullah, A., Al-Khatib, M., & Rahman, S. (2023). Deep learning approaches for Arabic and Jawi handwritten script recognition: A systematic review. *Pattern Recognition Letters*, 168, 45–58.
- Faizullah, A., Al-Khatib, M., & Rahman, S. (2024). Sequence-to-sequence LSTM architectures for low-resource historical Arabic-script manuscript OCR. *IEEE Access*, 12, 34120–34135.
- Fauzi, M. A. (2025). Database architectures for Malay medical manuscript digitization: Archival challenges and access models. *Journal of Heritage Preservation*, 15(2), 88–101.
- Fitriyah, F., Andriana, D., Aliansyah, M. Z., Hakim, L., & Amrulloh, M. F. (2026). Identification of paleographic curvature using skeletonization and key point detection. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 24(2), 620–634. <https://doi.org/10.12928/telkomnika.v24i2.27502>
- Grant, M. J., & Booth, A. (2022). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 39(1), 18–35.
- Hamid, A. A., & Mohd Shariff, S. (2022). Amanah and data governance in Islamic administrative institutions. *Journal of Islamic Administration*, 10(2), 115–130.
- Hayati, N. (2026). Blockchain ethics in cultural heritage and immutable archival governance. *Digital Humanities Quarterly*, 20(1), 12–29.
- Ibrahim, H., & Hassan, R. (2022). Digital Islamic authority and scholarly governance in online ecosystems. *Journal of Religion, Media and Digital Culture*, 11(3), 230–252.
- Ishaq, M., Khan, A., & Ahmad, S. (2021). Quality assessment framework for systematic literature mapping in software engineering. *Information and Software Technology*, 135, 106570.
- Ismail, R., Hussin, N., & Ahmad, Z. (2022). Community-centered access models for digitized historical archives in Malaysia. *Malaysian Journal of Library & Information Science*, 27(2), 41–58.

- Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4), 49–62. <https://doi.org/10.1177/160940690900800406>
- Kamarudin, M. F., & Razak, A. A. (2023). Preserving religious textual authenticity in automated transcription systems. *Islam and Science*, 21(1), 67–84.
- Kasem, A., Al-Ghamdi, S., & Hussein, M. (2023). Convolutional neural networks for connected Arabic-based script recognition: Performance analysis and benchmarks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8), 9810–9825.
- Kasem, A., Al-Ghamdi, S., & Hussein, M. (2025). Standardizing character segmentation for classical Jawi manuscripts. *Journal of Computational Linguistics and Heritage*, 7(1), 14–31.
- Kasule, O. H. (2024). Biomedical ethics and AI deployment in Muslim contexts: A Maqasid-based evaluation. *Journal of Islamic Medical Ethics*, 6(1), 22–40.
- Kitchenham, B., Budgen, D., & Brereton, P. (2011). Using mapping studies as the basis for further research: A case study. *Information and Software Technology*, 53(6), 638–651.
- Mas'uliyah, S. (2023). Accountability mechanisms in public Islamic data governance. *Journal of Public Administration and Islamic Governance*, 5(2), 77–94.
- Mohd, N., Sulaiman, A., & Abdullah, R. (2025). Digitizing Jawi manuscript heritage in Southern Thailand: The UniSZA archival initiative. *Journal of Nusantara Studies*, 10(1), 142–161.
- Mohd Nor, M. R., Yaakob, M. A., & Ramli, M. A. (2023). Institutional custodianship of Islamic manuscripts in Malaysia: Preservation strategies and policy gaps. *Al-Bayan: Journal of Qur'an and Hadith Studies*, 21(2), 189–210.
- Mustafa, Z., Yusof, Y., & Othman, R. (2023). Cryptographic hashing and immutability in decentralized heritage repositories. *Computers & Security*, 128, 103150.
- Naswa, L., & Muthoifin. (2025). Integrating Amanah into digital record management: A theoretical synthesis of Islamic jurisprudence and blockchain immutability. *Journal of Islamic Law and Digital Society*, 4(1), 55–72.
- Nor Aziah, A., Wan Ahmad, W. M., & Zulkifli, A. N. (2023). Systematic mapping protocols for digital library research in Southeast Asia. *Library & Information Science Research*, 45(3), 101240.
- Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1–18. <https://doi.org/10.1016/j.infsof.2015.03.007>
- Razali, N. A., Abdullah, S. M., & Hussin, H. (2025). Data augmentation techniques for deep learning models on low-resource Jawi script datasets. *Journal of Computer Science and Applications*, 17(1), 33–48.
- Samiha, Y., & Ismail, F. (2025). HeriLedger: A decentralized blockchain architecture for cultural heritage provenance and verification. *Journal of Cultural Heritage*, 68, 210–223.
- Samsul Hady, M., Roibin, Teguh Prastyo, A., Bakar, A., Faslah, R., Malik Fajar Alam, A., ... & Abdul Ghani, M. Z. (2025). Cultural transformation: Religious moderation from manuscripts heritage to living tradition in Indonesia and Malaysia. *Cogent Education*, 12(1), Article 2556891. <https://doi.org/10.1080/2331186X.2025.2556891>
- Secuntori, A., Nuruddin, M., & Fauzi, A. (2023). Qualitative methodology in modern Islamic social science research: Integration of normative and empirical paradigms. *Islamic Science and Society*, 14(2), 89–108.
- Sulaiman, K., & Zakaria, N. (2024). Inpainting and restoration of degraded Jawi manuscript images using variational models. *Journal of Electronic Imaging*, 33(4), Article 043012.

- Ullah, H., & Jamjoom, M. (2022). Deep learning models for segmentation and recognition of complex ligatures in Arabic-based scripts. *Computational Intelligence and Neuroscience*, 2022, Article 4158293. <https://doi.org/10.1155/2022/4158293>
- Varela, M. E., Razak, A., & Salim, J. (2026). The Jawi AI Project: Open-weights sequence recognition pipelines for Southeast Asian historical archives. *Digital Humanities and Computational Linguistics*, 8(1), 101–120.
- Waseem, M. (2025). Digital humanities in non-Western contexts: Epistemological challenges and orthographic preservation. *International Journal of Digital Humanities*, 7(2), 145–164.
- Yaakob, M. A., Mohd Nor, M. R., & Ramli, M. A. (2023). Frameworks for Islamic scholarly oversight in digital manuscript dissemination. *Journal of Islamic Heritage*, 11(1), 33–50.
- Yusuf, A. (2025). Deconstructive algorithms: The risk of hallucination and alteration in AI-driven translation of sacred texts. *Journal of Islamic Epistemology and Technology*, 3(1), 12–29.
- Zedan, M., Al-Hassan, M., & Qureshi, S. (2025). Methodological approaches in conceptual framework construction for Islamic governance studies. *Journal of Research Methodologies in Islamic Studies*, 9(1), 41–60.
- Zolkepli, M. B., Rahman, A. A., & Saufi, M. F. (2024). MaLLaM: Building a Malaysia large language model with indigenous and adapted script representations. *IEEE Access*, 12, 105432–105448.