INDONESIAN JOURNAL OF MANAGEMENT REVIEWS
AMUYA
e-ISSN: 3109-0974 | p-ISSN: 3090-5516 | DOI: 10.61860
Website: https://amuya.gemaperencana.id
Research Paper

Artificial Intelligence and Religious Policy: Ethical Governance in the Digital Public Sphere

Wildanul Hakim[1*]
Regional Development Planning Agency of Bima City, West Nusa Tenggara, Indonesia

Noerhayati Huda[2]
Regional Office of the Ministry of Religious Affairs, West Nusa Tenggara Province, Indonesia

Thariq Thieqiaq Thang[3]
Beihang University, Beijing, China

Corresponding Author's E-mail: wildanulh@gmail.com

Article History	Received	Revised	Accepted
	(August 2nd, 2026)	(September 4th, 2026)	(October 28th, 2026)

Keywords:
Algorithmic Governance; Artificial Intelligence; Digital Public Sphere; Religious Policy; Sociotechnical Systems;

Abstract

Artificial intelligence increasingly mediates religious communication, knowledge, and participation in the digital public sphere, yet research remains fragmented across digital religion, human rights, and AI governance. This study examines how AI-mediated religious communication shapes authority, visibility, and participation; identifies associated ethical and rights-based concerns; and analyzes the institutional conditions relevant to its governance. Using a qualitative, non-experimental, exploratory-analytical design, the study draws on academic publications, policy documents, regulatory instruments, institutional guidelines, human-rights reports, and documented analyses of digital religious practices. Materials were purposively selected for substantive relevance and examined through structured reading and extraction, open and focused coding, thematic synthesis, source triangulation, and iterative cross-source comparison. The analysis identified three interconnected domains: access, visibility, and religious authority; religious freedom, inclusion, and algorithmic risks; and accountability, contestability, and institutional responsibility. AI-mediated technologies can broaden access to religious resources while simultaneously shaping the visibility and perceived authority of religious actors and interpretations. Expanded access can coexist with unequal representation, privacy risks, profiling, and constraints on meaningful participation, while distributed responsibility can create accountability gaps when consequential outcomes are difficult to explain or challenge. These relationships are context-dependent rather than predetermined effects of AI. By integrating sociotechnical and algorithmic-governance perspectives, the study provides an analytical framework for understanding ethical governance as the intersection of technological mediation, religious pluralism, fundamental rights, participation, and identifiable institutional responsibility.

To cite this article: Wildanul Hakim, Noerhayati Huda, Thariq Thieqiaq Thang. (2026). Artificial Intelligence and Religious Policy: Ethical Governance in the Digital Public Sphere. *AMUYA: Indonesian Journal of Management Reviews, Volume 2 (2)*, Page: 689-712. DOI: https://doi.org/10.61860/amuya.v2i2.122

This article is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License ©2026 by author/s

INTRODUCTION

Research Background

Artificial intelligence (AI) is increasingly embedded in the digital infrastructures through which individuals access information, communicate, and participate in public life. Recommendation systems, automated moderation, personalization, conversational agents, and generative AI not only facilitate communication but also mediate the visibility, organization, and circulation of information, creating opportunities for broader access while raising concerns about selective exposure, bias, and unequal representation (Roche et al., 2023; Trattner et al., 2022). In Indonesia, about 221 million people (79.5%) used the internet in early 2024 (Bisnis.com, 2024), religious authority circulates through YouTube and digital Sufi networks (Harianto et al., 2025; Wahyudi et al., 2025), and a circular on AI ethics was issued in 2023 (Menteri Komunikasi dan Informatika Republik Indonesia, 2023); Indonesia is therefore used as the illustrative policy context.

These developments are particularly consequential in religious contexts, where AI increasingly intersects with religious communication, education, interpretation, leadership, and community practices (He, 2024; Tsuria & Tsuria, 2024). Religious leaders have begun incorporating generative AI into communication and leadership practices as an idea generator, research assistant, co-author, and translator, while simultaneously negotiating concerns about authenticity, spiritual depth, and the appropriate boundaries of human–machine collaboration (Cheong & Liu, 2025).

AI should therefore be understood not merely as a technological tool adopted by religious actors, but as part of a changing digital environment through which religious knowledge, representation, and participation are increasingly mediated.

This mediation becomes a religious-policy concern when algorithmic systems influence religious visibility, authority, autonomy, and fundamental rights. Recommendation systems and engagement-based personalization can shape which religious interpretations become visible and which actors acquire greater digital prominence, potentially shifting religious authority toward actors more adaptive to algorithmic visibility (Wahyudi et al., 2025). Generative AI presents an additional challenge because apparently coherent outputs do not necessarily capture the doctrinal, historical, and moral complexity of religious traditions (Tsuria & Tsuria, 2024). Religious policy is defined here as the public decisions, norms, and institutional arrangements through which states regulate religion-related communication, together with platform rules and religious self-regulation where these perform public functions or are shaped by law (European Parliament & Council of the European Union, 2024; Gillespie, 2018); the latter are distinguished from state policy in the analysis.

These developments also intersect with freedom of religion or belief because opaque AI systems may influence how individuals encounter, express, and practice religion online, while this right has received comparatively less attention in AI-related human-rights

scholarship than privacy and freedom of expression (Ashraf, 2022). Broader research further indicates that AI and religious freedom need to be examined alongside socioeconomic and institutional conditions rather than solely as matters of technological adoption (He, 2024).

Ethical governance in religiously plural digital environments consequently requires attention not only to technological performance but also to diversity and representation (Roche et al., 2023), meaningful human control (Siebert et al., 2023), and identifiable accountability across sociotechnical AI systems (Novelli et al., 2024).

Recent studies have begun to address different dimensions of this emerging relationship between AI and religion. Tsuria and Tsuria (2024) examined how generative AI represents religious traditions and demonstrated the difficulties these systems face in representing complex religious issues. Cheong and Liu (2025) investigated human–machine collaboration among religious leaders and showed how generative AI is being incorporated into religious communication and leadership practices.

From a human-rights perspective, Ashraf (2022) examined the implications of opaque AI systems for freedom of religion or belief online, while He (2024) investigated the relationship between AI, socioeconomic conditions, and religious freedom across countries. At the platform level, Wahyudi et al. (2025) examined algorithmic personalization on YouTube and identified implications for informational diversity and digitally mediated religious authority. In parallel, responsible-AI scholarship has emphasized diversity and inclusion (Roche et al., 2023), human-centered governance and stakeholder participation (Sigfrids et al., 2023), meaningful human control (Siebert et al., 2023), and institutional accountability (Novelli et al., 2024).

Despite these advances, the literature remains fragmented across AI ethics, digital religion, human rights, and platform governance. An integrated religious-policy account of how AI-mediated religious authority, freedom of religion or belief, inclusion, institutional accountability, and democratic participation interact within the digital public sphere remains underdeveloped. This gap is important because AI-mediated religious environments involve multiple actors and governance responsibilities that cannot be adequately understood through technological, theological, or rights-based analysis in isolation. Ashraf (2022) and Quirós-Fons (2025) already link AI and freedom of religion or belief, but they do not connect algorithmic visibility and authority with accountability, nor examine how specific instruments treat religion; this study addresses both.

Addressing these interconnected challenges requires an approach to AI and religious policy that integrates technological mediation, religious pluralism, fundamental rights, and institutional responsibility. AI operates within sociotechnical arrangements in which responsibility is distributed among developers, deployers, public institutions, religious organizations, civil society, and users, making accountability a multidimensional rather than exclusively technical requirement (Novelli et al., 2024).

Human-centered AI governance similarly emphasizes societal values, transparency, communication, and meaningful stakeholder participation (Sigfrids et al., 2023), while

meaningful human control remains important where autonomous systems complicate the attribution of responsibility (Siebert et al., 2023).

Accordingly, the present study examines how artificial intelligence intersects with religious policy in the digital public sphere, focusing on religious authority, freedom of religion or belief, inclusion, accountability, transparency, and democratic participation. Rather than treating AI as inherently beneficial or harmful, the study investigates the institutional and ethical conditions under which AI-mediated religious environments can be governed consistently with human dignity, religious pluralism, and the public interest. In this article, religious policy is used in a conceptual sense, referring to the public decisions, norms, and institutional arrangements through which states, platforms, and religious institutions shape religion-related communication; the study is not confined to a single jurisdiction and, it complements the scholarly synthesis with a focused reading of selected international and Indonesian policy and regulatory instruments (Table 3).

Problem Identification

The central problem addressed in this study is the absence of a sufficiently integrated religious-policy perspective for understanding and governing the multiple implications of AI-mediated religious environments. Existing scholarship has examined religious representation, digitally mediated authority, freedom of religion or belief, and responsible AI governance, yet these issues are often treated as separate technological, theological, human-rights, or governance concerns. Substantively, responsibility for AI-mediated religious content is unclear: language models show anti-Muslim bias (Abid et al., 2021), generative AI answers religious questions without an accountable author (Tsuria & Tsuria, 2024), and Indonesian law classifies religion as general rather than specific personal data (Republik Indonesia, 2022, Article 4).

A more integrated analysis is therefore needed to clarify how these interconnected concerns can be understood within religious policy and what institutional and ethical conditions are relevant to the governance of AI-mediated religious environments.

Research Questions

1. How is AI-mediated religious communication understood in the literature to reconfigure religious authority, visibility, and participation in the digital public sphere?
2. What ethical and human-rights concerns arise from AI-mediated religious environments, particularly in relation to freedom of religion or belief, inclusion, transparency, and accountability?
3. What institutional and ethical conditions do the literature and relevant policy instruments identify as relevant to governing AI-mediated religious environments in ways consistent with religious pluralism, democratic participation, and the public interest?

Research Objectives and Significance

This study aims to examine the intersection of artificial intelligence and religious policy in the digital public sphere by integrating questions of technological mediation, fundamental rights, and institutional responsibility. Specifically, it seeks to analyze how AI-mediated religious communication is understood to reconfigure religious authority, visibility, and participation; identify the ethical and human-rights concerns associated with AI-mediated religious environments, particularly those related to freedom of religion or belief, inclusion, transparency, and accountability; and examine the institutional and ethical conditions relevant to governing these environments in ways consistent with religious pluralism, democratic participation, and the public interest.

The significance of this study lies in its interdisciplinary examination of AI and religious policy within the digital public sphere. Theoretically and conceptually, it connects perspectives on digital religion and religious authority with human-rights and AI-governance scholarship, providing a broader analytical basis for understanding the relationships among AI-mediated religious authority, freedom of religion or belief, inclusion, accountability, and democratic participation.

The study is relevant to ongoing discussions on the governance of AI in religiously plural digital environments, particularly where technological mediation intersects with religious expression, institutional responsibility, and public participation. By organizing these concerns within a common analytical perspective, the study provides a foundation for further research and policy discussion on AI-mediated religious environments. Potential users include ministries of religious affairs (guidance on AI use), data-protection authorities (classification of religious data), platforms' trust-and-safety teams (moderation and recommendation of religious content), and religious councils (review of AI-generated religious answers).

LITERATURE REVIEW

AI-mediated religious environments increasingly connect religious authority with algorithmic visibility, participation, and fundamental rights. Digital mediation does not simply replace established religious authority but reconfigures it through interactions among institutional legitimacy, religious expertise, audience recognition, and platform mechanisms (Ergen, 2024; Harianto et al., 2025). Harianto et al. (2025) show for Indonesian digital Sufism that metrics and emotional resonance act as proxies for spiritual credibility, while legitimacy remains negotiated with the sanad and institutional recognition.

Scholarship on AI and religion similarly shows that digital technologies are becoming intertwined with religious practice, identity, community formation, agency, and power (Rähme & Prohl, 2025). As computational systems increasingly mediate religious knowledge, questions of authority also intersect with freedom of religion or belief, particularly where surveillance, automated classification, and digital infrastructures affect religious expression and participation (Quirós-Fons, 2025).

Governance scholarship complements this perspective by emphasizing the institutional conditions through which ethical principles become operational. Although transparency, accountability, fairness, human oversight, and social well-being are widely recognized in responsible AI frameworks, their implementation depends on organizational processes and clearly assigned responsibilities (Batool et al., 2025; Papagiannidis et al., 2025).

Accountability also requires contestability, enabling affected individuals and communities to question AI-supported decisions and the sociotechnical arrangements behind them (Alfrink et al., 2023; Cohen & Suzor, 2024). Participation is equally important, yet involvement in governance does not necessarily provide affected stakeholders with substantive influence over technological decisions (Denton & Erete, 2023).

This concern is particularly relevant when AI governance invokes the public interest, because defining the common good requires public deliberation rather than technical optimization alone (Coeckelbergh, 2025). Since responsibility for AI governance is distributed across public authorities, technology providers, institutions, and users (David et al., 2024), AI-mediated religious environments are best examined through an analytical framework that connects transparency, accountability, contestability, participation, and institutional responsibility with religious pluralism and freedom of religion or belief.

Three further strands complete the review: empirical evidence of religious bias, as a large language model persistently associated Muslims with violence (Abid et al., 2021); later work showing that digital creatives acquire religious authority alongside institutions (Campbell, 2020); and state regulation of online religious speech implemented through platform moderation (Gillespie, 2018). Appendix Table A1 summarizes the sources analysed.

THEORETICAL AND CONCEPTUAL FRAMEWORK

Theoretical Framework

This study draws first on a sociotechnical systems perspective, which understands technological and social arrangements as interdependent rather than treating technology as an autonomous determinant of social outcomes. Originating in the work of Trist and colleagues, this perspective emphasizes that technological systems operate within social structures, organizational arrangements, human practices, and wider environments that jointly shape their functioning and consequences (Pasmore & Khalsa, 1993; Trist, 1981; Trist & Bamforth, 1951). Classic sociotechnical systems theory is retained because it treats technical and social subsystems as jointly designed and thus directs attention to organizational responsibility; platform studies specify the platform layer (Gillespie, 2018). Each perspective is linked below to the concepts it supplies and the research question it supports.

Contemporary AI scholarship extends this logic by conceptualizing AI governance as a sociotechnical problem in which technical components interact with institutional rules, organizational practices, human judgment, and societal values (Lahusen et al., 2024; Zuiderwijk et al., 2021). This perspective is particularly relevant to the first two research questions because AI-mediated religious communication cannot be explained solely by

algorithmic capabilities: religious authority, visibility, participation, and rights emerge through interactions among AI systems, platforms, religious actors, institutions, and users.

To explain the governance implications of those interactions, the study complements the sociotechnical perspective with algorithmic governance, which focuses on how computational systems participate in ordering social environments through data collection, classification, prediction, recommendation, and other forms of automated or semi-automated mediation (Yeung, 2018).

Algorithmic systems are consequential because their design and operation can redistribute visibility, discretion, and decision-making power while simultaneously complicating conventional relationships of responsibility and democratic oversight (Sun & Medaglia, 2019). Recent public-governance research similarly shows that AI can reshape institutional power relations and create tensions involving public values, accountability, and justice rather than functioning merely as an efficiency-enhancing instrument (Hjaltalin & Sigurdarson, 2024).

This perspective therefore supports the third research question by directing analysis toward the institutional conditions under which algorithmic mediation is governed and responsibility is allocated. The two perspectives perform complementary analytical functions: the sociotechnical perspective identifies AI-mediated religious environments as configurations of mutually dependent technological and social elements, whereas algorithmic governance explains how computational mediation within those configurations can structure visibility, influence, and responsibility. Concepts such as religious pluralism, freedom of religion or belief, transparency, accountability, contestability, and participation are not presented here as components originating from either theory; they are integrated subsequently through the study's conceptual synthesis to operationalize the specific religious-policy concerns examined in this research.

Digital religion and mediatization supply the concepts of authority and visibility for the first research question: the Internet reconfigures religious authority, which is negotiated among institutions, individuals, and communities (Campbell, 2007; Campbell, 2020), and media act as agents of religious change (Hjarvard, 2008).

Human-rights law supplies the normative concepts for the second research question: freedom of thought, conscience, and religion, with limitations only when prescribed by law and necessary (Article 18), privacy (Article 17), expression (Article 19), and non-discrimination (Article 26) (United Nations, 1966). Because these obligations bind states while much AI mediation is performed by platforms, the UN Guiding Principles add the corporate responsibility to respect rights through due diligence and remedy (United Nations Human Rights Office of the High Commissioner, 2011), and states translate their duty to protect into platform governance through regulation.

Conceptual Framework

Building on the theoretical foundations and literature synthesis, this study conceptualizes AI-mediated religious environments as interconnected relationships among

technological mediation, religious authority, fundamental rights, and institutional governance. A sociotechnical understanding positions AI-mediated communication within interactions among technological systems, institutional arrangements, and human actors, while algorithmic governance highlights how computational mediation can structure social processes and distributions of responsibility (Trist, 1981; Yeung, 2018).

Within this environment, algorithmic mediation of religious knowledge and visibility is connected to religious authority and participation, while freedom of religion or belief, inclusion, and religious pluralism provide the rights-based dimensions through which these processes are assessed (Quirós-Fons, 2025).

Governance is subsequently examined through transparency, accountability, contestability, participation, and institutional responsibility, reflecting the need to connect ethical principles with mechanisms through which AI-mediated decisions can be questioned and governed (Alfrink et al., 2023; Cohen & Suzor, 2024).

Synthesizing these foundations, this study organizes the analysis into four interconnected domains: AI-mediated religious communication and visibility; religious authority and participation; freedom of religion or belief, inclusion, and pluralism; and transparency, accountability, contestability, and institutional responsibility. These domains are used as analytical categories rather than independent variables or predetermined causal relationships, providing an integrated framework for examining ethical governance in the digital public sphere. Figure 1 shows this initial framework; how the analysis revised it is reported in the Results.

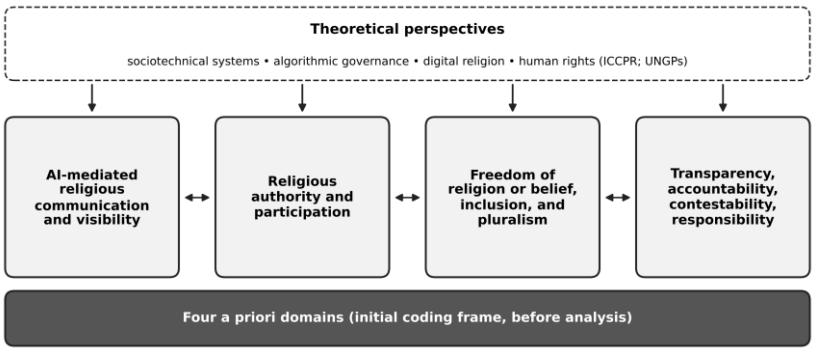


Figure 1. Initial Analytical Framework Defined Before the Analysis
Note. Initial framework defined before the analysis. Source: Authors' elaboration.

METHODOLOGY

This study employed a qualitative, non-experimental, exploratory-analytical design to examine the intersections of artificial intelligence, religious policy, and ethical governance in the digital public sphere. The study combined documentary analysis, conceptual interpretation, and thematic synthesis because its purpose was to interpret relationships among AI-mediated religious communication, religious authority, fundamental rights, participation, and institutional governance rather than to test statistical relationships or causal effects. In substance, the design is a conceptual synthesis based on an integrative review of scholarly literature (Jaakkola, 2020; Snyder, 2019), complemented by a focused reading of policy and regulatory instruments.

Documentary analysis is appropriate for systematically examining documents as sources of meaning, institutional discourse, and contextual evidence while allowing information from different documentary forms to be interpreted in relation to the research problem (Bowen, 2009; Morgan, 2022). The documentary material comprised scholarly publications on AI, religion, human rights, and governance, including analyses of digital religious practices relevant to the research questions. Scholarly sources were identified in using for, in English and Indonesian; after screening records and full texts, 26 substantive sources were retained, and two (Abid et al., 2021; Campbell, 2020) were added in revision through a search for counter-evidence (n = 28; Appendix Table A1); the other references are methodological or theoretical. Nine policy and regulatory instruments are analysed in Table 3.

Sources were purposively selected on the basis of their substantive relevance to AI and religion; engagement with ethics, rights, governance, or public participation; relevance to religious communication in the digital public sphere; and scholarly credibility. Purposive selection was used because qualitative inquiry requires information-rich materials that are directly relevant to the phenomenon and analytical objectives under investigation (Campbell et al., 2020). Substantive relevance meant addressing AI mediation in relation to religion or governance mechanisms applicable to religious contexts; scholarly credibility meant peer review or a recognized academic publisher. A targeted search for counter-evidence was conducted to limit selection bias (Table 2).

Data were collected through structured reading and extraction of information relevant to the research questions. Attention was directed to the conceptual arguments of each source, forms of AI-mediated religious activity, identified opportunities and risks, affected rights and interests, proposed safeguards or governance mechanisms, and implications for religious freedom and the public interest. Relevant material was organized around recurring concerns including religious freedom, equality and non-discrimination, privacy, dignity, transparency, oversight, accountability, contestability, pluralism, participation, misinformation, and institutional responsibility.

Analysis proceeded iteratively through open coding, focused coding, and thematic synthesis, moving from initial identification of recurring concepts and institutional relationships toward broader analytical categories and cross-source patterns. This iterative movement between coding, comparison, and conceptual interpretation is consistent with qualitative approaches in which codes are progressively refined into analytically meaningful patterns rather than treated as fixed categories from the outset (Saldaña, 2021). Focused codes were grouped into themes by shared mechanism; the resulting structure and the codes that did not fit the initial domains are reported at the start of the Results and in Table 2.

Thematic synthesis was subsequently used to integrate recurring patterns across heterogeneous documentary materials and relate them to the study's theoretical and conceptual framework (Thomas & Harden, 2008). The resulting themes supported analytical and conceptual interpretation rather than statistical generalization.

Analytical rigor was supported through cross-disciplinary comparison, iterative cross-source comparison, and reflexive attention to interpretation. This comparison was used

to set claims side by side across scholarly sources from different fields, including religious studies, human rights, AI ethics, and public governance, rather than relying on a single disciplinary or institutional perspective, while reflexive attention was maintained to distinguish source-based evidence from researchers' interpretation (Carter et al., 2014; Olmos-Vega et al., 2023).

Particular attention was also given to convergences, disagreements, and context-dependent differences across sources and to avoiding assumptions that religious communities constitute internally homogeneous groups. These procedures support the credibility and transparency expected in qualitative inquiry while recognizing that qualitative rigor depends on the coherence and transparency of the research process rather than statistical notions of reliability alone (Nowell et al., 2017). Because all sources are secondary, this comparison exposes claims to different disciplinary assumptions rather than confirming facts.

Because the study relied exclusively on documentary materials and did not involve interviews, surveys, experiments, or participant observation, no primary participant data were generated; consequently, participant consent and participant-level confidentiality procedures were not applicable.

RESULTS AND DISCUSSION

The documentary analysis identified three interconnected thematic patterns corresponding to the research questions and analytical framework: 1) AI-mediated access, visibility, and religious authority; 2) religious freedom, inclusion, and algorithmic risks; 3) accountability, contestability, and institutional responsibility. Across the documentary materials, AI did not emerge as an independently beneficial or harmful force.

Its implications varied according to the technological functions involved, the religious and institutional contexts in which they operated, and the actors participating in their development, deployment, and use. The principal patterns derived from the documentary analysis are consolidated in Table 1.

During coding, the four a priori domains (Figure 1) were consolidated into three themes because the sources treated visibility and authority as intertwined (Ergen, 2024; Harianto et al., 2025; Wahyudi et al., 2025). Open coding added material beyond the framework: the state as a source of risk, the legal classification of religious data, contradictory claims on whether AI democratizes or concentrates authority (Cheong & Liu, 2025; Wahyudi et al., 2025), and differences between lineage-based authority in Indonesian Sufism and creator-based authority in Türkiye. These elements are integrated below and in Table 2.

Table 1. Integrated Themes and Findings from the Documentary Analysis

Analytical Theme	Documentary Evidence/Pattern	Analytical Meaning
AI-mediated access, visibility, and religious authority	AI-supported platforms, translation tools, educational systems, and virtual religious services broaden access across geographical and physical boundaries. Search rankings, engagement	Expanded access does not necessarily produce equal visibility or legitimate authority. AI mediation

	metrics, recommendation systems, chatbots, and AI-generated religious content simultaneously influence which actors and interpretations receive greater visibility (Cheong & Liu, 2025; Ergen, 2024; He, 2024; Tsuria & Tsuria, 2024; Wahyudi et al., 2025).	changes the conditions through which religious knowledge is encountered, circulated, and recognized.
Religious freedom, inclusion, and algorithmic risks	Filtering, profiling, recommendation, and content moderation affect the visibility of religious beliefs and practices. Training data and design assumptions may reproduce stereotypes, misinformation, and unequal treatment, while religious beliefs, affiliations, locations, and conversational disclosures constitute sensitive personal information (Ashraf, 2022; Quirós-Fons, 2025; Roche et al., 2023; Tsuria & Tsuria, 2024).	Formal digital access can coexist with unequal representation, privacy risks, discriminatory profiling, and constraints on religious autonomy and meaningful participation.
Accountability, contestability, and institutional responsibility	Responsibility is distributed among designers, technology providers, institutions, moderators, religious authorities, and users. AI can facilitate participation while also contributing to misinformation, manipulation, and unequal visibility, whereas existing governance arrangements do not always address context-specific religious consequences (Alfrink et al., 2023; Cohen & Suzor, 2024; David et al., 2024; Denton & Erete, 2023; Novelli et al., 2024).	Distributed governance can create accountability gaps when affected actors cannot identify who is responsible for AI-mediated outcomes or meaningfully question consequential decisions.

Source: Authors' synthesis based on the documentary analysis.

Table 2. Number of Sources, Qualifying Evidence, and Governance Instruments by Theme

Theme	Corpus Sources (n)	Qualifying or Contrary Evidence	Relevant Governance Instruments
Access, visibility, and religious authority	8	Authority remains tied to lineage and institutional recognition (Harianto et al., 2025); religious leaders use AI selectively and keep human judgment (Cheong & Liu, 2025)	Platform recommendation policies; religious-council guidance on AI use; SE Kominfo 9/2023
Religious freedom, inclusion, and algorithmic risks	7	Risks also arise from state speech regulation (Presiden Republik Indonesia, 1965; Republik Indonesia, 2024); bias evidence concerns mainly Islam in English-language models (Abid et al., 2021)	GDPR Art. 9; UU PDP Art. 4; EU AI Act Art. 5(1)(g); ICCPR Arts. 17–19, 26
Accountability, contestability, and institutional responsibility	16	None of these sources concerns religion; participation often remains consultation (Denton & Erete, 2023)	EU AI Act Art. 27; UNGPs; UNESCO Recommendation; Permenkominfo 5/2020

Source: Authors' synthesis; n refers to corpus sources in Appendix Table A1.

AI-Mediated Access, Visibility, and Religious Authority

The first finding indicates that AI-mediated technologies simultaneously expand access to religious resources and alter the conditions under which religious content becomes visible and authoritative. As summarized in Table 1, the documentary analysis associated translation tools, educational systems, virtual religious services, and AI-supported

platforms with broader access to religious information and participation across geographical and physical boundaries (Cheong & Liu, 2025; He, 2024).

At the same time, search rankings, recommendation mechanisms, engagement metrics, chatbots, and AI-generated religious content emerged as mechanisms that influence which religious actors, teachings, and interpretations become more prominent (Ergen, 2024; Wahyudi et al., 2025). These patterns were not treated as evidence that AI independently determines religious authority; rather, they indicate that technological mediation increasingly participates in the processes through which religious knowledge is encountered and circulated.

The relationship between access and authority is therefore not straightforward. Greater availability of religious information can diversify opportunities for participation, yet increased visibility does not necessarily correspond to theological competence, institutional legitimacy, or communal recognition. The documentary material indicates that platform-based visibility may favor religious content that is extensively represented online, highly engaging, or readily compatible with recommendation mechanisms (Ergen, 2024; Wahyudi et al., 2025). This distinction is important because authority in digital religious environments is not simply transferred from established institutions to technological systems; instead, the conditions under which authority is recognized become increasingly intertwined with platform visibility.

This interpretation is consistent with Ergen's (2024) analysis of religious digital creators, which demonstrates how algorithmic visibility and audience metrics can become relevant to the formation of digital religious authority. Research on YouTube likewise indicates that personalization can shape exposure to religious content and has broader implications for digitally mediated religious authority (Wahyudi et al., 2025). Generative AI adds another dimension because systems capable of producing fluent religious explanations may appear authoritative even though their outputs can simplify or inadequately represent doctrinal and moral complexity (Tsuria & Tsuria, 2024).

The substantive implication is that AI-mediated religious communication affects religious life not only by expanding access to knowledge, but also by reorganizing the conditions of visibility through which religious knowledge and actors may acquire recognition. This finding addresses the first research question by locating the transformation of authority within the interaction among technological mediation, platform visibility, human interpretation, and existing forms of religious legitimacy. Concretely, creators who master platform metrics gain reach relative to institution-based scholars (Ergen, 2024; Wahyudi et al., 2025); chatbots become sources of religious answers without an identifiable author (Tsuria & Tsuria, 2024); and traditional legitimacy such as the sanad is not displaced but must be displayed online (Harianto et al., 2025). Religious policy must therefore address how platforms and AI systems rank and generate religious content, not only who is a recognized authority.

Religious Freedom, Inclusion, and Algorithmic Risks

The second finding concerns the tension between expanded digital participation and risks to freedom of religion or belief, inclusion, privacy, and equal representation. The documentary analysis identified filtering, profiling, recommendation, and automated content moderation as mechanisms capable of influencing the visibility of religious beliefs, practices, and interpretations. It also identified training data and design assumptions as possible sources of religious stereotypes, inaccurate representation, misinformation, and unequal treatment (Ashraf, 2022; Roche et al., 2023; Tsuria & Tsuria, 2024).

Religious beliefs, affiliations, practices, locations, search behavior, and conversational disclosures additionally emerged as sensitive forms of information that can be processed within AI-mediated environments. These findings indicate that participation in digital religious environments involves not only questions of access but also the conditions under which individuals and communities are represented, classified, and made visible. These concerns meet law in data protection: the GDPR treats data revealing religious beliefs as a special category (European Parliament & Council of the European Union, 2016, Article 9), whereas Indonesia's Personal Data Protection Law lists religion among general rather than specific personal data (Republik Indonesia, 2022, Article 4(3)), so religious inferences receive weaker protection, which matters for minority believers.

The documentary evidence further shows why formal accessibility should not automatically be interpreted as substantive inclusion. Personalized recommendation may direct users disproportionately toward dominant interpretations, highly engaging personalities, or commercially optimized religious content, while less digitally represented traditions may receive comparatively limited visibility.

Similarly, automated classification and moderation may encounter difficulties in distinguishing doctrinal language, internal theological disagreement, cultural expression, and potentially harmful content. The analytical issue is therefore broader than explicit censorship. Inclusion may also be affected by unequal representation, data practices, ranking mechanisms, and forms of profiling that operate without being immediately visible to users.

This interpretation corresponds with Ashraf's (2022) analysis showing that opaque AI systems create distinctive concerns for the enjoyment of freedom of religion or belief online, a dimension that has received less attention than privacy and freedom of expression in broader AI–human-rights debates. More recent work similarly places AI and surveillance within the emerging research agenda on religious freedom, emphasizing the relevance of technological infrastructures to the conditions under which religious freedom is exercised (Quirós-Fons, 2025). The diversity problem is also consistent with responsible-AI scholarship showing that questions of representation, intersectionality, and power are integral to evaluating AI policies and initiatives rather than peripheral considerations (Roche et al., 2023). Direct evidence supports this: a large language model persistently associated Muslims with violence more than other religious groups (Abid et al., 2021), which can distort what users learn about a tradition.

Accordingly, the substantive implication of the second theme is that protection of freedom of religion or belief in AI-mediated environments cannot be evaluated solely through formal access or the absence of explicit restrictions. It also concerns representational equality, informational privacy, religious autonomy, and the ability of diverse communities to participate without being systematically disadvantaged by algorithmic classification or visibility mechanisms.

The state plays a dual role. In Indonesia, the prohibition on information inciting religious hatred (Republik Indonesia, 2024, Article 28(2)) and the blasphemy law (Presiden Republik Indonesia, 1965) protect against hostility but can be applied to minority or dissenting interpretations, and takedown obligations on platforms (Menteri Komunikasi dan Informatika Republik Indonesia, 2020) can extend these effects at scale through automated moderation.

Accountability, Contestability, and Institutional Responsibility

The third finding concerns the distribution of responsibility across AI-mediated religious environments. The documentary analysis identified multiple actors involved in the relevant sociotechnical processes, including system designers, technology providers, platform operators, institutional deployers, moderators, religious authorities, public institutions, and users (David et al., 2024; Novelli et al., 2024).

This distribution creates a potential accountability problem because harmful or exclusionary outcomes may emerge through interactions among several actors rather than from the actions of a single identifiable decision-maker (Novelli et al., 2024). The documentary material also linked this problem to democratic participation and regulatory fragmentation: AI-mediated infrastructures can support communication and participation while simultaneously facilitating misinformation, manipulation, and unequal visibility, and existing governance arrangements do not always address the specific religious and cultural consequences of these processes. Such arrangements include platform community standards (Gillespie, 2018), Indonesia's AI-ethics circular (Menteri Komunikasi dan Informatika Republik Indonesia, 2023), and the EU AI Act's fundamental-rights impact assessment (European Parliament & Council of the European Union, 2024, Article 27); only the Act's prohibition of biometric categorization inferring religious beliefs (Article 5(1)(g)) names religion.

The analytical significance of this finding lies in the distinction between distributed responsibility and unidentifiable responsibility. The involvement of multiple actors does not itself eliminate accountability; the problem arises when individuals and communities affected by an AI-mediated outcome cannot determine who can explain the decision, who possesses authority to review it, or through which institutional process it can be challenged. Transparency alone therefore does not resolve the governance problem. Information about an AI system becomes substantively relevant when it enables affected actors to understand how a consequential outcome was produced and to identify an accountable institutional actor.

This interpretation is consistent with Novelli et al. (2024), who conceptualize AI accountability in terms of answerability and emphasize the sociotechnical relationships through which responsibility must be structured. Contestability complements this requirement because accountable AI governance also depends on opportunities to question and challenge consequential decisions rather than merely receive information about them (Alfrink et al., 2023).

Cohen and Suzor (2024) similarly identify public contestability as an important dimension of AI governance oriented toward the public interest. Participation, however, cannot be assumed to be meaningful simply because stakeholders are consulted; Denton and Erete (2023) found that participatory approaches to AI frequently provide opportunities to inform or consult rather than substantive partnership or delegated influence. Evidence from public perceptions of AI governance likewise shows that governance responsibility is understood across interdependent actors rather than being attributed exclusively to governments, corporations, or individuals (David et al., 2024).

Accountability is distinctive in religious contexts. When AI gives fatwa-like answers, responsibility lies with developers and deployers, but legitimacy depends on religious authorities outside the system (Cheong & Liu, 2025; Tsuria & Tsuria, 2024). This implies disclosure that an answer is machine-generated and of its sources, review by authorities of the relevant tradition, and complaint channels for misrepresentation.

The substantive implication is that ethical governance requires identifiable responsibility, meaningful opportunities for contestation, and forms of participation capable of connecting affected communities with the institutions that shape AI-mediated religious environments. This directly addresses the third research question without presuming that any single governance mechanism is universally applicable. The specific conditions are stated in the Conclusions.

Policy and Regulatory Instruments

Table 3 shows that religion is addressed explicitly mainly in data-protection and speech rules, while AI-ethics instruments treat it, if at all, within general non-discrimination principles; the reading is descriptive and selective.

Table 3. Treatment of Religion in Selected Policy and Regulatory Instruments

Instrument	Provision Relevant to Religion	Implication
UNESCO (2021)	Non-discrimination, diversity, ethical impact assessment	Religion only within general principles; non-binding
European Parliament & Council of the European Union (2024)	Prohibits biometric categorization to infer religious beliefs (Art. 5(1)(g)); fundamental-rights impact assessment for certain deployers (Art. 27)	Only AI rule naming religion; chatbots face transparency duties only
European Parliament & Council of the European Union (2016)	Data revealing religious beliefs are a special category (Art. 9)	Strong protection model for religious inferences

Instrument	Provision Relevant to Religion	Implication
United Nations Human Rights Office of the High Commissioner (2011)	Corporate responsibility; due diligence and remedy	Links state obligations to platforms
Republik Indonesia (2022)	Religion listed as general, not specific, personal data (Art. 4(3)); supervisory institution (Art. 58)	Weaker protection of religious data than under the GDPR
Menteri Komunikasi dan Informatika Republik Indonesia (2023)	Values incl. inclusivity, humanity, accountability	Non-binding; no religion-specific provision identified
Republik Indonesia (2024)	Prohibits disseminating information inciting hatred based on, among others, religion (Art. 28(2))	Protects against hostility but can restrict religious expression
Presiden Republik Indonesia (1965)	Penalizes deviant interpretation and defamation of recognized religions	Risk for minority and dissenting interpretations online
Menteri Komunikasi dan Informatika Republik Indonesia (2020)	Obliges private platforms to remove prohibited content on request	State-driven takedown at scale, often automated

Source: Authors' reading of the instruments.

Cross-Theme Synthesis

Across the three themes, the findings reveal an interconnected analytical structure rather than three independent sets of concerns. AI-mediated religious communication provides the common sociotechnical setting in which questions of access and visibility intersect with religious authority; these arrangements, in turn, are inseparable from questions of religious freedom, inclusion, privacy, and pluralism; and the consequences of these processes become governance concerns when institutional responsibility is fragmented or difficult to contest. The cross-theme relationships identified through the documentary synthesis are presented in Figure 2.

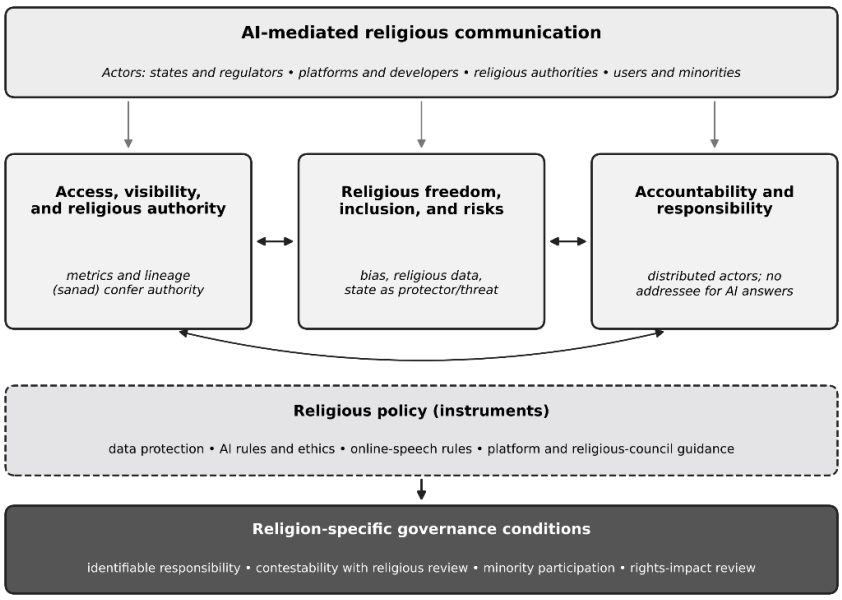


Figure 2. Integrated Analytical Framework for Ethical Governance of AI-Mediated Religious Environments
Note. The figure illustrates the analytical interconnections among the main domains identified through the documentary analysis. It does not represent a tested causal model. It adds to Figure 1 the actors, instruments, a pathway, and governance conditions. Source: Authors' elaboration.

Figure 2 synthesizes rather than duplicates the evidence summarized in Table 1. AI-mediated religious communication constitutes the shared sociotechnical context, while the three analytical domains represent distinct but interconnected dimensions identified through the documentary analysis. The reciprocal relationships between adjacent domains indicate analytical interdependence rather than a predetermined causal sequence: questions of visibility can become questions of rights and inclusion, while both acquire governance significance through the allocation of responsibility and opportunities for contestation. For example, low ranking of a minority tradition's content (visibility) limits its members' manifestation of belief online (rights), and because the ranking results from design, platform policy, and engagement together, no single actor is answerable (accountability).

Novelty and Contribution

Previous scholarship has largely examined AI and religion through separate perspectives, including religious representation by generative AI (Tsuria & Tsuria, 2024), digitally mediated religious authority (Ergen, 2024; Wahyudi et al., 2025), freedom of religion or belief in AI-mediated environments (Ashraf, 2022; Quirós-Fons, 2025), and accountability and contestability in AI governance (Alfrink et al., 2023; Novelli et al., 2024), leaving insufficient explanation of how these concerns interact as a unified religious-policy problem in the digital public sphere.

Addressing this gap, the present study contributes theoretically by connecting sociotechnical and algorithmic-governance perspectives to explain the interdependence of technological mediation, religious authority, fundamental rights, and institutional responsibility, and contributes conceptually by developing an integrated analytical framework comprising three interconnected domains: access, visibility, and religious authority; religious freedom, inclusion, and algorithmic risks; and accountability, contestability, and institutional responsibility. Relative to Quirós-Fons (2025) and Ashraf (2022), the insight that emerges only from connecting the perspectives is this pathway: visibility mechanisms create rights problems that accountability frameworks built around identifiable decisions cannot capture.

This integration shifts the analytical focus from treating access, rights, authority, and accountability as isolated concerns toward understanding their relationships within a common governance environment. Practically, the framework provides policymakers, technology providers, religious institutions, and civil-society actors with a structured basis for identifying where governance concerns arise when AI mediates religious communication and participation, without presenting the framework as a tested causal model or claiming empirical or methodological novelty beyond the documentary and conceptual scope of the study.

Limitations and Suggestions for Further Research

This study is limited by its qualitative documentary design, which captures patterns across existing scholarly and institutional sources rather than the direct experiences of religious communities, technology providers, or policymakers. The heterogeneous documentary corpus also limits systematic comparison across religious traditions, regions, languages, and digital platforms, while the non-experimental design does not support causal claims about the effects of AI on religious authority, rights, or participation. The corpus is also likely to reflect the dominance of English-language and Global North AI-ethics literature, with few sources on Indonesian or non-Abrahamic contexts, and the rapid change of AI systems may date the documents analyzed. These biases bear directly on the claims about religious pluralism, which should therefore be read as provisional. The instruments are read selectively rather than systematically, each theme rests on few sources, and the search was not systematic.

Future research should therefore test the analytical framework across specific religious and sociocultural contexts using interviews, comparative case studies, or platform-level evidence, while longitudinal studies could examine how evolving AI technologies and governance arrangements reshape these relationships over time.

CONCLUSIONS

This study argues that AI-mediated religious environments are best understood as an interconnected governance domain in which technological mediation shapes the conditions of religious visibility, authority, participation, and rights. The synthesis suggests that expanded digital access can coexist with unequal visibility, representational risks, and shifts in how religious authority is encountered, while these dynamics become particularly significant when freedom of religion or belief, inclusion, privacy, and pluralism intersect with fragmented institutional responsibility.

These relationships are context-dependent rather than inherently beneficial or harmful and should not be interpreted as predetermined causal effects of AI. By integrating sociotechnical and algorithmic-governance perspectives, the study conceptualizes ethical governance through the interconnection of three domains: access, visibility, and religious authority; religious freedom, inclusion, and algorithmic risks; and accountability, contestability, and institutional responsibility.

For the first research question, AI mediation shifts reach toward actors who master platform metrics, introduces chatbots as unaccountable sources of religious answers, and requires traditional legitimacy to be displayed online; for the second, the main concerns are documented religious bias, weak protection of religious data in Indonesian law, unequal visibility of minority traditions, and risks from state regulation of religious speech.

This synthesis shows that governing AI in the digital public sphere requires attention not only to technological performance or access, but also to the conditions under which religious knowledge gains visibility, diverse communities can participate without systematic disadvantage, and responsibility for consequential AI-mediated processes

remains identifiable and contestable. In response to the third research question, the synthesis points to four conditions: (1) identifiable responsible actors for consequential AI-mediated outcomes, so that affected communities can tell who can explain and review them; (2) contestability channels through which decisions can be questioned and challenged; (3) meaningful participation that goes beyond consultation and includes diverse and minority religious communities; and (4) rights-impact review attentive to freedom of religion or belief, privacy, and equal representation. In religious-policy practice, these conditions would apply, for example, when public authorities, religious institutions, and platforms define responsibilities in partnerships or guidelines on AI use, establish complaint and review mechanisms for decisions affecting religious content, involve a plurality of religious and non-religious actors in consultation, and assess AI deployments for their effects on religious freedom before and after adoption.

Specific to religion, these conditions require inclusion of minority, indigenous-belief, and non-religious communities, sensitivity to doctrinal diversity in moderation, protection of religious data, and review of AI religious answers by authorities of the relevant tradition; in Indonesia they could be implemented by the Ministry of Religious Affairs, the Ministry of Communication and Digital Affairs, the data-protection institution (Republik Indonesia, 2022, Article 58), and councils of all recognized religions. Given the reliance on secondary sources, the framework should be tested through document analysis of national AI guidance or a platform case study.

AUTHOR CONTRIBUTIONS STATEMENT

Wildanul Hakim contributed to the conceptualization, methodology development, and drafting of the manuscript; **Noerhayati Huda** contributed to data synthesis, analysis, and critical revision; and **Thariq Thieqiaq Thang** contributed to the construction of the policy framework, interpretation of the findings, and final review and approval of the manuscript.

CONFLICT OF INTEREST

The authors report no conflicts of interest that may have affected the conceptual development, documentary analysis, or presentation of the findings in this manuscript.

DECLARATION OF GENERATIVE AI USE

During manuscript preparation, generative artificial intelligence tools were used in a limited supporting capacity to assist with language editing, organization of arguments, and preparation of visual presentation. The tools were not treated as sources of evidence, and they did not replace the authors' responsibility for the research design, interpretation of the literature, verification of references, or final conclusions. All AI-assisted content was reviewed and edited by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript.

REFERENCES

Abid, A., Farooqi, M., & Zou, J. (2021). Persistent anti-Muslim bias in large language models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 298–306). Association for Computing Machinery. https://doi.org/10.1145/3461702.3462624

Alfrink, K., Keller, I., Kortuem, G., & Doorn, N. (2023). Contestable AI by design: Towards a framework. *Minds and Machines, 33*, 613–639. https://doi.org/10.1007/s11023-022-09611-z

Ashraf, C. (2022). Exploring the impacts of artificial intelligence on freedom of religion or belief online. *The International Journal of Human Rights, 26*(5), 757–791. https://doi.org/10.1080/13642987.2021.1968376

Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics, 5*(3), 3265–3279. https://doi.org/10.1007/s43681-024-00653-w

Bisnis.com. (2024, January 31). *Pertumbuhan pengguna internet Indonesia 2024 meningkat, tapi belum merata* [Indonesia's internet user growth in 2024 increases but remains uneven]. https://teknologi.bisnis.com/read/20240131/101/1737117/pertumbuhan-pengguna-internet-indonesia-2024-meningkat-tapi-belum-merata

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Campbell, H. A. (2007). Who's got the power? Religious authority and the Internet. *Journal of Computer-Mediated Communication, 12*(3), 1043–1062. https://doi.org/10.1111/j.1083-6101.2007.00362.x

Campbell, H. A. (2020). *Digital creatives and the rethinking of religious authority*. Routledge. https://doi.org/10.4324/9781003045625

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *Oncology Nursing Forum, 41*(5), 545–547. https://doi.org/10.1188/14.ONF.545-547

Cheong, P. H., & Liu, L. (2025). Generative artificial intelligence and collaboration: Exploring religious human–machine communication and tensions in leadership practices. *Human-Machine Communication, 11*, 171–190. https://doi.org/10.30658/hmc.11.9

Coeckelbergh, M. (2025). Artificial intelligence, the common good, and the democratic deficit in AI governance. *AI and Ethics, 5*, 1491–1497. https://doi.org/10.1007/s43681-024-00492-9

Cohen, T., & Suzor, N. P. (2024). Contesting the public interest in AI governance. *Internet Policy Review, 13*(3). https://doi.org/10.14763/2024.3.1794

David, P., Choung, H., & Seberger, J. S. (2024). Who is responsible? US public perceptions of AI governance through the lenses of trust and ethics. *Public Understanding of Science, 33*(5), 654–672. https://doi.org/10.1177/09636625231224592

Denton, R., & Erete, S. (2023). Power and public participation in AI. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (Article 8, pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3617694.3623228

Ergen, Y. (2024). Religious digital creatives in Türkiye: A research on the axis of algorithmic authority. *Türkiye İletişim Araştırmaları Dergisi*, *(46)*, 23–44. https://doi.org/10.17829/turcom.1488194

European Parliament & Council of the European Union. (2016). *Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*. http://data.europa.eu/eli/reg/2016/679/oj

European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. http://data.europa.eu/eli/reg/2024/1689/oj

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. https://doi.org/10.12987/9780300235029

Harianto, E., Lestari, R. D., Santoso, T. R., & Jamaludin. (2025). Algorithmic spiritual authority: Digital Sufism and the reconfiguration of sanad in Indonesia. *ESENSIA: Jurnal Ilmu-Ilmu Ushuluddin, 26*(2), 96–108. https://doi.org/10.14421/esensia.v26i2.7772

He, Y. (2024). Artificial intelligence and socioeconomic forces: Transforming the landscape of religion. *Humanities and Social Sciences Communications, 11*, Article 602. https://doi.org/10.1057/s41599-024-03137-8

Hjaltalin, I. T., & Sigurdarson, H. T. (2024). The strategic use of AI in the public sector: A public values analysis of national AI strategies. *Government Information Quarterly, 41*(1), 101914. https://doi.org/10.1016/j.giq.2024.101914

Hjarvard, S. (2008). The mediatization of religion: A theory of the media as agents of religious change. *Northern Lights, 6*(1), 9–26. https://doi.org/10.1386/nl.6.1.9_1

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

Lahusen, C., Maggetti, M., & Slavkovik, M. (2024). Trust, trustworthiness and AI governance. *Scientific Reports, 14*, Article 20752. https://doi.org/10.1038/s41598-024-71761-0

Menteri Komunikasi dan Informatika Republik Indonesia. (2020). *Peraturan Menteri Komunikasi dan Informatika Nomor 5 Tahun 2020 tentang Penyelenggara Sistem*

Elektronik Lingkup Privat [Minister of Communication and Informatics Regulation No. 5 of 2020 on private electronic system operators].

Menteri Komunikasi dan Informatika Republik Indonesia. (2023). *Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial* [Minister of Communication and Informatics Circular No. 9 of 2023 on artificial intelligence ethics].

Morgan, H. (2022). Conducting a qualitative document analysis. *The Qualitative Report, 27*(1), 64–77. https://doi.org/10.46743/2160-3715/2022.5044

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

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*, 1–13. https://doi.org/10.1177/1609406917733847

Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2023). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher, 45*(3), 241–251. https://doi.org/10.1080/0142159X.2022.2057287

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

Pasmore, W. A., & Khalsa, G. S. (1993). The contributions of Eric Trist to the social engagement of social science. *Academy of Management Review, 18*(3), 546–569. https://doi.org/10.5465/amr.1993.9309035150

Presiden Republik Indonesia. (1965). *Penetapan Presiden Nomor 1/PNPS Tahun 1965 tentang Pencegahan Penyalahgunaan dan/atau Penodaan Agama* [Presidential Decree No. 1/PNPS/1965 on the prevention of abuse and/or defamation of religion].

Quirós-Fons, A. (2025). Religious freedom in the age of AI and surveillance: Mapping the field and charting future research directions. *The Review of Faith & International Affairs, 23*(4), 104–116. https://doi.org/10.1080/15570274.2025.2545140

Republik Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi* [Law No. 27 of 2022 on personal data protection]. https://pasal.id/akn/id/act/uu/2022/27

Republik Indonesia. (2024). *Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik* [Law No. 1 of 2024 on the second amendment to the Electronic Information and Transactions Law].

Rähme, B., & Prohl, I. (2025). Religious studies approaches to the intersection of artificial intelligence and religion: Formations analogous to religion. *Religion, 55*(3), 573–595. https://doi.org/10.1080/0048721X.2025.2506893

Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics, 3*, 1095–1115. https://doi.org/10.1007/s43681-022-00218-9

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE.

Siebert, L. C., Lupetti, M. L., Aizenberg, E., Beckers, N., Zgonnikov, A., Veluwenkamp, H., Abbink, D., Giaccardi, E., Houben, G.-J., Jonker, C. M., van den Hoven, J., Forster, D., & Lagendijk, R. L. (2023). Meaningful human control: Actionable properties for AI system development. *AI and Ethics, 3*, 241–255. https://doi.org/10.1007/s43681-022-00167-3

Sigfrids, A., Leikas, J., Salo-Pöntinen, H., & Koskimies, E. (2023). Human-centricity in AI governance: A systemic approach. *Frontiers in Artificial Intelligence, 6*, Article 976887. https://doi.org/10.3389/frai.2023.976887

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly, 36*(2), 368–383. https://doi.org/10.1016/j.giq.2018.09.008

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*, Article 45. https://doi.org/10.1186/1471-2288-8-45

Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., Opdahl, A. L., Tessem, B., Borch, N., Fjeld, M., Øvrelid, L., De Smedt, K., & Moe, H. (2022). Responsible media technology and AI: Challenges and research directions. *AI and Ethics, 2*, 585–594. https://doi.org/10.1007/s43681-021-00126-4

Trist, E. L. (1981). *The evolution of socio-technical systems: A conceptual framework and an action research program* (Occasional Paper No. 2). Ontario Ministry of Labour, Ontario Quality of Working Life Centre. https://archive.org/details/39120320010110

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human Relations, 4*(1), 3–38. https://doi.org/10.1177/001872675100400101

Tsuria, R., & Tsuria, Y. (2024). Artificial intelligence's understanding of religion: Investigating the moralistic approaches presented by generative artificial intelligence tools. *Religions, 15*(3), Article 375. https://doi.org/10.3390/rel15030375

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137

United Nations. (1966). *International Covenant on Civil and Political Rights* (United Nations Treaty Series, Vol. 999, p. 171).

United Nations Human Rights Office of the High Commissioner. (2011). *Guiding principles on business and human rights*. https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf

Wahyudi, M. D. R., Hasan, N., & Fatwanto, A. (2025). From pulpit to platform: Algorithmic mediation and the transformation of religious authority on YouTube. *Profetik: Jurnal Komunikasi, 18*(2), 352–377. https://doi.org/10.14421/pjk.v18i2.3622

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance, 12*(4), 505–523. https://doi.org/10.1111/rego.12158

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly, 38*(3), 101577. https://doi.org/10.1016/j.giq.2021.101577

APPENDIX

Table A1. Sources analysed (n = 28); T = theme numbered as in Table 1.

Source	Context	Evidence	T	Source	Context	Evidence	T
Abid et al. (2021)	English LLM	Empirical	T2	Lahusen et al. (2024)	Global	Conceptual	T3
Alfrink et al. (2023)	Global	Conceptual	T3	Novelli et al. (2024)	Global	Conceptual	T3
Ashraf (2022)	Global	Conceptual	T2	Papagiannidis et al. (2025)	Global	Review	T3
Batool et al. (2025)	Global	Review	T3	Quirós-Fons (2025)	Global	Review	T2
Campbell (2020)	Global	Empirical	T1	Rähme & Prohl (2025)	Global	Conceptual	T1
Cheong & Liu (2025)	Leaders	Empirical	T1, T3	Roche et al. (2023)	AI policies	Review	T2
Coeckelbergh (2025)	Global	Conceptual	T3	Siebert et al. (2023)	Global	Conceptual	T3
Cohen & Suzor (2024)	Global	Conceptual	T3	Sigfrids et al. (2023)	Global	Conceptual	T3
David et al. (2024)	USA	Empirical	T3	Sun & Medaglia (2019)	China	Empirical	T3
Denton & Erete (2023)	Global	Review	T3	Trattner et al. (2022)	Global	Conceptual	T2
Ergen (2024)	Türkiye	Empirical	T1	Tsuria & Tsuria (2024)	GenAI outputs	Empirical	T1, T2
Harianto et al. (2025)	Indonesia	Conceptual	T1	Wahyudi et al. (2025)	Indonesia	Empirical	T1
He (2024)	Global	Empirical	T1, T2	Yeung (2018)	Global	Conceptual	T3
Hjaltalin & Sigurdarson (2024)	AI strategies	Documents	T3	Zuiderwijk et al. (2021)	Global	Review	T3