Proceedings Borneo International Islamic Conference; Vol. 17. 2026; 461-472
Published by Majmuah Enterprise
eISSN: 2948-5045

Maqasid-Driven Responsible Artificial Intelligence Framework (Mraif): An Integrated Governance Model for Higher Education Institutions

SULASTRI PUTIT[1, a], SAREHAN SADIKIN[2, b],LENNY YUSRINA BUJANG KHEDIF [3, c]

[1] Universiti Teknologi MARA, Sarawak Branch, Malaysia
[2] Universiti Teknologi MARA, Sarawak Branch, Malaysia
[3]Universiti Teknologi MARA, Sarawak Branch, Malaysia
[a]sulastri@uitm.edu.my , [b]sarehan@uitm.edu.my
& [c]lennykhedif@uitm.edu.my

Abstract

Artificial Intelligence (AI) is rapidly transforming higher education by enhancing teaching, learning, research, and institutional administration. Despite its considerable benefits, the widespread adoption of AI has introduced significant ethical and governance challenges, including concerns related to transparency, accountability, fairness, privacy, academic integrity, and human oversight. While internationally recognised Responsible Artificial Intelligence (Responsible AI) frameworks provide comprehensive governance principles, they are predominantly grounded in secular and regulatory perspectives, offering limited consideration of Islamic ethical values. This conceptual study addresses this gap by proposing the Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF), an integrated governance framework that aligns the objectives of Maqasid al-Shariah with globally accepted Responsible AI principles for higher education institutions. The study adopts an integrative literature review and conceptual framework development approach, synthesising contemporary literature on Responsible AI, AI governance, higher education, Islamic ethics, and Maqasid al-Shariah. The proposed framework consists of four interrelated components: (1) Ethical Foundation based on Maqasid al-Shariah, (2) Responsible AI Principles, (3) Institutional AI Governance Mechanisms, and (4) Trustworthy and Responsible AI Outcomes. The framework demonstrates how the five objectives of Maqasid al-Shariah can systematically strengthen internationally recognised Responsible AI principles by providing an ethical foundation for institutional policy, academic integrity, AI literacy, data governance, cybersecurity, and responsible decision-making. Unlike existing Responsible AI models, MRAIF establishes a sequential governance logic linking ethical values, governance principles, institutional implementation, and expected outcomes within a unified conceptual model. The study contributes to the Responsible AI literature by extending current governance discourse through the integration of Islamic ethical philosophy with contemporary AI governance, offering a culturally responsive framework that supports trustworthy, human-centred, and sustainable AI adoption in higher education. The framework also provides a theoretical foundation for future empirical validation and institutional policy development.

Keywords: Responsible Artificial Intelligence; Maqasid al-Shariah; Higher Education; AI Governance; Islamic AI Ethics; Conceptual Framework.

1. Introduction

Artificial Intelligence (AI) has become one of the most transformative technologies in higher education, significantly influencing teaching, learning, assessment, research, and institutional administration (Zawacki-Richter et al., 2019). The rapid adoption of Generative AI has accelerated AI integration across higher education institutions worldwide. Universities are increasingly incorporating AI-powered tools into teaching, assessment, research, and administrative functions, prompting the development of institutional AI governance policies (Chan, 2023; UNESCO, 2023). This rapid transformation highlights the growing need for comprehensive frameworks that ensure AI is implemented ethically, responsibly, and in alignment with institutional values. Generative AI applications can generate text, summarise information, assist with programming tasks, produce learning materials, and facilitate personalised learning experiences (Kasneci et al., 2023; Tlili et al., 2023). Consequently, universities are actively exploring strategies to integrate AI into academic environments to improve learning outcomes, teaching efficiency, and institutional effectiveness. Despite these advantages, the rapid adoption of AI also presents substantial ethical and governance challenges. Concerns relating to academic integrity, plagiarism, AI-generated misinformation, algorithmic bias, privacy violations, intellectual property, cybersecurity risks, and excessive dependence on AI-generated content have become increasingly prominent (Kasneci et al., 2023; Tlili et al., 2023; UNESCO, 2023). These challenges have shifted scholarly discussions from merely promoting AI adoption towards ensuring that AI technologies are implemented responsibly, transparently, and ethically. As a result, Responsible Artificial Intelligence (Responsible AI) has emerged as an important framework for balancing technological innovation with ethical responsibility (Floridi et al., 2018; Jobin et al., 2019). Responsible AI generally emphasises principles such as transparency, accountability, fairness, privacy, explainability, human oversight, and security (Floridi et al., 2018; Jobin et al., 2019; National Institute of Standards and Technology [NIST], 2023; Organisation for Economic Co-operation and Development [OECD], 2019; UNESCO, 2021). These principles have been adopted by international organisations, governments, and technology companies to encourage trustworthy AI development and deployment. Within higher education, Responsible AI supports academic integrity, protects students' rights, enhances institutional

governance, and promotes responsible digital citizenship (Chan, 2023; UNESCO, 2023). Nevertheless, most established Responsible AI frameworks are formulated as universal, rights-based, regulatory, or technology-governance models and do not explicitly articulate faith-based ethical foundations. Consequently, they provide limited contextual guidance for higher education institutions whose educational philosophies and governance practices are informed by Islamic values. For Islamic higher education institutions, AI adoption should not only satisfy technical and regulatory requirements but also align with Islamic ethical principles. Maqasid al-Shariah provides a comprehensive ethical framework centred on human welfare (maslahah), the prevention of harm (mafsadah), and the preservation of five essential interests: religion (Hifz al-Din), intellect (Hifz al-'Aql), life (Hifz al-Nafs), wealth (Hifz al-Mal), and lineage or posterity (Hifz al-Nasl) (Auda, 2008; Kamali, 2008). Contemporary Maqasid scholarship has extended these objectives beyond classical jurisprudence to address governance, development, social welfare, and emerging societal challenges (Auda, 2008; Chapra, 2008). However, their systematic application to Responsible AI governance in higher education remains limited and fragmented.

The absence of an integrated framework combining Responsible AI principles with Maqasid al-Shariah represents a significant theoretical and practical gap. Existing scholarship commonly examines AI ethics through universal, legal, technical, or educational perspectives, while Islamic ethical scholarship addresses moral purpose, welfare, and harm prevention through a different conceptual tradition (Auda, 2008; Jobin et al., 2019). Limited research has synthesised these domains into a coherent governance framework designed specifically for higher education. Consequently, educators and institutional leaders lack a structured model for implementing AI technologies in ways that are both technologically responsible and ethically aligned with Islamic values. Therefore, this conceptual paper proposes the Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF) for higher education. The proposed framework integrates internationally recognised Responsible AI principles with the five essential objectives of Maqasid al-Shariah to provide a holistic governance model for ethical AI adoption in higher education institutions. By bridging contemporary AI governance with Islamic ethical philosophy, this paper contributes to the growing discourse on AI ethics in education and offers a conceptual foundation for future empirical research and institutional policy development. The novelty of this study lies in the development of a four-stage conceptual framework that explicitly links Maqasid al-Shariah, Responsible AI principles, institutional AI governance mechanisms, and Responsible AI outcomes within a single governance model. Unlike existing Responsible AI frameworks that primarily define governance principles, the proposed MRAIF explains how Islamic ethical foundations can systematically guide AI governance and implementation within higher education institutions.

2. Literature review

2.1 Responsible Artificial Intelligence

The rapid advancement of Artificial Intelligence (AI) has transformed higher education by enhancing teaching, learning, assessment, research, and institutional management. The emergence of generative AI technologies, including ChatGPT, Google Gemini, Microsoft Copilot, and Claude, has further accelerated AI adoption because of their capacity to generate human-like text, summarise complex information, develop educational content, and provide personalised learning support. Although these technologies offer significant educational benefits, they also introduce ethical, legal, pedagogical, and governance challenges that require systematic management (Crompton & Burke, 2023; Kasneci et al., 2023; UNESCO, 2023; Zawacki-Richter et al., 2019). Responsible Artificial Intelligence (Responsible AI) has emerged as an ethical and governance approach intended to ensure that AI systems are designed, developed, deployed, and used in ways that are trustworthy, transparent, fair, accountable, safe, and respectful of human rights. International organisations and regulatory bodies, including UNESCO, the Organisation for Economic Co-operation and Development (OECD), the European Union, and the National Institute of Standards and Technology (NIST), have established principles and risk-management requirements to reduce potential harm while enabling socially beneficial AI innovation (European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021). Although terminology varies across institutions, major Responsible AI frameworks converge on several recurring principles, including transparency, fairness, accountability, privacy, explainability, security, and human oversight. (Floridi et al., 2018; Jobin et al., 2019; NIST, 2023; UNESCO, 2021). Transparency and explainability require meaningful information about how AI systems operate and how outputs or decisions are produced, enabling users to evaluate them critically. Fairness concerns the prevention and mitigation of discriminatory outcomes, whereas accountability requires clearly assigned responsibilities across developers, service providers, institutions, and users. Privacy and security protect personal and institutional data from unauthorised access, disclosure, and misuse, while human oversight ensures that consequential academic and administrative decisions remain subject to appropriate human judgement and control (European Parliament & Council of the European Union, 2024; OECD, 2019; UNESCO, 2021).

Within higher education, Responsible AI extends beyond technical implementation to encompass academic integrity, ethical assessment, digital and AI literacy, intellectual property, data governance, and institutional accountability. Universities are therefore expected to establish policies that clarify acceptable AI use, support educators and students,

maintain academic standards, and provide mechanisms for disclosure, review, and appeal (Chan, 2023; UNESCO, 2023). Consequently, Responsible AI in higher education is inherently multidisciplinary, connecting technology governance with pedagogy, ethics, law, and public policy. Despite the increasing maturity of Responsible AI frameworks, the global AI ethics landscape remains dominated by universalist, secular, legal, and technical approaches (Elmahjub, 2023; Jobin et al., 2019). These approaches provide important guidance on human rights, accountability, safety, and regulatory compliance, but they do not always articulate how faith-based moral traditions may shape educational purpose, institutional values, and ethical decision-making in religious higher education contexts (Raquib et al., 2022).

2.2 Artificial Intelligence in Higher Education

The integration of AI into higher education has substantially reshaped educational practices. AI-powered technologies support adaptive learning, intelligent tutoring, automated assessment, learning analytics, academic advising, student profiling, and administrative decision-making (Crompton & Burke, 2023; Zawacki-Richter et al., 2019). More recently, generative AI has enabled students and educators to produce essays, programming code, lesson plans, literature summaries, formative feedback, and research ideas within seconds, altering conventional approaches to learning, assessment, and knowledge creation (Kasneci et al., 2023; UNESCO, 2023). Existing literature reports a range of educational benefits associated with AI adoption. AI can support personalised learning, rapid feedback, accessibility, student engagement, and more responsive learning environments. Educators may also use AI to assist instructional design, assessment development, content generation, and administrative tasks, potentially allowing greater attention to higher-order pedagogical activities (Bond et al., 2024; Crompton & Burke, 2023; Kasneci et al., 2023). Nevertheless, research also identifies substantial risks associated with inappropriate or excessive reliance on AI. These include academic misconduct, undisclosed assistance, fabricated or inaccurate information, false references, weakened critical engagement, overdependence on generated content, and unequal access to advanced AI services. Privacy violations, cybersecurity risks, intellectual property concerns, algorithmic bias, and the misuse of AI-generated outputs further strengthen the need for institutional safeguards (Bond et al., 2024; Kasneci et al., 2023; Tlili et al., 2023; UNESCO, 2023). These challenges indicate that successful AI integration cannot be assessed solely through technological capability or efficiency. It must also promote ethical awareness, AI literacy, responsible use, pedagogical validity, and effective institutional governance. Accordingly, higher education institutions require policies and educational frameworks that regulate and guide AI use across teaching, learning, assessment, research, and administration (Chan, 2023; UNESCO, 2023). However, institutional responses to AI commonly prioritise academic integrity, legal compliance, data protection, and operational risk. Comparatively less attention has been given to the systematic incorporation of cultural, religious, and faith-based value systems into AI governance, particularly in Islamic higher education institutions where educational aims and institutional conduct may be explicitly connected to an Islamic worldview (Elmahjub, 2023; Raquib et al., 2022).

2.3 Maqasid al-Shariah and Ethical Governance

Maqasid al-Shariah refers to the higher purposes and objectives of Islamic law, which seek to realise human welfare (maslahah) and prevent harm (mafsadah). Classical formulations commonly identify five essential protections: religion (Hifz al-Din), life (Hifz al-Nafs), intellect (Hifz al-'Aql), wealth or property (Hifz al-Mal), and lineage or posterity (Hifz al-Nasl) (Auda, 2008; Chapra, 2008; Kamali, 2008). Collectively, these objectives provide a holistic ethical lens for evaluating individual conduct, institutional practice, public policy, and social consequences. Contemporary scholarship has extended Maqasid al-Shariah beyond classical jurisprudential discussion into fields such as development, finance, healthcare, governance, technology, and social policy. Its systems-oriented and welfare-centred character enables emerging issues to be evaluated through justice, responsibility, public interest, prevention of harm, and human flourishing rather than through technical legality alone (Auda, 2008; Chapra, 2008; Kamali, 2008). Islamic scholarship on AI similarly argues that religious ethical traditions can contribute distinctive moral reasoning to global AI governance (Elmahjub, 2023; Raquib et al., 2022). In the context of AI, Hifz al-'Aql may support intellectual integrity, critical reasoning, and protection from misinformation, manipulation, and excessive cognitive dependence. Hifz al-Mal can inform the protection of personal data, intellectual property, digital assets, and institutional information. Hifz al-Din is relevant to the accuracy and responsible communication of religious knowledge, while Hifz al-Nafs and Hifz al-Nasl can guide the protection of life, psychological well-being, human dignity, privacy, inclusion, and future societal welfare. These applications represent a conceptual interpretation of the classical objectives for contemporary AI governance, consistent with Maqasid-based and Islamic AI ethics scholarship (Auda, 2008; Elmahjub, 2023; Raquib et al., 2022). Although Maqasid al-Shariah demonstrates substantial conceptual compatibility with Responsible AI, Islamic AI ethics scholarship remains comparatively limited and is often presented through broad philosophical arguments or particular ethical concerns rather than through operational governance models for higher education (Elmahjub, 2023; Raquib et al., 2022). This leaves a need for a structured framework that systematically relates Maqasid objectives to recognised Responsible AI principles and institutional governance domains.

2.4 Comparative Analysis of Existing Responsible AI Frameworks

Several internationally recognised Responsible AI frameworks have been developed during the past decade, including

UNESCO's Recommendation on the Ethics of Artificial Intelligence, the OECD AI Principles, NIST's AI Risk Management Framework, and the European Union Artificial Intelligence Act (European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021). Their similarities, strengths, and limitations in relation to the proposed framework are summarised in Table 1.

Table 1. Comparative Analysis of Existing Responsible AI Frameworks and the Proposed MRAIF

Framework	Core Principles	Strengths	Limitations in Higher Education	Gap Addressed by MRAIF
UNESCO Recommendation on the Ethics of AI (2021)	Human rights, fairness, transparency, accountability, sustainability	Comprehensive global ethical guidance for AI governance	Provides universal ethical principles but does not explicitly incorporate religious or faith-based ethical values for AI governance in Islamic higher education	Integrates Maqasid al-Shariah as an ethical foundation while remaining aligned with Responsible AI principles
OECD AI Principles (2019)	Inclusive growth, human-centred values, transparency, robustness, accountability	Widely adopted by governments and policymakers	Focuses primarily on economic development and trustworthy AI with limited consideration of Islamic educational philosophy	Extends human-centred AI by incorporating Islamic ethical objectives and educational values
NIST AI Risk Management Framework (2023)	Govern, map, measure and manage AI risks	Strong emphasis on AI risk assessment and organisational governance	Concentrates on technical risk management without addressing religious ethics or educational moral development	Complements technical governance with ethical guidance derived from Maqasid al-Shariah
EU AI Act (2024)	Risk-based regulation, safety, transparency, accountability	Provides legal certainty and regulatory compliance	Regulatory in nature and designed mainly for legal compliance rather than educational ethics	Adds ethical governance and value-based decision-making beyond regulatory compliance
Existing Responsible AI Models in Higher Education	Academic integrity, fairness, explainability, privacy	Focus specifically on educational implementation of AI	Limited integration of Islamic ethics, AI governance and cybersecurity within a unified conceptual framework	Provides an integrated governance model linking Islamic ethics, Responsible AI principles and higher education practices
Proposed MRAIF (This Study)	Maqasid al-Shariah, transparency, accountability, fairness, explainability, privacy, security, human oversight	Integrates Islamic ethical values with internationally recognised Responsible AI principles specifically for higher education	Requires empirical validation in future studies	Provides a holistic governance framework for responsible AI adoption in Islamic higher education institutions

Critical Analysis

Table 1 shows that existing Responsible AI frameworks have made significant contributions by institutionalising principles such as transparency, accountability, fairness, explainability, privacy, security, safety, and human oversight. These frameworks provide influential references for governments, technology providers, organisations, and higher education institutions seeking to promote trustworthy AI adoption (European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021). Nevertheless, the principal international frameworks are formulated primarily through human-rights, legal, socioeconomic, and technical-risk perspectives. Although these foundations are essential, they provide limited explicit engagement with Islamic ethical concepts and the religiously informed educational purposes of Islamic higher education institutions. Similarly, much higher education guidance

emphasises academic integrity, institutional policy, and responsible use without systematically integrating Maqasid al-Shariah into a unified governance structure (Chan, 2023; Elmahjub, 2023; Raquib et al., 2022; UNESCO, 2023).

Figure 1. Evolution of Responsible AI Governance Towards the Proposed MRAIF

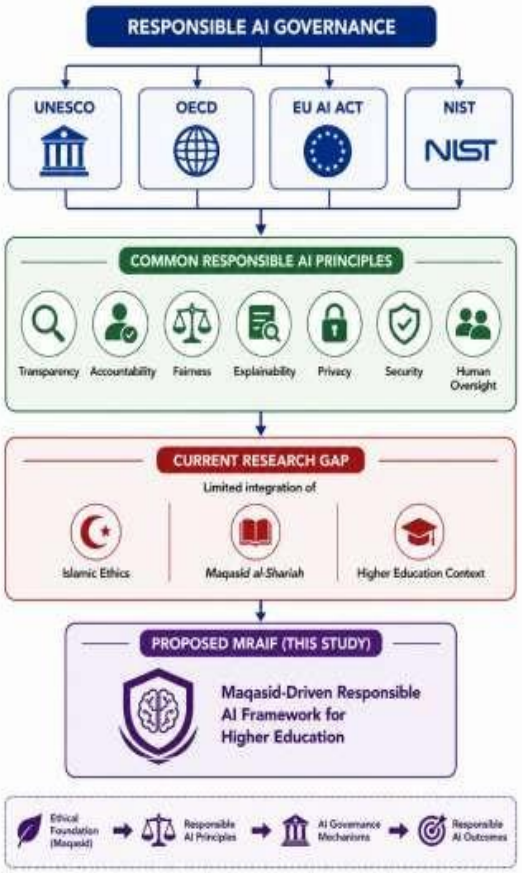


Figure 1 illustrates the conceptual progression from internationally recognised AI governance frameworks towards the proposed Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF). UNESCO, OECD, NIST, and the EU AI Act collectively establish influential principles and requirements relating to transparency, accountability, fairness, explainability, privacy, security, safety, and human oversight (European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021). However, these frameworks do not explicitly integrate Maqasid al-Shariah or address the distinctive value-based governance needs of Islamic higher education. This gap motivates MRAIF, which complements rather than replaces established Responsible AI frameworks by incorporating an Islamic ethical foundation into a structured higher education governance model

The reviewed literature indicates that Responsible AI has become increasingly important in higher education as AI technologies expand across teaching, learning, assessment, research, and administration (Bond et al., 2024; Chan, 2023; Crompton & Burke, 2023). Existing frameworks provide substantial guidance concerning transparency, fairness, accountability, privacy, explainability, safety, security, and human oversight, but their normative foundations are generally universalist, secular, legal, or technical (Jobin et al., 2019; NIST, 2023; UNESCO, 2021). At the same time, scholarship on Islamic AI ethics demonstrates that Islamic moral reasoning, including Maqasid-oriented analysis, can contribute to debates on human welfare, justice, privacy, autonomy, transparency, and responsible technological development (Elmahjub, 2023; Raquib et al., 2022). Nevertheless, this scholarship remains relatively new and has made limited progress in systematically connecting Maqasid al-Shariah with internationally recognised Responsible AI principles and the operational governance requirements of higher education. Consequently, a theoretical and practical gap remains between contemporary Responsible AI frameworks, Islamic ethical philosophy, and higher education governance. A holistic conceptual framework is needed to connect internationally accepted Responsible AI principles with Maqasid al-Shariah and translate them into institutional domains and mechanisms appropriate for higher education. Therefore, this paper proposes a Maqasid-Driven Responsible AI Framework that bridges these domains. The framework is intended to support higher education institutions in aligning technological innovation, internationally recognised AI governance principles, institutional responsibilities, and Islamic ethical values, thereby promoting responsible,

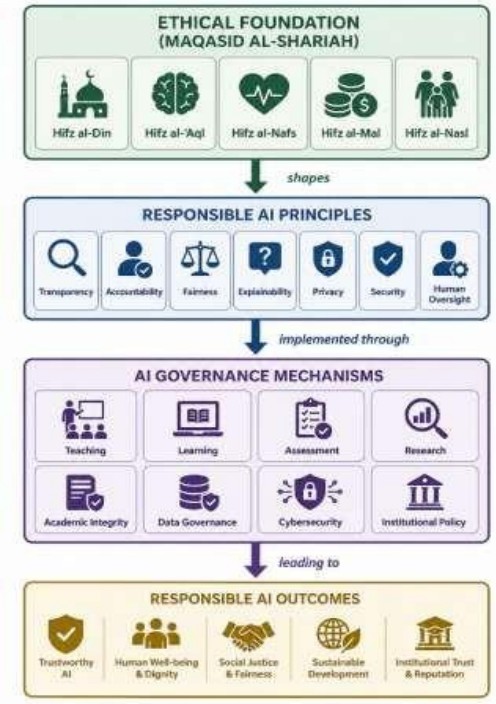
trustworthy, and sustainable AI adoption.

3. Conceptual Framework

3.1 Development of the Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)

Based on the identified research gap, this study proposes the **Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)** as a conceptual framework for ethical AI governance in higher education. Unlike existing Responsible AI frameworks that primarily focus on governance principles, MRAIF adopts a four-stage conceptual logic that explains why ethical AI is necessary, what principles should guide AI implementation, how these principles can be operationalised within higher education institutions, and what outcomes are expected from responsible AI adoption **(Jabareen, 2009; Snyder, 2019)**. Rather than presenting independent components, the framework illustrates a logical sequence in which **Maqasid al-Shariah** provides the ethical foundation that shapes Responsible AI principles. These principles are subsequently translated into institutional AI governance mechanisms that ultimately contribute to responsible, trustworthy, and sustainable AI adoption within higher education **(Auda, 2008; Floridi et al., 2018; Jobin et al., 2019)**. The proposed framework builds upon internationally recognised Responsible AI frameworks developed by UNESCO, the OECD, NIST, and the European Union while extending these frameworks through the integration of Islamic ethical principles derived from Maqasid al-Shariah **(European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021)**.

Figure 2 The Proposed Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)



3.2 Ethical Foundation (Why)

The first component of MRAIF explains why responsible AI governance is necessary within higher education institutions. This ethical foundation is established through the principles of **Maqasid al-Shariah**, which provide comprehensive guidance for promoting public interest (*maslahah*) while preventing harm (*mafsadah*) **(Auda, 2008; Chapra, 2008; Kamali, 2008)**. The five objectives of Maqasid al-Shariah collectively ensure that AI technologies support intellectual development, ethical decision-making, human dignity, social justice, and institutional responsibility.

Hifz al-Din (Protection of Religion)

The preservation of religion requires AI systems to promote truthful, authentic, and ethically appropriate religious knowledge. Within higher education, AI-generated religious content should be critically verified before being incorporated into teaching, learning, or scholarly activities. Human oversight remains essential to minimise religious misinformation, inaccurate interpretations, and fabricated references. Consequently, **Hifz al-Din** primarily strengthens the Responsible AI principles of transparency, explainability, and accountability **(Auda, 2008; Elmahjub, 2023)**.

Hifz al-'Aql (Protection of Intellect)

AI has enormous potential to improve learning effectiveness by providing personalised educational support and facilitating knowledge discovery. However, excessive dependence on AI-generated content may weaken students' critical

thinking, analytical reasoning, and independent learning abilities **(Kasneci et al., 2023; UNESCO, 2023)**. Therefore, AI should function as an intellectual assistant rather than a replacement for human cognition. **Hifz al-'Aql** supports explainability, human oversight, and AI literacy, encouraging educators and students to critically evaluate AI-generated information before accepting it as factual **(Auda, 2008; Raquib et al., 2022)**.

Hifz al-Nafs (Protection of Life and Well-being)

Responsible AI should contribute positively to human well-being by preventing psychological harm, discrimination, cyberbullying, and unsafe AI applications. Universities should ensure that AI systems minimise harmful outputs while supporting inclusive and respectful learning environments. Accordingly, **Hifz al-Nafs** reinforces fairness, safety, accountability, and responsible institutional governance **(Floridi et al., 2018; UNESCO, 2021)**.

Hifz al-Mal (Protection of Wealth and Property)

Within digital environments, wealth extends beyond financial assets to include personal information, research data, intellectual property, and institutional knowledge. The widespread use of cloud-based AI platforms raises important concerns regarding privacy, cybersecurity, confidential information, and ownership of AI-generated content **(NIST, 2023; OECD, 2019)**. Therefore, **Hifz al-Mal** directly supports Responsible AI principles related to privacy, security, data governance, and intellectual property protection **(Auda, 2008; UNESCO, 2021)**.

Hifz al-Nasl (Protection of Lineage and Human Dignity)

The preservation of lineage also encompasses the protection of human dignity, identity, reputation, and social harmony. AI systems should avoid discriminatory outputs, unethical profiling, deepfake misuse, and biased decision-making that undermine individual dignity **(European Parliament & Council of the European Union, 2024; UNESCO, 2021)**. Consequently, **Hifz al-Nasl** reinforces fairness, inclusiveness, transparency, and ethical governance within higher education institutions **(Raquib et al., 2022)**.

3.3 Responsible AI Principles (What)

The second component of MRAIF identifies what principles should govern AI adoption within higher education. Drawing upon internationally recognised Responsible AI frameworks, this study adopts seven core principles, namely **transparency, accountability, fairness, explainability, privacy, security, and human oversight (Floridi et al., 2018; Jobin et al., 2019; NIST, 2023; OECD, 2019; UNESCO, 2021)**. These principles translate the ethical aspirations of Maqasid al-Shariah into operational governance requirements for AI implementation. To operationalise the proposed framework, **Table 2** maps each objective of Maqasid al-Shariah to corresponding Responsible AI principles and their practical implications for higher education. The mapping demonstrates that Islamic ethical objectives are compatible with internationally recognised Responsible AI principles and can strengthen ethical AI governance within higher education institutions **(Auda, 2008; Jobin et al., 2019)**.

Table 2. Integration of Maqasid al-Shariah into Responsible AI

Maqasid al-Shariah	Responsible AI Principle	Higher Education Implications	Governance Example
Hifz al-Din	Transparency, Accountability	Verification of AI-generated religious knowledge	AI content verification policy
Hifz al-'Aql	Explainability, Human Oversight	AI literacy and critical thinking	AI literacy programme
Hifz al-Nafs	Fairness, Safety	Protection from harmful AI outputs	Ethical AI review
Hifz al-Mal	Privacy, Security	Data protection and intellectual property	Data governance policy
Hifz al-Nasl	Fairness, Inclusiveness	Human dignity and bias mitigation	Institutional ethical guidelines

3.4 AI Governance Mechanisms (How)

The third component explains how Responsible AI principles can be operationalised across major governance domains within higher education institutions. These governance domains comprise **teaching, learning, assessment, research, academic integrity, data governance, cybersecurity, and institutional policy**, as illustrated in Figure 2 **(Chan, 2023; UNESCO, 2023)**. Rather than remaining as abstract ethical values, Responsible AI principles are translated into institutional practices through governance mechanisms such as **AI policies, AI literacy programmes, academic integrity guidelines, ethical review procedures, AI risk assessment, data governance frameworks, cybersecurity management, and continuous monitoring (NIST, 2023; OECD, 2019; UNESCO, 2021)**. Collectively, these governance domains and supporting mechanisms enable higher education institutions to implement Responsible AI in a systematic, transparent, accountable, and ethically responsible manner. Effective AI governance therefore requires

collaboration among university leadership, academic staff, students, information technology units, and policymakers to ensure that AI technologies remain aligned with institutional values, regulatory requirements, and ethical standards **(European Parliament & Council of the European Union, 2024; Floridi et al., 2018)**. Furthermore, embedding AI governance within institutional policies encourages continuous evaluation, risk mitigation, and responsible innovation, enabling universities to respond proactively to the rapid evolution of AI technologies while safeguarding academic integrity, privacy, and public trust **(Chan, 2023; NIST, 2023; UNESCO, 2023)**.

3.5 Expected Outcomes

The final component describes the expected outcomes of responsible AI implementation. When AI governance is guided by **Maqasid al-Shariah**, operationalised through Responsible AI principles, and implemented across institutional governance domains, higher education institutions are expected to achieve **trustworthy AI implementation, enhanced academic integrity, improved human well-being, strengthened institutional trust, and sustainable AI adoption (Floridi et al., 2018; Jobin et al., 2019; UNESCO, 2021)**. These outcomes collectively reflect the ethical aspirations of both Responsible AI and Maqasid al-Shariah by promoting AI systems that are transparent, accountable, fair, secure, and centred on human welfare. The proposed framework therefore seeks to support higher education institutions in balancing technological innovation with ethical responsibility, institutional accountability, and long-term societal benefit **(Auda, 2008; Chapra, 2008; UNESCO, 2021)**.

4 Research Questions

This conceptual paper addresses the following research questions:

RQ1: How can the principles of **Maqasid al-Shariah** be systematically integrated into the Responsible Artificial Intelligence framework for higher education?

RQ2: What are the key ethical dimensions of a **Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)** for higher education institutions?

RQ3: How can the proposed framework support AI governance, ethical decision-making, and responsible AI implementation in teaching, learning, research, and institutional administration?

These research questions provide the conceptual foundation for developing the proposed MRAIF and guide the synthesis of existing literature on Responsible AI, higher education, and Islamic ethical governance **(Jabareen, 2009; Snyder, 2019)**. They also address the identified theoretical gap regarding the integration of internationally recognised Responsible AI principles with Maqasid al-Shariah within a unified governance framework for higher education **(Jobin et al., 2019; Elmahjub, 2023; Raquib et al., 2022)**.

5 Siginificance of Study

This study contributes to both theory and practice by proposing a holistic framework that integrates **Maqasid al-Shariah** with internationally recognised Responsible AI principles. From a theoretical perspective, the study extends the Responsible AI literature by introducing an Islamic ethical governance perspective that has received limited scholarly attention. While previous studies have discussed Responsible AI and Islamic ethics independently, few have attempted to integrate these domains into a comprehensive conceptual framework specifically designed for higher education **(Elmahjub, 2023; Jobin et al., 2019; Raquib et al., 2022)**.

From a practical perspective, the proposed framework offers guidance for higher education institutions seeking to develop AI governance policies that are ethically responsible and culturally relevant. The framework may assist university administrators, policymakers, educators, and students in promoting responsible AI practices while safeguarding academic integrity, data privacy, institutional accountability, and ethical decision-making **(Chan, 2023; UNESCO, 2021; UNESCO, 2023)**. The proposed MRAIF also contributes to institutional AI governance by providing a structured relationship between ethical foundations, Responsible AI principles, governance mechanisms, and expected outcomes. This structured approach may assist universities in translating abstract ethical principles into practical institutional policies and governance practices **(NIST, 2023; OECD, 2019)**. Furthermore, this study establishes a conceptual foundation for future empirical research. The proposed framework may be validated through quantitative, qualitative, or mixed-method approaches involving higher education stakeholders, thereby contributing to the continued development of Responsible AI governance within Islamic higher education institutions **(Jabareen, 2009; Snyder, 2019)**.

6 Research Methodology

This study adopts an **integrative literature review** and **conceptual framework development** approach to develop the proposed **Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)**. An integrative literature review enables the synthesis of knowledge from diverse theoretical and empirical studies, making it particularly suitable for emerging multidisciplinary topics such as Responsible Artificial Intelligence (Responsible AI), Islamic ethics, and higher education governance **(Snyder, 2019; Torraco, 2005)**. Unlike systematic reviews that primarily evaluate empirical

evidence, an integrative review facilitates the identification, comparison, and synthesis of concepts from different research domains to generate new theoretical insights and conceptual models **(Snyder, 2019; Webster & Watson, 2002)**. This approach is therefore appropriate for developing a governance framework that integrates internationally recognised Responsible AI principles with the ethical objectives of **Maqasid al-Shariah**. The conceptual framework was developed through four sequential stages, as illustrated in **Figure 3**.

Stage 1: Literature Identification

Relevant literature was identified from peer-reviewed journal articles, conference proceedings, books, policy documents, and international AI governance guidelines. The literature search focused on five major themes:

i. Responsible Artificial Intelligence
ii. Artificial Intelligence in Higher Education
iii. AI Governance
iv. Maqasid al-Shariah
v. Islamic AI Ethics

The primary sources included internationally recognised publications and policy documents from **UNESCO**, the **Organisation for Economic Co-operation and Development (OECD)**, the **National Institute of Standards and Technology (NIST)**, the **European Union AI Act**, and high-quality journal articles indexed in **Scopus** and **Web of Science (European Parliament & Council of the European Union, 2024; NIST, 2023; OECD, 2019; UNESCO, 2021)**.

Stage 2: Literature Analysis and Conceptual Synthesis

Following the literature identification process, the selected studies were analysed to identify recurring Responsible AI principles, governance domains, ethical values, and implementation challenges. Similar concepts from Responsible AI literature and Maqasid al-Shariah scholarship were compared and synthesised to establish conceptual relationships between Islamic ethical objectives and internationally recognised AI governance principles **(Jabareen, 2009; Snyder, 2019)**. The analysis emphasised similarities, complementarities, and conceptual gaps between existing Responsible AI frameworks and Islamic ethical governance literature. This synthesis formed the theoretical basis for constructing the proposed framework.

Stage 3: Framework Development

The proposed **Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)** was subsequently developed by integrating the findings obtained from the literature synthesis.

The framework consists of four interrelated components:

1. **Ethical Foundation (Maqasid al-Shariah)**
2. **Responsible AI Principles**
3. **AI Governance Mechanisms**
4. **Trustworthy and Responsible AI Outcomes**

These components are organised into a logical governance sequence in which ethical values shape governance principles, governance principles guide institutional implementation, and governance implementation contributes to responsible AI adoption within higher education institutions **(Auda, 2008; Floridi et al., 2018; Jobin et al., 2019)**.

Stage 4: Conceptual Validation

As this study is conceptual in nature, the proposed framework has not yet undergone empirical validation. Nevertheless, the framework was constructed by synthesising established theories, internationally recognised AI governance frameworks, and contemporary Maqasid al-Shariah scholarship to ensure conceptual consistency and theoretical validity **(Jabareen, 2009)**. Future studies may validate the framework using qualitative, quantitative, or mixed-method research involving AI experts, higher education practitioners, policymakers, and Islamic scholars. Methods such as the **Delphi Technique**, **Fuzzy Delphi Method**, **Structural Equation Modelling (SEM)**, or **Partial Least Squares Structural Equation Modelling (PLS-SEM)** may be employed to evaluate the relationships among the proposed constructs and assess the applicability of MRAIF in higher education institutions **(Hsu & Sandford, 2007)**.

Figure 3. Research Methodology

The methodology adopted in this study comprises four sequential stages:

1. **Literature Identification**
2. **Literature Analysis and Conceptual Synthesis**
3. **Framework Development**
4. **Conceptual Validation**

This structured methodology ensures that the proposed framework is systematically derived from existing knowledge

while providing a transparent process for future empirical refinement and validation **(Jabareen, 2009; Snyder, 2019)**

7 Findings

7.1 The Proposed Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)

The principal finding of this conceptual study is the development of the Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF). The framework demonstrates that Responsible AI governance can be strengthened by incorporating Islamic ethical principles into existing AI governance models. Unlike conventional Responsible AI frameworks that primarily emphasise technological governance, MRAIF introduces *Maqasid al-Shariah* as the ethical foundation that guides AI policy development and institutional decision-making. The proposed MRAIF consists of **four interconnected components** that collectively explain the logical process of responsible AI governance within higher education institutions. The framework begins with the ethical foundation provided by *Maqasid al-Shariah* (**Why**), followed by internationally recognised Responsible AI principles (**What**), institutional governance mechanisms that operationalise these principles (**How**), and the expected outcomes of responsible AI implementation (**Outcome**). This sequential relationship distinguishes MRAIF from existing Responsible AI frameworks that primarily focus on governance principles without explicitly connecting ethical foundations, implementation mechanisms, and institutional outcomes.

7.2 Practical Application of MRAIF

The proposed framework may be implemented at multiple organisational levels.
At the institutional level, universities may utilise MRAIF as a reference for developing AI governance policies, ethical guidelines, and strategic planning. At the academic level, educators may employ the framework to promote responsible AI usage during teaching, assessment, and research supervision. At the student level, MRAIF encourages AI literacy, ethical awareness, academic integrity, and responsible digital citizenship. Finally, at the policy level, the framework supports decision makers in balancing technological advancement with Islamic ethical values while ensuring sustainable AI adoption.

8 Discussion

8.1 Theoretical and Practical Implications

The increasing adoption of AI in higher education requires governance approaches that extend beyond technological capability. Existing Responsible AI frameworks provide valuable guidance regarding transparency, accountability, fairness, and privacy. However, these principles alone may not adequately address the ethical expectations of institutions whose educational philosophy is grounded in Islamic values. The proposed MRAIF addresses this limitation by positioning *Maqasid al-Shariah* as the ethical foundation for Responsible AI governance. Rather than introducing entirely new AI principles, the framework enriches existing governance models by contextualising them within an Islamic ethical worldview. This integration demonstrates that technological innovation and religious ethical values are complementary rather than contradictory. Unlike conventional Responsible AI frameworks that primarily describe governance principles, MRAIF introduces a sequential governance logic that explains the relationship between ethical foundations, governance principles, institutional implementation, and expected outcomes. This conceptual progression enables higher education institutions to translate ethical aspirations into practical governance strategies while remaining aligned with internationally recognised Responsible AI principles. One notable contribution of MRAIF is its recognition that responsible AI adoption should not only minimise technological risks but also promote holistic human development. By aligning AI governance with the preservation of religion, intellect, life, wealth, and human dignity, the framework broadens the objectives of Responsible AI beyond regulatory compliance to include moral responsibility and societal well-being. Another significant contribution concerns institutional governance. Universities increasingly require comprehensive AI policies addressing academic integrity, data protection, ethical assessment, intellectual property, cybersecurity, and responsible AI literacy. MRAIF provides a structured ethical reference that may assist higher education institutions in formulating AI governance strategies consistent with both international standards and Islamic educational values. Nevertheless, this study remains conceptual in nature. The proposed framework has not yet undergone empirical validation. Future research should therefore evaluate its applicability through expert consensus techniques such as Delphi or Fuzzy Delphi before examining its effectiveness using quantitative methods across different higher education contexts.

8.2 Comparison between MRAIF and Existing Responsible AI Frameworks

Although UNESCO, OECD, NIST, and the EU AI Act have established internationally recognised principles for Responsible AI governance, these frameworks primarily emphasise regulatory compliance, technical governance, and universal ethical values. In contrast, MRAIF complements these frameworks by incorporating *Maqasid al-Shariah* as an ethical foundation specifically relevant to Islamic higher education. Unlike existing frameworks that primarily describe

governance principles, MRAIF further proposes institutional governance mechanisms that explain how Responsible AI principles may be translated into educational practice. Consequently, MRAIF provides a more holistic governance model that links ethical values, governance principles, implementation strategies, and institutional outcomes within a single conceptual framework.

8.3 Limitations and Future Research

This study is conceptual in nature and has not yet undergone empirical validation. Consequently, the proposed framework should be interpreted as a theoretical governance model rather than an empirically validated implementation framework. Future studies should evaluate the framework through expert consensus techniques such as Delphi or Fuzzy Delphi before examining its effectiveness using quantitative or mixed-method approaches across different higher education institutions. Additional studies may also explore the applicability of MRAIF in non-Islamic educational settings and compare its effectiveness with existing Responsible AI governance frameworks.

9. Conclusion

Artificial Intelligence continues to reshape higher education by providing new opportunities to improve teaching, learning, research, and institutional management. Nevertheless, the rapid adoption of AI also introduces ethical challenges related to transparency, accountability, fairness, privacy, academic integrity, and responsible decision-making. These challenges highlight the need for comprehensive governance frameworks that not only regulate AI technologies but also ensure that AI applications remain aligned with human values and educational objectives. This study proposed the **Maqasid-Driven Responsible Artificial Intelligence Framework (MRAIF)** as a conceptual framework that integrates the objectives of **Maqasid al-Shariah** with internationally recognised Responsible AI principles. Through an integrative literature review and conceptual synthesis, the framework demonstrates how Islamic ethical values can complement existing Responsible AI governance by providing a stronger moral foundation for AI adoption within higher education institutions. The proposed framework organises AI governance into four interconnected components comprising the Ethical Foundation, Responsible AI Principles, AI Governance Mechanisms, and Trustworthy and Responsible AI Outcomes. Theoretically, this study extends the Responsible AI literature by introducing an Islamic ethical perspective into AI governance, an area that remains relatively underexplored in existing research. Practically, the framework provides useful guidance for university administrators, policymakers, educators, and researchers in developing AI governance policies that promote transparency, accountability, fairness, privacy, human oversight, and institutional responsibility while preserving the ethical objectives of Maqasid al-Shariah.

Despite these contributions, this study has several limitations. As a conceptual paper, the proposed framework has not yet been empirically validated. The relationships among the proposed components are derived from literature synthesis rather than statistical analysis or stakeholder evaluation. Therefore, the applicability of the framework across different institutional and cultural contexts requires further investigation.

Future research should empirically validate the proposed framework using qualitative, quantitative, or mixed-method approaches involving higher education stakeholders, AI practitioners, policymakers, and Islamic scholars. Techniques such as the Delphi Method, Fuzzy Delphi Method, Structural Equation Modelling (SEM), or Partial Least Squares Structural Equation Modelling (PLS-SEM) may be employed to examine the relationships among the proposed constructs and evaluate the effectiveness of MRAIF in supporting responsible AI governance within higher education institutions.

In conclusion, the proposed MRAIF offers a holistic governance framework that bridges Islamic ethical values with internationally recognised Responsible AI principles. By integrating Maqasid al-Shariah into AI governance, the framework contributes towards the development of trustworthy, ethical, human-centred, and sustainable AI implementation in higher education while providing a foundation for future empirical research and policy development.

References

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

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour.

International Journal of Educational Technology in Higher Education, 21, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3

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

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, Article 22. https://doi.org/10.1186/s41239-023-00392-8

Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. Philosophy

& Technology, 36, Article 73. https://doi.org/10.1007/s13347-023-00668-x
European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://data.europa.eu/eli/reg/2024/1689/oj
Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5
Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389–399.
https://doi.org/10.1038/s42256-019-0088-2
Kamali, M. H. (2008). Maqasid al-Shariah made simple. International Institute of Islamic Thought.
Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274
National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1
Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
Raquib, A., Channa, B., Zubair, T., & Qadir, J. (2022). Islamic virtue-based ethics for artificial intelligence. Discover Artificial Intelligence, 2, Article 11. https://doi.org/10.1007/s44163-022-00028-2
Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10, Article 15. https://doi.org/10.1186/s40561-023-00237-x
UNESCO. (2021). Recommendation on the ethics of artificial intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000381137 UNESCO. (2023). Guidance for generative AI in education and research. https://unesdoc.unesco.org/ark:/48223/pf0000386693 Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence
applications in higher education—Where are the educators? International Journal of Educational Technology in Higher
Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0
Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. Practical Assessment, Research and Evaluation, 12(10).https://doi.org/10.7275/pdz9-th90
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
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
Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. Human Resource Development Review, 4(3), 356–367. https://doi.org/10.1177/1534484305278283
Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii–xxiii.