AI and Ethics
https://doi.org/10.1007/s43681-025-00780-y

RESEARCH

Al-Ghazali's principle of trust (*Amanah*) as a framework for ethical AI governance in organizations

Bahar Serevan[1]

Received: 20 February 2025 / Accepted: 14 June 2025

Abstract
In a technology-driven, materialistic world, the need for timeless wisdom and spiritual-moral values becomes increasingly evident. This paper aims to provide a solid moral foundation for modern ethical frameworks of AI governance by drawing on the prominent Muslim theologian and philosopher, Al-Ghazali's principle of *Amanah* in ethical leadership as a tool to establish trust in AI systems and technologies. The research is theoretical and qualitative, utilizing secondary data from journal articles and books. Through a literature review and analysis, the study explores the role of Al-Ghazali's teachings in shaping ethical AI governance through ethical leadership. The findings reveal that incorporating *Amanah*, a core trait of ethical leadership from an Islamic perspective, into AI governance frameworks can effectively balance spiritual values with technological advancements. This approach offers a pathway to achieving business goals in a more humanistic, ethical, and sustainable manner. This theoretical paper has some limitations, primarily in lacking empirical evidence and focusing only on one principle of Al-Ghazali's ethical legacy; therefore, a more comprehensive study would help to enrich the research with empirical evidence and practical implications.

Keywords Artificial intelligence governance · AI ethics · Trust · Ethical leadership · *Amanah*

1 Introduction

The rapid advancement of artificial intelligence (AI) is transforming sectors such as healthcare, finance, communication, and logistics, enhancing efficiency, decision-making, and operational capacities [42, 49]. Yet, these innovations raise significant ethical concerns, including algorithmic bias, lack of accountability, and insufficient transparency [25, 32]. These challenges have intensified calls for robust AI governance frameworks that uphold not only technical standards but also moral and societal values.

Governments, corporations, and individuals increasingly recognize the need for ethical governance to harness AI's benefits while mitigating its risks, with trust being a foundational element for transparency and accountability [18, 47]. However, many contemporary governance models, such as AI4People (2018), the European Commission's Ethics Guidelines (2019), and IEEE's Ethically Aligned Design (2019), lack the philosophical and cultural depth needed to resolve emerging moral dilemmas [25, 32]. While these frameworks promote universal principles like fairness and transparency [15, 22], they often overlook culturally embedded ethical systems that are critical for societal acceptance [11, 21].

In culturally diverse societies, ethical systems rooted in spiritual and moral traditions, such as Islamic ethics grounded in *Maqasid al-Shari'ah* or Al-Ghazali's concept of *Amanah* (trustworthiness), can enhance the legitimacy and public acceptance of AI governance [45, 51]. This culturally grounded approach addresses not only what AI should do, but also why individuals and societies are motivated to adhere to its ethical use.

Al-Ghazali (1058–1111), a leading Islamic theologian and philosopher, emphasized moral responsibility in leadership through *Amanah*, integrating values like sincerity (*ikhlas*), accountability (*muhasabah*), and justice (*adl*) [3, 4]. These principles, articulated in *Ihya' Ulum al-Din*, offer both ethical depth and cultural resonance, making them highly relevant to contemporary AI governance discourse.

This paper argues that Al-Ghazali's principle of *Amanah* offers a culturally rooted ethical framework that

✉ Bahar Serevan
bahar.serevan@stu.ihu.edu.tr

[1] Ibn Haldun University, Istanbul, Turkey

Published online: 23 June 2025

complements and strengthens contemporary AI governance systems by embedding trust, sincerity, and accountability into their core, especially in morally diverse societies. In doing so, it contributes to current debates by responding to the growing call for non-Western ethical frameworks in AI, offering a timeless and inclusive approach to responsible innovation.

2 Methodology

This study employs a qualitative philosophical analysis rooted in conceptual and hermeneutic interpretation. The primary aim is to extract core ethical principles from Al-Ghazali's writings and explore their relevance and applicability to modern AI governance frameworks.

The process began with the selection of key texts by Al-Ghazali, primarily *Ihya Ulum al-Din*, chosen for their philosophical richness and relevance to the concept of *Amanah* (trust). Supplementary sources from Islamic jurisprudence and contemporary scholarship were also reviewed to support conceptual clarity.

A thematic analysis was applied to distill recurring ethical values linked to *Amanah*, including responsibility, transparency, accountability, and sincerity. These values were then cross-referenced with principles found in modern AI ethics literature, such as fairness, explainability, and harm prevention.

Next, a conceptual mapping exercise was conducted to align Al-Ghazali's ethical values with elements in contemporary AI governance frameworks, including the EU AI Act, IEEE's Ethically Aligned Design, and OECD AI Principles. This comparison identified areas of overlap, ethical gaps, and opportunities for augmentation.

Finally, a normative framework was synthesized to demonstrate how *Amanah* can serve as a guiding principle in AI ethics. This framework balances philosophical depth with practical integration in organizational governance. The study adopts a non-empirical approach grounded in normative ethical traditions, contributing to the enrichment of AI ethics discourse from underrepresented cultural perspectives.

3 Literature review

3.1 AI governance, balancing innovation and ethics

AI is one of the greatest achievements of the modern era, while improving business tools and conditions, it has also led to unethical weapons and immoral practices like algorithmic bias, privacy violations, and accountability failures [32].

The frameworks and policies that guide and regulate AI systems' design, development, and deployment are referred to as AI governance. These frameworks also involve various stakeholders such as governments, organizations, and citizens [10].

The vast adoption of AI technologies by various pivotal sectors, such as healthcare, security, finance, education, transportation, and communication, led to an urgent call for solid AI governance structures to eliminate risks associated with AI such as algorithmic bias, discrimination, privacy violence, and job displacement [32] and by ensuring the alignment of these technologies to ethical values and societal standards [18, 56].

Providing control and accountability to AI systems by ensuring the positive contribution of these technologies to society instead of magnifying existing moral dilemmas or generating novel ethical issues is crucial [25].

According to research, AI governance consists of three levels and three layers [39], the levels of AI governance according to [50] are team, organization, and industry, while the layers are social and legal layer which contains norms and regulations, the ethical layer consisting of criteria and principles, and the technical layer includes algorithm accountability and standards, and data governance [26].

In their article, Jobin et al. [32] found that most AI ethics guidelines reference several key ethical principles and values. These principles are trust, transparency, justice and fairness, responsibility, privacy, beneficence, freedom and autonomy, non-maleficence, dignity, sustainability, and solidarity. They also noticed that there is a significant divergence in these principles across communities, but the main overall emphasis is on preventing harm rather than promoting good.

These challenges demonstrate the need for ethical frameworks that not only regulate AI but also resonate with users' moral worldviews. *Amanah*, as proposed in this study, can fulfill this role by embedding ethical accountability and trust into the governance structure.

3.2 Core principle of AI governance

Robust AI governance frameworks are essential to foster public trust and to ensure the ethical and social appropriateness of AI systems, besides the right development and deployment of the technology systems [46]. This alignment will help to foster confidence in AI systems among all participants (users, developers, and regulators) and will facilitate the widespread adoption of AI technologies and maintain social harmony.

There are several core principles that aim to enhance the governance frameworks of AI; they aim to ensure that these technologies are developed and deployed in an ethical way (transparent, fair, responsible, accountable…) [25]. Governance principles function together to establish ethical and responsible AI systems and to guide organizations in creating innovative and socially beneficial technologies [32, 46].

Therefore, many groups and organizations in recent years have stressed the importance of AI governance by developing their guidelines and frameworks to stimulate the development and adoption of trustworthy AI like the European Commission's Ethics Guidelines for Trustworthy AI (2019), The Asilomar AI Principles (2017)b, UNESCO's Recommendation on the Ethics of AI (2021), and Microsoft's Responsible AI Principles (2021).

For instance, the OECD Principles focus on the importance of AI systems to be inclusive, transparent, and accountable, and to ensure, at the same time, that they operate in alignment with human rights and democratic values (OECD, 2019). Another example is Microsoft's Responsible AI Principles, these principles call for values like fairness, reliability, transparency, and privacy in AI systems, and they also focus on minimizing bias and maximizing the benefits for all parties and stakeholders (Microsoft, 2021).

Conversely, implementing ethical AI raises many challenges, including the variation in how ethical principles are interpreted and evaluated, and also the conflict in balancing them, e.g., the need for large data sets to reduce bias may conflict with individuals' privacy protection and autonomy. Another challenge is the required collaboration of regulatory parties, organizations, and designers of AI systems to develop practical tools and standards to ensure the ethical deployment of AI [32].

However, ethical principles embraced by AI governance structures can't be automatically translated into practice [42]. These governance frameworks also need guidelines sensitive to which context they are deployed [18]. Yet, regardless of the complications of cultural and social differences, various groups of stakeholders, governments, organizations, and citizens are involved in the concern for AI ethical governance. To achieve good governance, it's fundamental to understand what various stakeholders trust [56].

Overall, current AI governance frameworks are relatively effective regarding technical and regulatory aspects of AI governance, however, there remain some concerns about how should society integrate moral values and related ethical decisions to the use of technological advancements in a practical way [52] and a notable gap related to the need to integrate moral and spiritual values concerned with the human element of governance these concerns might be fulfilled by exploring the role of leadership in fostering a culture of trust and accountability within the organizations that developing and employing AI technologies as discussed by [33].

3.3 The role of trust in AI governance

Trust is commonly defined as a confident relationship with others [12], referring to the belief that someone (or something) will act reliably and ethically [40]. In the context of AI, trust is a core principle for harnessing the benefits of emerging technologies while mitigating risks associated with machine learning and autonomous decision-making [16]. This becomes especially crucial when AI is used in high-stakes areas such as healthcare, finance, or autonomous vehicles [55]. In such cases, the absence of trust can lead to resistance or even outright rejection of AI systems, ultimately limiting their benefits [30].

Ethical AI governance views trust as both a goal and a product of effective policy. It mediates the relationship between humans and AI systems [18] while promoting transparency, ensuring accountability, and supporting fairness and equality [32, 46, 47]. Establishing trust requires not only robust systems but also a deep understanding of what different stakeholders value and fear [56]. These stakeholders, developers, users, and regulators must believe that AI systems, despite their complexity and unpredictability, will be understandable, accountable, fair, and aligned with public interests [46].

Users typically trust AI when systems operate consistently and fairly. Increased transparency, where users understand how and why decisions are made, further enhances that trust [36]. Developers, in turn, must trust that their AI creations will be used ethically and not result in unintended harm [11]. Building trust, therefore, requires consistent ethical behavior, openness about limitations, and active engagement with all stakeholders to address their concerns and perspectives.

To ensure credibility and trustworthiness in AI governance, it is crucial to integrate ethical values into the roles of machines, humans, and institutions. This includes evaluating core factors such as reliability, resilience, safety, interpretability, explainability, fairness, transparency, and accountability [36]. Jobin et al. [32] emphasize tools like fairness certificates, stakeholder dialogues, awareness of data usage, and the avoidance of harm as essential mechanisms for fostering trust.

While some scholars argue that the human notion of trust may not directly apply to machines, since technology lacks free will or moral agency [18, 38, 41], trust in AI can still be framed in human terms by focusing on the intentions and ethics of those who design and deploy these systems. This includes adapting dimensions like functionality, reliability,

and helpfulness to meet human standards of trustworthiness [41].

Ultimately, trust is a cornerstone of effective AI governance. It fosters collaboration and public acceptance [32]. However, most existing trust models, such as those proposed by Choung et al. [18], focus primarily on technical reliability and transparency. These models often overlook spiritual or cultural dimensions of trust, which are vital in many communities. In such contexts, frameworks that ignore moral and spiritual underpinnings may fail to inspire confidence. Al-Ghazali's principle of *Amanah*, with its emphasis on sincerity (*ikhlas*) and accountability (*muhasabah*), addresses this gap by embedding spiritual ethics into trust-building. This offers a culturally grounded and spiritually resonant approach to enhancing AI acceptance and ethical compliance [3, 4].

3.4 Ethical leadership in the age of AI

Ethical leadership is described as a way of leading by example and role-modeling; it means that leaders must uphold moral principles and make decisions that are just, fair, and transparent [14]. It is believed to be one of the key foundations in creating organizational cultures based on values such as trust, respect, and justice, especially in today's technology-driven sectors [19].

Ethical leadership is characterized by integrity, empathy, accountability, transparency, fairness, vision, trustworthiness, and respect [1, 33]. In the age of AI, these traits assist leaders in managing AI risks while maintaining ethical standards, fostering innovation, enhancing decision-making, building trust, engaging stakeholders, and influencing policies and regulations [33]. Such leaders, adopting AI systems and automation to shape decision-making processes, by prioritizing ethics, can create a culture where trust is developed and maintained in the rapidly changing environments that surround new technologies.

Ethical leadership directly influences how stakeholders, such as the public, policymakers, and AI users, perceive the governance and deployment of these systems. Leaders who support diverse teams and unbiased algorithms help foster trust in AI technologies, ensuring they are viewed as ethical and beneficial to all [20].

3.4.1 The role of ethical leaders in shaping AI trust

AI technologies can be developed and deployed responsibly within an organization as long as a culture of trust within that organization is encouraged; therefore, ethical leadership in such organizations becomes crucial [37]. Ethical leaders are vital in fostering such environments where the use of AI technologies is responsible and beneficial for the well-being of humans as well as for governance, especially in situations where decision-making is blurred, and stakeholders have differing interests [1]. Ethical leaders earn the trust of their stakeholders through their proof of being accountable and fair, and by making decisions that are in the best interest of the stakeholders over technological advancement.

In the same context, ethical leadership is considered to foster trust in AI systems by ensuring decisions are morally sound and ethically justified, guaranteeing AI is used to benefit individuals without harm [35, 40]. Ethical leaders are role models who demonstrate exemplary behavior, helping to promote a culture of ethics and integrity. They guide AI governance when ethical concerns are integrated into the design and implementation of AI systems, which leads to trust, inclusivity, and fairness while calling for policies that align AI usage with societal good [16].

Achieving trust in AI governance frameworks requires people to trust the institutions and individuals responsible for these systems, and ethical leaders play a pivotal role in aligning technical processes with ethical frameworks and creating an environment where technology and regulatory bodies are considered trustworthy [16–36].

Trust in AI governance is not a static trait, but it evolves, influenced by continuous interactions between the developers and the users of AI [25]. This dynamic nature of trust is closely related to ethical leadership, which plays a pivotal role in reinforcing trust, as leaders must model ethical decision-making and ensure the development and use of AI systems align with legal standards, societal values, and organizational benefits.

Furthermore, by fostering trust through ethical leadership and inclusive governance practices, organizations can guarantee the responsible and reliable use of AI technologies for promoting the common welfare.

Reinforcing trust in AI systems through ethical leadership, besides inclusive governance practices, can ensure AI technologies adopted by various organizations are responsibly promoting the greater good of the community [32, 46].

3.5 Al-Ghazali's ethical legacy

Abu Hamid Al-Ghazali (1058–1111) was a highly influential Islamic philosopher, theologian, jurist, and mystic. Referred to as *Hujjat al-Islam* (The Proof of Islam), he is celebrated across Islamic scholarship for his mastery of diverse disciplines and enduring impact on Islamic thought [51].

Al-Ghazali's work deeply integrated philosophy, spirituality, and ethics. In *Tahafut al-Falasifa* (The Incoherence of the Philosophers), he critically examined Greek philosophy, while in his magnum opus, *Ihya' Ulum al-Din* (The Revival of the Religious Sciences), he fused Islamic law with spiritual ethics. These writings emphasize moral character,

particularly virtues like trustworthiness (Amanah), sincerity (ikhlas), accountability (muhasabah), and justice (adl) as essential traits of ethical leadership [28, 29].

In *Nasihat al-Muluk* (Counsel for Kings), Al-Ghazali outlined the ethical duties of rulers, stressing justice, compassion, and responsibility for public welfare. He envisioned an ideal society guided by spiritually grounded leaders (wise, just, and morally upright), balancing worldly concerns with spiritual integrity.

His thoughts have had a profound influence not only in the Islamic world but also in medieval Europe, where his writings shaped Western philosophical discourse [23, 29]. Today, his works remain widely studied and translated, reflecting their ongoing relevance across cultural and religious boundaries [8, 51].

Modern theories such as Social Contract Theory [31], Ethical Leadership [13], Transformational and Servant Leadership [9, 17, 27, 53], and frameworks of Ethical Governance and Accountability [19] reflect concepts that parallel Al-Ghazali's emphasis on the ethical and moral responsibility of leadership. Scholars increasingly recognize his principles as foundational to contemporary Islamic management theory and as a rich source of ethical guidance for addressing modern governance challenges [8]. These concepts are directly relevant to contemporary AI ethics and governance, where the emphasis on transparency, accountability, and trust echoes Al-Ghazali's values. His framework offers not only a religious or cultural foundation but also a universal ethical lens through which to evaluate and guide the development of trustworthy AI systems.

3.6 Al-Ghazali's principle of *Amanah* (Trustworthiness)

Al-Ghazali presented the concept of *Amanah* (trustworthiness) as a primary ethical and spiritual principle. In his writings, *Amanah* means the trust awarded to individuals by God, including the responsibility of the person to uphold moral and personal integrity, besides ensuring justice and fairness in the community (Al-Ghazali, *Ihya 'Ulum al-Din*, 11th century).

Amanah is one of humans' moral qualities that must be translated into everyday life, as it is intertwined with responsibility because a given trust should be undertaken with full responsibility. In al-Ghazali's perspective, *Amanah* is deemed central to the existence of humans because it affects every aspect of life (from personal relationships to leading in governance) [21]. It is not only about fulfilling contractual commitments but also about being responsible for the impact of one's actions on others and ensuring that one is acting with sincerity and integrity. For leaders, this concept underscores the duty to serve the greater good with humility and wisdom.

Therefore, *Amanah* is the core of Islamic leadership; it is about having the capabilities aligned with the responsibilities that leaders carry. It encompasses values like integrity, accountability, justice, and responsibility [44]. Amanah forms a type of psychological contract between the leader and his subordinates, so the leader does his best to support, guide, protect, and treat his followers fairly and justly [7].

The concept of *Amanah* encourages leaders to act with transparency and to ensure their decisions are serving the benefit of their subordinates by taking responsibility for the consequences of their actions. Leaders, through embedding the principles of *Amanah*, can cultivate trust within their organizations and society [5]. Leaders who are guided by *Amanah* values do not seek personal power, but they seek to serve others with fairness and justice, promoting harmony and ethical governance [3, 24]. They perform in a manner that chastens the unjust and protects the rights. *Amanah* was found to have a positive influence on the effectiveness of leadership [51].

This resonates with the principles of ethical leadership [13], which deeply reflect the values of *Amanah*. Ethical leaders hold the responsibility of guiding the design and use of AI systems in a way that reflects moral values and promotes society's well-being. This approach contrasts with how *Amanah* emphasizes the importance of fulfilling obligations and duties with integrity and honesty.

In this context, the *Amanah* principle, rooted in responsibility and integrity and in the fulfillment of obligations without the need for supervision, aligns closely with the core objectives of ethical AI governance.

In Islamic ethics, it is a fundamental moral principle in Islam, along with justice and righteousness. It is essential for human relationships as well as for the relationship with God. It's mentioned in the holy Quran (Al-Ahzab: 72) that betraying *Amanah* is equal to being unfaithful to God. It is a comprehensive concept that encompasses not only trust but also responsibility and moral integrity [2, 51]. *Amanah* is also considered a crucial virtue for developing Islamic science, and according to Islam, an individual must practice *Amanah* in all aspects of their life. While in the Western perspective, it may not always incorporate a direct concept of trust in the scientific and ethical frameworks.

It can be concluded that the concept of *Amanah* can shape the organizational structure and management used by humans [54]. This includes leaders and workers in organizations who practice through their constant ethical behavior [34]. *Amanah* offers a contextually relevant, value-based framework to complement existing AI governance models in morally diverse societies.

4 Theoretical framework

In the theoretical framework of this paper, theories and models related to modern ethical leadership and ethical AI governance that align with *Amanah*'s principle of ethical leadership as presented by al-Ghazali were considered.

4.1 Ethical leadership theory [13]

According to this theory, the key principles to be upheld by ethical leaders are: integrity, accountability, fairness, and transparency. Leaders who embrace these values are role models; they can ethically and practically fulfill their obligations toward their organizations, subordinates, stakeholders, and society in general. These leaders can foster a culture of trust and respect within their organization and among all stakeholders.

This theory could be considered as the bridge to connect *Amanah* and modern leadership practices; its principles align seamlessly with ethical leadership in AI governance.

4.2 Social trust theory [48]

This theory examines the role of trust within the governance system and between institutions, ensuring the cooperation and effective communication between stakeholders.

It resonates with the function of trust in ethical AI governance by ensuring that various stakeholders in the ecosystem of AI (users, developers, regulators) are working together to ensure the responsible use and development of AI technologies.

Al-Ghazali's concept of *Amanah* can enrich the understanding of trust within the environment of AI systems by extending the meaning of trust beyond the technical requirements to a more rational and moral value.

Leaders who acknowledge *Amanah* in their behavior can ensure their decisions are ethically made, so they foster collaboration between the stakeholders to build a trustworthy AI environment.

4.3 AFPTS framework [33]

The AFPTS model directly impacts trust and trustworthiness by promoting ethical practices in AI development and deployment. This framework is built on five key principles: accountability, fairness, privacy, transparency, and sustainability. These principles are considered to improve trust among users and stakeholders [20]. It offers a set of guidelines in addition to a key set of actionable strategies to help ethical leaders in implementing and fostering ethical practices in the age of AI.

This also resonates completely with the principles of Al-Ghazali that characterize ethical leaders, particularly the value of *Amanah*.

5 Conceptual model

In the developed conceptual model of this paper, al-Ghazali's concept of *Amanah* is considered a fundamental construct to foster trust in AI governance through ethical leadership.

In this model, Ethical Leadership bridges the gap between the technical governance of AI systems and the need for moral oversight. Leaders who adhere to ethical principles ensure that AI systems are designed and used in ways that align with broader societal values such as fairness, justice, and equity. Trust in AI systems is the outcome of the combined efforts of ethical leadership and the moral framework of *Amanah*. Trust is built when AI systems operate transparently, are free from bias, and are aligned with societal needs.

Ethical leaders, guided by *Amanah*, create policies and oversight mechanisms that ensure the ethical development and use of AI which leads as a result to fostering public and institutional trust in the AI systems, stakeholders will feel confident that the systems will operate in a fair, transparent, and accountable manner beside the technical competency.

6 Discussion

The rapid spread of AI applications across various organizations and institutions has generated growing interest in AI ethics among stakeholders and broader society. AI governance frameworks are intended not only to regulate technical operations but also to guide the ethical development and deployment of these technologies.

This paper explores Al-Ghazali's contributions to Islamic and humanistic philosophy, emphasizing the ongoing

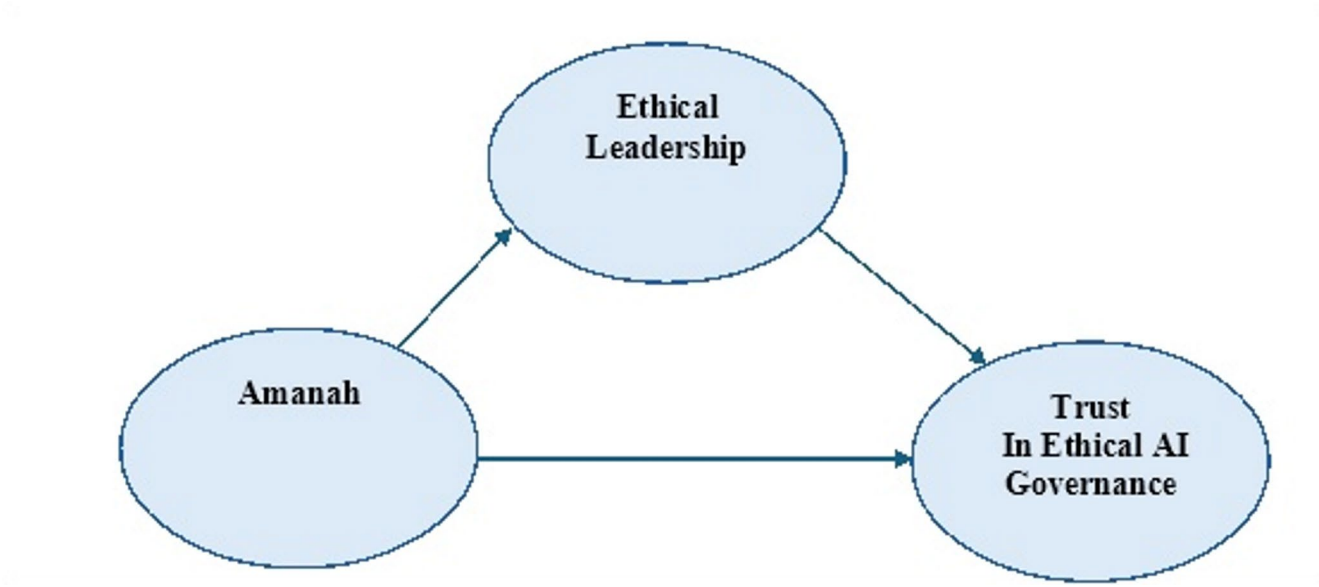

Fig. 1 Model Linking Amanah, Ethical Leadership, and Trust in AI Governance

relevance of his intellectual legacy. By reflecting on his principle of *Amanah* (trust), the study proposes a morally grounded approach to ethical leadership in the age of AI.

Amanah is an ethical value deeply rooted in Islamic tradition, as well as in many other religious and cultural systems. It emphasizes trust, trustworthiness, and the moral responsibilities of individuals and leaders toward each other and society. Integrating Al-Ghazali's concept of *Amanah* into AI governance offers a way to bridge the widening gap between technical standards and ethical obligations, particularly in AI-driven environments where decision-making is increasingly autonomous and detached from human oversight.

Incorporating authentic ethical values into AI governance frameworks would help ensure richer and more reliable ethical systems. Such integration encourages AI developers and users to prioritize ethical considerations, reduce risks, and promote societal well-being.

While contemporary AI governance frameworks such as the EU AI Act, IEEE Ethically Aligned Design, and OECD AI Principles establish foundational values like transparency, accountability, and fairness, they often operate within secular, technocratic paradigms that may lack resonance with spiritual and cultural traditions. This can lead to a disconnect between regulatory principles and the internal moral motivations required for ethical behavior, especially in contexts where trust is grounded in deeply held religious worldviews.

The *Amanah*-based framework proposed in this study addresses that gap by framing trust not merely as an institutional requirement, but as a moral obligation rooted in individual accountability before God (*muhasabah*), sincerity (*ikhlas*), and justice (*adl*). This deep moral anchoring fosters internal commitment to ethical conduct, something external rules alone may struggle to achieve.

Moreover, while existing frameworks prescribe *what* ethical values AI should follow, the *Amanah* approach adds depth by addressing *why* and *how* these values can be internalized within the conscience of AI developers, users, and policymakers. In culturally Muslim or values-oriented environments, the *Amanah* framework enhances legitimacy, motivation, and ethical alignment, complementing and reinforcing existing regulatory frameworks. Thus, this model is not a replacement for global governance standards but rather a culturally grounded ethical augmentation that supports responsible AI innovation in diverse sociotechnical contexts.

Amanah, as articulated by Al-Ghazali nearly a millennium ago, continues to inspire contemporary leadership by promoting foundational values such as honesty, integrity, fairness, and accountability. Embedding these values in AI governance frameworks can foster a culture of trust among stakeholders and ensure that AI technologies serve humanity ethically.

The findings of this paper highlight the potential of *Amanah* to inform ethical leadership and strengthen trust-based AI governance frameworks. These principles could be translated into actionable strategies, such as encouraging AI developers and organizational leaders to prioritize fairness and accountability through the use of audit trails, explainable AI, and mechanisms for ethical accountability consistent with *Amanah*.

This paper acknowledges limitations due to its theoretical nature and lack of empirical evidence. The analysis is based on a literature review rather than case studies or

data-driven investigations. Another limitation is the focus on a single ethical principle (*Amanah*) and one dimension of AI governance (trust).

Future research could include empirical investigations, such as interviews, surveys, or case studies, on the role of *Amanah* in shaping ethical leadership within technology-driven sectors. Additionally, broader studies exploring Al-Ghazali's wider philosophical and ethical legacy in light of modern global challenges could enrich the academic discourse around culturally rooted approaches to AI ethics.

7 Theoretical implications

Ethical leadership inspired by Al-Ghazali's moral values and ethical principles can significantly contribute to trust building within organizations, among stakeholders, and in society as a whole.

Ethical leaders, through their adoption of deep moral values, can further build confidence in the AI systems and technologies used and developed by their firms.

By practicing *Amanah* as defined and described in al-Ghazali's works, leaders can garner a competitive advantage for their organizations and differentiate themselves in the modern AI-driven marketplace.

Embracing *Amanah* and what it stands for (accountability, transparency, fairness, justice, self-control,…), leaders can ensure that their organizations contribute positively to the ethical deployment and development of AI technologies and promote valuable societal standards.

8 Conclusion

Even with the unprecedented technological advancements that have invaded our lives, the intellectual legacy of the leaders of thought who enriched humanity's history will always offer insights and guidance to navigate modern challenges, uphold ethical principles, and foster a more balanced and humane future.

Incorporating spiritual values, stemming from the wisdom of former prominent theologians like al-Ghazali, can strengthen the ethical foundation of AI systems and ensure they are used and developed in a way that respects ethical values and promotes the greater good of society. One of these ethical principles is *Amanah*, which exceeds the meaning of trust to a more holistic moral approach, inspiring leaders to behave ethically and fulfill their obligations with integrity, and prioritize the well-being of society over personal or organizational gains. This principle encourages accountability, fairness, and transparency in decision-making, which are essential for building trust in AI systems and ensuring their alignment with human-centric values.

Overall, as AI technologies become more integrated into society, the need for leaders who can navigate technical requirements and ethical issues becomes more evident. Ethical leaders help bridge the gap between technical capabilities and societal expectations by ensuring the right design and implementation of AI systems in alignment with ethical values.

This paper has argued that *Amanah*, as articulated by Al-Ghazali, provides a spiritually grounded ethical framework for AI governance. By embedding values of trust, sincerity, and accountability, it enhances the legitimacy, inclusivity, and practical acceptance of AI governance frameworks, especially in culturally diverse societies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43681-025-00780-y.

Author contributions B.S. (Bahar Serevan) conceptualized the study, conducted the literature review, developed the theoretical framework, wrote the manuscript, and prepared all figures. The author reviewed and approved the final version of the manuscript.

Funding The author did not receive support from any organization for the submitted work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval Not applicable- This article does not contain any studies with human participants or animals performed by the author.

References

1. Adeniyi, A.C., Okusi, O.: Ethical AI governance, financing, and human well-being in the 21st century. Afr. J. Humanit. Contemp. Educ. Res. **17**(1), 193–205 (2024). https://doi.org/10.62154/ajhcer.2024.017.010504
2. Ahmet, N., Akdogan, C.: Trust (Al-Amanah): A comparative study of its application in Islamic and Western science. In *Asian Conference on Ethics, Religion, and Philosophy*. (2012)
3. Al-Ghazali, A.: Ihya' Ulum al-Din (The Revival of the Religious sciences). Dar al-Kutub al-Ilmiyah, Beirut (2000)
4. Al-Ghazali, A.H.M.: Ihya Ulum al-din. Pustaka Nasional, Singapura (1988)
5. Al-Ghazali, M.: The Truth of Arab Nationalism, and the Myth of the Arab Baath, 3rd edn., pp. 13–14. Nahdat Misr Publishing, Cairo (2005)
6. Al-Hakim, A.A.A.N.: Al-Mustadrak 'ala al-Sahihayn. Dar al-Haramayn, Kaherah (1997)
7. Amirin, T.M.: Kepemimpinan Yang Amanah. Dinamika Pendidikan. **1**(14), 1–11 (2007)

8. Azwar, A., Usman, A.H., Noor, U.M., Abdullah, M.F.R.: Iṣlāḥ of al-Ghazālī's ideas and movements and their relevance to contemporary Islamic preaching. Ulumuna. **28**(1), 453–484 (2024). https://doi.org/10.20414/ujis.v28i1.513
9. Bass, B.M.: Leadership and Performance Beyond Expectations. Free, New York (1985)
10. Batool, A., Zowghi, D., Bano, M.: AI governance: A systematic literature review. AI Ethics. 1–15 (2025). https://doi.org/10.21203/rs.3.rs-4784792/v1
11. Binns, R.: On the ethics of artificial intelligence. Int. J. Hum. Comput. Stud. **58**(4), 374–391 (2018)
12. Botsman, R.: Who Can You Trust?? How Technology Brought Us together–and why It Could Drive Us Apart. Penguin UK, London (2017)
13. Brown, M.E., Treviño, L.K.: Ethical leadership: A review and future directions. Leadersh. Q. **17**(6), 595–616 (2006). https://doi.org/10.1016/j.leaqua.2006.10.004
14. Brown, M.E., Treviño, L.K., Harrison, D.A.: Ethical leadership: A social learning perspective for construct development and testing. Organ. Behav. Hum Decis. Process. **97**(2), 117–134 (2005). https://doi.org/10.1016/j.obhdp.2005.03.002
15. Brundage, M., et al.: Toward trustworthy AI development: mechanisms for supporting verifiable claims. ArXiv Preprint arXiv:2004 07213. (2020). https://doi.org/10.48550/arXiv.2004.07213
16. Bryson, J.J.: Patiency is not a virtue: The design of intelligent systems and systems of ethics. Ethics Inf. Technol. **20**(1), 15–26 (2018). https://doi.org/10.1007/s10676-018-9448-6
17. Burns, J.M.: Leadership. Harper & Row, New York (1978)
18. Choung, H., David, P., Ross, A.: Trust in AI and its role in the acceptance of AI technologies. Int. J. Human–Computer Interact. **39**(9), 1727–1739 (2022). https://doi.org/10.48550/arXiv.2203.12687
19. Ciulla, J.B.: What is good leadership? *[Add journal or book information if applicable]*. (2004)
20. Dastin, J.: Amazon scraps secret AI recruiting tool that showed bias against women. Reuters (2018)
21. Dignum, V.: Responsible autonomy. (2017). https://doi.org/10.48550/arXiv.1706.02513
22. Dignum, V.: *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way.*Springer. (2019)
23. Eichner, H.: Al-Ghazali: Philosopher of the Islamic Golden Age. Reaktion Books, London (2022)
24. Esposito, J.L.: What Everyone Needs To Know about Islam, 2nd edn. Oxford University Press (2011)
25. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V.,…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
26. Gasser, U., Almeida, V.A.: A layered model for AI governance. IEEE Internet Comput. **21**(6), 58–62 (2017). https://doi.org/10.1109/MIC.2017.4180835
27. Greenleaf, R.K.: Servant Leadership: A Journey into the Nature of Legitimate Power and Greatness. Paulist, New York (1977)
28. Griffel, F.: August 14; revised 2020, May 8). Al-Ghazali. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition). (2007)
29. Griffel, F.: Al-Ghazali's Philosophical Theology. Oxford University Press, New York (2009)
30. Grote, T., Berens, P.: On the ethics of algorithmic decision-making in healthcare. J. Med. Ethics. **46**(3), 205–211 (2020). https://doi.org/10.1136/medethics-2019-105586
31. Hobbes, T.: *Social contract theory*. (1964). Retrieved from https://iep.utm.edu/soc-cont
32. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nat. Mach. Intell. **1**(9), 389–399 (2019). https://doi.org/10.1038/s42256-019-0088-2
33. Kandasamy, U.C.: Ethical Leadership in the Age of AI Challenges, Opportunities and Framework for Ethical Leadership. (2024). https://doi.org/10.48550/arXiv.2410.18095
34. Kasim, A.C., Dien, M.I., Kasim, S.C., Manap, J.: The impact of *Amanah* on individual manners and the society. Int. J. Acad. Res. Bus. Social Sci. **10**(9), 629–640 (2020). https://doi.org/10.6007/IJARBSS/v10-i9/7854
35. Kizilcec, R.F.: How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2390–2395). (2016)., May https://doi.org/10.1145/2858036.2858402
36. Lahusen, C., Maggetti, M., Slavkovik, M.: AI governance and trust: Bridging the gap. J. AI Ethics. **1**(2), 45–59 (2024). https://doi.org/10.1038/s41598-024-71761-0
37. Maak, T., Pless, N.M., Voegtlin, C.: Business statesman or shareholder advocate? CEO responsible leadership styles and the micro-foundations of political CSR. J. Manage. Stud. **53**(3), 463–493 (2016). https://doi.org/10.1111/joms.12195
38. Madhavan, P., Wiegmann, D.A.: Similarities and differences between human–human and human–automation trust: An integrative review. Theoretical Issues Ergon. Sci. **8**(4), 277–301 (2007). https://doi.org/10.1080/14639220500337708
39. Mäntymäki, M., Minkkinen, M., Birkstedt, T., Viljanen, M.: Putting AI ethics into practice: The hourglass model of organizational AI governance. (2022). https://doi.org/10.48550/arXiv.2206.00335
40. Mayer, R.C.: An integrative model of organizational trust. Acad. Manage. Rev. (1995). https://doi.org/10.5465/AMR.2007.24348410
41. McKnight, D.H., Carter, M., Thatcher, J.B., Clay, P.F.: Trust in a specific technology: An investigation of its components and measures. ACM Trans. Manage. Inform. Syst. (TMIS). **2**(2), 1–25 (2011)
42. Mittelstadt, B.: Principles alone cannot guarantee ethical AI. Nat. Mach. Intell. **1**(11), 114 (2019). https://doi.org/10.1038/s42256-019-0114-4
43. Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L.: The ethics of algorithms: Mapping the debate. Big Data Soc. **3**(2), 2053951716679679 (2016). https://doi.org/10.1177/2053951716679679
44. Mukred Mohsen, N.R.: *Leadership From The Qur'an Relationship Between Taqwa, Trust And Business Leadership Effectiveness* (Doctoral dissertation, Universiti Sains Malaysia). (2007)
45. Nawi, A., Khamis, N.Y., Yaakob, M.F.M., Samuri, M.A.A., Zakaria, G.A.N.: Exploring opportunities and risks of artificial intelligence research for Islamic ethical guidelines. Afkar: Jurnal Akidah Dan. Pemikiran Islam. **25**(2), 1–34 (2023). https://doi.org/10.22452/afkar.vol25no2.1
46. O'Neil, C.: Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy. Broadway Books, New York, NY (2016)
47. Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., Wellman, M.: Machine behaviour. *Nature***568**(7753), 477–486 (2019). https://doi.org/10.1038/s41586-019-1138-y
48. Rothstein, B., Stolle, D.: The state and social capital: an institutional theory of generalized trust. Comp. Politics. **40**(4), 441–459 (2008). https://doi.org/10.5129/001041508X12911362383354
49. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach (4th ed). Pearson. (2021). https://doi.org/10.1109/MSP.2017.2765202

50. Shneiderman, B.: Human-centered artificial intelligence: reliable, safe & trustworthy. Int. J. Human–Computer Interact. **36**(6), 495–504 (2020). https://doi.org/10.48550/arXiv.2002.04087
51. Shuhari, M.H., Hamat, M.F., Ismail, M.S., Jaffar, Y., Mustafa, M.F., Abdullah, M.S.: Elements of integrity within Muslim individuals according to the thought of al-Ghazali. Int. J. Acad. Res. Bus. Social Sci. **8**(10), 270–278 (2018)
52. Siddiqui, A.: Christian-Muslim Dialogue in the Twentieth Century. Springer (1997). https://doi.org/10.1057/9780230378230
53. Spears, M.: Reflections on Robert K. Greenleaf and servant-leadership. Greenleaf Cent. Servant Leadersh. (1995). https://doi.org/10.1108/01437739610148367
54. Sulistiani, D.: The concept of Amanah metaphor in accounting: Meaning, guidelines, and purposes. Jurnal Akuntansi: Kajian Ilmiah Akuntansi. **11**(2), 286–300 (2024). https://doi.org/10.30656/jak.v11i2.7795
55. Vayena, E., Blasimme, A., Cohen, I.G.: Machine learning in medicine: Addressing ethical challenges. PLoS Med. **15**(11), e1002689 (2018). https://doi.org/10.1371/journal.pmed.1002689
56. Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M.C., Dafoe, A.: Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. J. Artif. Intell. Res. **71**, 591–666 (2021). https://doi.org/10.48550/arXiv.2105.02117