An-Nahdhah
JOURNAL OF EDUCATION AND CULTURE

Vol. 2 No. 1 (2026)

Available online at
https://journal.arate.id/index.php/An-Nahdhah

Human-AI Interaction in the Metaverse: A Maqasid al-Shariah Conceptual Framework for Autonomous Agents, Virtual Governance, and Moral Education

Zarqa Zulifqar[1]*

[1] M-Y University Islamabad, Pakistan

Article History

Submitted: 17/09/2026
Revised: 17/10/2026
Accepted: 28/10/2026
Published: 01/11/20026

Keywords:

Metaverse; Artificial Intelligence; Shariah; Maqasid al-Shariah; Virtual Governance; Moral Education.

Corresponding Author:
Zarqa Zulifqar

zarqazulifqar1986@gmail.com

The rapid convergence of artificial intelligence and immersive virtual environments has given rise to the Metaverse, a persistent digital space in which autonomous AI agents increasingly mediate human social, economic, and educational interaction. Existing Islamic legal scholarship has examined the legal capacity of AI and has proposed maqasid-based governance for Islamic finance and for national regulation, but it has not yet shown how the legal status of autonomous agents, governance in a space without a single jurisdiction, and moral education in immersive settings depend on one another. This study addresses that gap by asking how maqasid al-Shariah and the classical doctrine of legal capacity (al-ahliyyah) can ground a conceptual framework for autonomous agents, virtual governance, and moral education in the Metaverse. Using a qualitative library research method, the study analyzes classical fiqh concepts alongside contemporary Islamic legal scholarship on artificial intelligence, robotics, and digital ethics, and situates them within the secular literature on metaverse governance and virtual moral education. The study advances three conceptual propositions. First, autonomous AI agents cannot bear full legal or moral responsibility (taklif) because they lack the volition and consciousness required by Islamic legal theory, yet a bounded, functional capacity can be recognized for practical governance purposes, with liability ultimately resting with human developers, operators, and institutions. Second, the five maqasid can be mapped onto the governance gaps documented in secular metaverse scholarship to yield a jurisdiction-independent review checklist. Third, immersive moral education is compatible with tarbiyah only when its design is governed by explicit safeguards. The main contribution of this study is an integrated, maqasid-based framework that links these three areas, which have so far been examined separately. The findings are conceptual and normative rather than empirically tested, and they carry practical implications for Islamic financial and educational institutions designing metaverse platforms and AI agents intended for Muslim users.

DOI: *10.63199/an-nahdhah.v2i1.99*

Copyright © 2026 by the authors. This is an open access publication under the terms and conditions of the Creative Commons Attribution (CC BY NC 4.0) license.

1. Introduction

The Metaverse, understood as a persistent network of interconnected, immersive virtual environments (Kourtesis, 2024; Qin, Wang, & Hui, 2025), is increasingly populated by autonomous AI agents, ranging from conversational chatbots and non-player characters to AI-driven virtual companions. These agents are no longer passive tools but active participants in human interaction. This shift raises a question that classical Islamic legal scholarship never anticipated: when an autonomous agent acts in a space with no single jurisdiction, who bears responsibility for what it does, and by what normative standard should that responsibility be governed? This study treats that question as its central problem, and treats virtual governance and moral education as the two arenas in which the answer matters most.

Secular legal scholarship has already documented serious governance gaps in the Metaverse, including unresolved questions of jurisdiction, weak enforcement mechanisms, and the difficulty of classifying virtual misconduct as crime (Qin et al., 2025; Tyagi, 2025). At the same time, a rapidly growing body of Islamic legal research has begun to examine whether artificial intelligence systems can hold legal capacity (al-ahliyyah) comparable to that of a human being, a corporation, or an endowment (waqf) recognized in classical fiqh (Sitiris & Busari, 2024; Yunos, Wan Saifuddin, & Hamdan, 2024), and whether autonomous machines can be held morally or legally accountable at all (Hamsin, Anggriawan, & Jiatrahman, 2023). Parallel to these legal debates, education researchers have shown that immersive technology can be a powerful, double-edged tool for moral development, capable of cultivating empathy and virtue but also of producing manipulation, desensitization, or shallow moral learning if left unregulated (Harrison, 2024; Kurian & Hajir, 2024; Rueda & Lara, 2020).

Despite this activity, a more specific gap remains, and it has three parts. First, Islamic scholarship on AI legal capacity has examined AI systems, robots, and autonomous mobile robots in general (Hamsin et al., 2023; Sitiris & Busari, 2024; Yunos et al., 2024), but it has not asked what bounded status an agent can hold in a persistent, avatar-mediated space where harm may cross borders. Second, maqasid-based governance proposals have so far been developed for Islamic financial institutions or national regulatory systems (Qutieshat et al., 2026; Zafar & Ali, 2025), not for virtual environments in which no single jurisdiction applies. Third, research on moral education in virtual reality is largely secular (Harrison, 2024; Kurian & Hajir, 2024) and offers no Shariah-based criteria for judging whether immersive experiences build durable character.

These gaps are not only parallel but dependent on one another. The status assigned to an agent determines who is accountable within virtual governance, and the quality of that governance determines whether moral education in the same space is safe. What is missing, then, is not one more review of each field but an account of how the three connect. To the best of the author's knowledge, no existing study provides such an account within a single Shariah-based framework.

The present study therefore asks three interrelated questions. First, what standing can autonomous AI agents be given under Shariah, and who bears responsibility for their actions? Second, how can maqasid al-Shariah inform virtual governance in a space where conventional jurisdiction breaks down? Third, what does Shariah-informed moral education look like inside an immersive, AI-mediated environment?

The main contribution of this study is an integrated, maqasid-based conceptual framework that answers these questions together rather than separately. It differs from earlier work in three ways. It extends the debate on AI legal capacity from general AI and robots to autonomous agents in the Metaverse. It moves maqasid-based governance from Islamic finance and national regulation to jurisdiction-free virtual spaces. It adds tarbiyah-based safeguards that secular research on virtual moral education lacks. The framework is stated as three propositions, which are developed in the Results and tested for coherence against the literature. They are conceptual and normative. They offer a review checklist for Islamic financial institutions, educators, and platform designers, not tested or binding standards, and their empirical testing is left to future research.

Classical Islamic jurisprudence distinguishes between ahliyyat al-wujub, the capacity to hold rights and obligations, and ahliyyat al-ada, the capacity to perform legally binding acts, both of which are traditionally anchored in human reason (aql), maturity (bulugh), and moral responsibility (taklif) (Yunos et al., 2024). Contemporary Islamic legal scholars examining artificial intelligence largely agree that an autonomous system, however sophisticated its behavior, cannot satisfy the conditions of taklif because it lacks true consciousness, intention (niyyah), and moral discernment (Sitiris & Busari, 2024). Hamsin et al. (2023) reach a similar conclusion through a doctrinal analysis that classifies AI robots as jamadat, or inanimate objects, arguing that no matter how sophisticated its behavior, an artificial system lacks the biological vitality and spiritual dimension that Islamic law requires before civil or criminal liability can attach, so that only the humans behind the system can properly be sanctioned. At the same time, several scholars note that Islamic law already recognizes limited or artificial forms of legal personality for non-human entities such as corporations and waqf endowments, and argue that a similarly bounded, functional capacity could be extended to AI systems for

practical purposes such as contractual representation or liability allocation, without granting them the full legal status of a human being (Sitiris & Busari, 2024; Yunos et al., 2024). Under this view, responsibility for an autonomous agent's actions in the Metaverse would ultimately be traced back to its developers, operators, or the institutions that deploy it, in a manner resembling vicarious liability in Islamic tort doctrine.

Several recent reviews have mapped the wider field of Islamic legal and ethical scholarship on artificial intelligence, providing the backdrop against which the present study's three pillars can be situated. Applying the PRISMA method to sixty-seven peer-reviewed publications, Sudirman, Sutiah, and Supriyono (2025) identify five recurring themes in the literature, namely Shariah-aligned AI ethics, AI personhood and legal responsibility, the integration of AI in Islamic finance and judiciary systems, AI-assisted fatwa issuance, and regulatory gaps in aligning AI with maqasid al-shariah, while also noting persistent inconsistencies in theological interpretation and a general absence of policy frameworks across Muslim jurisdictions. Alamsyah, Yayuli, and Remanda (2026) reach a compatible conclusion through a narrower doctrinal lens, arguing that Islamic law possesses sufficient interpretive flexibility, through analogy (qiyas) and the legal maxims (qawaid fiqhiyyah), to respond to artificial intelligence even though no classical text addresses it directly. Complementing these reviews, Nawi, Khamis, Mohd Yaakob, Samuri, and Zakaria (2023) survey the opportunities and risks of AI research from an Islamic ethical standpoint and conclude that existing secular AI guidelines remain limited from an Islamic perspective, since they rarely engage with concepts such as khalifah, or human stewardship, and amanah, or trust, that Muslim scholars consider central to any legitimate governance of technology.

Not all Islamic scholars agree that maqasid al-Shariah, on its own, is the most adequate ethical lens for artificial intelligence. Ali, Bouzoubaa, Gelli, Hamzi, and Khan (2025) argue that maqasid-based and legal-maxim-based approaches were originally designed to resolve juristic rulings rather than to address the open-ended ethical complexity of AI, and propose instead a trusteeship, or i'timani, framework grounded in the Qur'anic concept of amana, or divine trust, structured around three covenants that describe human responsibility toward God, toward knowledge, and toward creation. Elmahjub (2023) offers a related but distinct contribution, proposing a pluralist ethical benchmarking approach that draws on multiple schools of Islamic ethical thought simultaneously rather than privileging any single framework, on the grounds that AI applications are too varied for one doctrinal lens to govern consistently. Al-Momani (2025) adds a cautionary note from within Islamic jurisprudence itself, arguing that AI tools used to interpret or generate Islamic legal opinions raise distinct ethical challenges

around authority and legitimacy that neither maqasid reasoning nor trusteeship ethics alone fully resolves. Taken together, this debate suggests that the maqasid-based framework adopted in the present study is best understood as one useful and practically tractable lens among several credible Islamic ethical approaches, rather than as an uncontested consensus position.

Despite the debate described above, maqasid al-Shariah remains one of the most frequently invoked frameworks in recent Islamic scholarship on artificial intelligence, precisely because it offers principles broad enough to address novel technologies that classical texts could not foresee (Mohammed, 2024). The five classical objectives, the protection of religion (din), life (nafs), intellect (aql), lineage and dignity (nasl), and property (mal), have been applied to digital contexts to argue that virtual platforms should support rather than undermine religious practice, that psychological wellbeing in virtual spaces deserves the same protection as physical safety, and that digital literacy is now a component of protecting the intellect from manipulation and misinformation. In the financial sector, this reasoning has already produced concrete proposals: Zafar and Ali (2025) develop a draft Shariah governance standard for generative AI in Islamic financial institutions that combines a Shariah Supervisory Board with a dedicated AI governance committee, embedding Shariah review at every stage of the AI lifecycle from data acquisition to deployment and monitoring. Najib, Basarud-din, and Fazial (2025) apply a similar maqasid al-Shariah perspective specifically to AI in Islamic finance, concluding that AI adoption should be evaluated against its contribution to maslahah, or public benefit, rather than judged purely on efficiency grounds. At a broader regulatory level, Qutieshat, Al Adwan, and Khater (2026) examine AI governance across several Muslim-majority countries and find that existing regulatory frameworks do not adequately address explainability standards, algorithmic auditing, judicial review, or legal liability for automated decisions, proposing instead a maqasid al-Shariah based governance model to close these normative gaps.

A closely related strand of research examines what happens when AI systems are used to generate Islamic legal content directly rather than merely being evaluated by Islamic ethics. Munshi, AlSabban, Farag, Rakha, Al Sallab, and Alotaibi (2022) present one of the earliest large-scale technical efforts in this area, building a natural language processing system trained on more than 850,000 questions and answers to automatically generate Islamic jurisprudential legal opinions, while explicitly cautioning that such systems function best as tools that assist a qualified mufti rather than as independent sources of religious authority. This distinction between AI as an assistive tool and AI as an autonomous authority recurs throughout the literature reviewed in this study and

reinforces the position, developed further below, that Shariah-informed principles for autonomous agents should deny independent legal personhood while still allowing a bounded, functional role. Alif Auladi and Muwahid (2025) extend this concern about authority and trust into the domain of cybersecurity, arguing from a philosophical-Islamic standpoint that digital trust cannot be reduced to technical safeguards alone and must be anchored in moral responsibility (amanah) distributed across developers, institutions, and users, a claim that applies with equal force to autonomous agents operating inside the Metaverse.

Beyond conceptual and doctrinal analysis, some studies examine how specific Muslim-majority states are already responding to AI governance challenges. Gorian and Osman (2024) compare digital ethics regulation for artificial intelligence in Saudi Arabia and the United Arab Emirates, finding that both states frame data protection and AI ethics partly through Islamic conceptions of privacy and human dignity, yet neither has produced a comprehensive Shariah-specific AI governance code comparable to the sector-specific standard proposed by Zafar and Ali (2025) for Islamic finance. This uneven regulatory landscape mirrors, at the national level, the jurisdictional fragmentation that secular scholars have documented at the global level for the Metaverse, and it is precisely this fragmentation that a maqasid-based, jurisdiction-independent governance logic is intended to address.

Independent of any religious framework, secular legal and technical scholarship has identified serious structural weaknesses in how the Metaverse is currently governed. Qin et al. (2025) show that the Metaverse's borderless, multi-jurisdictional nature makes it difficult to determine which country's laws should apply when harm occurs, and that existing categories of crime do not map cleanly onto virtual misconduct such as avatar harassment or identity theft. Tyagi (2025) similarly documents unresolved questions of enforceability, consumer protection, and financial regulation across virtual economies, noting that jurisdictions such as the United States, the European Union, and China are pursuing incompatible regulatory paths. Kourtesis (2024) broadens this picture by cataloguing the wider risks associated with immersive extended reality technologies inside the Metaverse, including data privacy vulnerabilities, cybersickness, addiction, dissociation, harassment, and misinformation, and argues that governance frameworks have generally lagged behind the pace of technical deployment. These findings matter for Islamic legal scholarship because they show that the absence of unified governance is not merely a theoretical problem but a documented, practical one, and that any Shariah-based framework for the Metaverse must be designed to function without relying on a single national jurisdiction or enforcement authority.

A separate but converging strand of research examines what immersive technology does to moral development. Harrison (2024) argues that virtual reality can offer character educators four distinctive learning opportunities, described as expansive, immersive, embodied, and autonomous, that are difficult to replicate in a traditional classroom, while also cataloguing risks such as desensitization and the blurring of fictional and real moral responsibility. Rueda and Lara (2020) focus specifically on the use of virtual reality to cultivate empathy, concluding that immersive perspective-taking exercises can enhance empathic response but require careful ethical design to avoid manipulation or superficial engagement. Kurian and Hajir (2024) offer a more cautionary perspective, questioning whether virtual reality experiences marketed as empathy machines actually produce durable moral change, and highlighting that decisions about whose suffering is dramatized, and by whom, carry their own ethical weight. Dunivan, Mann, Collins, and Wittmer (2024) extend this line of inquiry beyond empathy to compassion and moral reasoning more broadly, suggesting that the moral effects of virtual reality are multidimensional and not yet fully understood. Taken together, this literature suggests that immersive, AI-mediated environments are neither inherently beneficial nor inherently harmful for moral formation, but depend heavily on how the experience is designed and governed, which is precisely the point at which Islamic educational principles (tarbiyah) can offer additional normative guidance beyond secular pedagogy.

The moral education literature above sits within a wider body of scholarship on AI in education generally. Kamalov, Santandreu Calonge, and Gurrib (2023) review the applications, advantages, and challenges of AI across schools and universities, concluding that institutions have little practical alternative but to embrace the technology while building guardrails against its misuse, a conclusion that applies directly to Islamic educational institutions considering AI-mediated instruction inside metaverse platforms. Read alongside the digital trust concerns raised by Alif Auladi and Muwahid (2025), this literature reinforces the view that moral and educational safeguards for immersive AI cannot be left to informal norms but require deliberate institutional design.

Building on the strands above, this study adopts a theoretical framework that pairs maqasid al-Shariah with the secular five-principle ethical framework for artificial intelligence proposed by Floridi and Cowls (2019), which synthesizes beneficence, non-maleficence, autonomy, justice, and explicability from prominent AI ethics guidelines. The maqasid provide the normative anchor and ultimate purpose, protecting religion, life, intellect, lineage, and property, while Floridi and Cowls's principles supply an operational vocabulary, particularly explicability, understood as both intelligibility and

accountability, that translates smoothly into Islamic legal concepts of transparency and liability. Where the trusteeship and pluralist alternatives reviewed above (Ali et al., 2025; Elmahjub, 2023) emphasize the metaphysical and spiritual dimensions of AI ethics, the maqasid-Floridi synthesis adopted here is deliberately oriented toward practical governance, since it allows the study to evaluate autonomous agents, virtual governance, and moral education in the Metaverse using a single, coherent set of normative and operational criteria rather than treating Islamic and secular literatures as separate, non-communicating traditions. The three propositions presented in the Results are derived through this combined lens.

2. Methods

This study employs a qualitative library research method, sometimes described as doctrinal or normative legal research, which is appropriate for questions that require the interpretation and synthesis of textual sources rather than the collection of new empirical data (Sitiris & Busari, 2024; Yunos et al., 2024). Primary materials consist of the general principles of Islamic legal theory concerning legal capacity (al-ahliyyah) and moral responsibility (taklif) as discussed in the classical fiqh concepts referenced by contemporary scholars, together with the five classical objectives of maqasid al-Shariah. Secondary materials consist of twenty-five peer-reviewed journal articles, working papers, and conference proceedings published between 2019 and 2026 that address artificial intelligence and Islamic law, Metaverse governance, and moral or character education through virtual reality.

Sources were identified through targeted searches of academic databases and search engines using combinations of the terms Shariah, Islamic law, artificial intelligence, autonomous agents, Metaverse, virtual governance, moral education, and virtual reality. Studies were included if they were published in a peer-reviewed journal or a recognized academic repository, directly addressed at least one of the three thematic pillars of this study, and were available in English (one source, Yunos et al., 2024, has an English abstract but its full text is in Malay). Studies that addressed artificial intelligence in Islamic finance without any bearing on legal personality, governance, or moral education were excluded unless they offered a governance model, such as a Shariah supervisory structure, that could be generalized beyond the financial sector.

Once collected, sources were read closely and organized thematically into the three pillars identified in the research questions: the legal status of autonomous agents, the governance of virtual space, and moral education within immersive environments.

Within each theme, the study applied a form of qualitative content analysis, comparing how different authors defined key concepts such as legal capacity, jurisdiction, and moral formation, and identifying points of convergence and disagreement across the Islamic and secular literatures. This comparative synthesis forms the basis for the integrated framework and the findings presented in the following sections. The propositions stated in the Results are the product of this comparative synthesis. They are analytical claims reasoned from the sources, not empirical findings.

3. Result and Discussion

3.1. Shariah-Informed Principles for Autonomous Agents

The reviewed literature converges on the position that autonomous AI agents operating in the Metaverse cannot be granted full legal personality equivalent to that of a human being, because they lack the reason, intention, and moral consciousness that Islamic legal theory requires for taklif (Hamsin et al., 2023; Sitiris & Busari, 2024; Yunos et al., 2024). At the same time, the literature does not treat AI agents as ordinary property either. Sitiris and Busari (2024) argue that AI systems possess a form of artificial personality that Islamic jurisprudence can accommodate by analogy with existing legal fictions such as the corporation or the waqf, provided that this artificial personality is understood as a practical, bounded construct rather than genuine moral agency. Yunos et al. (2024), examining autonomous mobile robots specifically, reach a similar conclusion through the classical distinction between ahliyyat al-wujub and ahliyyat al-ada, finding that such systems cannot hold ahliyyah in the full human sense but may still be assigned a functional status sufficient to determine, for example, who is liable when the system causes harm. Munshi et al. (2022) provide indirect support for this bounded view from an engineering perspective, insisting that even a highly capable AI system generating Islamic legal content should function as a tool assisting a qualified human authority rather than as an independent source of religious or legal judgment.

Synthesizing these positions, the present study proposes that Shariah-informed principles for autonomous agents in the Metaverse should rest on three components: first, a clear denial of independent moral or legal personhood to the agent itself; second, a chain of traceable human responsibility running from the end user through the platform operator to the original developer; and third, a Shariah-compliance layer, similar to the Shariah Supervisory Board and AI Governance Committee model proposed by Zafar and Ali (2025) for Islamic financial institutions, adapted to review autonomous agents deployed in virtual environments for compliance with maqasid-based principles before and during deployment. This yields Proposition 1: because an autonomous agent lacks

the capacities that taklif requires, responsibility for its acts must remain traceable to an identifiable human or institutional actor.

3.2. Virtual Governance in the Metaverse from a Shariah Perspective

The secular governance literature reviewed in this study shows that the Metaverse currently lacks a unified jurisdiction, that enforcement mechanisms are weak, and that many forms of virtual misconduct fall into legal gray zones (Kourtesis, 2024; Qin et al., 2025; Tyagi, 2025). This study finds that these gaps are not simply obstacles that a Shariah framework must work around, but are themselves opportunities for maqasid al-Shariah to contribute a governance logic that does not depend on territorial jurisdiction. Because the maqasid are organized around the protection of universal human interests rather than around geographic boundaries, they can in principle be applied consistently to any Muslim user or Muslim-oriented platform regardless of which country's servers or courts happen to be involved, an approach broadly consistent with the maqasid-based regulatory model that Qutieshat et al. (2026) propose for AI governance across Muslim-majority states more generally.

Applying this logic, the study maps each maqsad onto a corresponding governance concern already documented in the secular literature, as summarized in Table 1: protection of intellect corresponds to safeguards against misinformation and manipulation inside immersive environments; protection of life corresponds to safeguards against psychological harm, harassment, and avatar-based abuse of the kind discussed by Qin et al. (2025) and Kourtesis (2024); protection of property corresponds to Shariah-compliant rules for virtual assets and digital transactions; protection of lineage and dignity corresponds to safeguards for family relationships, modesty norms, and the treatment of minors in virtual spaces; and protection of religion corresponds to ensuring that virtual environments accommodate rather than obstruct religious practice and identity. This mapping offers platform designers and regulators a maqasid-based checklist that can be applied even in the absence of a single governing jurisdiction, and it is consistent with the uneven but partly Islamically informed data protection regimes that Gorian and Osman (2024) describe in Saudi Arabia and the United Arab Emirates. This yields Proposition 2: because the maqasid attach to protected interests rather than to territory, they can serve as a jurisdiction-independent basis for reviewing virtual platforms.

Table 1. Mapping Maqasid al-Shariah to Metaverse Governance Concerns

Maqasid (Objective)	**Metaverse Governance Concern**	**Illustrative Governance Measure**
Protection of Religion (Hifz al-Din)	Accommodating religious practice and identity in virtual spaces	Virtual prayer spaces, halal content filters, safeguards against religious misinformation
Protection of Life (Hifz al-Nafs)	Psychological harm, harassment, and avatar-based abuse	Reporting mechanisms, avatar safety zones, wellbeing monitoring
Protection of Intellect (Hifz al-Aql)	Misinformation, manipulation, and addictive design	Digital literacy tools, explicability and transparency requirements, usage safeguards
Protection of Lineage and Dignity (Hifz al-Nasl)	Family relationships, modesty, and protection of minors	Age verification, modesty-respecting avatar design, parental oversight tools
Protection of Property (Hifz al-Mal)	Virtual assets, digital transactions, and virtual economies	Shariah-compliant contract structures, riba-free transaction rules, dispute resolution

3.3. Moral Education in Immersive Virtual Environments

The moral education literature reviewed in this study confirms that immersive virtual environments can support meaningful moral development, but only under specific conditions. Harrison (2024) identifies expansive, immersive, embodied, and autonomous learning as genuine opportunities for character education, while Rueda and Lara (2020) show that carefully designed perspective-taking scenarios can enhance empathic response. However, Kurian and Hajir (2024) caution that empathy-focused virtual reality experiences can just as easily produce shallow or even exploitative engagement with others' suffering if the underlying design choices are not scrutinized, and Dunivan et al. (2024) find that moral outcomes extend beyond empathy into compassion and moral reasoning in ways that are not yet fully understood. Kamalov et al. (2023) situate these specific findings within the broader observation that AI-driven educational tools generally cannot be adopted safely without deliberate institutional guardrails, a conclusion that reinforces the case for explicit safeguards in AI-mediated moral education rather than passive reliance on the technology's own design.

Read through a Shariah lens, these findings suggest that immersive moral education is compatible with the Islamic concept of tarbiyah, or holistic character formation, only when the same maqasid-based safeguards identified above are applied to educational content specifically: protecting the intellect from manipulative design, protecting dignity from exploitative or objectifying scenarios, and protecting the psychological wellbeing that the secular literature increasingly treats as inseparable from physical safety. Where these safeguards are absent, the same immersive features that make virtual reality attractive for moral education, its vividness and emotional intensity, become precisely the features that make manipulation and desensitization possible. This yields Proposition 3: immersive moral education is consistent with tarbiyah only when its content and design are reviewed against the same maqasid-based safeguards.

3.4. Discussion

Taken together, the three propositions of this study extend existing Islamic legal scholarship on artificial intelligence by applying it specifically to the Metaverse rather than to AI in general, and by connecting it explicitly to secular governance and moral education research that has so far developed largely in isolation from Islamic legal debates. The conclusion that autonomous agents cannot bear full taklif but can be assigned a bounded, functional capacity is consistent with the broader consensus documented in the Islamic legal literature on AI personhood (Hamsin et al., 2023; Sitiris & Busari, 2024; Yunos et al., 2024), and it resolves a practical tension: without any legal status at all, an autonomous agent's actions would be difficult to regulate coherently, whereas full legal personhood would inappropriately dissolve human accountability. This is the core of Proposition 1: a bounded functional status gives regulators something to work with, while a traceable chain of human responsibility prevents accountability from being displaced onto the agent. The dual Shariah-technical governance model proposed by Zafar and Ali (2025) for Islamic finance, echoed in the maqasid-based regulatory proposals of Najib et al. (2025) and Qutieshat et al. (2026), appears sufficiently general to be adapted for Metaverse platforms, suggesting that Islamic financial institutions that have already begun building generative AI governance structures may be well positioned to extend the same institutional expertise to Metaverse-based products and services. Proposition 1 therefore has an institutional reading as well: the Shariah Supervisory Board and AI governance committee model already exists in practice and needs adaptation rather than invention.

The comparison between maqasid al-Shariah and the documented governance gaps in the Metaverse (Kourtesis, 2024; Qin et al., 2025; Tyagi, 2025) suggests that a maqasid-based approach has a specific comparative advantage in exactly the situation

where secular governance struggles most: the absence of a single jurisdiction. Because the maqasid attach to the interests being protected rather than to territorial boundaries, a maqasid-based framework can, in principle, travel with the Muslim user or the Muslim-oriented platform across borders in a way that national law cannot, a pattern already visible in the partial, jurisdiction-specific efforts documented by Gorian and Osman (2024). This does not resolve the enforcement problem entirely, since Shariah compliance still depends on institutional will and community accountability rather than state coercion, but it does offer a normatively coherent standard that Muslim-majority jurisdictions, Islamic financial institutions, and voluntary industry bodies could adopt even before formal international agreements on Metaverse jurisdiction are reached. This is the practical meaning of Proposition 2, and it is why the mapping in Table 1 is offered as a review checklist rather than a legal code. It is also worth noting that the maqasid-based approach favored here is not the only credible Islamic ethical option: the trusteeship framework advanced by Ali et al. (2025) and the pluralist benchmarking approach proposed by Elmahjub (2023) offer alternative normative vocabularies that emphasize spiritual stewardship over practical governance mechanics, and future Shariah-based frameworks for the Metaverse may ultimately need to draw on all three traditions rather than treating maqasid reasoning as sufficient on its own, a concern echoed in Al-Momani's (2025) caution about the limits of AI-assisted legal reasoning.

The moral education findings carry a similarly practical implication. The secular literature's own internal debate, between optimistic accounts of virtual empathy training (Rueda & Lara, 2020) and more skeptical accounts of its limits (Kurian & Hajir, 2024), maps closely onto long-standing Islamic educational concerns about the difference between emotional stimulation and genuine character formation (tarbiyah). This suggests that Islamic educators do not need to import an entirely new pedagogical theory to engage with virtual reality; existing tarbiyah-based criteria for judging whether an experience builds durable virtue or merely produces a passing emotional reaction can be applied directly to immersive AI-mediated content, provided that educators and platform designers actively evaluate content against those criteria rather than assuming that immersion alone produces moral benefit, a caution that aligns with Kamalov et al.'s (2023) broader call for deliberate guardrails around AI in education and Alif Auladi and Muwahid's (2025) emphasis on distributed moral responsibility for digital trust. These points support Proposition 3: the value of immersive moral education depends on a deliberate review of its content and design against tarbiyah-based criteria.

This study has several limitations. It relies entirely on secondary literature and does not include empirical data from actual Metaverse platforms, Muslim users, or AI

systems, so its conclusions describe a normative and conceptual framework rather than a tested intervention. Propositions 1 to 3 should therefore be read as claims open to refinement, not as settled rulings. The literature on Shariah and AI personhood is still young and largely produced within a small number of Muslim-majority jurisdictions, particularly Malaysia, Indonesia, and the Gulf states, which may limit the diversity of juristic opinion represented. Future research could test the proposed maqasid-based governance checklist against real Metaverse platforms and AI agents, conduct empirical studies of Muslim users' moral experiences in immersive environments, and pursue comparative fiqh analysis across different schools of Islamic law, as well as across the maqasid, trusteeship, and pluralist ethical traditions reviewed here, to determine whether the conclusions reached about legal capacity and liability hold consistently across juristic and ethical traditions.

4. Conclusion

The main contribution of this study is an integrated, maqasid-based conceptual framework that links the legal status of autonomous agents, the governance of virtual space, and moral education in the Metaverse, three areas previously examined separately. The framework rests on three propositions. First, autonomous agents should be denied independent moral and legal personhood but assigned a bounded, functional capacity that keeps developers, operators, and institutions accountable. Second, the five maqasid offer a jurisdiction-independent checklist for the governance gaps documented in secular scholarship. Third, immersive moral education is compatible with tarbiyah only under explicit safeguards protecting intellect, dignity, and psychological wellbeing.

These propositions are conceptual and have not been tested empirically, and the maqasid approach is one credible lens alongside trusteeship and pluralist ethics. Even so, Islamic financial institutions, religious education providers, and Muslim-oriented platforms can adapt the Shariah Supervisory Board and AI governance committee model to review autonomous agents, virtual transactions, and educational content before they reach Muslim users. Future research should test the framework on real platforms and with Muslim users.

References

Al-Momani, I. M. (2025). Ethical challenges for using artificial intelligence in understanding Islamic jurisprudence. *Salud, Ciencia y Tecnologia - Serie de Conferencias*, *4*, Article 1519. https://doi.org/10.56294/sctconf20251519

Alamsyah, M. F., Yayuli, & Remanda, A. (2026). Ethics of the use of artificial intelligence (AI) in the paradigm of Islamic law. *Sicopus: Solo International Collaboration and Publication of Social Sciences and Humanities, 4*(1), 91-104. https://doi.org/10.61455/sicopus.v4i01.393

Ali, F., Bouzoubaa, K., Gelli, F., Hamzi, B., & Khan, S. (2025). Islamic ethics and AI: An evaluation of existing approaches to AI using trusteeship ethics. *Philosophy & Technology, 38*(3), 1-34. https://doi.org/10.1007/s13347-025-00922-4

Alif Auladi, G., & Muwahid, F. (2025). Cybersecurity and moral responsibility: A philosophical-Islamic approach to digital trust. *TechComp Innovations: Journal of Computer Science and Technology, 2*(1), 53-65. https://doi.org/10.70063/techcompinnovations.v2i1.93

Dunivan, D. W., Mann, P., Collins, D., & Wittmer, D. P. (2024). Expanding the empirical study of virtual reality beyond empathy to compassion, moral reasoning, and moral foundations. *Frontiers in Psychology, 15*, Article 1402754. https://doi.org/10.3389/fpsyg.2024.1402754

Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy & Technology, 36*(4), Article 73. https://doi.org/10.1007/s13347-023-00668-x

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.8cd550d1

Gorian, E., & Osman, N. D. (2024). Digital ethics of artificial intelligence (AI) in Saudi Arabia and United Arab Emirates. *Malaysian Journal of Syariah and Law, 12*(3), 583-597. https://doi.org/10.33102/mjsl.vol12no3.798

Hamsin, M. K., Anggriawan, R., & Jiatrahman, F. (2023). Unveiling ethical implications: AI robot accountability in Islamic context. *Jurnal Media Hukum, 30*(2), 117-135. https://doi.org/10.18196/jmh.v30i2.18524

Harrison, T. (2024). Virtual reality and character education: Learning opportunities and risks. *Journal of Moral Education, 53*(2), 219-239. https://doi.org/10.1080/03057240.2023.2206553

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), Article 12451. https://doi.org/10.3390/su151612451

Kourtesis, P. (2024). A comprehensive review of multimodal XR applications, risks, and ethical challenges in the Metaverse. *Multimodal Technologies and Interaction, 8*(11), Article 98. https://doi.org/10.3390/mti8110098

Kurian, N., & Hajir, B. (2024). Should we stop talking about empathy? Virtual reality, refugee children and the ethics of technology-mediated moral education. *Journal of Moral Education, 53*(4), 699-716. https://doi.org/10.1080/03057240.2023.2300503

Mohammed, T. A. S. (2024). A scientometric study of Maqasid al-shariah research: Trending issues, hotspot research, and co-citation analysis. *Frontiers in Research Metrics and Analytics, 9*, Article 1439407. https://doi.org/10.3389/frma.2024.1439407

Munshi, A. A., AlSabban, W. H., Farag, A. T., Rakha, O. E., Al Sallab, A., & Alotaibi, M. (2022). Automated Islamic jurisprudential legal opinions generation using artificial intelligence. *Pertanika Journal of Science & Technology, 30*(2), 1135-1156. https://doi.org/10.47836/pjst.30.2.16

Najib, N. M., Basarud-din, S. K., & Fazial, F. (2025). Artificial intelligence (AI) in Islamic finance: A Maqasid al-Shariah perspective. *International Journal of Law, Government and Communication, 10*(40), 41-50.

Nawi, A., Khamis, N. Y., Mohd Yaakob, M. F., Samuri, M. A. A., & Zakaria, G. A. N. (2023). Exploring opportunities and risks of artificial intelligence research for Islamic ethical guidelines. *Afkar: Jurnal Akidah dan Pemikiran Islam, 25*(2), 1-34. https://doi.org/10.22452/afkar.vol25no2.1

Qin, H. X., Wang, Y., & Hui, P. (2025). Identity, crimes, and law enforcement in the Metaverse. *Humanities and Social Sciences Communications, 12*(1), Article 194. https://doi.org/10.1057/s41599-024-04266-w

Qutieshat, E., Al Adwan, M., & Khater, M. (2026). Artificial intelligence in Islamic law: Ethics, governance, and accountability. *Journal of Sustainable Development and Regulatory Issues, 4*(2). https://doi.org/10.53955/jsderi.v4i2.236

Rueda, J., & Lara, F. (2020). Virtual reality and empathy enhancement: Ethical aspects. *Frontiers in Robotics and AI, 7*, Article 506984. https://doi.org/10.3389/frobt.2020.506984

Sitiris, M., & Busari, S. A. (2024). The legal capacity (al-ahliyyah) of artificial intelligence from an Islamic jurisprudential perspective. *Malaysian Journal of Syariah and Law, 12*(1), 31-42. https://doi.org/10.33102/mjsl.vol12no1.453

Sudirman, S., Sutiah, S., & Supriyono, S. (2025). Islamic law in the era of artificial intelligence: A systematic literature review. *International Journal of Engineering, Science and Information Technology, 5*(4), 321-328. https://doi.org/10.52088/ijesty.v5i4.1072

Tyagi, A. (2025). *Regulating the Metaverse: Legal challenges in virtual worlds and digital assets* [Working paper]. SSRN. https://doi.org/10.2139/ssrn.5321877

Yunos, A. S., Wan Saifuddin, W. M. S., & Hamdan, M. N. (2024). The concept of al-'ahliyyah and takyif fiqhiyy for autonomous mobile robots according to syariah. *Al-Qanatir: International Journal of Islamic Studies, 33*(7), 14-32. https://al-qanatir.com/aq/article/view/1035

Zafar, M. B., & Ali, H. (2025). *Shariah governance standard on generative AI for Islamic financial institutions* [SSRN Scholarly Paper]. SSRN. https://doi.org/10.2139/ssrn.5143165