

METAPHYSICAL FOUNDATIONS OF AI ETHICS: A THOMISTIC-ISLAMIC DIALOGUE ON HUMAN DIGNITY AND MORAL AGENCY IN THE DIGITAL AGE

Patricius Neonnub¹, Octovianus Kosat²

^{1,2}Fakultas Filsafat, Universitas Katolik Widya Mandira Kupang, Indonesia

Email: patrisn@yahoo.com¹, kosatkote11@gmail.com²

ISBN: 978-623-97987-1-0

Received: 01 October 2025

Accepted: 25 October 2025

Published: 21 January 2026

Abstract:

This paper explores the ontological foundations of moral agency in the context of Artificial Intelligence (AI) by engaging two metaphysical traditions: Thomistic natural law and Islamic metaphysical ethics. While contemporary AI ethics is often shaped by secular frameworks such as utilitarianism and deontology, these approaches frequently lack the metaphysical depth needed to address fundamental questions of human dignity, intentionality, and responsibility. Drawing from *actus essendi*, *lex aeterna*, *fitrah*, and *'aql*, the study demonstrates that both traditions affirm moral agency as rooted in the human being's participation in a transcendent moral order. In contrast, AI, as a non-rational artifact lacking teleology and metaphysical substance, cannot be considered a moral agent. Using a comparative-philosophical method, the paper proposes a transcultural ethical framework that re-centers AI governance on human dignity, ecological responsibility, and the common good. This framework contributes to global sustainability discourse by integrating Islamic knowledge and classical philosophy to guide ethical innovation in the digital age.

Keywords: *AI Ethics, Moral Agency, Islamic Metaphysics, Thomism, Global Sustainability*

Abstrak:

Makalah ini mengeksplorasi dasar ontologis dari agensi moral dalam konteks kecerdasan buatan (AI) melalui dialog antara dua tradisi metafisika: hukum kodrat Thomistik dan etika metafisika Islam. Di tengah dominasi paradigma etika AI sekuler seperti utilitarianisme dan deontologi, pendekatan ini menawarkan kerangka kerja yang lebih dalam dan transendental. Dengan menganalisis konsep seperti *actus essendi*, *lex aeterna*, *fitrah*, dan *'aql*, studi ini menunjukkan bahwa kedua tradisi menolak pemberian status agensi moral pada AI karena ketiadaan substansi rasional, teleologi, dan partisipasi dalam tatanan moral ilahi. Melalui pendekatan filosofis-komparatif dan hermeneutik, artikel ini merumuskan sebuah kerangka etika transkultural yang menekankan martabat manusia, tanggung jawab ekologis, dan keadilan sosial. Kontribusi ini diharapkan memperkaya wacana etika global dan mendorong pengembangan teknologi yang berorientasi pada kemaslahatan dan keberlanjutan.

Kata Kunci: *AI Etika, Agensi Moral, Metafisika Islam, Thomisme, Keberlanjutan Global*

INTRODUCTION

Artificial Intelligence (AI) has emerged as a transformative force in the digital age, offering unprecedented possibilities for advancing global sustainability. However, it also presents significant challenges to human dignity, moral responsibility, and social justice. In the context of accelerating technological globalization, the imperative for ethical frameworks that are both philosophically rigorous and culturally inclusive has never been more pressing. Despite growing interest in AI ethics, current discourse remains shaped mainly by Western secular paradigms—most notably utilitarianism, deontology, and virtue ethics. While each offers valuable normative insights, these approaches often lack sufficient metaphysical grounding to address more profound questions concerning intentionality, moral agency, and the ontological status of the human person. (Floridi & Cowls, 2019; Jobin et al., 2019). This methodological reductionism risks confining ethics to a set of procedural or technocratic norms, thereby neglecting the profound philosophical and spiritual dimensions that underpin human action and accountability.

In response, recent scholarly efforts have begun to integrate religious and transcendental perspectives into the discourse. For instance, Elmahjub (2023a, 2023b) introduces an Islamic pluralist benchmarking model for AI ethics, while Laracy, Kirova, and Ku (2025; 2025) draw on Catholic social teaching to propose a complementary ethical framework. The *I'timāni* approach, developed by Elmahjub et al. (2025), grounds ethical inquiry in the Qur'anic concept of *amanah* (divine trusteeship), offering a corrective to models rooted exclusively in consequentialist reasoning. Likewise, Rueda (2025a, 2025 b) critically interrogates the possibility of attributing dignity or moral agency to AI systems, cautioning against conflating functional performance with genuine moral responsibility. Within Islamic scholarship, nascent studies in the theology of technology have begun to explore how technological systems intersect with human existence within a broader metaphysical horizon (MDPI, 2023). Nevertheless, much of this literature remains confined to normative or theological analysis, with limited engagement in the ontological foundations that underlie moral agency itself.

This study seeks to address that gap by initiating a structured dialogue between two rich and underexplored metaphysical traditions: Thomistic natural law and Islamic metaphysical ethics. Thomas Aquinas's notions of *lex aeterna* (eternal law) and *actus essendi* (act of being) provide a realist ontological account of morality, rooted in participation in divine being. (Aquinas, 2012; E. Gilson, 1940) Islamic philosophy, by contrast, contributes complementary concepts such as *fitrah* (primordial human nature), *'aql* (intellect), and *maqāṣid al-sharī'ah* (higher objectives of divine law), which together construct a theocentric view of human purpose and ethical responsibility. (Adamson, 2016). Despite their shared metaphysical depth, these traditions have rarely been brought into direct conversation—particularly in relation to AI ethics.

This paper argues that such a dialogue can produce a transcultural, metaphysically grounded framework for AI ethics—one that reaffirms human uniqueness, moral accountability, and orientation toward the transcendent good. The originality of this study lies in three key contributions:

1. It presents a comparative analysis of the ontological assumptions underlying moral agency in Thomistic and Islamic traditions.
2. It demonstrates how both traditions reject the attribution of moral agency to

AI systems, based on their lack of rational substance and ontological participation.

3. It proposes an integrative ethical framework aligned with the conference theme—"Islamic Knowledge and Innovation for Global Sustainability"—by placing human dignity, ecological responsibility, and social justice at the core of technological governance.

The remainder of the paper proceeds as follows: after outlining the methodological approach, the study examines the ontological foundations of ethics in Thomism and Islamic philosophy, identifies key areas of convergence and divergence, and considers their implications for AI ethics and policy within a globalized technological landscape.

RESEARCH METHOD

The methodological foundation of this study is grounded in qualitative inquiry, informed by the traditions of comparative philosophy and hermeneutics. The research emerged from a fundamental concern: that ethical responses to artificial intelligence (AI) often lack engagement with more profound metaphysical questions surrounding moral agency, human dignity, and the ontological status of persons. In seeking to address this gap, the study turned to two profound philosophical traditions—Thomistic natural law and Islamic metaphysical ethics—to construct a conceptual framework for AI ethics that is both transcultural and ontologically grounded.

Rather than approaching these traditions through an empirical or data-driven model, the research unfolded as a comparative textual exploration. The methodological decision to conduct a philosophical case study was driven by the distinctiveness and metaphysical coherence of the two traditions under consideration. Both Thomism and Islamic philosophy provide richly developed accounts of the human person, moral reasoning, and ethical purpose. By examining these traditions side-by-side, the study aimed to uncover not only points of convergence but also meaningful divergences that could inform a nuanced ethical vision for AI governance.

The analytical process began with immersive engagement with foundational philosophical texts, including *Summa Theologica* and *Summa contra Gentiles* by Thomas Aquinas, with particular attention to his teachings on *lex aeterna* (eternal law), *lex naturalis* (natural law), and the metaphysical structure of the human soul. Parallel to this, the study drew upon seminal works in Islamic metaphysics, such as Ibn Sina's *al-Shifa'* and *al-Najat*, Al-Ghazali's *Ihya' Ulum al-Din*, and Mulla Sadra's *al-Asfar al-Arba'a*. These texts offered key insights into concepts such as *fitrah* (primordial nature), *'aql* (intellect), *wujūd* (existence), and *maqāṣid al-sharī'ah* (higher objectives of divine law).

Complementing these primary sources were several contemporary scholarly commentaries and recent contributions in the field of AI ethics, including the works of Elmahjub (2023), Rueda (2025), and Laracy et al. (2025), all of which attempt to engage religious traditions in modern ethical discourse. These materials provided context and contrast, allowing the philosophical insights from Aquinas and the Islamic thinkers to be situated within today's global conversation about technology and ethics.

The process of data collection and interpretation followed a three-stage hermeneutic approach. The first stage involved close reading, during which key

metaphysical and ethical themes were identified through multiple, careful readings of the texts. This was followed by textual exegesis, allowing the study to distill the philosophical assumptions embedded within each tradition. Finally, the material was subjected to thematic coding, through which recurring motifs—such as teleology, rationality, freedom, and moral accountability—were identified and organized for comparison.

These steps culminated in a structured analytical model. Initially, each tradition was studied on its own terms, allowing for an understanding of its internal logic and metaphysical coherence. This intra-traditional analysis was then followed by a comparative phase, in which points of intersection and divergence were mapped out in relation to key philosophical concerns. A comparative matrix was constructed to capture these insights in both conceptual and visual terms. Finally, the study entered a phase of integrative synthesis, weaving the insights from Thomistic and Islamic thought into a unified framework that could speak meaningfully to the ethical dilemmas posed by AI.

Ensuring the philosophical integrity of this process required several strategies. The study employed triangulation, cross-referencing interpretations across primary texts, secondary commentaries, and AI ethics literature. (Sheikhalizadeh & Piralaiy, 2017). Additionally, a process of peer debriefing was undertaken, with emerging findings discussed among scholars of philosophy, ethics, and technology to surface assumptions and refine conceptual clarity. Throughout, care was taken to maintain thick description, offering richly textured accounts of the philosophical traditions in their intellectual and spiritual contexts. (Levina et al., 2016).

In sum, this study does not merely juxtapose two philosophical systems. It invites them into dialogue—a conversation not for comparison's sake alone, but in pursuit of a deeper, metaphysically grounded, and spiritually resonant framework for addressing one of the most urgent ethical challenges of our time: how we, as humans, ought to design, deploy, and govern intelligent technologies in a way that safeguards dignity, justice, and the common good.

RESULTS AND DISCUSSION

This comparative analysis reveals a deep resonance between Thomistic and Islamic metaphysical traditions regarding the ontological foundations of ethics. Both assert that moral agency is inherently tied to the metaphysical constitution of the human person—something that Artificial Intelligence (AI), by its very nature, does not possess. The following subsections present the key findings from each tradition, synthesize them in a comparative framework, and discuss their implications for the ethics and governance of AI in a global context.

1. Ontological Foundations of Ethics in Thomistic Philosophy

Thomas Aquinas's ethical system is grounded in a hierarchical metaphysics of being that begins with the concept of *lex aeterna* (eternal law)—the rational order of the universe as emanating from the divine intellect. (Davies, 1992; É. Gilson, 1994). The human person, as a rational creature, participates in this eternal law through *lex naturalis* (natural law), which is discernible via reason (Porter, 2005). Moral agency, in this framework, is inseparable from *actus essendi* (the act of being), which signifies a creature's metaphysical participation in God's being.

Aquinas identifies the human being as a rational substance endowed with

intellect (*intellectus*) and free will (*voluntas*), oriented toward the attainment of the good. This teleological orientation culminates in *beatitudo*—the end of human existence, understood as union with God. (MacIntyre, 2016). Importantly, dignity and moral responsibility are not social constructs but metaphysical realities derived from this participation in the divine order.

Artificial Intelligence, lacking *actus essendi*, a rational soul, or a final cause (*telos*), cannot partake in this metaphysical order. While it may simulate cognitive functions or ethical decision-making, it lacks ontological depth and moral intentionality. Therefore, AI cannot be considered a moral agent but remains an artifact—an extension of human reason to be governed under the principles of natural law and directed toward the *bonum commune* (common good).

2. Ontological Foundations of Ethics in Islamic Metaphysical Thought

Islamic metaphysical ethics, as articulated by classical thinkers such as Ibn Sina, Al-Ghazali, and Mulla Sadra, similarly grounds moral agency in the metaphysical constitution of the human soul. Central to this framework is the concept of *fitrah*—the innate, primordial disposition toward truth, goodness, and divine recognition (Nasr, 2007). This *fitrah* is actualized through ‘*aql* (intellect) and *ikhtiyār* (free will), which together form the epistemic and volitional basis of ethical responsibility.

Mulla Sadra’s metaphysical principle of *aṣālat al-wujūd* (primacy of existence) introduces a dynamic ontology in which all beings participate in existence to varying degrees (Rahman, 1975). The human soul, as an immaterial and evolving substance, journeys through existential intensification (*tashkīk al-wujūd*) toward *kamāl* (perfection), culminating in divine proximity. Moral action, therefore, is meaningful only within this graded and teleologically oriented metaphysical landscape.

AI, by contrast, is a manufactured artifact (*makhlūq maṣnūʿ*) devoid of *fitrah*, ‘*aql*, and the spiritual telos that defines human personhood. It cannot engage in the moral struggle (*jihād al-naḥs*) that characterizes ethical formation, nor can it ascend the ontological hierarchy toward perfection. Accordingly, Islamic ethics preclude AI from being considered a moral subject. It is to be understood as a tool entrusted to humans under the doctrine of *amanah* (divine trusteeship). It must be regulated in accordance with *maqāṣid al-sharīʿah*—the higher objectives of divine law, including the protection of life, intellect, faith, lineage, and property (Auda, 2008).

3. Convergences and Divergences: A Comparative Matrix

To synthesize the findings, the table below presents a comparative matrix that outlines the central points of convergence and divergence between Thomistic and Islamic metaphysical ethics, highlighting their collective relevance for AI governance.

Table 1: Comparative Matrix of Thomistic and Islamic Metaphysical Ethics in Relation to AI

Aspect	Thomism (Thomas Aquinas)	Islamic Metaphysical Ethics	Implications for AI Ethics
Ontological	Participation in <i>actus</i>	Rooted in <i>fitrah</i> and	AI lacks metaphysical

Basis	<i>essendi</i> and <i>lex aeterna</i>	dynamic <i>wujūd</i>	participation; it is a tool, not a moral agent
Human Nature	Rational animal oriented toward the good	<i>Khalīfah</i> endowed with <i>‘aql</i> and <i>ikhtiyār</i>	Only humans possess moral accountability and spiritual teleology
Source of Moral Knowledge	Natural law discerned by reason, completed by revelation	<i>Fitrah</i> and <i>‘aql</i> guided by Qur’an and Sunnah	AI cannot access moral knowledge through reason or revelation
Teleology Purpose	<i>Beatitudo</i> —union with God	<i>Kamāl</i> —perfection of the soul and realization of <i>maqāṣid</i>	AI must serve human flourishing, not pursue autonomous ends
View of Technology	A tool to be governed by natural law and ordered to the common good	Artifact under <i>amanah</i> must align with divine purposes	AI governance is a human ethical responsibility
Human Dignity	Derived from rationality and participation in divine being	Inherent in <i>fitrah</i> and <i>khalīfah</i> status (Qur’an 17:70)	Dignity is exclusive to humans; AI cannot be a bearer of rights
Moral Status of AI	Artifact without <i>actus essendi</i> or rational soul	Artifact lacking <i>fitrah</i> or moral agency (<i>taklīf</i>)	AI is not a moral agent; ethical responsibility lies with its creators/users.

4. Analytical Discussion: Toward a Transcultural Framework for AI Ethics

The comparative findings yield three critical insights that contribute to the development of a transcultural and metaphysically grounded framework for AI ethics:

First, both traditions unequivocally root moral agency in ontological constitution, not in functional performance. This challenges popular functionalist models that suggest attributing rights or moral standing to AI systems based on behavioral sophistication or simulated rationality (Rueda, 2025b). Without a metaphysical grounding—*actus essendi* or *fitrah*—AI remains ontologically inert, disqualifying it from moral subjectivity.

Second, the convergences between Thomism and Islamic metaphysics are not superficial but substantive. Both affirm an objective moral order, the centrality of reason, and a teleological conception of human existence. Thomism provides a framework of natural law accessible through reason, whereas Islamic ethics emphasizes divine revelation and the concept of trusteeship. Together, these traditions form a philosophically resilient and spiritually grounded basis for resisting technocratic or utilitarian reductionism.

Third, the points of divergence—such as the emphasis on revelation in Islam and the rational discernibility of natural law in Thomism—are not contradictory but

mutually enriching. They offer a balanced foundation for ethical reasoning that is both accessible to secular discourse and resonant with religious traditions. This is crucial for inclusive AI governance in culturally plural societies.

5. Implications for Global Sustainability and Digital Transformation

The findings have direct implications for global efforts to ensure ethical and sustainable digital transformation, particularly in light of the United Nations Sustainable Development Goals (SDGs) and the conference theme of *Islamic Knowledge and Innovation for Global Sustainability*:

- Human-Centered Innovation: AI systems should be designed in service of the *bonum commune* (common good) and *maṣlaḥah* (public benefit). Ethical innovation must prioritize human flourishing, not technological autonomy.
- Guardianship of Creation: The Islamic concept of *khalīfah* and the Thomistic principle of natural law demand that AI be deployed responsibly in environmental contexts—e.g., climate modeling, ecological preservation, and sustainable agriculture.
- Safeguarding Human Dignity: As algorithmic systems increasingly impact rights, freedoms, and identities, metaphysical ethics rooted in dignity and moral agency offer critical safeguards. Frameworks informed by *maqāṣid al-sharīʿah* and natural law provide normative tools to evaluate the ethical limits of AI applications.

CONCLUSION

This study presents a comparative philosophical inquiry into the ontological foundations of moral agency within Thomistic natural law theory and Islamic metaphysical ethics, addressing a critical gap in contemporary AI ethics: the absence of a metaphysically grounded framework capable of responding to the ethical and existential challenges posed by advanced artificial intelligence.

Amid rapid digital transformation and the growing integration of AI into social, political, and economic systems, ethical debates often remain confined to procedural, functionalist, or utilitarian paradigms. These approaches, while useful for regulatory purposes, often fail to address the more profound metaphysical questions concerning what it means to be a moral agent and what ontological conditions make dignity, responsibility, and justice possible. This paper has argued that such questions cannot be meaningfully answered without returning to realist metaphysical traditions that affirm the intrinsic worth and teleological orientation of the human person.

Through a detailed comparative analysis, this study has shown that both Thomistic and Islamic metaphysical frameworks locate moral agency in the ontological structure of human existence—in *actus essendi* and *lex aeterna* in the Thomistic tradition, and in *fitrah*, *ʿaql*, and *amanah* in Islamic thought. AI, by contrast, lacks the metaphysical depth required for genuine agency: it does not possess a rational soul, a final cause, or any participation in divine or moral order. It may imitate human behavior, but it cannot bear moral responsibility. Accordingly, both traditions categorically preclude the attribution of moral status or dignity to artificial systems.

The original contribution of this research lies in the development of a transcultural, metaphysically grounded framework for AI ethics—one that integrates the realist ontologies of Thomism and Islamic philosophy. This

framework surpasses the limitations of secular ethical models by grounding technological governance in a vision of the human being as an agent endowed with reason, purpose, and accountability before a transcendent moral order. It affirms that AI ethics must begin not with machines, but with a renewed understanding of the human person.

Moreover, the implications of this study extend directly to the theme of "Islamic Knowledge and Innovation for Global Sustainability." In a world increasingly shaped by technological innovation, the recovery of metaphysical wisdom is not an academic luxury but a practical necessity. By re-centering ethics on human dignity and moral responsibility, this framework offers guidance for AI development that advances the common good, protects the environment, and promotes justice—core concerns in both Islamic ethical thought and global sustainability discourse.

Future research may extend this dialogue by engaging other metaphysical traditions, such as Confucian, Hindu, or Buddhist ethics, further enriching the global moral conversation on AI. Additionally, the framework proposed here can be applied to concrete domains, such as algorithmic bias, AI in healthcare, or environmental technology, to generate actionable ethical guidelines rooted in this integrated metaphysical perspective.

Ultimately, this study issues a call to scholars, developers, and policymakers to pursue technological innovation that is not only intelligent but also wise—innovation that is informed by philosophical depth, spiritual vision, and an unwavering commitment to human flourishing in an age of machines.

REFERENCES

1. Journal

- Elmahjub, E. (2023a). Artificial Intelligence in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI. *Philosophy & Technology*, 36(4), 1–21. <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>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Laracy, D., Finnis, J., & Chandler, M. (2025). Human Dignity and the Ethics of Artificial Intelligence: A Framework Based on Catholic Social Teaching. *Studies in Christian Ethics*.
- Laracy, N., Kirova, T., & Ku, J. (2025). Artificial intelligence and the common good: A framework rooted in Catholic social teaching. *Technology and Society*.
- MDPI (Ed.). (2023). Special Issue: The Theology of Technology in the Anthropocene. In *Religions* (Vol. 14, Issue 5).
- Rueda, J. (2025a). Can we coherently attribute dignity to artificial intelligence? *Science and Engineering Ethics*.
- Rueda, J. (2025b). The Problem with Dignity in AI Ethics. *Journal of Ethics and Information Technology*, 27(1), 34–52.
- Sheikhalizadeh, M., & Piralaizadeh, S. (2017). The role of documentary research in social science studies. *International Journal of Humanities and Cultural*

2. Book

- Adamson, P. (2016). *Philosophy in the Islamic World: A History of Philosophy Without Any Gaps, Volume 3*. Oxford University Press UK.
- Aquinas, T. (2012). *Summa Theologiae: Latin-English Edition* (revised by the A. I. for the S. of S. D. Dominican Fathers, Ed.). Emmaus Academic.
- Auda, J. (2008). *Maqasid al-Shariah as Philosophy of Islamic Law: A Systems Approach*. The International Institute of Islamic Thought.
- Davies, B. (1992). *The Thought of Thomas Aquinas*. Clarendon Press.
- Elmahjub, E. (2023b). Pluralist benchmarking for artificial intelligence ethics: An Islamic perspective. In *Ethics, Governance, and Policies in Artificial Intelligence* (pp. 255–273). Springer.
- Gilson, E. (1940). *The Christian Philosophy of St. Thomas Aquinas*. Random House.
- Gilson, É. (1994). *The Christian Philosophy of St. Thomas Aquinas*. University of Notre Dame Press.
- Levina, N., Vaast, E., & Williams, C. (2016). Bridging the digital divide: A qualitative framework for studying knowledge reuse. In *Routledge Companion to Management Information Systems* (pp. 345–360). Routledge.
- MacIntyre, A. (2016). *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning, and Narrative*. Cambridge University Press.
- Nasr, S. H. (2007). *The Essential Seyyed Hossein Nasr*. World Wisdom.
- Porter, J. (2005). *Nature as Reason: A Thomistic Theory of the Natural Law*. Wm. B. Eerdmans Publishing.
- Rahman, F. (1975). *The Philosophy of Mulla Sadra*. State University of New York Press.

About the Author:

Chief Researcher
Patricius Neonnub
Researcher Member
Octovianus Kosat