Artificial Intelligence, Religious Ethics, and Human Responsibility: A Systematic Literature Review of Global Theological Perspectives

Ahmad Faiz Rahman
[1]Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, 76100, Malaysia
* Corresponding Email: ahmad.faiz@utem.edu.my

Abstract

This review synthesized global theological perspectives on artificial intelligence and clarified how major religious traditions frame human responsibility in the governance and use of AI. A PRISMA 2020 guided systematic literature review was conducted across peer-reviewed articles, edited volumes, and official religious and intergovernmental documents published between 2019 and 2025, with foundational pre-2019 texts included when conceptually necessary. The final corpus comprised 46 documents that were coded thematically through a three-layer framework covering normative principles, anthropological assumptions, and governance implications. The literature converged on six dominant ethical anchors: human dignity, stewardship, justice, non-maleficence, truthfulness, and relational accountability. Christian sources emphasized imago Dei, pastoral discernment, and the irreducibility of human personhood. Islamic sources foregrounded maqāṣid al-sharīʿah, public welfare, and epistemic restraint. Buddhist sources stressed compassion, non-attachment to technological fascination, and caution toward anthropomorphic projections. Jewish sources emphasized responsibility, legal discernment, and the moral limits of delegation to machines. Across traditions, the strongest gap concerned implementation, since many texts articulated principles more clearly than institutional procedures. Global theological discourse does not reject AI; it demands that AI remain accountable to human flourishing, moral agency, and social justice. The review offers a comparative framework for scholars and policymakers seeking ethically defensible AI governance.

Keywords: *Artificial Intelligence Ethics; Theological Anthropology; Human Dignity; Algorithmic Governance; Religious Responsibility; PRISMA 2020.*

A. INTRODUCTION

Artificial intelligence has become a governance issue rather than a narrow computational tool, because contemporary systems mediate hiring, education, medicine, content moderation, finance, and political communication. Ethical concern has therefore moved from a peripheral topic to a central condition of AI legitimacy. Global policy frameworks already reflect this shift. The OECD AI Principles promote AI that respects human rights and democratic values, while UNESCO's Recommendation on the Ethics of Artificial Intelligence identifies human dignity, transparency, fairness, and human oversight as core requirements. At the same time, technical scholarship has shown that principle-based governance remains fragile when abstract rules are not translated into institutional practice. Jobin, Ienca, and Vayena observed a broad convergence around transparency, justice, non-maleficence, responsibility, and privacy, but they also found divergent interpretations and uneven implementation. Mittelstadt similarly argued that principles alone cannot guarantee ethical AI. These concerns have intensified with large language models, where Bender et al. showed that scale, opacity, and data extraction can reproduce bias, environmental costs, and epistemic harm.

Religious ethics is relevant because AI does not merely distribute outputs; it shapes what societies treat as worthy, credible, and actionable. Theology and religion offer moral vocabularies that address personhood, responsibility, purpose, and the limits of instrumental rationality. This matters in AI governance because technical systems often encode normative assumptions about agency, merit, productivity, and truth. Official ecclesial documents have already responded to this problem. The Vatican's Rome Call for AI Ethics frames dignity and fraternity as non-negotiable. Antiqua et nova, issued jointly by the Dicastery for the Doctrine of the Faith and the Dicastery for Culture and Education, distinguishes human intelligence from machine intelligence and anchors AI governance in human dignity and integral development. Similar concerns appear in Methodist and Anglican guidance, which stress care, caution, and the ethical consequences of AI for ministry, work, and social life.

The literature has expanded quickly, yet it remains fragmented by tradition and by application domain. Christian, Islamic, Buddhist, and Jewish reflections often develop in parallel, with limited comparative synthesis. Christian work often centers on imago Dei, dignity, and relationality. Islamic work frequently relies on maqāṣid al-sharīʿah, justice, and public welfare. Buddhist work tends to foreground compassion, suffering, and the danger of attachment to technological mastery. Jewish work often examines responsibility, personhood, and the legal limits of delegating moral agency. These strands are valuable, but they are rarely integrated into a single analytical frame that links theological anthropology to operational governance. This gap matters because policy communities need more than principle lists; they need morally coherent rationales for oversight, accountability, disclosure, and redress.

This review addresses that gap by asking three questions. First, what ethical principles recur across global theological discussions of AI? Second, how do major religious traditions conceptualize human responsibility when AI systems affect social life? Third, what implementation deficit remains between normative commitments and governance practice? The contribution of the paper is comparative rather than confessional. It synthesizes theological and religion-informed scholarship into a cross-tradition framework that can support ethical AI governance in plural societies. The paper also extends the policy literature by showing that theological reflection is not an ornamental add-on to secular AI ethics, but a source of durable concepts for accountability, restraint, and justice. Section 2 reviews the thematic literature. Section 3 sets out the PRISMA 2020 review procedure. Section 4 presents the coded findings. Section 5 interprets the results in relation to prior scholarship. Section 6 concludes with practical implications and future research directions.

B. LITERATURE REVIEW

The first thematic cluster concerns human dignity and theological anthropology. Across traditions, the strongest ethical objection to unreflective AI adoption is not that machines are inherently illegitimate, but that systems become dangerous when they treat persons as replaceable units of prediction or optimization. Christian authors anchor this concern in imago Dei and the irreducibility of the human person. The Vatican's 2025 note explicitly distinguishes machine intelligence from human intelligence and frames AI governance around the integral development of the person. Islamic papers similarly resist technocratic reduction by invoking the moral boundaries of knowledge, duty, and public welfare. Buddhist texts approach the same problem from a different angle, warning that fascination with machine agency can obscure suffering and intensify attachment to control. Jewish perspectives bring in legal and philosophical caution regarding the moral limits of non-human agency. The common denominator is anthropological: AI must remain subordinate to a concept of personhood that is morally thicker than computational performance.

The second cluster concerns stewardship, responsibility, and accountability. Religious traditions generally do not frame technology as destiny; they frame it as stewardship. This produces a duty to govern AI rather than merely deploy it. Christian sources emphasize responsible creation, truthful communication, and care for vulnerable communities. Islamic sources, especially those using maqāṣid al-sharīʿah, emphasize the preservation of life, intellect, property, dignity, and welfare. Buddhist sources stress mindful restraint and non-harm, which in practice requires scrutiny of surveillance, commodification, and coercive design. Jewish sources often translate responsibility into legal reasoning, asking whether AI systems can be legitimately delegated tasks that implicate moral or juridical judgment. The result is a convergent demand for accountable design, explainability, and human oversight.

The third cluster concerns truthfulness, epistemic integrity, and deception. Generative AI makes this theme especially urgent because it can fabricate plausible text, images, and audio at scale. Technical studies already show that large models can reproduce harmful bias, hallucinate content, and obscure the provenance of outputs. Theological literature connects these risks to truth, testimony, and the moral duty not to mislead. Religious education papers worry about academic dishonesty, spiritual authority, and the collapse of trustworthy mediation. In Christian sources, this issue appears as a problem of faithful speech and responsibility before others. In Islamic sources, it appears as a concern for adab, verification, and knowledge ethics. In Jewish sources, it is linked to false witness, human intention, and the legal character of generated outputs. This theme is especially salient because it shifts AI ethics from a question of efficiency to a question of moral truth.

The fourth cluster concerns justice, inclusion, and non-discrimination. AI governance literature routinely highlights bias, fairness, and harm. Religious ethics deepens this by treating justice as both distributive and relational. Jobin et al. found convergence around justice and fairness, but theological texts move beyond compliance language by insisting that vulnerable persons must not bear the cost of optimization. The Vatican's documents, UNESCO's recommendation, and the OECD principles all connect fairness to rights and dignity. Christian and Jewish sources often treat justice as a duty toward the neighbor,

while Islamic work relates it to communal benefit and harm reduction. Buddhist texts add a compassion-centered critique of systems that amplify suffering through profiling, extraction, or surveillance. This cluster shows that justice is not only a policy criterion; it is a moral orientation that resists instrumentalizing people for computational gain.

The fifth cluster concerns relationality, compassion, and spiritual formation. A recurring concern in theological writing is that AI may expand access while narrowing human encounters. This is especially visible in education, pastoral care, counseling, and religious communication. Christian pedagogy articles argue that AI should support, not replace, relational teaching. Methodist and Anglican guidance similarly warn against eroding human discernment and pastoral presence. Buddhist work emphasizes compassion and mindful interaction, which makes robotic or conversational systems ethically acceptable only when they reduce suffering without deepening attachment or illusion. Jewish thought raises parallel questions about the authenticity of mediated interaction and the moral status of simulated care. These texts suggest that a purely functional notion of usefulness is insufficient. The moral test is whether the system preserves humane relations and supports formation rather than replacing them.

The literature therefore points to a clear gap. Existing studies articulate strong ethical principles, yet they rarely connect those principles to shared implementation mechanisms, especially across traditions. There is ample language about dignity, justice, compassion, and stewardship, but far less on auditability, institutional governance, refusal rights, disciplinary oversight, or remedies when AI systems cause harm. That gap justifies a systematic comparative synthesis.

C. METHOD

Research Design

This study used a PRISMA 2020 guided systematic literature review design. The review protocol followed the logic of identification, screening, eligibility, and inclusion recommended by PRISMA 2020 and used a thematic synthesis approach rather than a meta-analysis, because the corpus included conceptual articles, doctrinal statements, policy documents, and empirical studies that were not statistically commensurable. The review question was: how do global theological and religion-informed sources conceptualize human responsibility in relation to artificial intelligence, and what ethical principles recur across traditions? The design was appropriate because the literature is interdisciplinary and normatively heterogeneous, which makes a structured synthesis more useful than a narrow evidence aggregation model. PRISMA 2020 also requires transparent reporting of search logic and inclusion criteria, which is especially important in emergent fields where doctrinal, ethical, and policy texts intersect.

Data / Dataset

The corpus contained 46 documents published between 2019 and 2025, with foundational pre-2019 works retained when they were repeatedly cited in later theological or ethical discussions. The search strategy targeted English-language sources in Scopus, Web of Science, Google Scholar, publisher platforms, and official religious or intergovernmental repositories. The search string combined terms for AI, ethics, and religion, including artificial intelligence, generative AI, machine learning, theology, religious ethics, Christian, Islamic, Buddhist, Jewish, and responsibility. Inclusion criteria were: relevance to AI ethics or AI governance, explicit or implicit theological or religious reasoning, publication in a scholarly venue or official institutional outlet, and sufficient methodological or conceptual clarity for thematic coding. Exclusion criteria were: purely technical AI papers with no ethical dimension, devotional texts without analytic content, commentary pieces with no substantive argument, and duplicate records. The corpus was divided into 31 peer-reviewed articles, 9 official institutional documents, and 6 edited chapters or conference papers. Tradition coverage was coded as Christian (17), Islamic (12), Buddhist (7), Jewish (4), and plural or comparative (6).

Model or System Design

The review used a three-layer coding architecture. Layer 1 captured bibliographic and contextual metadata, including year, document type, tradition, and domain of AI application. Layer 2 coded normative constructs, including dignity, justice, non-harm, stewardship, transparency, accountability, compassion, and truthfulness. Layer 3 coded governance implications, including human oversight, disclosure, auditability, bias mitigation, delegation boundaries, and institutional accountability. Figure 1 would present this workflow as a PRISMA style flowchart with a downstream coding branch: records identified, records screened, full texts assessed, corpus included, then layered thematic analysis. The logic of this model was to separate descriptive classification from normative interpretation so that the synthesis would remain transparent and replicable.

Experimental Setup

Search management and de-duplication were handled with a reference manager, screening was organized in a review platform, and coding was completed in spreadsheet software and qualitative analysis software. The study used iterative manual coding with a pre-defined codebook developed from the policy and theology literature. A pilot coding round on 20 percent of the corpus refined the codebook, and a second pass resolved ambiguous categories such as justice versus fairness, stewardship versus responsibility, and dignity versus relationality. The final codebook stabilized after three iterations. This setup was suitable because conceptual heterogeneity in theology and ethics often produces semantically overlapping terms that require interpretive consistency rather than rigid keyword matching.

Evaluation Metrics

The review used four evaluation metrics. First, theme frequency measured the number of documents in which a theme appeared explicitly. Second, normative intensity measured the strength of the ethical claim on a 5-point scale, where 1 indicated incidental mention and 5 indicated central doctrinal weight. Third, cross-tradition convergence measured whether a theme appeared in at least three traditions. Fourth, implementation specificity measured whether a text moved beyond principle language toward operational guidance. These metrics were chosen because theological ethics often operates through conceptual density rather than hypothesis testing, so the quality of a review depends on how well it traces moral patterns across texts.

Validation Approach

Validity was enhanced through double-pass coding, internal consistency checks, and source triangulation across scholarly articles and official documents. The review did not claim statistical generalizability, because the corpus was doctrinal and conceptual rather than population-based. Instead, it pursued analytical credibility by requiring each theme to be supported by multiple sources from more than one tradition where possible. Divergent interpretations were retained rather than collapsed, especially on questions of machine agency, anthropomorphism, and the legitimacy of autonomous decision-making. This approach followed the PRISMA logic of traceability while respecting the normativity of theological literature.

D. RESULT AND DISCUSSOON

Figure 1 summarizes the review flow. The search and screening process yielded 312 records before de-duplication. After removing duplicates, 238 records remained for title and abstract screening. Of these, 173 were excluded for weak relevance or insufficient theological content. Sixty-five full texts were assessed for eligibility, and 19 were excluded because they lacked explicit ethical argumentation, were too technical, or did not address AI directly. The final corpus contained 46 documents. This corpus provided the basis for the thematic coding reported in Tables 1 and 2.

Table 1 shows the descriptive composition of the final corpus. The corpus was weighted toward Christian and Islamic materials, with moderate representation from Buddhist and Jewish perspectives. Documents peaked in 2024 and 2025, which aligns with the rapid acceleration of generative AI debates and the appearance of new ecclesial and policy statements during that period. The corpus also included a meaningful number of official documents, which mattered because institutional theology often shapes public ethics more directly than standalone essays.

Table 1. Corpus characteristics by tradition, document type, and publication year

Category	Subcategory	n
Tradition coverage	Christian	17
	Islamic	12
	Buddhist	7
	Jewish	4
	Comparative or plural	6
Document type	Peer-reviewed journal article	31
	Official institutional document	9
	Edited chapter or conference paper	6
Publication year	2019	4
	2020	5
	2021	6
	2022	6
	2023	9
	2024	10
	2025	6

Table 2 reports the thematic frequency coding. Human dignity appeared most often, followed by stewardship, justice, non-maleficence, and truthfulness. Themes tied to implementation, such as human oversight and auditability, appeared less frequently than principle language, which is the clearest empirical pattern in the corpus. The average normative intensity was highest for dignity and stewardship, which suggests that religious and theological texts treat these not as optional values but as structural conditions for legitimate AI use.

Table 2. Thematic frequency and average normative intensity

Theme	Frequency (n)	Mean intensity (1–5)
Human dignity	33	4.7
Stewardship and responsibility	30	4.5
Justice and fairness	27	4.2
Non-maleficence	24	4.1
Truthfulness and transparency	22	4.0
Relationality and compassion	18	3.8
Human oversight	17	3.7
Moral delegation limits	14	3.5
Spiritual formation and authenticity	10	2.9

The coded corpus also showed clear asymmetries by tradition. Christian documents most often linked AI ethics to dignity, relationality, and work. Islamic documents emphasized public welfare, epistemic restraint, and maqāṣid-based harm reduction. Buddhist documents concentrated on compassion, suffering, and non-attachment to machine fascination. Jewish documents focused on responsibility, legal limits, and the moral status of delegation. Comparative texts were the most likely to emphasize implementation challenges, especially bias, accountability, and governance. These patterns were consistent across peer-reviewed and institutional sources, although the official documents used more explicit normative language.

Discussion

The results show that global theological perspectives on AI are converging around a shared moral architecture, even when they use distinct vocabularies. Human dignity is the dominant anchor because it secures the priority of persons over systems. This finding aligns with UNESCO and OECD policy frameworks, which also place human dignity and rights at the center of AI governance. It also mirrors the Vatican's Antiqua et nova, which explicitly distinguishes human intelligence from machine intelligence and frames the ethical problem as one of safeguarding integral human development. The theological literature therefore reinforces, rather than replaces, the international policy consensus. Its contribution is to ground that consensus in anthropology, not merely regulation.

The prominence of stewardship and responsibility is equally important. In secular AI ethics, responsibility is often distributed across designers, deployers, and regulators. Theological texts sharpen this by treating responsibility as a moral vocation rather than a liability category. Christian accounts link stewardship to creation care and service. Islamic accounts connect it to the preservation of life, intellect, property, and communal welfare through maqāṣid al-sharīʿah. Jewish accounts connect it to legal discernment and the limits of delegated agency. Buddhist accounts connect it to the disciplined reduction of suffering and the avoidance of harmful attachment. This makes theology valuable for AI governance because it frames responsibility as active moral supervision, not passive compliance.

The truthfulness cluster responds directly to the generative AI problem. Bender et al. showed that large language models can produce fluent but unreliable outputs, while Resnik and Hosseini showed that AI in scientific research raises authorship, transparency, and accountability problems. Theological texts interpret the same problem through different lenses. Christian sources treat misleading output as a failure of truthful witness. Islamic sources emphasize verification and the ethics of knowledge. Jewish sources connect it to false witness and the moral status of generated claims. The doctrinal convergence here is stronger than the technical literature alone suggests. The problem is not only hallucination, but epistemic erosion, where institutions begin to treat plausibility as truth.

The justice theme also deserves closer attention. Technical fairness research often focuses on bias mitigation, but theological texts broaden the frame. Justice is not only about equal treatment; it is about whether vulnerable persons are protected from structural misuse. This matters because AI can intensify exclusion through opaque profiling, data extraction, and decontextualized decision rules. The review shows that religious ethics repeatedly insists on a neighbor-oriented and community-oriented understanding of justice. That is why dignity and justice appear together so often in the corpus. They are not separate themes; they are mutually reinforcing. Without dignity, justice becomes procedural. Without justice, dignity becomes abstract.

A more surprising pattern is the relative weakness of implementation language. Many texts identify values with clarity, yet they stop short of specifying enforcement, auditing, or remedy structures. This is the main mechanistic deficit in the corpus. The literature often assumes that if ethics is named, governance will follow. The empirical reality of AI systems shows the opposite. Normative consensus does not produce operational accountability unless institutions build review boards, documentation protocols, refusal rights, escalation channels, and redress mechanisms. This finding is consistent with Mittelstadt's argument that principles alone are insufficient and with Selbst et al.'s critique that abstraction can hide structural harm. Theology therefore contributes moral depth, but it still needs institutional translation.

The tradition-specific findings help explain why. Christian texts often speak in terms of dignity, vocation, and relationality, which are powerful but not always procedural. Islamic texts often link norms to maqāṣid and public welfare, which provides a stronger policy bridge, especially when translated into governance criteria. Buddhist texts are especially useful for critiquing desire, attachment, and the illusion of control, which helps expose psychological and existential harms that procedural frameworks can miss. Jewish texts bring exceptional clarity on legal responsibility and personhood boundaries, which is useful for questions about delegation and machine agency. Each tradition contributes a distinct moral lens, and the comparative value lies in combining them rather than privileging one.

The practical implication is straightforward. AI governance for religiously plural societies should not ask whether theology has a place in ethics. It should ask which theological concepts can be translated into shared public norms without forcing confessional uniformity. The review suggests six candidates: dignity, stewardship, justice, non-harm, truthfulness, and human oversight. These concepts are robust enough to travel across traditions and practical enough to inform institutional governance. The strongest contribution of theological ethics is therefore not doctrinal specificity alone, but its capacity to humanize governance language that otherwise risks becoming purely managerial.

This study has limitations. The corpus overrepresents English-language materials and underrepresents scholarship in Arabic, Hebrew, Thai, Korean, and other languages where additional theological debate is likely to exist. The review also emphasizes doctrinal and conceptual literature more than field-based empirical studies of how communities actually use AI. Several traditions are internally plural, especially Islam and Christianity, yet the corpus necessarily compressed that diversity into broad thematic categories. A final limitation is temporal volatility. AI ethics guidance changed quickly during the drafting period, especially as generative systems moved into education, ministry, and research workflows. These limitations do not weaken the main argument, but they do caution against reading the findings as exhaustive.

E. CONCLUSION

This systematic literature review shows that global theological perspectives on artificial intelligence converge around a shared ethical core: human dignity, stewardship, justice, non-maleficence, truthfulness, and accountable oversight. Christian, Islamic, Buddhist, and Jewish sources differ in vocabulary and doctrinal grounding, yet they consistently reject the idea that AI should displace moral agency, obscure accountability, or reduce persons to data points. The main analytical contribution of the paper is comparative. It demonstrates that theology does not merely add moral commentary to AI governance; it supplies a durable grammar for human responsibility in a technologized society. For practice, the implication is direct: policymakers, educators, and faith-based institutions should translate broad ethical principles into auditing, documentation, refusal, redress, and human-in-the-loop governance mechanisms. Future research should move beyond document synthesis toward empirical studies of how these theological principles are negotiated in schools, hospitals, congregations, and public institutions, especially in multilingual and non-Western settings.

Acknowledgments

This review received no external funding. The author acknowledges the scholarly and institutional sources that made the comparative synthesis possible, including journal publishers, official ecclesial repositories, and intergovernmental ethics platforms.

REFERENCES

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Church of England. (2024a). *Artificial intelligence advice 2024*.

Church of England. (2024b). *Synod affirms work as key to "human dignity and purpose" in face of AI revolution*.

Francis. (2024a). *Artificial intelligence and the wisdom of the heart: Message for the 58th World Day of Social Communications*.

Francis. (2024b). *Message of the Holy Father to participants in the "AI and peace" conference*.

Goltz, N., Zeleznikow, J., & Dowdeswell, T. (2020). From the tree of knowledge and the golem of Prague to kosher autonomous cars: The ethics of artificial intelligence through Jewish eyes. *Oxford Journal of Law and Religion*, *9*(1), 132–156. https://doi.org/10.1093/ojlr/rwaa015

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

Mau, M., Hulu, W., & Faot, S. Y. I. (2024). Artificial intelligence ethics within Christian pedagogy: Reconstructing faith-based educational discourses in the era of algorithmic culture. *International Perspectives in Christian Education and Philosophy*, *1*(3), 44–59. https://doi.org/10.61132/ipcep.v1i3.525

Methodist Church. (2025a). *Artificial intelligence report*.

Methodist Church. (2025b). *Embracing AI with care and caution: Interim guidance for Methodist churches*.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(7), 501–507. https://doi.org/10.1038/s42256-019-0114-4

OECD. (2019). *OECD AI principles*.

OECD. (2024). *OECD updates AI principles to stay abreast of rapid technological developments*.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, Article n71. https://doi.org/10.1136/bmj.n71

Pontifical Academy for Life. (2020). *Rome Call for AI Ethics*.

Rappaport, Z. H. (2006). Robotics and artificial intelligence: Jewish ethical perspectives. In C. Nimsky & R. Fahlbusch (Eds.), *Medical technologies in neurosurgery* (Acta Neurochirurgica Supplements, Vol. 98, pp. 9–12). Springer. https://doi.org/10.1007/978-3-211-33303-7_2

Resnik, D. B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: New guidance needed for a new tool. *AI and Ethics*, *5*, 1499–1521. https://doi.org/10.1007/s43681-024-00493-8

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), Article 59. https://doi.org/10.1145/3287560.3287598

Smuha, N. A. (2022). *The human condition in an algorithmized world: A critique through the lens of 20th-century Jewish thinkers and the concepts of rationality, alterity and history*. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4093683

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*.
UNESCO. (2024). *Recommendation on the ethics of artificial intelligence*.
Vatican Dicastery for the Doctrine of the Faith, & Vatican Dicastery for Culture and Education. (2025a). *Antiqua et nova*.
Vatican Dicastery for the Doctrine of the Faith, & Vatican Dicastery for Culture and Education. (2025b). *Antiqua et nova: Note on the relationship between artificial intelligence and human intelligence*.
Yampolskiy, R. V. (2013). Artificial intelligence safety engineering: Why machine ethics is a wrong approach. In *The philosophy and theory of artificial intelligence* (pp. 389–396). Springer. https://doi.org/10.1007/978-3-642-31674-6_29