HUMANUS DISCOURSE Vol. 7. NO 1. 2027
ISSN 2787-0308 (ONLINE)

Digital immorality, Artificial Intelligence, and Religious Ethics: Toward an Interreligious Moral Framework

Oyeniyi Temitayo Lydia
Department of Religious Studies
Redeemer's University Ede
Osun State Nigeria

George O. Olatayo
Department of Religious Studies
Redeemer's University Ede
Osun State Nigeria

Abstract

Artificial intelligence (AI) has moved from laboratory curiosity to everyday infrastructure, and with it has emerged a set of moral wrongs that existing ethical vocabularies struggle to name. This article introduces "digital immorality" as an analytic category for harms and vices that are mediated, scaled, or concealed by algorithmic systems, including deception through synthetic media, surveillance and the commodification of persons, discriminatory automation, the evasion of responsibility, and the erosion of human attention and relationship. Drawing on Christian, Jewish, Islamic, Buddhist, and Hindu sources, as well as the growing body of institutional statements on AI by religious bodies, the article asks what religious ethics can contribute to the diagnosis and governance of these wrongs. It argues that religious traditions supply three resources that secular, principle-based frameworks often lack: thick anthropologies that ground human dignity in more than rational agency, virtue and formation-centered accounts of moral failure that attend to character rather than compliance alone, and communal practices of restraint, confession, and accountability. The article also identifies tensions within and among traditions, including disagreement over personhood, authority, and the permissibility of machine participation in ritual life, and cautions against instrumentalizing religion as a legitimating gloss for corporate or state AI agendas. It concludes by proposing five normative commitments for an interreligious ethics of AI and by outlining directions for further research. Methodologically, the article combines conceptual analysis with comparative textual interpretation, treating religious sources as living moral resources rather than museum pieces.

Keywords: artificial intelligence; digital immorality; religious ethics; human dignity; virtue ethics; interreligious dialogue; algorithmic accountability

Introduction

In January 2024 Pope Francis devoted his annual World Day of Peace message to artificial intelligence, warning that technologies of this kind can be turned

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

toward uses that diminish human dignity and destabilize peace.[1] A year later, two Vatican dicasteries jointly issued *Antiqua et Nova*, a lengthy note on the relationship between artificial intelligence and human intelligence.[2] These texts are no longer anomalies. Evangelical Protestants issued a statement of principles on AI in 2019,[3] and the Pontifical Academy for Life's *Rome Call for AI Ethics* (2020) has since drawn signatories beyond the Catholic Church, including representatives of other religious traditions at a 2024 gathering in Hiroshima.[4] Religious communities, it seems, have concluded that AI is not merely a technical matter. It is a moral and, in some sense, a spiritual one.

The secular literature points in the same direction, though in different language. Scholars have documented how commercial data extraction converts human experience into raw material for prediction,[5] how search and scoring systems reproduce racial and economic hierarchies,[6] and how facial analysis tools have performed unequally across skin tones and genders.[7] Meanwhile the dominant response within the field of AI ethics has been to articulate principles such as transparency, fairness, non-maleficence, and accountability. A survey of eighty-four guideline documents found striking convergence on these themes but also considerable divergence in how they are interpreted and implemented.[8] Critics have observed that principles alone cannot guarantee ethical AI, because they lack enforcement mechanisms, professional norms, and an account of the kind of person or institution that would actually honor them.[9]

This article argues that something is missing from the conversation, and that religious ethics is positioned to supply part of it. The missing element is a robust account of *wrongdoing* in digital environments: not only what rules a system should satisfy, but what sorts of vices, injustices, and betrayals of the human person such systems enable, amplify, or hide. To name this territory the article uses the term "digital immorality." The term is not established in the

[1]Pope Francis, "Artificial Intelligence and Peace," Message for the Fifty-Seventh World Day of Peace, January 1, 2024, https://www.vatican.va.

[2]Dicastery for the Doctrine of the Faith and Dicastery for Culture and Education, *Antiqua et Nova: Note on the Relationship between Artificial Intelligence and Human Intelligence* (Vatican City, January 28, 2025).

[3]Ethics and Religious Liberty Commission of the Southern Baptist Convention, "Artificial Intelligence: An Evangelical Statement of Principles," 2019.

[4]Pontifical Academy for Life, *Rome Call for AI Ethics* (Rome, February 28, 2020). Subsequent signatories have included representatives of other faith communities, notably at an interreligious event held in Hiroshima in July 2024.

[5]Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2019).

[6]Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018); Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016).

[7]Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of Machine Learning Research* 81 (2018): 77–91.

[8]Anna Jobin, Marcello Ienca, and Effy Vayena, "The Global Landscape of AI Ethics Guidelines," *Nature Machine Intelligence* 1 (2019): 389–99.

[9]Brent Mittelstadt, "Principles Alone Cannot Guarantee Ethical AI," *Nature Machine Intelligence* 1 (2019): 501–7.

scholarly literature, and it is offered here as a heuristic rather than a settled concept. Its purpose is to direct attention to moral failures that are partly constituted by digital architecture itself.

The argument proceeds in several steps. Section 2 clarifies the concept and the method. Section 3 offers a taxonomy of five forms of digital immorality. Section 4 explains why religious ethics offers distinctive resources. Sections 5 through 8 examine Christian, Jewish, and Islamic perspectives, followed by Buddhist and Hindu perspectives in Section 8. Section 9 compares them, Section 10 considers objections and risks, and Section 11 proposes five commitments for an interreligious ethics of AI before the conclusion.

Conceptualizing Digital Immorality

A Working Definition

This article defines *digital immorality* as morally culpable conduct, or morally corrosive structures, in which digital and especially algorithmic systems are constitutive rather than incidental. The qualifier matters. A fraudster who sends a letter commits fraud, and the postal service is incidental. A fraudster who clones a relative's voice with a generative model to extort money from a grandparent commits a wrong whose very shape, including its believability, its cheapness, and its scalability, depends on the technology. In the second case the digital layer is part of what makes the act what it is.

Three features distinguish digital immorality from older forms of wrongdoing. The first is *mediation*: human relationships, including relationships of trust, are increasingly conducted through systems that interpret, rank, filter, or generate communication on our behalf. Luciano Floridi has argued that we now inhabit an "infosphere" in which informational entities and processes are not merely tools but part of the environment in which agents form their identities and obligations.[10] The second is *scale*: a single design decision, whether a ranking criterion, a training dataset, or a default setting, can affect millions of persons simultaneously. The third is *opacity*: the operations of machine-learning systems are often inscrutable not only to affected persons but even to their developers, which makes it difficult to locate intention, negligence, or responsibility.

The last feature has attracted particular attention. Andreas Matthias coined the phrase "responsibility gap" to describe situations in which learning automata act in ways that neither their designers nor their operators could foresee or control, so that traditional ascriptions of blame seem to fail.[11] Mark Coeckelbergh has responded that responsibility in technological contexts is always distributed and relational, and that the task of ethics is to develop practices of answerability suited to that distribution rather than to hunt for a

[10]Luciano Floridi, *The Ethics of Information* (Oxford: Oxford University Press, 2013), 3–25.

[11]Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata," *Ethics and Information Technology* 6, no. 3 (2004): 175–83.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

single culpable agent.[12]Religious ethics, as will be seen, has long traditions of thinking about collective and structural sin that anticipate this move.

What the Term Does Not Mean

Two clarifications guard against misreading. First, "digital immorality" is not a synonym for cybercrime. Many harms discussed below are lawful in some jurisdictions, and some legal acts are morally gravely wrong. Second, the term does not imply that digital technologies are inherently corrupting. Religious traditions have often embraced new media, from the printing press to broadcasting, as vehicles for their message, and the field of digital religion has documented the creativity with which believers adapt online environments for worship, study, and community.[13] The claim is narrower: that certain patterns of harm are intensified or newly possible in algorithmic environments, and that they merit moral analysis in their own right.

Method

The article employs conceptual analysis and comparative textual interpretation. Comparative religious ethics faces well-known hazards, among them the temptation to flatten distinct traditions into a lowest common denominator and the risk of treating the texts of one tradition as if they could be read independently of the interpretive communities that sustain them. Jeffrey Stout has argued that moral discourse in pluralistic settings must proceed by honest engagement with particular traditions rather than by appeal to a neutral standpoint.[14] The article accepts that caution. It treats each tradition as internally plural, notes where scholars disagree, and refrains from claiming to speak for any community. The selection of five traditions reflects the availability of substantial primary sources and secondary literature, not a judgment that other traditions, including Sikh, Confucian, Indigenous, and African traditional religions, are less significant. Their omission is a limitation addressed in the conclusion.

A Taxonomy of Digital Immorality

Deception and Synthetic Reality

The first form concerns deception. Generative models can produce fluent text, realistic images, and convincing audio and video at negligible cost. Robert Chesney and Danielle Citron's influential analysis of "deep fakes" identified harms ranging from individual defamation and non-consensual sexual imagery to the corrosion of public trust in any recorded evidence, a phenomenon they described as a "liar's dividend," in which genuine evidence can be dismissed as

[12]Mark Coeckelbergh, *AI Ethics* (Cambridge, MA: MIT Press, 2020), chap. 8.

[13]Heidi A. Campbell, ed., *Digital Religion: Understanding Religious Practice in New Media Worlds* (London: Routledge, 2013).

[14]Jeffrey Stout, *Ethics after Babel: The Languages of Morals and Their Discontents* (Boston: Beacon Press, 1988).

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

fabricated.[15] Large language models raise a related concern. Emily Bender and her co-authors argued that such systems stitch together linguistic forms without grounding in meaning or communicative intent, and that readers readily mistake fluency for understanding.[16] The moral problem is not only that falsehoods are produced, but that the conditions for testimony, the practice by which communities transmit knowledge on the basis of trust, are degraded.

Surveillance and the Commodification of Persons

The second form concerns the extraction of personal data and the treatment of human beings as sources of behavioral surplus. Shoshana Zuboff's analysis of "surveillance capitalism" describes an economic logic in which private experience is unilaterally claimed as raw material, processed into predictions, and sold to parties interested in shaping future behavior.[17] Whatever one thinks of her broader historical claims, the pattern she identifies raises a clear moral question: whether persons may be treated primarily as means to the ends of prediction and influence. Facial recognition, emotion-inference technologies, and workplace monitoring extend this logic into domains where individuals have little practical power to refuse.

Discrimination and Automated Injustice

The third form is the encoding of unjust hierarchies into automated decision-making. Safiya Umoja Noble showed how commercial search results reinforced racist and sexist stereotypes, and Cathy O'Neil described opaque scoring models that sorted people in hiring, credit, education, and criminal justice in ways that often punished the poor.[18] Joy Buolamwini and Timnit Gebru's audit of commercial gender-classification systems found markedly higher error rates for darker-skinned women than for lighter-skinned men.[19] Kate Crawford has situated such cases within a wider critique of AI as an extractive industry dependent on labor, minerals, energy, and data gathered with little consent.[20] What makes these harms distinctive is their appearance of neutrality: injustice arrives dressed as calculation.

Evasion of Responsibility

The fourth form is the use of technical systems to launder accountability. When an institution says "the algorithm decided," it invokes the responsibility gap discussed above as an alibi. This is not a mere philosophical puzzle. It is a

[15]Robert Chesney and Danielle Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security," *California Law Review* 107, no. 6 (2019): 1753–1819.

[16]Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2021), 610–23.

[17]Zuboff, *Age of Surveillance Capitalism*.

[18]Noble, *Algorithms of Oppression*; O'Neil, *Weapons of Math Destruction*.

[19]Buolamwini and Gebru, "Gender Shades," 77–91.

[20]Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (New Haven, CT: Yale University Press, 2021).

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

recurrent organizational vice, in which actors exploit opacity to avoid being answerable for outcomes they were positioned to foresee and prevent. The concern extends to the gravest decisions. Pope Francis has urged that lethal autonomous weapons not be permitted to make decisions over human life, on the ground that no algorithm can substitute for moral judgment and responsibility.[21]

Erosion of Attention, Relationship, and Character

The fifth form is less visible but perhaps more pervasive. Sherry Turkle has documented the ways in which simulated companionship, from social robots to chatbots, can offer the appearance of intimacy while relieving users of the demands of reciprocity, leaving them "alone together."[22] Shannon Vallor has argued that technologies are not merely instruments but environments that shape habits, and that sustained exposure to systems optimized for engagement can erode capacities such as patience, honesty, empathy, and self-control, the very "technomoral virtues" needed to live well with technology.[23] Here the wrong is not a discrete act but a gradual deformation of the self, a pattern religious traditions have long described under the heading of vice, distraction, or idolatry.

These five forms are analytically distinct but practically entangled. A deepfake campaign may exploit surveillance data, target a marginalized group, be defended as an algorithmic accident, and deepen cynicism in the audience. A moral framework adequate to such cases must address acts, structures, and character together.

Why Religious Ethics?

The Limits of Principle-Based Frameworks

Principle-based AI ethics has made a genuine contribution by establishing shared reference points. Luciano Floridi and Josh Cowls, for example, proposed a unified framework of five principles, namely beneficence, non-maleficence, autonomy, justice, and explicability, drawn from bioethics and adapted to AI.[24] Yet the weaknesses of such frameworks are widely acknowledged. Mittelstadt points out that AI development lacks the common aims, professional history, and enforceable accountability that give bioethical principles traction in medicine.[25] Moreover, principles tend to be abstract, and organizations can endorse them without changing practice. Elettra Bietti has described the

[21]Francis, "Artificial Intelligence and Peace."

[22]Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other* (New York: Basic Books, 2011).

[23]Shannon Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016).

[24]Luciano Floridi and Josh Cowls, "A Unified Framework of Five Principles for AI in Society," *Harvard Data Science Review* 1, no. 1 (2019), https://doi.org/10.1162/99608f92.8cd550d1.

[25]Mittelstadt, "Principles Alone," 501–7.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

resulting phenomenon as "ethics washing," in which ethical language is deployed to forestall regulation and legitimate existing business models.[26]

Three Distinctive Resources

Religious ethics does not replace principles, but it offers at least three resources that complement them.

First, thick anthropologies. Religious traditions typically ground the dignity of the person not merely in rationality, autonomy, or utility, but in a relation to the sacred: the creature made in the image of God, the bearer of a divine trust, the possessor of Buddha-nature, or the manifestation of the self. Such accounts resist the reduction of persons to data profiles and offer reasons for treating the vulnerable, including those who cannot exercise rational agency, as inviolable.

Second, formation and virtue. Most religious ethics is not primarily a code of permissible and impermissible acts. It is a discipline of becoming a certain kind of person through repeated practices, such as prayer, fasting, study, almsgiving, meditation, and confession. This orientation fits the fifth form of digital immorality in particular, because it takes seriously the possibility that technologies train desire and habit. It also addresses a gap in compliance-driven frameworks: the question of who the designers, executives, and users of AI are becoming.

Third, communal accountability. Religious communities maintain institutions, from councils and courts of law to monastic orders and congregations, that interpret norms, rebuke wrongdoers, and rehabilitate offenders. They also possess vocabularies of repentance, restitution, and forgiveness that technical ethics seldom provides. In a field where regulation lags and corporate self-governance is suspect, such moral communities can serve as sites of formation and, at times, of prophetic critique.

Religion as Part of the AI Story

Religion is not only a commentator on AI but also a participant in its cultural imagination. Robert Geraci has shown how contemporary visions of superintelligence and digital immortality reproduce the structures of apocalyptic and millenarian thought, promising transcendence of the body and the renewal of the world through technology.[27] Recognizing these quasi-religious narratives is important for two reasons. It reveals that debates about AI are never purely technical, and it invites religious ethics to ask whether certain techno-utopian projects function as idols, objects of ultimate trust that

[26]Elettra Bietti, "From Ethics Washing to Ethics Bashing: A View on Tech Ethics from within Moral Philosophy," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2020), 210–19.

[27]Robert M. Geraci, *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality* (New York: Oxford University Press, 2010).

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

cannot bear the weight placed upon them. With this in view, the article turns to the traditions themselves.

Christian Perspectives

The Image of God and the Irreducibility of the Person

Christian reflection on AI begins, almost invariably, with the doctrine that human beings are created in the image of God (Latin *imago Dei*).[28] The doctrine has been interpreted in at least three ways: as a substantive capacity (reason, will, moral agency), as a vocation (a royal or priestly task of stewardship), and as a relationship (a capacity for communion with God and neighbor).[29] Each shapes the assessment of AI differently. If the image is chiefly rational capacity, then machines that perform cognitive tasks appear to threaten human uniqueness. Noreen Herzfeld has argued that this substantive reading has quietly governed much AI research, and she proposes instead a relational understanding, drawing on Karl Barth, according to which the image consists in the capacity for genuine otherness and mutual relationship, something no current system can offer.[30] Anne Foerst, who worked with robotics researchers at MIT, similarly suggests that Christian theology can welcome humanoid machines as creations of human creativity without conferring on them the status of persons.[31]

The Catholic note *Antiqua et Nova* develops a related position. It insists that human intelligence is embodied, relational, and oriented toward truth and goodness, and that AI, however impressive, remains a product of human ingenuity that simulates certain functions of intelligence without possessing the person's depth.[32] A moral consequence follows: to anthropomorphize AI systems, or to design them so as to encourage users to believe they are persons or friends, is a form of deception that corrupts the user's understanding of what a person is. This concern speaks directly to the first and fifth forms of digital immorality.

Sin, Structures, and Idolatry

Christian ethics has a developed vocabulary for the wrongs that exceed individual acts. Catholic social teaching speaks of "structures of sin," meaning institutional arrangements that, though built from human decisions, come to embody and perpetuate injustice and to draw participants into complicity.[33]

---

[28]Gen. 1:26–27.

[29]On these three families of interpretation, see J. Richard Middleton, *The Liberating Image: The Imago Dei in Genesis 1* (Grand Rapids, MI: Brazos Press, 2005).

[30]Noreen Herzfeld, *In Our Image: Artificial Intelligence and the Human Spirit* (Minneapolis: Fortress Press, 2002).

[31]Anne Foerst, *God in the Machine: What Robots Teach Us about Humanity and God* (New York: Dutton, 2004).

[32]Dicastery for the Doctrine of the Faith and Dicastery for Culture and Education, *Antiqua et Nova*.

[33]John Paul II, *Sollicitudo Rei Socialis* (1987), §36.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

This notion offers a useful frame for the responsibility gap. The designers, investors, deployers, and users of a biased scoring system may each contribute only a small share, yet the system as a whole can inflict serious harm, and the tradition insists that diffuse causation does not dissolve moral responsibility. It assigns a duty to dismantle such structures and to refuse cooperation with them.

Equally important is the biblical critique of idolatry. The psalmist says of idols made by human hands that those who make them will become like them, a line that anticipates contemporary worries about technologies that shape their makers.[34] Benedict XVI warned that technology becomes ideological when it is regarded as self-sufficient and as the answer to every human need.[35] Jacques Ellul's earlier argument that modern technique imposes its own logic of efficiency on every sphere of life, regardless of human ends, belongs to the same family of critiques.[36] On this view, the uncritical pursuit of ever more powerful AI, justified solely by what is technically possible, is not neutral progress but a form of misplaced trust.

Truthfulness, Neighbor-Love, and the Problem of Distance

The Decalogue's prohibition of false witness is among the oldest ethical norms directly relevant to synthetic media.[37] The Christian tradition has generally read it broadly, as a prohibition of lying, slander, and any practice that damages a neighbor's reputation or the common life of truthful speech. Deepfakes and large-scale automated disinformation are, in this light, not simply regrettable side effects of innovation but violations of a basic obligation of community.

The parable of the Good Samaritan raises a different issue.[38] The Samaritan acts because he sees and approaches a wounded stranger, and the parable redefines "neighbor" as whoever is placed in need before me. Algorithmic systems tend to reverse this dynamic by increasing the distance between the decision maker and the person affected, converting persons into rows of data. A Christian ethic of AI would therefore press for design and governance practices that preserve proximity and encounter, particularly in contexts such as healthcare, welfare administration, and criminal justice. The Evangelical statement of principles points in a compatible direction when it affirms human dignity, treats AI as a tool rather than a moral agent, and calls for care about privacy, bias, and the effects of automation on work.[39]

[34]Ps. 115:4–8.

[35]Benedict XVI, *Caritas in Veritate* (2009), §70.

[36]Jacques Ellul, *The Technological Society*, trans. John Wilkinson (New York: Knopf, 1964).

[37]Exod. 20:16.

[38]Luke 10:25–37.

[39]Ethics and Religious Liberty Commission, "Artificial Intelligence: An Evangelical Statement of Principles."

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

Jewish Perspectives

Halakhic Method and the Image of God

Jewish ethics is distinguished by its legal-interpretive method. Rather than deriving norms from a single principle, rabbinic tradition reasons by analogy from textual precedents (halakha) and so has characteristically responded to new technologies, from electricity to genetic engineering, by asking how old categories apply to new facts. The foundational anthropological claim is shared with Christianity: the human being is created *betzelem Elohim*, in the image of God.[40] Yet the rabbinic emphasis tends to fall on obligation (*mitzvot*) rather than on metaphysical definition. The relevant question is less "what is a person?" than "what is demanded of me in this situation?"

Deception, Slander, and the Theft of Mind

Rabbinic literature contains a remarkably rich set of categories for wrongs of speech and perception. *Geneivat da'at*, literally "theft of the mind," prohibits creating a false impression in another person even where no explicit lie is told.[41]*Ona'at devarim* prohibits verbal oppression and the deliberate causing of distress by words.[42]*Lashon hara*, or derogatory speech, is treated by Maimonides as a grave offense even when the statement is true, because of the damage it does to the subject and to the social fabric.[43] These categories map naturally onto digital immorality. A synthetic video that implies a person said what they never said, an algorithm that amplifies rumor because it generates engagement, and a chatbot that is designed to be mistaken for a human interlocutor each create a false impression in the mind of another. The concept of theft of mind is especially suggestive because it locates the wrong in the manipulation of understanding rather than in the narrow falsity of a proposition.

Privacy, Stumbling Blocks, and Safety Design

The Talmud develops the concept of *hezek re'iyah*, damage caused by sight, to require that a person not construct a window overlooking a neighbor's courtyard.[44] Contemporary writers have invoked this principle in discussions of surveillance, since it recognizes that mere observation, without any further use of the information, can injure the observed. The prohibition against placing a stumbling block before the blind, traditionally extended to those who are blind in matters of knowledge or vulnerable to temptation, provides a second resource.[45] Designing systems that exploit cognitive weakness, such as interfaces engineered to maximize compulsive use among adolescents, can be understood as placing such a stumbling block.

---

[40]Gen. 1:27.
[41]b. Ḥullin 94a.
[42]b. Bava Metzia 58b.
[43]Moses Maimonides, *Mishneh Torah*, Hilkhot De'ot 7.
[44]b. Bava Batra 2b–3a.
[45]Lev. 19:14; see also b. Pesaḥim 22b.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

The Torah's requirement that a householder build a parapet around the roof, so that no one will fall and bring bloodguilt on the house, is perhaps the most apt precedent for the responsibility gap.[46] The law does not wait for harm to occur and then search for a culpable agent. It imposes a positive duty on the person who controls a dangerous space to anticipate and prevent foreseeable injury. Applied to AI, the parapet principle implies that developers and deployers who introduce powerful, opaque systems into public life bear a duty of foresight and protective design, and cannot discharge it by pointing to unpredictability. Alongside this, the principle of mutual responsibility (*arevut*) holds that members of a community are answerable for one another.[47]

The Golem and the Limits of Artificial Creation

Jewish tradition also contains a long narrative reflection on artificial beings in the golem legends.[48] A Talmudic passage reports that the sage Rava created a man and sent him to Rabbi Zeira, who addressed him and received no answer, and then declared that the creature had been made by a colleague and should return to dust.[49] The story has often been read as an early meditation on the difference between a being that imitates human form and one that speaks, that is, one that can answer and be answerable. Whether or not the reading is historically precise, it illustrates a Jewish intuition that artificial creation is not forbidden but is bounded, and that the capacity for responsive, accountable speech marks a threshold that imitation does not cross. Heschel's reflections on the Sabbath add a complementary lesson about limits: the weekly rest functions as a disciplined refusal to be defined by productive mastery over the world, a practice that offers an instructive counterweight to always-on, optimization-driven technologies.[50]

Islamic Perspectives

Khilafah, Amanah, and Human Dignity

Islamic ethics begins from the premise that the human being is God's steward (*khalifah*) on earth and the bearer of a trust (*amanah*) that other creatures declined to carry.[51] The Qur'an also affirms that God has honored the children of Adam, which grounds a doctrine of inherent human dignity (*karamah*) independent of merit or status.[52] Because the human being is a trustee, power over nature and over other persons is never absolute but is held conditionally and will be examined. This framework supports a stewardship ethic of

[46]Deut. 22:8.

[47]b. Shevu'ot 39a.

[48]Moshe Idel, *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid* (Albany: State University of New York Press, 1990).

[49]b. Sanhedrin 65b.

[50]Abraham Joshua Heschel, *The Sabbath: Its Meaning for Modern Man* (New York: Farrar, Straus and Giroux, 1951).

[51]Qur'an 2:30 and 33:72.

[52]Qur'an 17:70.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

technology that is neither technophobic nor triumphalist. Seyyed Hossein Nasr has argued, from within the Islamic intellectual tradition, that modern technology becomes spiritually dangerous when it is divorced from the sense that nature and the human being are signs of the sacred.[53]

Verification, Truthfulness, and Privacy

The Qur'an and hadith literature offer especially pointed guidance for the information environment. Believers are commanded to verify reports brought by an unreliable person before acting on them, lest they harm others out of ignorance.[54] Another verse prohibits pursuing what one has no knowledge of, because hearing, sight, and the heart will each be questioned.[55] The same chapter that commands verification also forbids suspicion, spying (*tajassus*), and backbiting (*ghibah*).[56] Entering homes without permission is likewise prohibited, establishing a robust norm of domestic and personal privacy.[57] Read together, these texts supply a ready-made ethic for synthetic media and surveillance. They place on each person a duty of epistemic care, which includes not circulating unverified content, and they place on those with power a duty not to penetrate private life. A well-known prophetic saying states that whoever deceives is not one of the community, a statement that applies with force to the design of systems that deliberately mislead.[58]

Maqasid al-Shari'ah and Precautionary Governance

The classical framework of the higher objectives of Islamic law (*maqasid al-shari'ah*), traditionally enumerated as the protection of religion, life, intellect, lineage or dignity, and property, offers a structured way to evaluate technologies. Jasser Auda has argued that the *maqasid* should be reconceived as a systems-oriented, human-development framework rather than a static list, and his approach is well suited to AI governance because it asks about the overall effects of a practice on human flourishing.[59] The protection of intellect (*'aql*) takes on new meaning in an environment of mass-produced misinformation and addictive design. The related principles of public interest (*maslahah*) and the blocking of the means to harm (*sadd al-dhara'i'*) support precautionary regulation: where a technology predictably opens the way to serious harm, restraint is warranted even before harm is proven.

---

[53]Seyyed Hossein Nasr, *Religion and the Order of Nature* (New York: Oxford University Press, 1996).

[54]Qur'an 49:6.

[55]Qur'an 17:36.

[56]Qur'an 49:12.

[57]Qur'an 24:27–28.

[58]Ṣaḥīḥ Muslim, Kitāb al-Īmān, on the saying "whoever deceives us is not of us."

[59]Jasser Auda, *Maqasid al-Shariah as Philosophy of Islamic Law: A Systems Approach* (London: International Institute of Islamic Thought, 2008).

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

Accountability and the Limits of Delegation

Islamic ethics is also emphatic about individual and collective accountability. A frequently cited saying holds that each person is a shepherd and is responsible for those under their care.[60] This tradition resists the idea that responsibility can be transferred to a tool. Debates about whether AI systems may issue religious opinions illustrate the point. The classical practice of giving a legal opinion (*fatwa*) presupposes a qualified person who can be questioned, who stands within a chain of scholarly transmission, and who bears moral responsibility before God for the ruling. Many Muslim scholars and commentators have therefore expressed caution about treating machine-generated outputs as authoritative, even while acknowledging their usefulness for retrieval and study. The underlying concern is that authority without answerability is a moral hazard.

Buddhist and Hindu Perspectives

Buddhist and Hindu traditions are treated in one section for reasons of space and because they share a family of concepts, notably dharma, karma, and the analysis of craving and attachment. The grouping is expository only. The two traditions differ sharply on the nature of the self, the authority of scripture, and the status of the divine, and the label "Dharmic," which is not universally accepted by adherents or scholars, is therefore avoided as a category. Each tradition is examined on its own terms in 8.1 and 8.2.

Buddhist Ethics: Intention, Interdependence, and Craving

Buddhist ethics offers distinctive resources in three areas. The first is the primacy of intention. The Buddha is reported to have said that it is intention (*cetana*) that he calls karma, and Buddhist ethics accordingly evaluates actions by their motivating states, whether rooted in greed, hatred, and delusion or in generosity, loving-kindness, and wisdom.[61] Applied to AI, this focus shifts attention from the machine's outputs to the motivations of those who build and deploy it. A system optimized to exploit craving for attention is morally shaped by the mental states of its commissioners, even if no individual intends a specific harm.

The second resource is the doctrine of dependent origination, the teaching that all phenomena arise in dependence on conditions. This offers a way of thinking about the responsibility gap that does not require locating a single author. Harms arise from webs of conditions, including training data, incentives, institutions, and users' habits, and ethical practice involves recognizing and

[60]Ṣaḥīḥ al-Bukhārī, Kitāb al-Jumʿa.

[61]Anguttara Nikaya 6.63. See the discussion in Peter Harvey, *An Introduction to Buddhist Ethics: Foundations, Values and Issues* (Cambridge: Cambridge University Press, 2000), 16–20.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

modifying these conditions. The perspective is close to Coeckelbergh's relational account but is grounded in a metaphysics of interdependence. It also supports attention to environmental and labor costs along the supply chain of AI.

The third resource is a sophisticated psychology of craving (*tanha*) and attachment. The Buddhist diagnosis of suffering as rooted in clinging resonates with analyses of engagement-optimizing platforms that cultivate compulsive checking, social comparison, and fear of missing out. Contemplative practices such as mindfulness and the cultivation of restraint, together with the precept of refraining from false speech and the broader discipline of Right Speech, which includes avoiding divisive and idle talk, supply a practical counterpart to the vice-language of Section 3.5.[62]

Buddhist views on AI personhood are more open than many Abrahamic ones, because the tradition denies a permanent, independent self even in humans and focuses instead on sentience and the capacity to suffer. Damien Keown's account of Buddhist ethics stresses that moral status is tied to living, sentient beings, which suggests that an artifact that merely processes information would not qualify, though the question remains debated.[63] Some Buddhist communities have also experimented with robotic or android figures in temple settings, which raises contested questions about the authenticity of ritual and teaching mediated by machines.

Hindu Ethics: Dharma, Action, and the Welfare of the World

Hindu traditions are plural, and no single text speaks for all. The *Bhagavad Gita*, however, has been among the most influential ethical sources. Its teaching of disciplined action without attachment to results (*nishkama karma*) is sometimes misunderstood as indifference to consequences.[64] In context, it asks the agent to act from duty and compassion rather than from ego-driven desire for success or recognition. For AI, this principle can be turned as a critique of development driven by the desire for dominance or profit. The Gita also articulates the ideal of acting for the maintenance of the world (*lokasamgraha*), implying that persons with power and knowledge bear a heightened responsibility for the common welfare.[65]

The principle of non-harm (*ahimsa*) and the commitment to truthfulness (*satya*) are central to Hindu ethical thought and apply naturally to technologies that harm bodily or psychological integrity or that deceive. The notion of context-sensitive duty (*dharma*, and individually *svadharma*) also suggests that ethical evaluation of AI must attend to roles: the duties of an engineer, a physician, a judge, or a parent differ and cannot be collapsed into a uniform rule. Furthermore, in Vedanta the self (*atman*) is understood as consciousness, not as

[62]Harvey, *Introduction to Buddhist Ethics*, chap. 2.

[63]Damien Keown, *The Nature of Buddhist Ethics* (London: Macmillan, 1992).

[64]*Bhagavad-Gita* 2.47. See *The Bhagavad-Gita: Krishna's Counsel in Time of War*, trans. Barbara Stoler Miller (New York: Bantam Books, 1986).

[65]*Bhagavad-Gita* 3.20–25.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

a bundle of computable functions, so that machine simulation of intelligence is not a simulation of the self. A modern expression of Hindu-inflected technological critique appears in Mohandas Gandhi's warning that the worship of machinery can displace human labor and moral self-rule.[66] One need not share his conclusions to appreciate his insistence that technological systems must be judged by their effect on the moral and social life of the community.

Convergences across Traditions

The preceding sections reveal significant convergence amid real difference. The table below summarizes the principal anthropological anchor, characteristic moral concept, and digital wrong most directly addressed by each tradition.

Tradition	Anthropological anchor	Characteristic moral concept	Digital wrong most directly addressed
Christian	Image of God; relational personhood	Structures of sin; idolatry; false witness	Deception; automated injustice; anthropomorphic simulation
Jewish	Image of God; covenantal obligation	Theft of mind (*geneivat da'at*); parapet duty; mutual responsibility	Manipulation; surveillance; evasion of foresight
Islamic	Vicegerency and trust; inherent dignity	Verification; *maqasid*; accountability; blocking harm	Misinformation; privacy invasion; unaccountable authority
Buddhist	Sentient beings; interdependence	Intention; craving; Right Speech	Attention exploitation; structural harm; deceptive speech
Hindu	Self as consciousness; dharmic role	Non-harm; truthfulness; *lokasamgraha*	Violence; deception; power without welfare orientation

Five convergences stand out. First, *dignity* is treated as a given rather than an earned status, and this supports resistance to dehumanizing classification and scoring. Second, *truthfulness* is a communal as well as an individual obligation, and traditions offer detailed taxonomies of deception that exceed the simple lie. Third, *answerability* cannot be delegated to tools: whether through the parapet law, the shepherd's responsibility, or the Buddhist focus on intention and conditions, each tradition insists that someone must answer. Fourth, *justice toward the vulnerable* is a test of any social arrangement, and this is a theme

[66]M. K. Gandhi, *Hind Swaraj and Other Writings*, ed. Anthony J. Parel (Cambridge: Cambridge University Press, 1997).

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

common to prophetic, juridical, and compassion-centered traditions alike. Fifth, *restraint and formation* are essential: moral life is cultivated through practices that limit appetite and train attention.

These convergences align in notable ways with institutional statements. The *Rome Call*, for instance, names transparency, inclusion, responsibility, impartiality, reliability, and security and privacy as core principles.[67] The principles are phrased in the idiom of public ethics, but the traditions examined here provide the deeper grounds that make them intelligible and binding for believers. Jürgen Habermas has argued that secular citizens and religious citizens share an obligation of mutual translation in democratic deliberation, with religious contributions open to being rendered in publicly accessible terms.[68] The convergences identified above suggest that such translation is feasible for much of the ethics of AI.

Tensions, Objections, and Risks

Personhood and Moral Status

The traditions diverge on whether and how artifacts might be moral patients or agents. Many Christian and Islamic thinkers treat personhood as dependent on a created soul or spirit that machines lack, while Buddhist approaches center on sentience, and Jewish thought focuses on responsive, accountable speech. These differences matter if future systems appear to exhibit behaviors associated with consciousness. An interreligious ethics should therefore be prepared to articulate a shared minimum, which is that no AI should be treated in ways that degrade human beings or promote deception about its nature, while leaving the ultimate metaphysical questions to the respective traditions.

Machines in Religious Life

A further tension concerns the use of AI in religious practice, including chatbots that answer theological questions, generate sermons, or simulate spiritual advisers. Defenders point to access and education, particularly for communities with few trained leaders. Critics worry that such systems sever teaching from lived witness, hollow out the relationship between teacher and learner, and blur lines of authority and responsibility. Heidi Campbell's work on religious authority in digital contexts suggests that such developments will be negotiated differently by communities depending on how they understand legitimate leadership and the nature of the sacred.[69] In Christian sacramental traditions, embodiment is theologically central, so that machine-mediated rites raise questions that do not arise to the same degree in some other settings.

---

[67]Pontifical Academy for Life, *Rome Call for AI Ethics*.

[68]Jürgen Habermas, "Religion in the Public Sphere," *European Journal of Philosophy* 14, no. 1 (2006): 1–25.

[69]Campbell, *Digital Religion*.

© humanusdiscourse@gmail.com , http://humanusdiscourse.website2.me

Instrumentalization and Ethics Washing

There is a danger that religious voices will be co-opted. Corporations and governments may invite religious leaders to endorse AI initiatives in order to claim moral legitimacy, a religious analogue of the ethics washing criticized by Bietti.[70] Religious communities should therefore insist on conditions for participation, including transparency about funding, independence to criticize, and willingness to say no. Prophetic traditions, in particular, remind believers that proximity to power is a test of fidelity, not a measure of influence.

Religious Communities as Perpetrators and Beneficiaries

Finally, religious communities are not innocent bystanders. They too can commit digital immorality: through surveillance of members, through the spread of disinformation and hate speech, through the manipulation of vulnerable seekers, or through the adoption of engagement-maximizing tools for fundraising. Traditions also contain exclusionary or hierarchical elements that, if uncritically encoded into automated systems, could reproduce injustice. A credible religious ethics of AI must therefore be self-critical, applying its own standards to its own institutions. The argument of this article is not that religion is a reservoir of untainted wisdom, but that the practices of self-examination, confession, and reform that religious traditions cultivate are precisely what an age of powerful, opaque technologies requires.

Five Commitments for an Interreligious Ethics of AI

Building on the comparative analysis, the article proposes five commitments. They are intended not as a creed but as a basis for cooperation among traditions and with secular partners.

1. The inviolability of the person. Systems that classify, rank, or predict human beings must not be permitted to deny the fundamental dignity of those affected. In practice this implies meaningful human review for decisions affecting liberty, livelihood, health, or life; prohibition of uses that rely on demeaning inference; and strict limits on the ability of automated systems to determine who is worthy of help.

2. The integrity of truth and testimony. Persons have a right not to be deceived about whether they are interacting with a machine or viewing synthetic content, and communities have a duty to protect the conditions of trustworthy testimony. This entails disclosure and provenance standards, strong sanctions for malicious impersonation, and educational efforts that cultivate verification as a moral habit.

3. Human answerability. For every consequential system there must be identifiable human beings and institutions that answer for its effects, and no system should be entrusted with decisions over life and death. The parapet duty and the shepherd's responsibility converge in a principle of anticipatory accountability: those who introduce risk bear the burden of foresight.

[70]Bietti, "From Ethics Washing to Ethics Bashing," 210–19.

4. Preferential concern for the vulnerable and the common good. Evaluation of AI should begin with its effects on the poor, the marginalized, children, workers, and the natural environment. This implies independent auditing for bias, protections against exploitative data practices, attention to the labor and ecological costs of AI supply chains, and fair distribution of benefits.

5. Formation and restraint. Technologies should be designed and used in ways that support the flourishing of human capacities rather than exploiting weakness. This requires the education of engineers and executives in moral formation as well as in compliance, cultural practices of rest and limits comparable to the Sabbath, and openness to refusing or retiring technologies that deform the people they serve.

Religious communities can advance these commitments in at least four ways: by educating clergy, teachers, and members in digital moral literacy; by adopting internal policies governing their own use of AI; by participating, under conditions of independence, in public deliberation and regulation; and by offering spaces of rest, silence, and face-to-face encounter that counter the pressures of a frictionless digital environment.

Conclusion

This article has proposed digital immorality as a category for understanding wrongs that are constituted by algorithmic mediation, scale, and opacity, and has argued that religious ethics offers valuable resources for addressing them. Through comparative analysis of Christian, Jewish, Islamic, Buddhist, and Hindu sources, it has shown that these traditions possess rich vocabularies for deception, surveillance, injustice, evasion of responsibility, and the deformation of character, and that they converge on commitments to dignity, truthfulness, answerability, justice for the vulnerable, and formation through restraint.

This article has limitations. It has examined only five traditions, has relied largely on textual and institutional sources rather than on ethnographic data about how believers actually use and judge AI, and has necessarily generalized across internally diverse communities. The concept of digital immorality also requires further refinement, including empirical testing against concrete case studies. Future research might therefore examine lived religious responses to generative AI through interviews and ethnography, extend the comparative framework to additional traditions including Sikh, Confucian, and Indigenous perspectives, and investigate how the commitments proposed here might be translated into concrete professional codes, regulatory proposals, and educational curricula.

The deepest contribution of religious ethics may be a reminder that the central question about AI is not only what machines can do but what we are becoming as we build and use them. The traditions surveyed here, for all their differences, agree that human beings are accountable, that truth is a shared trust, and that power without love or compassion is dangerous. In an age when many decisions are quietly delegated to systems no one fully understands, the insistence that someone must answer, and that someone must also learn to live well, remains a vital contribution to the common good.