DIYĀ' AL-AFKĀR
JURNAL STUDI AL-QUR'AN DAN AL-HADIS
VOLUME 14, NOMOR 1, (JUNI 2026): 147-162
ISSN 2303-0453 I E-ISSN 2442-987
https://journal1.uinssc.ac.id/index.php/diya/index
https://doi.org/10.24235/x0wf1t53

Toward A Qur'anic Human-Centered AI Framework: Epistemic Humility, Human Dignity, and Technological Responsibility in Qur'ān Sūrah al-'Alaq [96]:1-5

Yasser Muda Lubis*[1], Topikurohman[2], Muhammad Irsan Barus[3], Abdulroman Mahae[4]
[12]Universitas PTIQ Jakarta, [3]Sekolah Tinggi Agama Islam Negeri Mandailing Natal, [4]Jamiah Islam Syeikh Daud Al-Fathoni, Thailand
Email: yassermudalubis@ptiq.ac.id

Abstract: Artificial intelligence (AI) increasingly mediates knowledge production, learning, and decision-making, intensifying debates about human agency, epistemic authority, and moral accountability. This article develops a normative-conceptual Qurʾanic Human-Centered Artificial Intelligence (QHCAI) framework through a disciplined contemporary reading of Qurʾān 96:1–5 (al-ʿAlaq). Rather than claiming that the verses contain a technical theory of AI, the study distinguishes textual-exegetical meaning from contemporary significance. Methodologically, it proceeds through three stages: exegetical reconstruction using major classical tafsīr works, comparative conceptual mapping with established Human-Centered AI and responsible-AI literature, and normative synthesis. The analysis argues that the passage supports five interrelated dimensions: transcendental orientation, creaturely dignity and epistemic humility, ethical beneficence, instrumental mediation under meaningful human control, and epistemic responsibility under Divine sovereignty. The proposed model does not replace established requirements such as fairness, transparency, accountability, safety, and human oversight. Its distinctive contribution is to reframe human-centeredness as theocentric human responsibility: human beings remain the locus of moral and epistemic accountability without becoming the ultimate source of value or truth. The article operationalizes the framework through evaluative questions and governance indicators and explicitly treats the analogy between al-qalam and contemporary digital technologies as functional rather than predictive. The QHCAI framework is therefore offered as a normative lens for assessing AI systems, not as an empirically validated governance standard.

Keywords: Human-Centered Artificial Intelligence; Qurʾanic Epistemology; AI Ethics; Epistemic Humility; Human Responsibility

Abstrak: Kecerdasan buatan (AI) semakin memediasi produksi pengetahuan, pembelajaran, dan pengambilan keputusan sehingga memperkuat perdebatan tentang agensi manusia, otoritas epistemik, dan tanggung jawab moral. Artikel ini mengembangkan kerangka normatif-konseptual Qurʾanic Human-Centered Artificial Intelligence (QHCAI) melalui pembacaan kontemporer yang terkontrol terhadap Qurʾān 96:1–5 (al-ʿAlaq). Penelitian ini tidak mengklaim bahwa ayat-ayat tersebut memuat teori teknis tentang AI, melainkan membedakan makna tekstual-eksegetis dari signifikansi kontemporernya. Analisis dilakukan melalui tiga tahap: rekonstruksi eksegetis berdasarkan karya-karya tafsīr klasik utama, pemetaan konseptual komparatif dengan literatur Human-Centered AI dan responsible AI, serta sintesis normatif. Hasil analisis merumuskan lima dimensi yang saling berkaitan, yaitu orientasi transendental, martabat makhluk dan kerendahan hati epistemik, kemaslahatan etis, mediasi instrumental di bawah meaningful human control, serta tanggung jawab epistemik di bawah kedaulatan Ilahi. Model ini tidak menggantikan prinsip fairness, transparency, accountability, safety, dan human oversight yang telah berkembang dalam etika AI. Kontribusi khasnya adalah membingkai human-centeredness sebagai tanggung jawab manusia yang teosentris: manusia tetap menjadi locus tanggung jawab moral dan epistemik tanpa diposisikan sebagai sumber nilai dan kebenaran yang tertinggi. Kerangka QHCAI dioperasionalkan melalui pertanyaan evaluatif dan indikator tata kelola, sedangkan analogi antara al-qalam dan teknologi digital diperlakukan secara fungsional, bukan prediktif. Dengan demikian, QHCAI ditawarkan sebagai lensa normatif untuk menilai sistem AI, bukan sebagai standar tata kelola yang telah teruji secara empiris.

Kata Kunci: Human-Centered Artificial Intelligence; Epistemologi Qurʾani; Etika AI; Kerendahan Hati Epistemik; Tanggung Jawab Manusia.

Introduction

Artificial intelligence has moved from a specialized computational technology into a general infrastructure for communication, education, research, health care, administration, and decision support. Contemporary systems can generate text and images, rank information, recommend actions, summarize evidence, and automate parts of professional judgment. These capacities create substantial opportunities, yet they also complicate the distribution of agency and responsibility between human users, institutions, and computational systems. Human-Centered Artificial Intelligence (HCAI) emerged in response to this tension by emphasizing designs that augment human capabilities while preserving human control, responsibility, and trust.[1]

The contemporary AI-ethics landscape already contains a substantial normative vocabulary. Global guidelines converge around transparency, justice and fairness, non-maleficence, responsibility, privacy, human rights, and oversight, even though they differ considerably in interpretation and implementation. This convergence is important, but it does not settle the prior question of why these values should govern technological development or how competing values should be ordered in concrete contexts. Mittelstadt has therefore warned that high-level ethical principles cannot by themselves guarantee ethical AI because implementation depends on institutional practices, professional norms, accountability structures, and a defensible conception of the purposes that technology ought to serve.[2]

This normative underdetermination creates space for religious and philosophical traditions to contribute without requiring a rejection of existing human-rights or responsible-AI frameworks. The relevant question is not whether a Qur‘anic approach should replace contemporary HCAI, but what additional normative and epistemological resources it can bring to debates on human agency, knowledge, and accountability. Recent work on AI in Islamic ethics has begun to articulate precisely this possibility by emphasizing plural ethical benchmarking, the authority of revelation, and the need to preserve human responsibility in technologically mediated contexts.[3]

Within Qur‘anic studies, Qur‘ān 96:1–5 has most frequently been discussed in relation to literacy, education, the pursuit of knowledge, and Islamic epistemology. Recent scholarship continues to reconstruct *iqra‘* as an integrative knowledge paradigm connecting revelation,

[1] Ben Shneiderman, “Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy,” *International Journal of Human–Computer Interaction* 36, no. 6 (2020): 495–504, esp. 495–499; Tara Capel and Margot Brereton, “What Is Human-Centered about Human-Centered AI? A Map of the Research Landscape,” *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023): 1–23, esp. 3–8.

[2] Brent Mittelstadt, “Principles Alone Cannot Guarantee Ethical AI,” *Nature Machine Intelligence* 1 (2019): 501–507, esp. 501–504; Anna Jobin, Marcello Ienca, and Effy Vayena, “The Global Landscape of AI Ethics Guidelines,” *Nature Machine Intelligence* 1 (2019): 389–399, esp. 389–394; Luciano Floridi et al., “AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations,” *Minds and Machines* 28 (2018): 689–707, esp. 700–704.

[3] Ezieddin Elmahjub, “Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI,” *Philosophy & Technology* 36 (2023): Article 73; Muhammad Hazim Mohd Azhar et al., “Ethics and Limits of Artificial Intelligence (AI) in Quranic Exegesis According to the Epistemological Framework of Islamic Knowledge,” *International Journal of Quranic Research* 17, no. 2 (2025): 97–124, esp. 101–112.

reason, and engagement with reality. A separate but increasingly visible body of scholarship examines the use of AI in Qur‘anic interpretation, where the principal concerns are not merely efficiency but also interpretive authority, decontextualization, epistemic reliability, and the inability of current AI systems to assume autonomous religious responsibility.[4]

The gap addressed by this article therefore lies at the intersection of three literatures that are often discussed separately: exegetical studies of Qur‘ān 96:1–5, Islamic reflection on AI and knowledge, and international HCAI scholarship. Existing studies frequently move directly from a Qur‘ānic value to a contemporary technological recommendation. Such a move is methodologically risky because it can blur the difference between what a verse meant within the exegetical tradition and what contemporary readers construct as its present significance. Conversely, HCAI scholarship often specifies desirable governance values without exploring a transcendental account of knowledge, human limitation, or ultimate moral accountability.

This article addresses that double gap by asking three questions. First, what epistemological and anthropological meanings can be reconstructed from Qur‘ān 96:1–5 through comparison of major classical tafsīr works? Second, how do those meanings enter into critical dialogue with established HCAI principles such as human control, dignity, fairness, accountability, and epistemic reliability? Third, what specifically Qur‘anic contribution can be formulated without presenting the Qur‘ān as a technical blueprint for artificial intelligence?

The central argument is that the distinctive contribution of Qur‘ān 96:1–5 is not a set of technical instructions for AI, nor even a wholly new list of ethical principles. Its contribution is a reconfiguration of the normative location of the human subject. The passage links knowing to the Creator, reminds the human being of creaturely origin, presents knowledge as taught rather than self-generated, and identifies the qalam as an instrument by which knowledge is mediated. When placed in dialogue with HCAI, these elements support what this article calls theocentric human-centeredness: human beings remain the responsible moral and epistemic agents within technological systems, but they are not treated as the ultimate source of value, truth, or accountability.

This argument also requires a deliberate limitation. The article focuses on the first five verses because they form the initial revelatory unit traditionally associated with the beginning of Qur‘anic revelation. It does not claim to offer a complete thematic Qur‘anic theology of technology. Other verses, including Qur‘ān 17:70 on human honor and Qur‘ān 2:30 on human stewardship, would be indispensable for a comprehensive Qur‘anic anthropology. At the same time, Qur‘ān 96:6–7 immediately warns that the human being may transgress when imagining himself self-sufficient. This adjacent warning functions here as a hermeneutical control against reading the first five verses as an uncomplicated celebration of knowledge or technological power.[5]

---

[4] Nehru Millat Ahmad, “Artificial Intelligence in Qurʾanic Interpretation: A Theological and Ethical Analysis,” *Journal of Islamic Thought and Philosophy* 5, no. 1 (2026): 60–79, esp. 65–72; Hartono Hartono and Ismail, “Iqraʾ as Knowledge Paradigm: Islamic Epistemology Reconstruction from QS Al-Alaq 1–5,” *Maghza: Jurnal Ilmu Al-Qur’an dan Tafsir* 11, no. 1 (2026): 1–27, esp. 3–10.

[5] Al-Rāzī, *Mafātīḥ al-Ghayb*, 32:218–219; Muḥammad ibn Jarīr al-Ṭabarī, *Jāmiʿ al-Bayān ʿan Taʾwīl Āy al-Qurʾā*n, ed. ʿAbd Allāh ibn ʿAbd al-Muḥsin al-Turkī (Cairo: Dār Hajr, 2001), 24:517–520.

Method

This study is a qualitative, normative-conceptual inquiry based on library research. It does not claim to conduct a systematic literature review in the technical sense, because the study does not employ a PRISMA-style exhaustive search and screening protocol. Instead, it uses a structured and purposive review designed to support conceptual comparison across two corpora: Qurʾanic exegesis and contemporary HCAI/AI-ethics scholarship. The methodological aim is analytical transparency rather than bibliometric comprehensiveness.

The primary scriptural corpus is Qurʾān 96:1–5. The exegetical reconstruction gives priority to four major classical commentaries representing different interpretive emphases: al-Ṭabarī‘s Jāmi’ al-Bayān, Ibn Kathīr’s Tafsīr al-Qur‘ān al-’Aẓīm, al-Qurṭubī’s Al-Jāmi’ li-Aḥkām al-Qur‘ān, and al-Rāzī's Mafātīḥ al-Ghayb. Their discussions of the passage are located in al-Ṭabarī, vol. 24, pp. 517–520; Ibn Kathīr, vol. 8, pp. 436–437; al-Qurṭubī, vol. 20, pp. 119–121; and al-Rāzī, vol. 32, pp. 215–218. These sources are not treated as interchangeable. Al-Ṭabarī foregrounds transmitted interpretation, Ibn Kathīr integrates report-based exegesis with the narrative of first revelation, al-Qurṭubī develops the civilizational significance of writing, and al-Rāzī offers sustained linguistic and philosophical reflection on the movement from human origin to knowledge.[6]

The contemporary comparison corpus consists of influential and directly relevant works on HCAI, responsible AI, AI ethics, meaningful human control, algorithmic accountability, and Islamic AI ethics. Sources were selected because they either define HCAI, synthesize global AI-ethics principles, criticize principle-based governance, articulate meaningful human control, or address AI through an Islamic epistemological lens. The study therefore uses benchmark concepts rather than attempting to represent every strand of AI ethics.

The analysis proceeds in three stages. The first stage is exegetical reconstruction. It identifies the lexical and thematic functions of *iqra‘*, *bi-smi rabbika*, *khalaqa al-insāna min ’alaq*, *al-akram*, *al-qalam*, and *’allama*, and compares how the selected exegetes explain them. This stage asks what claims can reasonably be attributed to the passage within the exegetical tradition. The second stage is comparative conceptual mapping. Reconstructed meanings are placed in dialogue with HCAI categories such as human control, dignity, accountability, fairness, epistemic reliability, and technological mediation. The third stage is normative synthesis, in which the article formulates QHCAI dimensions and operational questions for evaluating AI systems.

A crucial methodological distinction is maintained between textual meaning and contemporary significance. When the article relates al-qalam to digital technologies, for example, it does not claim that the Qur‘ān predicts computers or artificial intelligence. The relationship is a functional contemporary analogy: just as the qalam is interpreted as an

[6] Muḥammad ibn Jarīr al-Ṭabarī, *Jāmiʿ al-Bayān ʿan Taʾwīl Āy al-Qurʾān*, ed. ʿAbd Allāh ibn ʿAbd al-Muḥsin al-Turkī (Cairo: Dār Hajr, 2001), 24:517–520; Ismāʿīl ibn ʿUmar Ibn Kathīr, *Tafsīr al-Qurʾān al-ʿAẓīm*, ed. Sāmī ibn Muḥammad Salāmah, 2nd ed. (Riyadh: Dār Ṭayyibah, 1999), 8:436–437; Muḥammad ibn Aḥmad al-Qurṭubī, *Al-Jāmiʿ li-Aḥkām al-Qurʾān*, ed. Aḥmad al-Bardūnī and Ibrāhīm Aṭfaysh (Cairo: Dār al-Kutub al-Miṣriyyah, 1964), 20:119–121; Fakhr al-Dīn al-Rāzī, *Mafātīḥ al-Ghayb*, 3rd ed. (Beirut: Dār Iḥyāʾ al-Turāth al-ʿArabī, 1999), 32:215–218.

instrument that records, preserves, and communicates knowledge, contemporary digital systems can be assessed as technologically different instruments that mediate knowledge. The analogy therefore operates at the level of function and ethical evaluation, not lexical identity or historical prediction.

To reduce dependence on a single interpretive voice, the study uses source triangulation through comparison among the four classical tafsīr works and relevant contemporary scholarship. It also applies a negative control: a proposed contemporary principle is not treated as a direct Qurʾanic meaning merely because it is normatively attractive. Where the textual basis is indirect—as in the move from human creatureliness to modern language of human dignity—the article explicitly marks that move as a conceptual synthesis rather than a lexical conclusion. This procedure is intended to minimize eisegesis while preserving the possibility of responsible contemporary interpretation.

Results and Discussion

Exegetical Reconstruction: Knowledge, Creatureliness, and Mediation in Qur'ān 96:1–5

The first five verses of al-ʿAlaq form a compact sequence in which command, creation, generosity, mediation, and learning are tightly connected. The passage begins with *iqraʿ bi-smi rabbika alladhī khalaq*, continues by specifying the creation of the human being from 'alaq, repeats the command to read in relation to the Lord's superlative generosity, and then explains that God teaches by the qalam and teaches the human being what he did not know. Read as a sequence, the verses do not present knowledge as an isolated human achievement. Knowing is framed by dependence, mediation, and Divine teaching.

The command iqraʾ is traditionally connected with recitation and the opening event of revelation. Al-Ṭabarī reads the imperative as God's command to the Prophet to recite what is revealed, while Ibn Kathīr situates the verse within the account of the first revelation in the cave of Ḥirāʿ. Al-Rāzī gives particular attention to bi-smi rabbika and considers several grammatical and theological possibilities, including reading in reliance upon or for the sake of the Lord. The important point for this study is not to inflate iqraʾ into a modern theory of information. Rather, the command embeds cognitive activity within an orientation: reading and knowing are not normatively self-sufficient activities.[7]

This provides the first basis of the proposed framework, Transcendental Orientation. The contemporary significance of the verse is not that every technological action must be verbally sacralized, but that knowledge production is subject to a purpose beyond technical capability. In AI terms, the question 'Can a system perform this task?' is therefore normatively incomplete. It must be accompanied by questions concerning the purposes for which a system is built, the goods it is expected to serve, and the responsibilities that remain with human actors. This move from capability to purpose parallels a major concern within AI ethics, where technical performance alone is insufficient to establish social legitimacy.

The second verse, khalaqa al-insāna min ʿalaq, is equally important but should be interpreted with greater restraint than in the earlier version of this article. The verse does not lexically state a theory of human dignity. Al-Ṭabarī and al-Qurṭubī explain ʿalaq in relation to

[7] Al-Rāzī, *Mafātīḥ al-Ghayb*, 32:215–216; Al-Ṭabarī, *Jāmiʿ al-Bayān*, 24:517–518; Ibn Kathīr, *Tafsīr al-Qurʾān al-ʿAẓīm*, 8:436–437.

the human being's created origin, while al-Rāzī draws a striking contrast between the lowly beginning of human material existence and the later capacity for knowledge. The more direct exegetical implication is therefore creatureliness and dependence, not modern rights language.[8]

Nevertheless, creatureliness has a significant contemporary ethical implication when placed in dialogue with HCAI. It resists technological hubris by refusing to equate human worth with computational performance, productivity, or informational mastery. Human dignity enters the framework here as a normative correlate, not as a direct synonym of 'alaq: because the human subject is more than a bundle of cognitive functions, neither human value nor moral status can be reduced to a competition between biological and artificial intelligence. The adjacent verses 6-7 sharpen this point by warning that the human being may transgress when perceiving himself as self-sufficient. In the context of AI, epistemic humility therefore applies first to humans—the designers, institutions, and users who might mistake technical power for moral authority.[9]

The repetition iqraʾ wa-rabbuka al-akram in verse 3 changes the tone from command to Divine generosity. Classical exegetes connect al-akram with God's generosity in teaching and elevating the human being through knowledge. Al-Rāzī in particular relates the movement from created origin to knowledge to the superlative generosity of God. The verse does not itself articulate a complete theory of justice, fairness, or beneficence. Yet within the sequence, knowledge appears as a gift and responsibility rather than a morally neutral possession. This supports a contemporary principle of Ethical Beneficence: knowledge and technology should be evaluated by the human goods they enable and the harms they prevent, not solely by novelty, profit, or efficiency.[10]

Verse 4, *alladhī ʿallama bi-l-qalam*, provides the clearest basis for discussing technological mediation. Al-Qurṭubī interprets the *qalam* in relation to writing and emphasizes the enormous social importance of recording knowledge: without writing, religious and worldly affairs, transmitted wisdom, and the preservation of learning would be severely impaired. Al-Rāzī likewise understands the qalam as a means by which absent knowledge is preserved and transmitted, describing writing as an instrument that makes knowledge endure across distance and time. The exegetical point is therefore mediation: knowledge is not only possessed inwardly but is externalized through tools and practices.[11]

The analogy to AI must remain functional. Contemporary AI is not 'the qalam' in a hidden or prophetic sense, and the verse is not a forecast of machine learning. AI can, however, be assessed as one among many historically contingent knowledge-mediating instruments. This analogy immediately imposes a hierarchy: the tool serves a humanly accountable process of knowing; it does not become the autonomous source of moral authority. The analogy also exposes a difference. A pen does not rank, infer, or generate content, whereas contemporary AI systems can transform the information environment

[8] Al-Qurṭubī, *Al-Jāmiʿ li-Aḥkām al-Qurʾān*, 20:119; al-Rāzī, *Mafātīḥ al-Ghayb*, 32:216–217.

[9] Al-Rāzī, *Mafātīḥ al-Ghayb*, 32:218–219.

[10] Fakhr al-Dīn al-Rāzī, *Mafātīḥ al-Ghayb*, 3rd ed. (Beirut: Dār Iḥyāʾ al-Turāth al-ʿArabī, 1999), 32:215–218.

[11] Al-Rāzī, *Mafātīḥ al-Ghayb*, 32:217–218; Al-Qurṭubī, *Al-Jāmiʿ li-Aḥkām al-Qurʾān*, 20:120.

through probabilistic outputs, recommendations, and automated classifications. For that reason, the ethical demands of oversight, provenance, verification, and accountability are stronger, not weaker.

Verse 5, *'allama al-insāna mā lam ya'lam*, completes the sequence by emphasizing that the human being comes to know what was previously unknown. Al-Ṭabarī links the verse to learning through writing and other forms of Divine teaching, while al-Qurṭubī emphasizes the transition from ignorance to knowledge; al-Rāzī reads the verse as completing the relation between teaching by the qalam and human acquisition of previously unknown knowledge. The verse therefore destabilizes any image of the human knower as epistemically self-originating. Knowledge is acquired, mediated, contingent, and subject to limits.[12]

From this comes the fifth dimension, Epistemic Responsibility under Divine Sovereignty. Its contemporary implication is not that every AI output should be rejected because it is machine-generated. Rather, no machine output acquires independent moral or epistemic authority merely by being produced at scale or with high statistical confidence. Human beings and institutions remain responsible for verification, interpretation, contextual judgment, and the consequences of acting on machine-generated information. In religious interpretation, this limit is especially important because computational fluency cannot by itself supply historical consciousness, moral intention, or responsibility for authoritative interpretation.[13]

The exegetical reconstruction therefore produces a more differentiated result than the previous five-principle formulation. Qur'ān 96:1–5 directly supports a sequence of transcendental orientation, creaturely dependence, knowledge as Divine generosity, technological mediation through the qalam, and human learning under Divine teaching. Human dignity, fairness, accountability, and meaningful human control enter at the second stage of analysis, where these reconstructed meanings are placed in dialogue with contemporary HCAI rather than projected back into the lexical content of the verses.

Critical Dialogue with Human-Centered AI: From Anthropocentric Control to Theocentric Human Responsibility

HCAI is not a single theory but a family of design and governance approaches that seek to ensure that AI systems enhance rather than displace meaningful human agency. Shneiderman's influential formulation rejects the assumption that high automation necessarily requires low human control. Instead, reliable, safe, and trustworthy systems can combine high levels of automation with high levels of human control when design preserves responsibility and user competence. This is a useful benchmark for the Qurʾanic framework because it treats technology as augmentative rather than inherently substitutive.[14]

[12] Al-Ṭabarī, *Jāmiʿ al-Bayān*, 24:519–520; al-Qurṭubī, *Al-Jāmiʿ li-Aḥkām al-Qurʾān*, 20:120–121; al-Rāzī, *Mafātīḥ al-Ghayb*, 32:218.

[13] Mohd Azhar et al., "Ethics and Limits of Artificial Intelligence," 112–120; Ahmad, "Artificial Intelligence in Qurʾanic Interpretation," 70–76.

[14] Shneiderman, "Human-Centered Artificial Intelligence," 495–504; Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens, "A Model for Types and Levels of Human Interaction with Automation," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 30, no. 3 (2000): 286–297.

A second benchmark is meaningful human control. Santoni de Sio and van den Hoven argue that autonomous systems can be compatible with responsibility only when human actors remain meaningfully connected to the reasons and capabilities that guide system behavior. This concern is reinforced by research on 'moral crumple zones,' where human operators may be blamed for failures even when system design and institutional structures have deprived them of meaningful control. QHCAI therefore cannot be satisfied by a ceremonial human-in-the-loop. It requires traceable authority, practical ability to intervene, and institutional allocation of responsibility.[15]

A third benchmark concerns the convergence and limitations of AI-ethics principles. International guidelines repeatedly invoke fairness, transparency, privacy, accountability, non-maleficence, and human dignity. Yet the existence of principles does not guarantee their application. Mittelstadt's critique is especially relevant: abstract principles can conceal disagreements over professional duties, institutional incentives, and enforcement mechanisms. QHCAI should therefore be evaluated not by the religiosity of its vocabulary but by whether it can generate questions, constraints, and accountability practices that matter in actual governance.[16]

The Qurʾanic contribution proposed here is best described as theocentric human-centeredness. This formulation preserves the central ethical insight of HCAI—that systems should be designed around human dignity, agency, and responsibility—while rejecting anthropocentrism understood as the assumption that human preference is the ultimate source of value. In a Qurʾanic epistemic horizon, the human subject is central as the accountable user, designer, interpreter, and bearer of consequences, yet human agency is itself morally bounded. This distinction allows religious ethics to supplement rather than negate rights-based and procedural AI governance.[17]

Theocentric human-centeredness also reframes epistemic authority. Contemporary AI systems can produce linguistically persuasive outputs while lacking the social position, intentional responsibility, and accountable judgment normally expected of expert knowers. Research on explainability likewise shows that explanation is not merely a technical property of a model; useful explanation depends on social expectations, contrastive questions, and human interpretation. QHCAI therefore distinguishes computational output from warranted knowledge. An output becomes actionable knowledge only through human processes of

[15] Madeleine Clare Elish, "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction," Engaging Science, Technology, and Society 5 (2019): 40–60; Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account," Frontiers in Robotics and AI 5 (2018): Article 15.

[16] Mittelstadt, "Principles Alone," 501–507; Jobin, Ienca, and Vayena, "Global Landscape," 389–399; UNESCO, Recommendation on the Ethics of Artificial Intelligence (Paris: UNESCO, 2022), paras. 10–14, 35–40.

[17] Elmahjub, "Artificial Intelligence (AI) in Islamic Ethics," Article 73; Khaled Abou El Fadl, "The Epistemology of the Truth in Modern Islam," *Philosophy & Social Criticism* 41, nos. 4–5 (2015): 473–486, esp. 475–478.

verification, provenance checking, domain interpretation, and responsibility for consequences.[18]

This reframing is particularly important because fairness and accountability are sociotechnical rather than purely computational properties. Selbst and colleagues have shown how fairness interventions can fail when technical abstraction removes the social context in which a system is deployed. Similarly, algorithmic auditing must include the full organizational pipeline rather than only model behavior. These findings converge with the QHCAI emphasis on accountable human mediation: ethical responsibility cannot be delegated to an algorithm because the relevant choices include problem formulation, data selection, institutional incentives, deployment conditions, and post-deployment responses.[19]

The relationship between QHCAI and existing HCAI is therefore complementary rather than competitive. Existing frameworks contribute mature governance vocabulary and implementation practices. QHCAI contributes a normative orientation toward purpose, epistemic humility, and accountability beyond human preference. The distinction can be summarized in the following conceptual map.

Existing HCAI / AI-Ethics Concern	**Qurʾanic Interpretive Lens**	**Distinctive QHCAI Contribution**	**Governance Implication**
Human agency and oversight	*Iqraʿ bi-smi rabbika*	Human agency is responsible but morally bounded by a purpose beyond capability.	Define legitimate purpose before automation; preserve substantive human authority.
Human dignity and non-displacement	*khalaqa al-insāna min ʾalaq*	Creatureliness produces epistemic humility; human worth is not reducible to performance.	Avoid designs that erode agency, induce dependency, or treat people as merely data points.
Beneficence, fairness, and non-maleficence	*wa-rabbuka al-akram*	Knowledge is a gift/responsibility whose use must be ethically directed.	Assess distribution of benefits and harms; include affected stakeholders and bias review.
Meaningful human control	*ʾallama bi-l-qalam*	Technology is a knowledge-mediating instrument, not an autonomous moral subject.	Maintain override, traceability, auditability, fallback procedures, and accountable owners.

[18] Tim Miller, "Explanation in Artificial Intelligence: Insights from the Social Sciences," *Artificial Intelligence* 267 (2019): 1–38, esp. 19–27.

[19] Inioluwa Deborah Raji et al., "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York: ACM, 2020), 33–44; Andrew D. Selbst et al., "Fairness and Abstraction in Sociotechnical Systems," in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (New York: ACM, 2019), 59–68.

Epistemic reliability and accountability	*ʾallama al-insāna mā lam yaʿlam*	Human knowledge is acquired and limited; AI output does not become authority by default.	Require provenance, verification, uncertainty communication, and domain-expert validation.

Table 1. Conceptual mapping between established HCAI concerns and the QHCAI contribution.

The QHCAI Framework as an Analytical and Governance Model

The revised QHCAI framework is intended to function as an analytical lens rather than a decorative summary of Qurʾanic values. Its five dimensions are connected but not identical. Transcendental Orientation concerns the purpose and moral direction of knowledge; Creaturely Dignity and Epistemic Humility concern the status and limits of the human subject; Ethical Beneficence concerns the ends toward which knowledge is applied; Instrumental Mediation and Meaningful Human Control concern the proper position of technology; and Epistemic Responsibility concerns validation, authority, and accountability in the production and use of knowledge.

First, Transcendental Orientation asks a prior question that is often absent from technical design: what human and moral good justifies this system? In practice, this dimension requires a purpose statement that is more substantive than efficiency, market share, or novelty. The purpose should identify the human need being addressed, the persons affected, the foreseeable harms, and the conditions under which the technology should not be deployed. This dimension is compatible with risk-management approaches such as the NIST AI Risk Management Framework, which emphasizes governance, mapping context, measuring risk, and managing impacts. The Qurʾanic addition is the insistence that capability remains subordinate to accountable purpose.[20]

Second, Creaturely Dignity and Epistemic Humility establishes a double constraint. The system must protect human dignity and agency, but humans must also resist the fantasy that technological capacity eliminates dependence, uncertainty, or moral limits. This is important in generative AI, where fluent outputs can invite both automation bias and exaggerated claims of machine competence. Human-centered design should therefore support users' ability to question, contest, correct, and refuse AI recommendations. It should also avoid design patterns that conceal uncertainty or intentionally induce over-reliance.

Third, Ethical Beneficence requires that the benefits of AI be judged distributively rather than only aggregatively. A system may produce overall efficiency while systematically burdening marginalized groups, reducing access, or shifting risk to those least able to contest it. Responsible-AI scholarship increasingly frames AI in relation to human flourishing rather than isolated technical metrics. QHCAI places this concern within the moral logic of knowledge as a trust: useful knowledge should contribute to justice, welfare, and the

[20] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1* (Gaithersburg, MD: NIST, 2023), 1–10.

avoidance of harm. This principle requires stakeholder analysis, bias assessment, accessibility review, and ongoing evaluation of actual rather than promised benefits.[21]

Fourth, Instrumental Mediation and Meaningful Human Control places AI firmly within an instrumental order. This is not a claim that AI is simple or passive. Contemporary AI systems can shape options, classify persons, prioritize information, and influence behavior. Precisely because the instrument is powerful, human control must be substantive. Governance should identify who may override the system, who bears responsibility for errors, what logs are retained, how decisions can be appealed, and what fallback procedures apply when confidence is low or the system fails. Ethical reflection on technology also warns against proliferating technology-specific principles without connecting them to institutional responsibility.[22]

Fifth, Epistemic Responsibility concerns the movement from output to warranted judgment. AI-generated content should disclose relevant provenance, uncertainty, and limitations where feasible. High-stakes domains require independent validation and domain expertise; religious interpretation requires additional attention to textual authority, interpretive tradition, and scholarly responsibility. Recent Islamic scholarship on AI and Qurʾanic exegesis similarly insists that AI may assist linguistic search, intertextual mapping, and retrieval while interpretive responsibility remains human. The principle therefore rejects both uncritical automation and a simplistic refusal of computational assistance.[23]

These five dimensions operate reciprocally. A morally attractive purpose cannot excuse unreliable knowledge; accurate information cannot justify the erosion of dignity; human oversight is meaningless if institutions deny users the authority or information necessary to intervene; and transcendental language cannot compensate for unfair or opaque design. The framework is therefore strongest when used as a set of cross-checks rather than as a linear checklist.

Figure 1 represents this revised structure. The arrows indicate conceptual dependence rather than a chronological sequence. Purpose frames the human subject; human humility conditions beneficence; beneficence constrains instrumental use; instrumental use requires human control; and epistemic responsibility validates the knowledge claims that feed back into subsequent decisions. The result is not an autonomous 'Islamic AI' system, but an evaluative architecture for human responsibility in AI-mediated environments.

[21] Bernd Carsten Stahl et al., "Artificial Intelligence for Human Flourishing—Beyond Principles for Machine Learning," *Journal of Business Research* 124 (2021): 374–388.

[22] Henrik Skaug Sætra and John Danaher, "To Each Technology Its Own Ethics: The Problem of Ethical Proliferation," *Philosophy & Technology* 35 (2022): Article 93; Lambèr Royakkers et al., "Societal and Ethical Issues of Digitization," *Ethics and Information Technology* 20, no. 2 (2018): 127–142.

[23] Mohd Azhar et al., "Ethics and Limits of Artificial Intelligence," 112–120; Ahmad, "Artificial Intelligence in Qurʾanic Interpretation," 70–76.

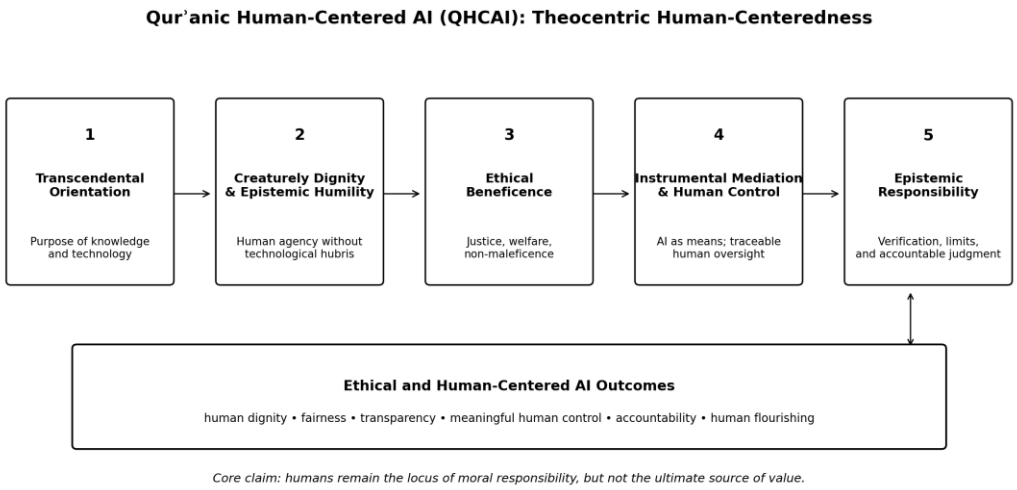

Figure 1. Revised Qurʾanic Human-Centered Artificial Intelligence (QHCAI) framework.

To make the framework usable beyond conceptual description, each dimension can be translated into evaluative questions and observable governance practices. The following matrix does not claim to be a universal certification standard. It is an operationalization for future empirical testing and adaptation in specific sectors.

QHCAI Dimension	Core Evaluative Question	Illustrative Governance Indicators
Transcendental Orientation	What legitimate human good is the system designed to serve, and when should it not be used?	Documented purpose; impact assessment; prohibited-use conditions; stakeholder justification.
Creaturely Dignity & Epistemic Humility	Does the system preserve human agency while communicating the limits of both human and machine judgment?	Contestability; user choice; uncertainty disclosure; anti-dependency design; non-deceptive interface.
Ethical Beneficence	Who benefits, who bears risk, and how are injustice and harm identified and corrected?	Bias and accessibility audits; distributional impact review; grievance mechanism; stakeholder participation.
Instrumental Mediation & Meaningful Human Control	Are humans genuinely able and authorized to supervise, override, and assume responsibility?	Named accountable owner; override/fallback; audit logs; escalation rules; appeal pathway.
Epistemic Responsibility	How are provenance, reliability, and contextual validity verified before outputs become decisions?	Source traceability; expert validation; confidence/uncertainty reporting; documentation; post-deployment monitoring.

Table 2. Operational evaluation matrix for the QHCAI framework.

Theoretical Contribution, Boundaries, and Potential Applications

The theoretical contribution of QHCAI should be stated narrowly but clearly. The framework does not discover fairness, accountability, transparency, or human oversight for the first time; these are well established in international AI ethics. Nor does it claim that Qur'ān 96:1–5 contains a complete governance code. Its contribution is to alter the normative architecture within which those principles are understood. Human-centeredness becomes theocentric human responsibility: humans are central because they are accountable agents who design, interpret, authorize, and bear consequences, not because human preference is metaphysically ultimate.[24]

This shift produces two concepts that can enrich contemporary HCAI. The first is epistemic humility. AI ethics often addresses uncertainty at the level of model performance, but QHCAI adds a moral anthropology of limited knowers. Designers, experts, religious scholars, and users are all fallible, and technological sophistication does not abolish that fallibility. Epistemic humility therefore supports practices of verification, openness to correction, transparent uncertainty, and refusal to overstate system competence. The second is transcendental accountability. Existing accountability mechanisms ask who is legally, organizationally, or professionally responsible. QHCAI accepts these layers but adds that responsibility is also moral: the goals pursued through AI require justification beyond institutional convenience or user desire.

These additions do not make QHCAI anti-modern or anti-technological. The qalam motif itself supports a positive account of mediated knowledge, while the warning against human self-sufficiency prevents technological enthusiasm from becoming an ideology of mastery. This balance is important because digitalization produces both social benefit and new ethical risks, including surveillance, dependency, informational asymmetry, and the redistribution of responsibility. The framework therefore treats technological development as conditionally valuable: it is justified insofar as it enhances responsible human flourishing under morally defensible purposes.[25]

The potential applications are deliberately presented as future domains of inquiry rather than empirically demonstrated outcomes. In AI-assisted education, QHCAI would support systems that augment learning while preserving students' critical reasoning, authorship, and capacity to contest machine feedback. In Qurʾanic studies, it would permit AI-assisted search, text alignment, linguistic analysis, and retrieval while requiring scholars to retain interpretive responsibility and verify sources. In public-sector or institutional decision systems, it would require that consequential judgments remain contestable and traceable to identifiable human and organizational actors.[26]

The framework also sets boundaries on delegation. Low-stakes repetitive tasks may permit extensive automation, whereas decisions affecting rights, religious authority, educational evaluation, or access to essential services require stronger human control and evidentiary standards. This is consistent with contemporary governance approaches that scale

[24] Jobin, Ienca, and Vayena, "Global Landscape," 389–399; Mittelstadt, "Principles Alone," 501–507.

[25] Sætra and Danaher, "To Each Technology Its Own Ethics," Article 93; Royakkers et al., "Societal and Ethical Issues," 127–142.

[26] Mohd Azhar et al., "Ethics and Limits of Artificial Intelligence," 112–120; Ahmad, "Artificial Intelligence in Qurʾanic Interpretation," 70–76.

oversight according to risk. QHCAI adds that the threshold for delegation is not only a technical question about model accuracy but also a normative question about which responsibilities should remain irreducibly human.[27]

Finally, QHCAI should not be treated as a substitute for sector-specific law, engineering standards, or professional ethics. A Qurʾanic framework can orient purposes and responsibilities, but it cannot independently specify cybersecurity controls, statistical validation thresholds, data-protection obligations, or institutional procedures. Responsible implementation therefore requires dialogue among Islamic ethics, HCAI, law, social science, and technical disciplines. This interdisciplinary boundary is a strength rather than a weakness because it prevents religious concepts from being used as labels that bypass established governance expertise.

Conclusion

This study proposes a revised Qurʾanic Human-Centered Artificial Intelligence (QHCAI) framework through a methodologically controlled reading of Qurʻān 96:1–5. The exegetical reconstruction shows that the passage connects knowledge to Divine orientation, human creatureliness, generosity, mediated learning through the qalam, and the acquisition of previously unknown knowledge. The study deliberately distinguishes these textual-exegetical meanings from contemporary normative synthesis. Accordingly, concepts such as human dignity, fairness, meaningful human control, and algorithmic accountability are not claimed as direct lexical meanings of the verses; they emerge through critical dialogue between the reconstructed Qurʾanic horizon and established HCAI scholarship. The resulting framework consists of five dimensions: Transcendental Orientation, Creaturely Dignity and Epistemic Humility, Ethical Beneficence, Instrumental Mediation and Meaningful Human Control, and Epistemic Responsibility under Divine Sovereignty. Its central theoretical contribution is the concept of theocentric human-centeredness. QHCAI does not replace existing responsible-AI principles but reframes their normative foundation: human beings remain the locus of moral and epistemic responsibility while neither human preference nor machine output is treated as the ultimate source of value or truth. This reframing yields practical evaluative questions concerning purpose, dignity, distribution of benefit and harm, substantive human control, provenance, verification, and accountability. The study has clear limitations. Its Qurʾanic corpus is intentionally restricted to the first five verses of al-ʾAlaq, the framework is normative-conceptual rather than empirically validated, and it does not provide sector-specific technical standards. Future research should therefore test the QHCAI matrix in concrete settings such as AI-assisted education, Qurʾanic research tools, public-sector decision support, and institutional governance. Comparative work should also place Qurʻān 96:1–5 in dialogue with a wider Qurʾanic anthropology of human dignity, stewardship, justice, consultation, and accountability. The value of QHCAI will ultimately depend not on the religious resonance of its terminology but on whether it can discipline technological judgment, preserve meaningful human responsibility, and contribute to ethically defensible AI practice.

[27] NIST, *Artificial Intelligence Risk Management Framework*, 1–10.

Bibliography

Abou El Fadl, Khaled. "The Epistemology of the Truth in Modern Islam." *Philosophy & Social Criticism* 41, nos. 4–5 (2015): 473–486. https://doi.org/10.1177/0191453715577739.

Ahmad, Nehru Millat. "Artificial Intelligence in Qurʾanic Interpretation: A Theological and Ethical Analysis." *Journal of Islamic Thought and Philosophy* 5, no. 1 (2026): 60–79. https://doi.org/10.15642/jitp.2026.5.1.60-79.

Capel, Tara, and Margot Brereton. "What Is Human-Centered about Human-Centered AI? A Map of the Research Landscape." *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023): 1–23. https://doi.org/10.1145/3544548.3580959.

Dignum, Virginia. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Cham: Springer, 2019.

Elish, Madeleine Clare. "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction." *Engaging Science, Technology, and Society* 5 (2019): 40–60. https://doi.org/10.17351/ests2019.260.

Elmahjub, Ezieddin. "Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI." *Philosophy & Technology* 36 (2023): Article 73. https://doi.org/10.1007/s13347-023-00668-x.

European Commission, High-Level Expert Group on Artificial Intelligence. *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission, 2019.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28 (2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Hartono, Hartono, and Ismail. "Iqraʾ as Knowledge Paradigm: Islamic Epistemology Reconstruction from QS Al-Alaq 1–5." *Maghza: Jurnal Ilmu Al-Qur'an dan Tafsir* 11, no. 1 (2026): 1–27. https://doi.org/10.24090/maghza.v11i1.16050.

IEEE. *Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems*. 1st ed. Piscataway, NJ: IEEE, 2019.

Ibn Kathīr, Ismāʿīl ibn ʿUmar. *Tafsīr al-Qurʾān al-ʿAẓīm*. Edited by Sāmī ibn Muḥammad Salāmah. 2nd ed. Riyadh: Dār Ṭayyibah, 1999.

Jobin, Anna, Marcello Ienca, and Effy Vayena. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1 (2019): 389–399. https://doi.org/10.1038/s42256-019-0088-2.

Miller, Tim. "Explanation in Artificial Intelligence: Insights from the Social Sciences." *Artificial Intelligence* 267 (2019): 1–38. https://doi.org/10.1016/j.artint.2018.07.007.

Mittelstadt, Brent. "Principles Alone Cannot Guarantee Ethical AI." *Nature Machine Intelligence* 1 (2019): 501–507. https://doi.org/10.1038/s42256-019-0114-4.

Mohd Azhar, Muhammad Hazim, Mohamad Fikri Mohd Bakri, Khadher Ahmad, and Muhammad Ikhlas Rosele. "Ethics and Limits of Artificial Intelligence (AI) in Quranic Exegesis According to the Epistemological Framework of Islamic Knowledge." *International Journal of Quranic Research* 17, no. 2 (2025): 97–124. https://doi.org/10.22452/quranica.vol17no2.5.

Mujahidin, Anwar. "Epistemologi Islam: Kedudukan Wahyu Sebagai Sumber Ilmu." *Ulumuna* 17, no. 1 (2017): 41–64. https://doi.org/10.20414/ujis.v17i1.171.

National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1*. Gaithersburg, MD: NIST, 2023. https://doi.org/10.6028/NIST.AI.100-1.

Nizamani, Mir Muhammad, Hai-Li Zhang, and Zhongping Lai. "Human-Centered AI: Advancing Ethical, Transparent, and Context-Aware Systems for Sustainable Development." *Technology in Society* 84 (2026): 103121. https://doi.org/10.1016/j.techsoc.2025.103121.

Parasuraman, Raja, Thomas B. Sheridan, and Christopher D. Wickens. "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 30, no. 3 (2000): 286–297. https://doi.org/10.1109/3468.844354.

Al-Qurṭubī, Muḥammad ibn Aḥmad. *Al-Jāmiʿ li-Aḥkām al-Qurʾān*. Edited by Aḥmad al-Bardūnī and Ibrāhīm Aṭfaysh. Cairo: Dār al-Kutub al-Miṣriyyah, 1964.

Raji, Inioluwa Deborah, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." *In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. New York: ACM, 2020. https://doi.org/10.1145/3351095.3372873.

Al-Rāzī, Fakhr al-Dīn. *Mafātīḥ al-Ghayb*. 3rd ed. Beirut: Dār Iḥyāʾ al-Turāth al-ʿArabī, 1999.

Royakkers, Lambèr, Jelte Timmer, Linda Kool, and Rinie van Est. "Societal and Ethical Issues of Digitization." *Ethics and Information Technology* 20, no. 2 (2018): 127–142. https://doi.org/10.1007/s10676-018-9452-x.

Sætra, Henrik Skaug, and John Danaher. "To Each Technology Its Own Ethics: The Problem of Ethical Proliferation." *Philosophy & Technology* 35 (2022): Article 93. https://doi.org/10.1007/s13347-022-00591-7.

Santoni de Sio, Filippo, and Jeroen van den Hoven. "Meaningful Human Control over Autonomous Systems: A Philosophical Account." *Frontiers in Robotics and AI* 5 (2018): Article 15. https://doi.org/10.3389/frobt.2018.00015.

Selbst, Andrew D., danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. "Fairness and Abstraction in Sociotechnical Systems." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. New York: ACM, 2019. https://doi.org/10.1145/3287560.3287598.

Shneiderman, Ben. "Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy." *International Journal of Human–Computer Interaction* 36, no. 6 (2020): 495–504. https://doi.org/10.1080/10447318.2020.1741118.

Stahl, Bernd Carsten, Andreas Andreou, Paul Brey, Tally Hatzakis, Ana Kirichenko, Kevin Macnish, Shaelagh Laulhé Shaelou, Rowena Patel, Mark Ryan, and David Wright. "Artificial Intelligence for Human Flourishing—Beyond Principles for Machine Learning." *Journal of Business Research* 124 (2021): 374–388. https://doi.org/10.1016/j.jbusres.2020.11.030.

Al-Ṭabarī, Muḥammad ibn Jarīr. *Jāmiʿ al-Bayān ʿan Taʾwīl Āy al-Qurʾān*. Edited by ʿAbd Allāh ibn ʿAbd al-Muḥsin al-Turkī. Cairo: Dār Hajr, 2001.

UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO, 2022.