CONCEPTUAL PAPER

Leading in the AI Era: The HUMAN Framework

Jawad Syed, PhD | Professor of Leadership, Lahore University of Management Sciences (LUMS)

Citation: Syed, J. (2026). Leading in the AI Era: The HUMAN Framework. *SAAM Working Paper 202608.* South Asian Academy of Management.

Abstract

Artificial intelligence (AI) is increasing the speed, scale and apparent precision of organisational decisions, but it cannot determine which purposes are legitimate, whose interests deserve protection, or who should answer for consequences. This paper proposes a distinctive formulation of a HUMAN leadership framework for AI-supported organisational decision-making. HUMAN refers to five connected leadership responsibilities: Human dignity first; Understand AI and uncertainty; Moral judgment and meaningful purpose; Accountability and authenticity; and Nurture trust, voice and connection. Drawing on ethical and responsible leadership, human-centred AI, psychological safety, and research on human–AI decision making, the framework converts broad principles into five linked leadership obligations. Its central proposition is that AI may inform action, but leaders must retain responsibility for purpose, judgment, contestability and consequences. The paper explains the framework, distinguishes it from technical governance models, and proposes a practical decision protocol and research agenda for organisations across sectors.

Keywords: artificial intelligence; ethical leadership; human-centred AI; psychological safety; responsible leadership

1. Why AI makes leadership more consequential

AI is becoming embedded in hiring, healthcare, finance, education, public administration, emergency management and other areas of organisational life. It can search large datasets, recognise patterns, generate alternatives and compress decision time. Yet its outputs are probabilistic products of data, objectives and design choices; they are not independent moral judgments. As AI capability grows, the leadership question is therefore not whether humans or machines are superior in the abstract. It is: what must remain human when consequential decisions are AI-supported?

This question matters because speed is not wisdom, prediction is not judgment, and technical capability is not moral legitimacy. A credit model may predict default without deciding whether its proxies are fair. A clinical system may rank diagnoses without understanding a patient's life or bearing responsibility for harm. An early-warning model may forecast a flood, but only leadership can ensure that an intelligible warning reaches vulnerable communities and triggers evacuation. In each case, a technically accurate output can coexist with a leadership failure.

The demand for AI capability and human leadership is rising simultaneously. The World Economic Forum (2025) reports that employers expect 39% of workers' core skills to change by 2030. Compared with the 2023 survey, the proportion of employers identifying leadership and social influence as core skills increased by 22 percentage points; the corresponding increase for AI and big data was 17 percentage points. The implication is not that leadership compensates for technical ignorance. Leaders need both AI literacy and the capacity to make ethically defensible choices under uncertainty.

2. The human–AI performance paradox

The popular formula "human plus AI" is too optimistic. In a preregistered meta-analysis of 106 experiments and 370 effect sizes, Vaccaro et al. (2024) found that human–AI combinations performed better than humans alone, on average, but worse than the better of either humans or AI alone (Hedges' g = −0.23). Losses were particularly evident in decision tasks; results were more promising for content-creation tasks. The finding does not justify rejecting AI. It shows that complementarity must be deliberately designed rather than assumed.

Several mechanisms explain failure. People may reject algorithmic advice after observing it make an error, even when the algorithm still outperforms human judgment (Dietvorst et al., 2015). Conversely, users may rely too heavily on automated recommendations or fail to understand what a model's confidence score represents (Lee & See, 2004). Models may be trained on stale, incomplete or historically biased data. Workflows may assign the wrong subtask to the machine, while accountability becomes diffuse because "the algorithm" appears to have made the decision. Raisch and Krakowski's (2021) automation–augmentation paradox likewise shows that augmentation and automation are interdependent organisational choices, not clean alternatives. Shrestha et al. (2019) argue that decision structures must be matched to task specificity, interpretability, scale, speed and replicability.

A useful division of labour may be as follows. AI is generally advantaged in volume, consistency, pattern recognition, option generation and speed. Human leaders remain responsible for selecting what matters, interpreting context, defining legitimate purpose, protecting verification time, resolving value conflicts and owning consequences (Jarrahi, 2018). The proposed division may therefore be understood as a normative allocation of responsibility, rather than as a permanent claim about the relative capabilities of humans and machines.

3. The HUMAN Leadership Framework

The HUMAN framework, proposed in this paper, synthesises five literatures that are too often separated: human-centred AI and human rights; AI literacy and uncertainty; ethical and responsible leadership; accountability and authentic conduct; and psychological safety and employee voice.

TABLE 1. HUMAN Framework of Leadership in the AI Era

	Leadership obligation	Diagnostic question
H	**Human dignity first**	Whose dignity, rights, safety or livelihood is affected—and who may be invisible in the data?
U	**Understand AI and uncertainty**	What is the model good at, how uncertain is this output, and what could the system or I be missing?
M	**Moral judgment and meaningful purpose**	Is the action necessary, proportionate, lawful and consistent with the organisation's legitimate purpose?
A	**Accountability and authenticity**	Who owns the decision, can explain it, records it, and will answer for its consequences?
N	**Nurture trust, voice and connection**	Who can challenge the recommendation safely, and whose experiential or technical knowledge must be heard?

The framework originated in research on ethical and human-centred leadership and was refined through discussions with executives and public-sector professionals confronting AI-supported decisions. These discussions repeatedly returned to five concerns: the human consequences of a decision, uncertainty in the system, the purpose being served, responsibility for the outcome, and whether people felt able to question an AI-supported recommendation. The framework was first presented in August 2026 during my guest lecture at a leadership seminar hosted by the Pakistan Navy, where it was applied to AI-supported decision-making in naval and wider military settings. Although initially presented in a defence context, its underlying leadership dimensions are relevant to consequential decisions across public, private and non-profit organisations.

Extant research and international reports already emphasise dignity, human oversight, transparency, fairness, safety and accountability (NIST, 2023; OECD, 2024; UNESCO, 2021; WHO, 2021). Leadership research explains how values are modelled, relationships are governed and dissent is enabled (Brown et al., 2005; Edmondson, 1999; Maak & Pless, 2006). The HUMAN framework integrates these principles into a memorable, leader-facing architecture for decisions before, during and after AI use. Unlike technical risk frameworks, HUMAN is explicitly oriented toward the leader's lived decision process, connecting each governance principle to a concrete question that can be asked at the point of action.

3.1 H — Human dignity first

Human dignity requires leaders to see persons behind data points and to treat fundamental rights, autonomy and equitable treatment as design constraints rather than after-the-fact ethical decoration. Dignity asks who bears risk, who receives benefit, who is excluded from training data, and whether a person can understand and contest a consequential decision. This obligation is grounded in human-centred values in the UNESCO (2021) and OECD (2024) principles and in responsible leadership's stakeholder orientation (Maak & Pless, 2006). It also broadens operational awareness: empathy can reveal consequences that aggregate performance metrics hide.

Leadership practice: require an impact map identifying affected groups, plausible harms and inequalities, distributional effects and accessible routes for appeal. For high-impact decisions, prohibit reduction of individuals to a score without contextual review.

3.2 U — Understand AI and uncertainty

AI literacy for leaders is neither coding expertise nor blind confidence in technical specialists. It is the capacity to ask what data were used, how current and representative they are, what the output actually means, where performance degrades, and what information the system cannot observe. A high confidence score is not high certainty that the contemplated action is right. Leaders must distinguish predictive accuracy from causal understanding, model confidence from real-world uncertainty, and explanation from justification. For example, NIST's (2023) govern–map–measure–manage logic provides a useful structure for identifying and managing AI risks. However, leaders must also examine their own limitations, assumptions and incentives—for example, whether pressure for speed, efficiency or favourable results is encouraging them to overlook uncertainty, contrary evidence or potential harm.

Leadership practice: establish thresholds for independent verification, monitor drift and subgroup performance, document known limitations, and design calibrated reliance rather than a ceremonial "human in the loop." Meaningful oversight requires time, authority, competence and the option to pause or override.

3.3 M — Moral judgment and meaningful purpose

An AI system optimises a specified objective; it cannot establish that the objective is worthy. Leaders must ask whether the use and proposed action are necessary, proportionate and consistent with law, professional duties and the organisation's mission. This is the domain of practical judgment: reconciling competing goods, considering longer-term and system-wide effects, and refusing technically feasible actions that violate legitimate purpose. Ethical leadership is enacted through both the "moral person" and "moral manager"—personal conduct, communication, reinforcement and decision processes (Brown et al., 2005).

Leadership practice: articulate the purpose before viewing the recommendation; test alternatives against necessity, proportionality, reversibility and precedent; and preserve explicit "do not automate" boundaries for decisions involving irreducible moral or relational responsibility.

3.4 A — Accountability and authenticity

Accountability cannot be delegated to an algorithm, vendor or data-science team. It requires a named decision owner, a traceable record of evidence and dissent, an explanation appropriate to affected stakeholders, and mechanisms for remedy. Authenticity adds behavioural consistency: leaders must not publicly claim human control while privately rewarding unquestioning reliance, speed at any cost or plausible deniability. Algorithmic systems can redistribute control and obscure responsibility, making clear governance essential (Kellogg et al., 2020).

Leadership practice: assign an accountable executive and operational owner; record model version, material inputs, human modifications and rationale; audit outcomes; disclose material AI use where appropriate; and maintain incident, correction and redress procedures.

3.5 N — Nurture trust, voice and connection

AI-supported decisions often fail not because no one saw the anomaly, but because the person who saw it could not speak, was not heard, or assumed the system outranked experience. Psychological safety—the shared belief that interpersonal risk-taking is safe—supports learning behaviour (Edmondson, 1999). In the AI era, voice must extend to junior staff, technical specialists, frontline professionals and affected communities. Trust in an AI-supported process should not depend on an assumption that the system is always right. It should depend on whether errors can be identified, challenged and corrected.

Leadership practice: create protected challenge channels, invite a pre-decision dissenting view, separate respectful challenge from insubordination, and close the feedback loop by explaining what was accepted, rejected and learned.

4. HUMAN as a decision and governance protocol

HUMAN is designed for use at three moments. Before deployment, leaders define purpose, affected stakeholders, prohibited uses, required evidence and ownership. During a consequential decision, they run the five diagnostic questions and apply a controlled challenge-and-verification process. After action, they review outcomes, distributional effects, near misses, overrides and stakeholder feedback. The framework therefore complements rather than replaces technical testing, legal compliance, cybersecurity, model documentation and risk management.

A practical protocol is: (1) pause when stakes or uncertainty exceed the approved threshold; (2) surface the recommendation, confidence and missing data; (3) seek an independent or dissenting assessment; (4) compare at least two feasible actions, including delay or non-use; (5) decide with named ownership and recorded reasons; and (6) review consequences and provide remedy. The protocol preserves speed for routine, reversible decisions while increasing deliberation as stakes, irreversibility and vulnerability rise.

Consider three cross-sector applications. In recruitment, HUMAN shifts attention from predictive fit to discriminatory proxies, candidates' right to explanation, legitimate selection criteria, documented ownership and recruiter voice. In healthcare, it combines diagnostic assistance with patient autonomy, expert interpretation, clinical purpose, named responsibility and a nurse's or patient's ability to challenge. In disaster response, it treats a forecast as incomplete until warnings are verified, translated into accessible communication, connected to evacuation capacity and owned by an accountable authority. Across settings, the same lesson holds: an accurate prediction that does not lead to responsible action is not leadership success.

5. Conceptual contribution and questions for future research

The HUMAN framework seeks to make four contributions. First, it moves beyond the minimal 'human-in-the-loop' metaphor. Merely retaining a human decision-maker does not ensure meaningful oversight; that person must also have the competence, authority, time and independence to question the system and accept responsibility for the decision. Second, HUMAN connects moral and relational approaches to leadership with AI governance. Technical assurance can establish whether a model performs as specified, but it cannot determine whether the purpose being pursued is legitimate or whether organisational pressures are discouraging people from raising concerns. Third, the framework treats trust as a reasoned judgment about whether an AI-supported decision process is competent, transparent and open to challenge and correction—not as unquestioning confidence in the technology. Fourth, the five obligations are connected: dignity is undermined without accountability; accountability is hollow without voice; voice is ineffective without psychological safety. The framework therefore resists treating these as separate boxes.

Questions for future research

As a conceptual framework, HUMAN has not yet been operationalised or empirically validated. Its value therefore remains an open question rather than an established effect. Future research could examine whether its five dimensions are sufficiently distinct to be measured separately and whether, taken together, they offer explanatory value beyond established concepts such as ethical leadership, responsible leadership, psychological safety and AI literacy.

Several practical questions also merit investigation. Does using the HUMAN framework help leaders rely on AI more appropriately, avoiding both uncritical acceptance and reflexive rejection? Does its usefulness vary according to the consequences, ambiguity and reversibility of a decision? Under what conditions do psychological safety and structured dissent help teams identify questionable AI recommendations? Research might also examine whether clearly assigned accountability strengthens stakeholder trust following an AI-related error, and whether attention to human dignity helps leaders identify harms that conventional measures of accuracy or efficiency may overlook.

These questions could be examined through leadership simulations, field experiments, case studies and comparative research across occupations, sectors, levels of hierarchy and national contexts. An important first step would be to develop and validate a multidimensional HUMAN Leadership Scale. Relevant outcomes would include not only decision accuracy and speed, but also fairness, voice, learning, explainability, trust, access to redress and the longer-term legitimacy of AI-supported decisions.

Conclusion

AI changes the instruments of leadership, not its moral centre. Leaders may use machines for computation, pattern recognition, consistency and speed while retaining human responsibility for context, purpose, proportionality, relationships and consequences. The HUMAN framework offers a compact but substantive compass: put Human dignity first; Understand AI and uncertainty; exercise Moral judgment in service of meaningful purpose; preserve Accountability through authentic conduct; and Nurture trust, voice and connection. Its promise is not that every difficult decision will become easy or error-free. It is that consequential decisions will become more transparent, challengeable, responsible and worthy of trust. AI informs. HUMAN leads.

References

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97(2), 117–134. https://doi.org/10.1016/j.obhdp.2005.03.002

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. https://doi.org/10.1037/xge0000033

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. https://doi.org/10.2307/2666999

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons,* 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. Academy of Management Annals, 14(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Maak, T., & Pless, N. M. (2006). Responsible leadership in a stakeholder society: A relational perspective. *Journal of Business Ethics*, 66, 99–115. https://doi.org/10.1007/s10551-006-9047-z

NIST (National Institute of Standards and Technology). (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

OECD (Organisation for Economic Co-operation and Development). (2024). OECD AI principles. https://oecd.ai/en/ai-principles

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66–83. https://doi.org/10.1177/0008125619862257

UNESCO (United Nations Educational, Scientific and Cultural Organization). (2021). Recommendation on the ethics of artificial intelligence. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

Vaccaro, M., Almaatouq, A., & Malone, T. W. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8, 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

World Economic Forum. (2025). The future of jobs report 2025. https://www.weforum.org/publications/the-future-of-jobs-report-2025/

WHO (World Health Organization). (2021). Ethics and governance of artificial intelligence for health: WHO guidance. https://www.who.int/publications/i/item/9789240029200