



Human-Centred Leadership in the Age of Artificial Intelligence: A Multidisciplinary Framework for Trust, Judgment and Organizational Well-Being

Ignacio Bonasa Alzuria

Universidad de la Felicidad y del Bienestar, Mexico

DOI: <https://doi.org/10.68050/JAMS.2026.590>

Received: 21 Sept 2026; Accepted: 22 Sept 2026; Published: 23 September 2026

ABSTRACT

Artificial intelligence is rapidly changing how organizations allocate attention, interpret data and make decisions. Yet the growing technical capacity of organizations does not automatically produce better leadership. This conceptual review develops a human-centred framework for integrating artificial intelligence into leadership practice without weakening judgment, trust, responsibility or employee well-being. Drawing on research in leadership, organizational psychology, human-AI collaboration and responsible AI, the article distinguishes between tasks that can be delegated to computational systems and responsibilities that remain irreducibly human. It proposes four mutually reinforcing leadership capabilities: augmented judgment, relational trust, ethical accountability and purposeful adaptation. The paper argues that AI creates the greatest organizational value when it expands the leader's capacity to listen, anticipate and learn, rather than when it is used simply to accelerate control. Practical implications are presented for executive teams, educators and organizations designing AI-enabled work. The contribution is a multidisciplinary model that treats technological adoption and human flourishing as complementary design objectives rather than competing priorities.

Keywords: human-centred leadership; artificial intelligence; organizational well-being; trust; ethical judgment; human skills

Manuscript type: Conceptual review / integrative review

INTRODUCTION

Artificial intelligence has moved from a specialist technology to an everyday organizational infrastructure. Generative systems draft documents, summarize meetings, support recruitment, detect patterns in customer behaviour and assist managers with decisions. The managerial temptation is therefore understandable: if machines can process more information faster, perhaps leadership itself can become faster, more consistent and more scalable. The difficulty is that leadership is not reducible to information processing. It also involves meaning, moral responsibility, uncertainty, identity, emotion and the quality of relationships.

This distinction matters because the adoption of AI can create a paradox. An organization may become more efficient while its people experience less autonomy, less clarity and less trust. Conversely, technology can remove repetitive work and create time for coaching, creativity and strategic reflection. The difference lies less in the technology itself than in the way leaders frame, govern and use it. Recent scholarship on human-AI collaboration and responsible AI suggests that effective adoption depends on complementary human and machine capabilities, transparent governance and continuing human oversight (Jarrahi, 2018; Raisch & Krakowski, 2021; UNESCO, 2021).

This article asks a practical question: what should remain distinctly human when leadership becomes increasingly AI-enabled? Rather than treating AI as either a threat or a universal solution, the paper develops a



framework for leadership that uses computational capability while protecting the relational and ethical foundations of organizational life.

METHODOLOGY

The study uses an integrative conceptual review. Literature was selected from four intersecting domains: leadership theory, organizational psychology, human-AI collaboration and responsible AI governance. Seminal works were combined with recent institutional frameworks to identify recurring tensions around decision quality, trust, accountability and employee experience. The analysis followed three steps. First, human and computational capabilities were differentiated. Second, recurrent organizational risks associated with automation were grouped thematically. Third, the themes were synthesized into a four-capability framework designed for practical application.

No primary human-subject data were collected. The purpose is theory building and multidisciplinary synthesis rather than statistical generalization. This approach is appropriate where a rapidly changing phenomenon requires concepts from fields that are often studied separately.

RESULTS: A FOUR-CAPABILITY FRAMEWORK

The synthesis produced four capabilities.

Augmented judgment. AI can widen the evidence available to a leader, but evidence is not identical to judgment. Leaders must understand context, challenge outputs, notice what is absent from a dataset and decide when a recommendation should not be followed. The most useful question is not “What does the model say?” but “What decision becomes better because the model has helped us see something we might otherwise miss?”

Relational trust. Leadership is experienced through interactions. Employees judge not only the accuracy of decisions but whether they were heard, respected and treated fairly. When AI mediates evaluation, scheduling or performance management, leaders need to explain its role and preserve meaningful channels for dialogue and appeal. Trust is strengthened when technology is visible enough to be understood but does not become a shield behind which human responsibility disappears.

Ethical accountability. Delegation to a system does not eliminate accountability. Leaders remain responsible for the purposes for which AI is deployed, the quality of data, foreseeable harms and the consequences of decisions. Responsible leadership therefore requires governance before deployment, monitoring during use and correction when outcomes diverge from organizational values.

Purposeful adaptation. AI adoption changes roles and identities, not only workflows. Leaders need to connect technological change to a credible purpose and create learning conditions in which employees can experiment without being humiliated by initial uncertainty. Adaptation becomes more sustainable when people understand both what is changing and why their contribution still matters.

DISCUSSION

The framework challenges two simplistic narratives. The first assumes that human-centred leadership requires resistance to automation. The second assumes that greater automation is inherently progressive. A more productive position is complementarity. Machines are strong at scale, pattern detection, consistency and rapid recombination. Humans remain essential for moral responsibility, embodied empathy, contextual interpretation, political negotiation and the construction of shared meaning.

This has consequences for leadership development. Traditional executive education often separates “digital skills” from “people skills.” AI makes that separation increasingly artificial. A leader who understands technology but cannot create psychological safety may accelerate disengagement. A leader who is highly empathic but unable to interrogate algorithmic systems may unintentionally delegate consequential decisions to tools they do not understand. Human-centred AI leadership therefore requires hybrid literacy.



The framework also reframes well-being. Employee well-being should not be treated as a compensatory programme introduced after technological change. It is a design criterion. Work systems should be evaluated by whether they reduce unnecessary cognitive load, preserve agency, enable learning and support meaningful contribution. This perspective is consistent with self-determination theory, which emphasizes autonomy, competence and relatedness as fundamental psychological needs (Deci & Ryan, 2000).

PRACTICAL IMPLICATIONS

Organizations can operationalize the framework through four routines. First, require a named human decision owner for consequential AI-supported decisions. Second, establish explanation and appeal mechanisms for employees affected by algorithmic recommendations. Third, include well-being and trust indicators alongside productivity metrics in AI pilots. Fourth, train leaders to test outputs, identify bias, communicate uncertainty and conduct difficult conversations that cannot be outsourced.

A simple implementation sequence is: clarify the human purpose; map the decision and its stakeholders; identify what AI can augment; define non-delegable human responsibilities; pilot with feedback; and review both performance and human consequences. This sequence shifts adoption from a tool-first mentality to a purpose-first mentality.

The framework is also consistent with research showing that algorithmic systems redistribute organizational control and reshape decision structures rather than merely automating isolated tasks (Shrestha et al., 2019; Kellogg et al., 2020). Evidence from generative-AI field settings further suggests that productivity gains can be substantial but uneven, reinforcing the need to examine augmentation together with capability development and human oversight (Brynjolfsson et al., 2023). Psychological safety remains relevant because employees must be able to question AI-supported recommendations without interpersonal penalty (Edmondson, 1999).

Scholarly Contribution and Testable Propositions

Table 1 summarizes the proposed framework and its principal analytical dimensions.

Table 1. Analytical framework for human-centred AI leadership

Capability	AI contribution	Non-delegable human responsibility	Indicative outcome
Augmented judgment	Pattern recognition, synthesis, prediction	Contextual interpretation; challenge; final accountability	Decision quality
Relational trust	Information access; communication support	Listening; explanation; dialogue; appeal	Trust and psychological safety
Ethical accountability	Monitoring; anomaly detection	Purpose setting; fairness judgments; responsibility for consequences	Legitimacy
Purposeful adaptation	Personalized learning; workflow support	Meaning-making; role redesign; humane change leadership	Adaptive capacity

Research propositions

P1. AI-enabled leadership will be associated with higher perceived decision quality when algorithmic recommendations are combined with explicit human challenge and contextual interpretation.

P2. Transparency and meaningful appeal mechanisms will strengthen the relationship between AI use and employee trust.

P3. The positive effects of AI augmentation on organizational performance will be more sustainable when well-being indicators are included in governance alongside productivity indicators.



LIMITATIONS AND FUTURE RESEARCH

The article is conceptual and therefore does not establish causal relationships. The proposed capabilities should be tested empirically across sectors, cultures and levels of leadership. Future research could develop validated measures of human-centred AI leadership, examine its relationship with psychological safety and performance, and compare employee reactions when AI use is transparent versus opaque. Longitudinal studies would be particularly valuable because trust may change as employees gain experience with AI-enabled work.

CONCLUSION

AI does not make human leadership obsolete; it makes the quality of human leadership more visible. As computational systems assume a larger role in analysis and execution, leaders become more accountable for purpose, interpretation, trust and ethical boundaries. The central challenge is therefore not choosing between people and technology. It is designing organizations in which technology increases human capability without diminishing human dignity. Leadership with AI is strongest when speed is balanced by reflection, prediction by judgment and automation by responsibility.

REFERENCES

1. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. NBER Working Paper No. 31161.
2. Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
3. Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383.
4. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586.
5. 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.
6. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210.
7. 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.
8. UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. UNESCO.
9. World Health Organization. (2022). World mental health report: Transforming mental health for all. WHO.

DECLARATIONS

Funding: No external funding was received for this study.

Conflict of interest: The author declares no conflict of interest.

Ethics statement: This conceptual/review study did not involve human participants, identifiable personal data, or animal subjects.

Data availability: No new dataset was generated or analysed for this conceptual study.

Originality statement: This manuscript presents an original conceptual synthesis. No fabricated empirical data are reported. All external ideas and sources are attributed in the text and reference list.