E-ISSN: 3108-4192

APSSHS

Academic Publications of Social Sciences and Humanities Studies
2026, Volume 6, Issue 1, Page No: 1-9
Available online at: https://apsshs.com/

Journal of Applied Organizational Systems and Behavior

Responsible Leadership and Human-Centered Governance in Technology-Mediated Organizations

Lucas Meyer[1]*, Anna Schmid[2], Stefan Braun[1], Laura Keller[3]

1. Department of Responsible Leadership and Governance, Faculty of Business, ETH Zurich, Zurich, Switzerland.
2. Department of Human-Centered Management, Faculty of Management, University of Bern, Bern, Switzerland.
3. Department of Technology-Mediated Organizations, Faculty of Psychology, EPFL Lausanne, Lausanne, Switzerland.

Abstract

Technology-mediated organizing increasingly redistributes decision activity, visibility, expertise, and control across managers, employees, and computational systems. Yet prevailing debates often separate responsible leadership, algorithmic accountability, worker inclusion, and human-centered design, leaving organizations without a coherent conceptual account of what leadership responsibility means when consequential work is technologically mediated. This Original Conceptual Framework Article develops a proposed Responsible-Leadership Framework that integrates four governance obligations: ethical orientation, retained accountability, consequential inclusion, and human-centered governance. The framework treats technological capability as distinct from legitimate authority, delegation as distinct from responsibility transfer, formal participation as distinct from consequential influence, and nominal human oversight as distinct from substantively informed review. It further proposes a governance lifecycle linking authorization, configuration, delegation and use, review and contestation, and learning and revision. The framework suggests that responsible leadership operates not as a single leader trait or ethics statement but as an organizational integrating function that shapes how decision rights, review duties, stakeholder influence, and work design are connected. Its propositions remain conditional on task characteristics, expertise, opacity, worker power, organizational readiness, and human–AI coordination. The article therefore offers a testable conceptual architecture rather than a validated causal model or universal prescription. Future research should examine whether the proposed obligations and lifecycle relationships explain differences between symbolic governance and enacted responsibility across organizational contexts. Its value therefore lies in clarifying what must be distinguished, observed, and tested when leadership responsibility is distributed through technological systems.

Keywords: Leadership under technological mediation, Ethical foundations of responsible leadership, Accountability for delegated decisions, Human-centered governance principles, Research implications, Responsible

How to cite this article: Meyer L, Schmid A, Braun S, Keller L. Responsible Leadership and Human-Centered Governance in Technology-Mediated Organizations. J Appl Organ Syst Behav. 2026;6(1):1-9. https://doi.org/10.51847/vjZ12wp6RV

Received: 14 January 2026; **Revised:** 25 March 2026; **Accepted:** 27 March 2026
Corresponding author: Lucas Meyer
E-mail ✉ lucas.meyer@ethz.ch

Introduction

Algorithmic technologies increasingly mediate organizational control, the distribution of automation and augmentation, and the configuration of human-technological agency [1–3]. This mediation matters because technologies do not merely accelerate pre-existing decisions; they can alter who frames problems, supplies information, exercises discretion, evaluates performance, and contests outcomes. The resulting organizational question is therefore not whether leaders remain involved, but how responsibility should be organized when action is distributed across human and technological actors. Treating technology as a neutral instrument understates these changes, whereas treating technology as an autonomous decision maker risks obscuring the continuing organizational choices through which systems are selected, configured, authorized, and governed.

Learning algorithms further complicate organizing because their outputs can interact with professional knowledge, routines, and evolving forms of expertise [4]. Leadership under technological mediation consequently concerns more than technology adoption. It concerns the allocation of authority, the conditions under which human judgment is expected to engage with computational output, and the organizational arrangements through which affected actors can question or reshape technologically mediated decisions. These issues are especially important where the system influences employment opportunities, task allocation, evaluation, scheduling, monitoring, or other consequential aspects of work.

Empirical scholarship on artificial intelligence in organizations is already multilevel and heterogeneous, spanning collaboration, control, attitudes, work practices, and organizational capability [5]. That heterogeneity cautions against assuming that technological mediation produces uniformly beneficial or harmful outcomes. It also exposes a conceptual fragmentation: responsible-leadership scholarship emphasizes stakeholder responsibility, algorithmic-accountability research emphasizes attribution and oversight, work-design research emphasizes job characteristics, and human-centered technology research emphasizes control and usability. Each literature illuminates part of the problem, but none by itself specifies how leadership responsibility should connect these domains. A purely technical framing can therefore mistake system accuracy or visibility for legitimate governance, while a purely leader-centered framing can overlook decisions already embedded in data practices, system configuration, workflow architecture, and delegated authority.

This article develops an original, non-empirical conceptual framework for responsible leadership and human-centered governance in technology-mediated organizations. Its proposed contribution is to integrate ethical orientation, retained accountability, consequential inclusion, and human-centered governance into one responsibility architecture. The framework does not assume that these obligations are sufficient, equally important in every context, or empirically validated as a bundle. Instead, it distinguishes conceptual plausibility from demonstrated organizational effect and specifies relationships that can be challenged through future multilevel, longitudinal, comparative, and experimental research. This orientation also keeps employee experience separate from organizational-level inference: perceptions of fairness, trust, or control may diagnose governance conditions without by themselves demonstrating that the organization is responsible, effective, or legitimate.

Ethical foundations of responsible leadership

Responsibility-oriented leadership scholarship places organizational action within broader stakeholder, relational, and innovation-governance responsibilities rather than reducing responsibility to leader conduct alone [6–8]. This foundation is useful for technology-mediated organizations because consequential technological decisions rarely affect only the formal decision maker. Employees may experience changes in visibility, autonomy, evaluation, or access to opportunities, while managers, developers, vendors, and senior leaders can occupy different positions in the chain through which a system is authorized and used. Responsible leadership is therefore best treated here as a relational and organizational responsibility orientation rather than as a claim that one individual leader can personally control every technological consequence.

The ethical foundation proposed in this article has three elements. First, affected stakeholders are morally relevant to the governance problem even when they do not possess formal technical authority. Second, leadership responsibility extends to the design of organizational arrangements through which technological decisions are made, reviewed, and revised. Third, responsible intent is analytically distinct from responsible practice. A leader may endorse fairness, transparency, or human-centeredness while organizational incentives, information asymmetries, or weak review mechanisms prevent those commitments from shaping actual decisions.

This distinction is reinforced by analyses showing that AI-ethics guidelines can remain weak when principles are not translated into enforceable or operational practices [9]. The implication is not that principles are useless, but that their organizational significance depends on how they are connected to roles, decision rights, review routines, and consequences for noncompliance. Ethical orientation is therefore proposed as the normative starting condition of the framework, not as evidence that governance has succeeded.

Alternative explanations remain important. Apparent responsible conduct may reflect regulation, reputational pressure, professional norms, or risk management rather than leadership ethics alone. Moreover, individual leader intentions cannot establish team- or organizational-level responsibility. The framework therefore treats ethical orientation as one necessary object of inquiry while leaving its causal influence, relative importance, and interaction with institutional conditions open to empirical testing. This boundary prevents a common category error: observable statements of values are descriptive evidence about formal commitments, whereas claims about how those values alter technological decisions require evidence about enacted practices and consequences.

Accountability for delegated decisions

Technological delegation should be distinguished from ordinary tool use. Agentic information systems can become recipients of delegated activity when organizational actors authorize them to perform tasks or decision functions that were previously carried out by humans [10]. This shift changes the architecture of action, but it does not logically follow that responsibility is

transferred in the same way. The proposed framework therefore separates operational delegation from retained accountability: an organization may delegate analysis, ranking, recommendation, or even bounded decision execution while still needing identifiable human and organizational responsibility for authorization, review, escalation, and correction.

Different configurations of human and artificial intelligence can allocate decision authority in different ways depending on the characteristics of the decision problem [11]. Capability, however, is not identical to legitimate authority. A system may outperform a person on a narrowly defined prediction task without thereby acquiring moral standing, organizational answerability, or authority to determine how competing values should be balanced. This distinction is central to responsible leadership because authority allocation is itself a governance choice.

Algorithmic systems also raise accountability questions that cannot be reduced to the technical properties of their outputs [12]. Responsibility may become diffuse when designers, vendors, managers, users, and executives each control only part of a decision process. The proposed concept of retained accountability addresses this risk by requiring that delegation be accompanied by clearly identifiable responsibility for the conditions under which the technology is authorized and relied upon.

Field evidence from AI-assisted medical diagnosis further shows why nominal human presence is an insufficient proxy for meaningful oversight: professionals' ability to interrogate opaque outputs and incorporate them into expert judgment varied with local practices [13]. This evidence does not establish that human review always improves decisions. It instead supports a narrower boundary: human oversight is substantive only when reviewers possess relevant competence, information, authority, and practical opportunity to engage.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, constructs, and conceptual boundaries for responsible leadership and human-centered governance in technology-mediated organizations.

Table 1. Definitions, constructs, and conceptual boundaries for Responsible Leadership and Human-Centered Governance in Technology-Mediated Organizations.

Construct or component	Definition	Problem addressed	Distinction from adjacent concepts	Proposed role	Uncertainty	Boundary statement
Responsible leadership	Relational responsibility toward affected stakeholders and wider organizational consequences	Narrow leader-centered ethics	Not equivalent to ethical personality	Proposed integrating orientation	Role configuration varies	No single leadership form is established as universally sufficient
Ethical orientation	Explicit attention to responsibility in technological innovation and use	Symbolic or fragmented ethics	Principle is not enacted practice	Proposed normative starting condition	Causal influence remains untested	Ethical commitment does not demonstrate ethical outcomes
Delegated technological agency	Authorized performance of tasks or decision functions by agentic systems	Unclear allocation of action	Delegation is not responsibility transfer	Proposed trigger for accountability analysis	Boundaries of "agency" vary	Technical capability does not establish legitimate authority
Retained accountability	Identifiable responsibility for authorization, review, escalation, and correction after delegation	Responsibility diffusion	Not the same as nominal human oversight	Proposed governance obligation	Effective allocation requires testing	Human presence alone does not prove meaningful control
Consequential inclusion	Capacity of affected actors to influence design, use, review, or contestation	Token participation	Voice opportunity is not consequential influence	Proposed inclusion principle	Influence depends on power and context	Participation does not guarantee changed decisions
Human-centered work design	Organization of autonomy, demands, discretion, and responsibilities around technology	Technology-centered implementation	Autonomy is not responsibility transfer	Proposed governance pathway	Direction of effects is context-dependent	No universal work-design effect is assumed
Human-centered governance	Organizational arrangements connecting human control, contestability, accountability, and learning	Governance reduced to transparency or human presence	Broader than "human in the loop"	Proposed integrative governance domain	Necessity and sufficiency remain unvalidated	No component alone demonstrates responsible governance

Inclusion in technology design and use

Inclusion becomes consequential when affected actors can influence how technology is designed, configured, used, reviewed, or contested. This definition is deliberately narrower than a general claim that participation is always beneficial and broader than a procedural requirement to consult employees. Algorithmic human-resource systems can create tensions between efficiency-oriented decision making and employees' personal integrity, particularly when standardized representations displace context or moral judgment [14]. Such concerns justify examining participation, but they do not establish that any particular participatory mechanism will prevent harm.

Experimental evidence also shows why inclusion cannot be reduced to technical debiasing. Algorithmic procedures may be perceived as less procedurally fair when people believe relevant qualitative information has been excluded, even where the procedure is intended to reduce human bias [15]. The appropriate inference is not that algorithmic evaluation is inherently unfair. Rather, formal consistency and bias reduction are distinct from whether affected actors can understand, contextualize, or challenge how a consequential judgment is produced.

Work-design scholarship provides a second route into inclusion. Digital technologies can alter autonomy, task variety, demands, monitoring, and other job characteristics through the way work is designed [16]. Employees may therefore be affected by choices made long before a final algorithmic decision appears. The proposed framework accordingly treats inclusion as relevant not only at the moment of review but also during configuration and work redesign, where assumptions about acceptable discretion and responsibility become embedded.

Finally, capacity to influence technological change is conditioned by structural power. Qualitative evidence from algorithmic management in manufacturing indicates that workplace regime and union power shape conflict and worker influence over digitalization [17]. Voice opportunity should therefore not be mistaken for substantive power. Consequential inclusion is proposed as a governance obligation whose realization depends on access, information, representation, timing, and the capacity to affect decisions. Its scope therefore includes both upstream influence over assumptions and work design and downstream capacity to seek explanation, correction, or reconsideration. These are analytically different from consultation: employees may be invited to speak while lacking access to relevant information or any route through which objections can alter system use. Whether consequential inclusion improves fairness, legitimacy, work quality, or error detection remains an empirical question rather than an assumed outcome.

Human-centered governance principles

Human-centered governance requires more than attaching ethical principles to a technological system after design decisions have been made. Work on AI-ethics operationalization shows the need to connect principles with tools, methods, and practices across development and use [18]. For organizational governance, the relevant implication is procedural: responsibility must become observable in how requirements are specified, decision rights assigned, exceptions handled, and systems revised. The framework therefore treats operationalization as a bridge between ethical orientation and enacted governance, while not assuming that adopting a tool or procedure demonstrates ethical effectiveness.

Human-centered artificial intelligence also challenges a simplistic trade-off between automation and human control. High levels of automation can coexist conceptually with meaningful human control when systems and work are designed around complementary capabilities [19]. The proposed governance principle is therefore not “keep a human in the loop.” It is to preserve usable human authority where judgment, contextual interpretation, contestation, or value trade-offs remain consequential. Whether a particular arrangement provides such authority depends on competence, timing, information access, and the practical capacity to intervene.

Transparency likewise requires a bounded interpretation. Making an algorithm visible or explainable does not by itself establish accountability because observers may lack the knowledge, authority, resources, or institutional position needed to act on what is disclosed [20]. Visibility is therefore distinct from performance and from answerability. A transparent process can remain substantively unchallengeable, while a less technically transparent process may contain robust review and escalation arrangements.

Organizational evidence from HR and people-analytics projects further suggests that ethics becomes consequential through roles, practices, and project arrangements rather than abstract principles alone [21]. Human-centered governance is consequently proposed as the organizational configuration through which ethical orientation, accountability, inclusion, work design, review, and learning are connected. It is not an empirically established bundle of sufficient conditions.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for proposed mechanisms, relationships, and organizational consequences in responsible leadership and human-centered governance in technology-mediated organizations.

Table 2. Proposed mechanisms, relationships, and organizational consequences in Responsible Leadership and Human-Centered Governance in Technology-Mediated Organizations.

Mechanism or relationship	Antecedent	Mediator or process	Moderator	Expected consequence	Rival explanation	Validation requirement
Delegation with retained accountability	Transfer of bounded decision activity	Explicit authorization, review, escalation, correction	Decision type; reviewer competence	Proposed reduction in responsibility diffusion	Existing hierarchy may already clarify responsibility	Compare matched delegation arrangements with different accountability structures
Opacity and human engagement	Difficult-to-interpret system output	Interrogation and contextual professional judgment	Expertise; local practice	Proposed improvement in substantive rather than nominal oversight	Strong professional routines may explain engagement independently	Observe review behavior, not only formal human presence
Reductionistic evaluation and procedural justice	Standardized algorithmic evaluation	Perceived exclusion of relevant contextual information	Decision procedure and context	Lower perceived procedural fairness under specified conditions	General distrust of algorithms	Manipulate contextual-information inclusion and procedure separately
Algorithmic control through work design	Digital allocation, evaluation, or monitoring	Changes in autonomy, demands, discretion, and task structure	Workplace regime; worker power	Context-dependent changes in control and contestation	Concurrent restructuring may drive outcomes	Longitudinally separate technological and organizational redesign
Principles-to-practice translation	Formal ethical commitments	Tools, roles, routines, review, and revision	Organizational incentives and readiness	Proposed movement from symbolic to enacted governance	Regulation or reputation may produce similar practices	Compare formal commitments with observed decision routines over time
Consequential inclusion and contestability	Technology affecting work or decisions	Contextual input, challenge, reconsideration	Access, timing, worker power	Proposed increase in substantive rather than symbolic influence	Consultation may primarily serve legitimacy or communication	Trace whether objections can change configuration, use, or review outcomes

Proposed responsible-leadership framework

Human–AI collaboration is conditional on calibrated trust, cognitive and task fit, and explicit coordination of human and machine roles [22–24]. These conditions matter because neither reliance nor human intervention is intrinsically responsible. Excessive trust can normalize uncritical acceptance, while excessive distrust can prevent useful delegation. Similarly, retaining a human decision maker has limited governance value if the person lacks information, expertise, time, or authority to challenge the system. Responsible leadership must therefore organize the relationship between technological capability and human responsibility rather than maximizing either automation or human control.

Algorithmic systems can also influence work by functioning as work designers, changing how tasks, autonomy, demands, and managerial functions are structured [25]. This extends the governance problem beyond discrete "AI decisions." A system may alter responsibility and experience through scheduling, performance evaluation, task assignment, or monitoring even when no single output appears to constitute a major decision. The proposed framework therefore places work design inside governance rather than treating it as an implementation consequence external to leadership.

The original Responsible-Leadership Framework integrates four proposed obligations: ethical orientation, retained accountability, consequential inclusion, and human-centered governance. Ethical orientation identifies whose interests and potential harms require consideration. Retained accountability keeps identifiable responsibility attached to organizational authorization and review despite delegation. Consequential inclusion asks whether affected actors possess meaningful routes to shape assumptions, configuration, use, and contestation. Human-centered governance connects these obligations to work design, review competence, escalation, and organizational learning. Their integration is proposed rather than empirically demonstrated; organizations may satisfy one obligation while failing another.

The framework further proposes a governance lifecycle of authorize, configure, delegate or use, review or contest, and learn or revise. Authorization concerns legitimate purpose and decision scope; configuration concerns data, rules, interfaces, and work arrangements; delegation or use concerns actual allocation of decision activity; review or contest concerns substantive

capacity to question outcomes; and learning or revision concerns whether experience changes future configurations. The sequence is analytical rather than necessarily chronological, and recursive movement is expected.

Two additional boundaries prevent the lifecycle from becoming a managerial checklist. First, formal permission to intervene should be distinguished from substantive power to do so: employees, reviewers, or managers may possess nominal override rights while facing information, time, hierarchy, or incentive constraints. Second, governance quality cannot be inferred from the visibility of a procedure. Documented review may establish that a process exists without showing that it detects errors, changes disputed decisions, or protects affected actors. These distinctions make the framework observable without presuming that visible governance structures perform their stated function.

Figure 1 presents the original conceptual synthesis for responsible-leadership framework, showing how the article's principal constructs, mechanisms, uncertainties, and decision boundaries are connected.

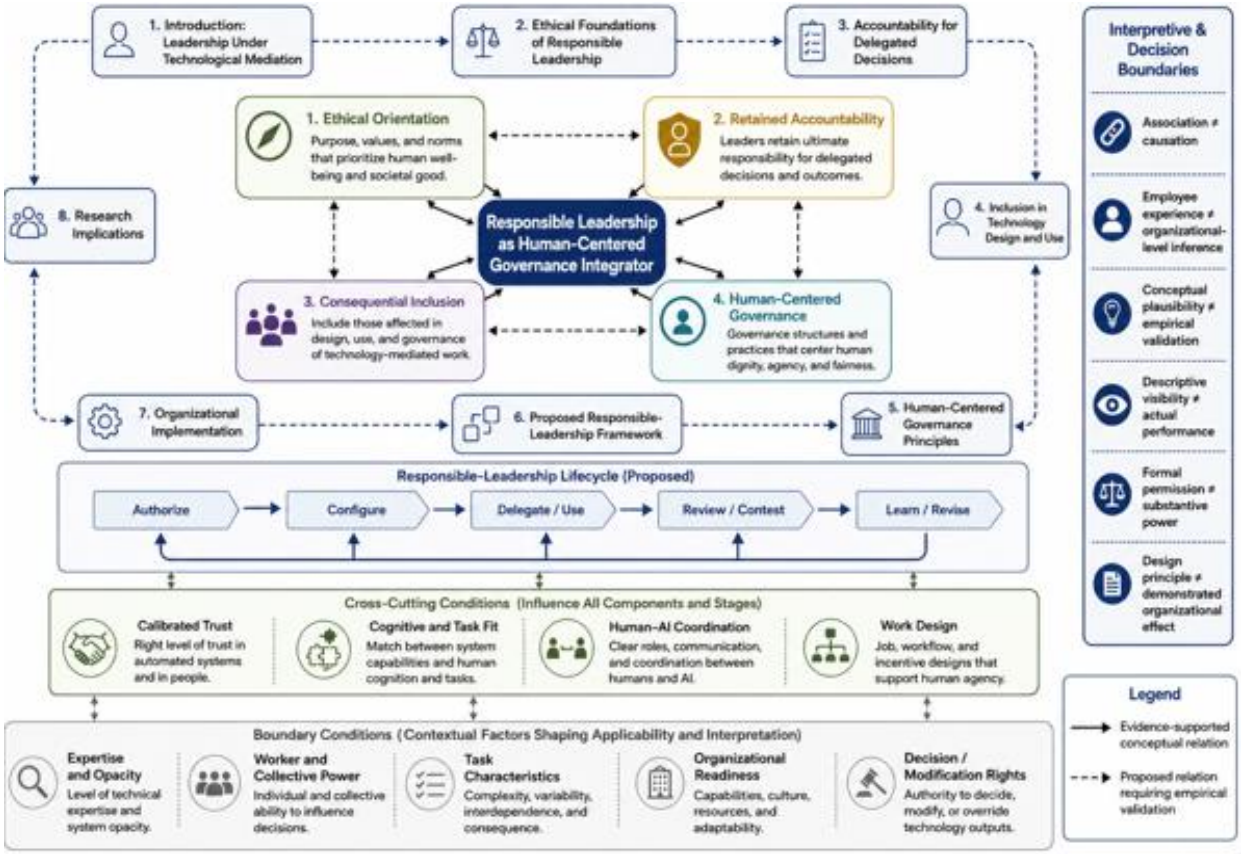

Figure 1. Responsible-Leadership Framework.

The figure is an original conceptual synthesis developed for "Responsible Leadership and Human-Centered Governance in Technology-Mediated Organizations." Arrows and grouping indicate evidence-supported or proposed relationships rather than measured effect sizes. The figure does not represent a validated causal model, universal organizational prescription, regulatory determination, or deployment-ready system.

Organizational implementation

Implementation should be understood as organizational enactment rather than adoption of a technology or publication of an ethics statement. Interview research on organizational AI readiness indicates that implementation depends on organizational factors extending beyond the availability of technical capability [26]. For the proposed framework, readiness therefore includes the ability to assign responsibility, supply competent review, coordinate relevant functions, and create routes for escalation and revision. These are proposed governance requirements, not a validated readiness scale.

Practitioner research on responsible AI similarly indicates that organizational culture, structures, incentives, and workflows can enable or obstruct responsible practices [27]. This evidence cautions against locating responsibility in a single ethics officer or committee. A governance arrangement can appear complete formally while incentives discourage challenge, review occurs too late to matter, or operational teams cannot alter a deployed system. The proposed implementation principle is therefore distributed responsibility with identifiable ownership: multiple actors may participate, but particular decisions should remain answerable to designated organizational roles.

Human-resource applications illustrate why such enactment is difficult. AI use in HRM raises intertwined problems involving data, fairness, accountability, causal reasoning, and employee responses [28]. These concerns make implementation context-specific. A review mechanism suitable for low-consequence task recommendations may be inadequate for hiring, discipline, or performance evaluation, while highly restrictive controls may be disproportionate for other uses. Responsible implementation therefore requires explicit scoping of decision consequences rather than assuming one governance intensity for all technologies. The relevant unit is not simply the technology but the technology-in-use: the same technical system may carry different governance implications when it recommends rather than decides, informs experts rather than novices, or affects reversible operational choices rather than employment consequences. Implementation should therefore preserve the ability to revisit scope as uses, data, or organizational dependencies change.

Table 3 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for testable propositions, evaluation criteria, and practical responsibilities.

Table 3. Testable propositions, evaluation criteria, and practical responsibilities.

Proposition or criterion	Unit and context	Observable indicator	Falsification condition	Organizational responsibility	Misuse risk	Readiness boundary
P1 Proposed retained-accountability proposition	Delegated organizational decisions	Named authorization, review, escalation, correction roles	Equivalent governance without identifiable retained responsibility	Assign answerable decision ownership	Treating documentation as accountability	Roles must possess usable authority
P2 Proposed consequential-inclusion proposition	Affected employees or users	Evidence that input can alter design, use, or review	Nominal and consequential inclusion produce indistinguishable processes	Provide timely information and contestation routes	Token consultation	Influence must be observable, not merely invited
P3 Proposed meaningful-control criterion	Human–AI decision settings	Reviewer competence, information, time, intervention rights	Human presence alone performs equivalently across relevant conditions	Resource substantive review	Rubber-stamp oversight	Review authority must match decision consequence
P4 Proposed operationalization proposition	Projects and governance systems	Ethical commitments linked to routines and revision	Formal commitments alone consistently generate equivalent practice	Connect principles to roles and workflows	Compliance theatre	Organization must be capable of changing practice
P5 Proposed context-sensitive delegation proposition	Tasks with varying characteristics	Changes in reliance across task and capability conditions	One delegation configuration dominates across contexts	Define decision scope and reassessment triggers	Universal automation rule	Task and capability boundaries must be specified
P6 Proposed governance-readiness proposition	Organizational implementation	Evidence of competence, ownership, escalation, and revision capacity	Technical readiness predicts equivalent responsible enactment without governance capacity	Assess organizational capacity before expanding delegation	Treating adoption readiness as governance readiness	Capacity must be evaluated separately from technological availability

Research implications

Future research should treat reliance on algorithmic systems as an empirical variable rather than assume generalized resistance. Experimental evidence demonstrates that people can display algorithm appreciation under some conditions [29]. This creates an important rival explanation for governance outcomes: high use of a system may reflect perceived competence or convenience rather than legitimate authority, responsible design, or organizational trustworthiness. Studies should therefore measure why actors rely on technological recommendations and separate voluntary reliance from dependence created by workflow or managerial control.

Algorithm attitudes also vary with task characteristics, including perceived objectivity or subjectivity [30]. Tests of the proposed framework should consequently model task type, decision consequence, expertise, opacity, and organizational power as moderators. Cross-sectional employee perceptions alone are insufficient because positive attitudes can coexist with weak accountability, while skepticism can coexist with robust governance. Multi-level research should connect employee experiences to observable organizational arrangements such as decision rights, review practices, escalation events, and changes made after contestation.

Limited opportunities to modify algorithmic forecasts can increase people's willingness to use imperfect algorithms [31]. This finding makes decision rights a particularly useful experimental variable, but greater adoption should not be interpreted as better governance. Future studies could vary modification, appeal, explanation, or override rights and test whether these

arrangements affect error detection, perceived procedural justice, responsible reliance, and actual contestation. Longitudinal field research could examine whether formal rights remain usable after implementation and whether repeated challenges produce learning or merely procedural closure.

The framework is falsifiable at several levels. Research may find that retained accountability adds little where responsibilities are already institutionalized, that consequential inclusion does not alter relevant decisions, or that different lifecycle stages matter in different settings. Researchers should also test interaction rather than merely additive assumptions. Ethical orientation may matter little without accountability, while accountability may become ceremonial where affected actors cannot surface contextual information or initiate review. Comparative case designs can identify alternative configurations, while field experiments can test narrower mechanisms without claiming validation of the whole framework. Research should also distinguish outcomes at individual, team, and organizational levels and examine whether apparent improvements arise from regulation, professional norms, workforce power, technology maturity, or selection rather than responsible leadership itself.

Conclusion

Technology-mediated organizing creates a leadership problem because decision activity can be distributed without eliminating organizational responsibility. This article proposes a Responsible-Leadership Framework that integrates ethical orientation, retained accountability, consequential inclusion, and human-centered governance across authorization, configuration, delegation and use, review and contestation, and learning and revision. Its central claim is conceptual: technological capability should not be equated with legitimate authority, delegation should not be equated with responsibility transfer, participation should not be equated with consequential influence, and human presence should not be equated with meaningful oversight.

The framework also places work design and organizational power inside governance rather than treating them as peripheral implementation issues. This widens responsible leadership from leader intention to the organizational arrangements through which technological systems become consequential. The contribution remains deliberately bounded. The proposed obligations and lifecycle are not validated causal mechanisms, sufficient conditions, regulatory standards, or universal prescriptions. Their value depends on whether future research can test, refine, reconfigure, or reject them across different technologies, tasks, occupations, institutional settings, and levels of analysis. Responsible leadership in technology-mediated organizations is therefore best approached as an empirically testable problem of organizing responsibility, not as a label conferred by technological sophistication or ethical aspiration.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Kellogg KC, Valentine MA, Christin A. Algorithms at work: The new contested terrain of control. Acad Manag Ann. 2020;14(1):366-410. doi:10.5465/annals.2018.0174
2. Raisch S, Krakowski S. Artificial intelligence and management: The automation-augmentation paradox. Acad Manag Rev. 2021;46(1):192-210. doi:10.5465/amr.2018.0072
3. Murray A, Rhymer J, Sirmon DG. Humans and technology: Forms of conjoined agency in organizations. Acad Manag Rev. 2021;46(3):552-71. doi:10.5465/amr.2019.0186
4. Faraj S, Pachidi S, Sayegh K. Working and organizing in the age of the learning algorithm. Inf Organ. 2018;28(1):62-70. doi:10.1016/j.infoandorg.2018.02.005
5. Bankins S, Ocampo ACG, Marrone M, Restubog SLD, Woo SE. A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. J Organ Behav. 2024;45(2):159-82. doi:10.1002/job.2735
6. Miska C, Mendenhall ME. Responsible leadership: A mapping of extant research and future directions. J Bus Ethics. 2018;148(1):117-34. doi:10.1007/s10551-015-2999-0
7. Voegtlin C, Frisch C, Walther A, Schwab P. Theoretical development and empirical examination of a three-roles model of responsible leadership. J Bus Ethics. 2020;167(3):411-31. doi:10.1007/s10551-019-04155-2

8. Voegtlin C, Scherer AG. Responsible innovation and the innovation of responsibility: Governing sustainable development in a globalized world. J Bus Ethics. 2017;143(2):227-43. doi:10.1007/s10551-015-2769-z
9. Hagendorff T. The ethics of AI ethics: An evaluation of guidelines. Minds Mach. 2020;30(1):99-120. doi:10.1007/s11023-020-09517-8
10. Baird A, Maruping LM. The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. MIS Q. 2021;45(1):315-41. doi:10.25300/MISQ/2021/15882
11. Shrestha YR, Ben-Menahem SM, von Krogh G. Organizational decision-making structures in the age of artificial intelligence. Calif Manage Rev. 2019;61(4):66-83. doi:10.1177/0008125619862257
12. Martin K. Ethical implications and accountability of algorithms. J Bus Ethics. 2019;160(4):835-50. doi:10.1007/s10551-018-3921-3
13. Lebovitz S, Lifshitz-Assaf H, Levina N. To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. Organ Sci. 2022;33(1):126-48. doi:10.1287/orsc.2021.1549
14. Leicht-Deobald U, Busch T, Schank C, Weibel A, Schafheitle S, Wildhaber I, et al. The challenges of algorithm-based HR decision-making for personal integrity. J Bus Ethics. 2019;160(2):377-92. doi:10.1007/s10551-019-04204-w
15. Newman DT, Fast NJ, Harmon DJ. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. Organ Behav Hum Decis Process. 2020;160:149-67. doi:10.1016/j.obhdp.2020.03.008
16. Parker SK, Grote G. Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. Appl Psychol. 2022;71(4):1171-204. doi:10.1111/apps.12241
17. Dupuis M. Algorithmic management and control at work in a manufacturing sector: Workplace regime, union power and shopfloor conflict over digitalisation. New Technol Work Employ. 2025;40(1):81-101. doi:10.1111/ntwe.12298
18. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Sci Eng Ethics. 2020;26(4):2141-68. doi:10.1007/s11948-019-00165-5
19. Shneiderman B. Human-centered artificial intelligence: Reliable, safe & trustworthy. Int J Hum Comput Interact. 2020;36(6):495-504. doi:10.1080/10447318.2020.1741118
20. Ananny M, Crawford K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media Soc. 2018;20(3):973-89. doi:10.1177/1461444816676645
21. Bar-Gil O, Ron T, Czerniak O. AI for the people? Embedding AI ethics in HR and people analytics projects. Technol Soc. 2024;77:102527. doi:10.1016/j.techsoc.2024.102527
22. Glikson E, Woolley AW. Human trust in artificial intelligence: Review of empirical research. Acad Manag Ann. 2020;14(2):627-60. doi:10.5465/annals.2018.0057
23. Fügener A, Grahl J, Gupta A, Ketter W. Cognitive challenges in human-artificial intelligence collaboration: Investigating the path toward productive delegation. Inf Syst Res. 2022;33(2):678-96. doi:10.1287/isre.2021.1079
24. Seeber I, Bittner E, Briggs RO, de Vreede T, de Vreede GJ, Elkins A, et al. Machines as teammates: A research agenda on AI in team collaboration. Inf Manag. 2020;57(2):103174. doi:10.1016/j.im.2019.103174
25. Parent-Rocheleau X, Parker SK. Algorithms as work designers: How algorithmic management influences the design of jobs. Hum Resour Manag Rev. 2022;32(3):100838. doi:10.1016/j.hrmr.2021.100838
26. Jöhnk J, Weißert M, Wyrtki K. Ready or not, AI comes—An interview study of organizational AI readiness factors. Bus Inf Syst Eng. 2021;63(1):5-20. doi:10.1007/s12599-020-00676-7
27. Rakova B, Yang J, Cramer H, Chowdhury R. Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. Proc ACM Hum Comput Interact. 2021;5(CSCW1):1-23. doi:10.1145/3449081
28. Tambe P, Cappelli P, Yakubovich V. Artificial intelligence in human resources management: Challenges and a path forward. Calif Manage Rev. 2019;61(4):15-42. doi:10.1177/0008125619867910
29. Logg JM, Minson JA, Moore DA. Algorithm appreciation: People prefer algorithmic to human judgment. Organ Behav Hum Decis Process. 2019;151:90-103. doi:10.1016/j.obhdp.2018.12.005
30. Castelo N, Bos MW, Lehmann DR. Task-dependent algorithm aversion. J Mark Res. 2019;56(5):809-25. doi:10.1177/0022243719851788
31. Dietvorst BJ, Simmons JP, Massey C. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Manage Sci. 2018;64(3):1155-70. doi:10.1287/mnsc.2016.2643