International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society - https://journalviral.org
DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

From Human-Centered Firms to Hybrid Human-AI Organizations: Rethinking Organizational Design, Decision Rights, and Distributed Agency

Vladimir Aurelian Enachescu, Razvan Barbulescu
Bucharest University of Economic Studies (ROMANIA)

Abstract: *The emergence of autonomous artificial intelligence (AI) agents challenges established assumptions about organizational design, managerial authority, and the distribution of decision-making responsibilities. Unlike conventional AI applications that primarily support human decision-making, agentic AI systems can initiate tasks, coordinate workflows, and execute bounded decisions with varying degrees of human supervision. This conceptual article examines how organizations should be designed and governed when autonomous AI agents become active participants in organizational processes. Drawing on organizational design theory, bounded rationality, sociotechnical systems, human-centered AI, and emerging perspectives on agentic management, the study develops the Hybrid Organizational Agency Architecture (HOAA). The proposed framework distinguishes four organizational configurations: human-directed AI assistance, supervised AI delegation, coordinated human–AI workflows, and bounded agent-led operations. It introduces distributed operational agency to describe how task initiation, information processing, coordination, and execution capabilities are allocated between human actors and autonomous AI systems. The article further identifies decision rights, autonomy boundaries, coordination mechanisms, human oversight, and accountability structures as fundamental dimensions of hybrid organizational design. Six theoretical propositions explore the relationships among AI autonomy, organizational coordination, decision quality, adaptability, trust, and accountability. The analysis suggests that effective hybrid organizational arrangements depend not simply on increasing technological autonomy but on aligning AI capabilities with organizational objectives, clearly defined authority structures, and appropriate governance mechanisms. Importantly, operational autonomy is distinguished from human moral agency and legal responsibility, preserving the central role of human and institutional accountability. The article contributes to organizational theory by extending human-centered AI research beyond technological augmentation toward the structural and governance implications of autonomous AI participation. It also offers a conceptual foundation for future empirical investigations and practical guidance for organizational leaders seeking to balance technological innovation, managerial control, and responsible human–AI collaboration.*

Keywords: Hybrid human–AI organizations; agentic artificial intelligence; organizational design; distributed operational agency; decision rights; human-centered AI; sociotechnical systems; AI governance.

1. Introduction

The evolution of artificial intelligence (AI) from analytical and decision-support applications toward increasingly autonomous, agent-based systems raises fundamental questions about the organization of work, managerial authority, and the distribution of decision-making responsibilities. Traditional organizational structures have generally been designed around human actors who establish objectives, interpret information, allocate resources, coordinate activities, and assume responsibility for organizational outcomes. Although digital technologies have progressively transformed these functions, they have typically been conceptualized as instruments supporting human decisions or

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

automating predefined tasks. The emergence of autonomous AI agents challenges this conventional distinction by introducing technological systems capable of initiating activities, planning sequences of actions, interacting with organizational resources, and executing bounded decisions within delegated operational environments.

This transformation may signal a transition from predominantly human-centered firms toward **hybrid human–AI organizations**, in which human actors and autonomous AI systems jointly contribute to organizational processes through differentiated forms of operational participation. Such a transition does not imply that artificial agents possess the same intentionality, moral standing, or institutional responsibilities as human organizational members. Rather, it suggests that organizational design must increasingly account for the distribution of task initiation, information processing, coordination, and execution capabilities across human and technological components. Consequently, questions concerning organizational hierarchy, decision rights, accountability, and coordination acquire renewed theoretical importance.

Established organizational theories provide valuable foundations for understanding these developments. Simon's (1947) concept of bounded rationality challenged assumptions of unlimited managerial information-processing capacity, while March and Simon (1958) emphasized the organizational mechanisms through which decision-making is structured and coordinated. Galbraith's (1974) information-processing perspective subsequently conceptualized organizational design as a response to uncertainty and information-processing requirements. These perspectives remain relevant because AI systems may expand organizational information-processing capabilities while simultaneously introducing new forms of uncertainty, technological interdependence, and governance complexity. The theoretical challenge is therefore not merely to determine whether AI improves managerial decisions, but to understand how organizational structures should evolve when elements of operational decision-making are delegated to autonomous technological systems.

Research on AI in management has increasingly recognized the limitations of treating automation and human augmentation as mutually exclusive organizational strategies. Raisch and Krakowski (2021) conceptualized automation and augmentation as interdependent processes that generate persistent tensions for management. Their analysis suggests that organizations must consider the complementarities and contradictions between technological substitution and human collaboration rather than assuming that one approach will consistently produce superior outcomes. From the perspective of hybrid organizational design, this insight is particularly important because autonomous AI agents may simultaneously automate specific activities, augment managerial capabilities, and participate in coordination processes that previously depended on human organizational roles.

Human-centered AI scholarship offers a complementary theoretical foundation. Margetis et al. (2021) emphasized the importance of designing AI systems around human needs, capabilities, and interactions, while Ozmen Garibay et al. (2023) identified major interdisciplinary challenges associated with developing AI that supports human values and meaningful participation. Herrmann and Pfeiffer (2023) further argued for a sociotechnical understanding of human-centered AI that incorporates organizational arrangements rather than restricting attention to individual human–machine interactions. This organizational perspective is especially relevant because the consequences of AI autonomy depend not only on technological performance but also on the institutional structures within which AI systems operate.

Wilkens et al. (2023) advanced this discussion by identifying seven configurations of human-centered AI at work, illustrating the importance of relationships between organizational actors and structural conditions. Their contribution indicates that human-centered AI cannot be understood through a universal design principle detached from organizational context. Similarly, Townsend and Romme (2024) examined the emerging concept of the human-centered organization, providing a broader perspective on organizational arrangements that prioritize human development, participation, and wellbeing. These contributions establish an important normative and organizational foundation for examining AI-enabled transformation. However, the increasing operational autonomy of AI agents raises a further question: how can organizations preserve human-centered objectives while redistributing elements of operational agency to nonhuman systems?

Emerging research on agentic management begins to address this issue. Ajmal and Islam (2025) explicitly frame the transition from human-centered to agentic management as a challenge for organizational theory, while Chen (2026) approaches managerial rationality through the perspective of human–machine collaboration in organizational decision-making. These perspectives suggest that the relationship between managerial judgement and AI-enabled decision processes deserves renewed theoretical attention. Nevertheless, the conceptual distinction between human agency and technologically delegated operational autonomy remains essential. The capacity of an AI system to

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

execute organizational actions does not, by itself, establish that the system possesses independent normative authority or responsibility for the consequences of those actions.

A multilevel sociotechnical perspective is particularly useful for clarifying this distinction. Schoenherr and Michael (2026) position human-centered AI within interconnected technological, organizational, and social arrangements. In a related direction, Winby and Xu (2026) address human-centered AI maturity from an organizational design perspective, suggesting that the organizational capacity to integrate AI responsibly cannot be reduced to the sophistication of individual technologies. Xu (2026) further contributes to the foundational discussion of human-centered AI, while Kossowska-Lai et al. (2026) connect human-centered design with organizational resilience. Together, these perspectives provide a basis for investigating hybrid organizations as evolving sociotechnical systems in which operational autonomy must be aligned with institutional objectives and governance capabilities.

The governance implications of autonomous AI participation are increasingly visible across organizational settings. Studies addressing AI agents in higher education and educational leadership, for example, raise questions about the delegation of activities, the transformation of professional roles, and the preservation of institutional oversight (Costache, 2026a, 2026b). Although educational institutions differ from commercial firms in their objectives and accountability arrangements, they offer relevant contexts for examining how autonomous technologies may reshape established professional and managerial processes. Similarly, discussions of human-centered talent management and Society 5.0 emphasize the continuing importance of human capabilities, organizational adaptation, and responsible technological integration (Atay et al., 2025).

Despite these contributions, an important theoretical challenge remains insufficiently resolved: **how should organizational decision rights, coordination mechanisms, and accountability structures be configured when autonomous AI agents are authorized to perform organizational actions?** Existing research provides substantial insight into AI adoption, human-centered design, human–AI collaboration, and algorithmic decision-making. However, these perspectives do not automatically yield an integrated organizational architecture specifying how different degrees of AI autonomy should correspond to different arrangements of human authority, supervision, coordination, and accountability.

This article addresses that challenge by developing the concept of **distributed operational agency**. The concept refers to the allocation of task initiation, information processing, coordination, and bounded execution capabilities across human organizational actors and autonomous AI systems. It is deliberately narrower than philosophical or sociological accounts of human agency. In this article, operational agency describes what organizational participants and technological systems are enabled to do within defined institutional arrangements; it does not imply equivalence in consciousness, moral judgement, legal status, or responsibility. Human actors and organizations remain responsible for establishing legitimate objectives, defining acceptable delegation boundaries, and maintaining appropriate oversight.

Building on this distinction, the article introduces the **Hybrid Organizational Agency Architecture (HOAA)**, a conceptual framework designed to explain how organizations may structure relationships between human decision-makers and autonomous AI agents. The proposed architecture distinguishes four organizational configurations: (1) human-directed AI assistance, (2) supervised AI delegation, (3) coordinated human–AI workflows, and (4) bounded agent-led operations. These configurations are differentiated according to the distribution of operational initiative, the scope of delegated decision rights, the intensity of human oversight, and the mechanisms used to coordinate human and artificial activities. They are presented as analytical organizational forms rather than as an empirically validated maturity sequence or an assumption that greater autonomy necessarily represents superior organizational development.

The article is guided by three research questions:

RQ1. How does the emergence of autonomous AI agents challenge established assumptions about organizational design, managerial authority, and decision-right allocation?

RQ2. What organizational configurations can structure the distribution of operational agency between human actors and autonomous AI systems?

RQ3. Under what theoretical conditions might different configurations of human–AI agency influence organizational coordination, decision quality, adaptability, trust, and accountability?

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

The article makes three principal contributions to organizational and management theory. First, it extends human-centered AI scholarship by shifting analytical attention from the design of individual AI systems and human–AI interactions toward the organizational architecture of distributed operational agency. Second, it integrates established organizational design concepts with emerging research on autonomous AI agents to develop a typology of hybrid organizational configurations. Third, it advances theoretical propositions concerning the relationships among delegation boundaries, coordination structures, organizational outcomes, and accountability mechanisms, thereby establishing a foundation for subsequent empirical investigation.

The practical relevance of this inquiry extends beyond technological implementation. Organizations introducing autonomous AI agents must determine which activities may be delegated, how agent actions should be monitored, when human authorization is required, and how responsibility should be assigned when decisions involve multiple human and technological contributors. These choices are fundamentally organizational rather than exclusively technical. Consequently, the development of hybrid human–AI organizations requires a reconsideration of managerial roles, organizational structures, and governance practices.

2. Theoretical Foundations and Literature Review

2.1. Organizational Design, Bounded Rationality, and Information Processing

Organizational design theory provides a fundamental starting point for understanding the emergence of hybrid human–AI organizations. Classical perspectives conceptualize organizations as coordinated systems in which activities are differentiated, authority is allocated, and decision-making processes are structured to achieve collective objectives. From this standpoint, organizational effectiveness depends not only on the capabilities of individual participants but also on the relationships among roles, information flows, coordination mechanisms, and institutional arrangements.

Simon (1947) introduced bounded rationality as a central limitation of organizational decision-making, emphasizing that human decision-makers operate under constraints involving information, cognitive capacity, and time. March and Simon (1958) subsequently demonstrated that organizations respond to these limitations through routines, communication structures, authority relationships, and decision procedures. These contributions remain relevant in AI-enabled environments because autonomous AI agents may alter organizational information-processing capacities without eliminating uncertainty, incomplete knowledge, or the need for judgement.

Galbraith (1974) extended organizational design theory by explaining how organizations develop information-processing mechanisms to manage uncertainty and interdependence. According to this perspective, organizational structures should be aligned with the information-processing demands generated by organizational tasks and environments. Autonomous AI agents potentially change this relationship by enabling more extensive data analysis, continuous monitoring, and automated coordination. However, expanded technological processing capacity may also create new information demands associated with evaluating AI outputs, interpreting system behaviour, monitoring delegated actions, and resolving conflicts between human and machine recommendations.

The organizational implications are therefore multidimensional. AI systems may reduce the cognitive burden associated with routine analysis while increasing the complexity of oversight and coordination. For example, a procurement agent may autonomously compare suppliers and prepare purchasing decisions, yet managers must still establish contractual boundaries, assess exceptional circumstances, and determine responsibility for inappropriate transactions. Consequently, technological autonomy does not necessarily reduce the need for organizational design; it changes the activities that organizational structures must coordinate.

Chen (2026) addresses managerial rationality from a human–machine collaboration perspective, providing a contemporary point of reference for reconsidering how decision-making is organized when AI becomes integrated into managerial processes. The present article extends this discussion by focusing on the organizational distribution of decision rights rather than solely on the cognitive relationship between human managers and AI-supported analysis.

From this perspective, hybrid organizational design involves two interconnected problems. The first concerns the allocation of information-processing and execution capabilities between humans and AI systems. The second concerns the allocation of authority, supervision, and responsibility over the activities performed through those capabilities. These problems are related but analytically distinct:

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

a technological system may possess the capability to execute a decision without possessing legitimate authority to determine whether that decision should be executed.

This distinction forms an essential foundation for the proposed Hybrid Organizational Agency Architecture.

2.2. From Human-Centered Organizations to Human-Centered AI

Human-centered organizational theory emphasizes the importance of human wellbeing, meaningful participation, development, and institutional arrangements that recognize people as more than instruments of organizational performance. Townsend and Romme (2024) examine the emerging concept of the human-centered organization, offering a conceptual foundation for considering how organizational structures can reflect human-oriented objectives.

Within the technological domain, human-centered AI has developed as an interdisciplinary approach concerned with aligning AI design and implementation with human capabilities, values, needs, and societal expectations. Margetis et al. (2021) address human-centered AI design from a human factors perspective, while Ozmen Garibay et al. (2023) identify major challenges involving human–AI interaction, trust, responsible design, and broader societal implications.

These contributions suggest that AI integration should not be evaluated exclusively through efficiency, productivity, or technological sophistication. Human participation, transparency, appropriate control, and the consequences of AI deployment for organizational members also require consideration.

Wilkens et al. (2023) advance this perspective by examining configurations of human-centered AI at work. Their systematic review identifies seven actor–structure engagements, demonstrating that human-centeredness emerges through combinations of organizational practices, structural arrangements, and relationships among actors. This configuration-based approach is particularly relevant because it challenges the assumption that a single human-centered AI model can be applied uniformly across organizational contexts.

At the organizational level, the central question is not simply whether AI applications are designed around human users, but whether organizational arrangements enable people to exercise meaningful judgement, challenge technological recommendations, and retain appropriate influence over consequential decisions.

Xu (2026) provides a contemporary foundational perspective on human-centered AI, while Winby and Xu (2026) approach human-centered AI maturity through organizational design. These contributions can be interpreted as supporting the need to consider organizational readiness, governance capabilities, and institutional arrangements alongside technological implementation.

The transition toward autonomous AI agents, however, introduces a conceptual tension. Human-centered AI traditionally emphasizes the design of systems that support human activities, whereas agentic AI may perform activities with limited direct human involvement. The resulting challenge is to preserve human-centered organizational objectives without assuming that every operational action must be manually performed or continuously approved by a human.

This article therefore distinguishes **human-centered governance** from **human-exclusive execution**. Human-centered governance refers to institutional arrangements through which legitimate human authorities establish organizational objectives, define acceptable boundaries, and retain responsibility for consequential outcomes. Human-exclusive execution, by contrast, would require human actors to perform or directly authorize every organizational activity.

These arrangements are not equivalent. An organization may delegate bounded operational tasks to AI agents while preserving human-centered governance, provided that authority remains appropriately defined, oversight is meaningful, and accountability can be established.

2.3. Sociotechnical Systems and the Organizational Embeddedness of AI

Sociotechnical systems theory offers an important perspective for examining organizations in which technological and social components are increasingly interdependent. Rather than treating technological implementation as an isolated intervention, this perspective emphasizes the relationships among work processes, organizational structures, human capabilities, and technical systems.

Herrmann and Pfeiffer (2023) argue for keeping the organization within the analytical scope of human-centered AI, highlighting the importance of sociotechnical arrangements beyond individual interactions between users and algorithms. Their perspective suggests that AI integration should be

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

evaluated in relation to organizational roles, decision-making practices, work processes, and institutional conditions.

Schoenherr and Michael (2026) similarly adopt a multilevel sociotechnical perspective on human-centered AI. This orientation is particularly valuable for hybrid human–AI organizations because agentic systems may influence several organizational levels simultaneously. At the individual level, AI agents may alter managerial cognitive demands and professional responsibilities. At the team level, they may change coordination processes and patterns of interdependence. At the organizational level, they may reshape decision rights, reporting structures, and governance mechanisms.

The organizational consequences of AI autonomy therefore cannot be inferred solely from the technical characteristics of individual agents. The same AI capability may produce different outcomes depending on organizational context. For example, autonomous scheduling may support coordination in a standardized operational environment but generate conflict in a setting where scheduling decisions involve complex interpersonal considerations, negotiated priorities, or professional discretion.

Kossowska-Lai et al. (2026) connect AI and organizational resilience through a human-centered Industry 5.0 perspective. This relationship is relevant because resilience in hybrid organizations may depend on the capacity to detect failures, reassign responsibilities, preserve human expertise, and restore organizational functions when AI-supported processes become unreliable.

Similarly, Atay et al. (2025) discuss human-centered technology and talent management within the Society 5.0 paradigm. Their perspective broadens the discussion from technological efficiency toward the organizational capabilities and human competencies required for sustainable digital transformation.

A sociotechnical approach also suggests that organizational redesign should accompany AI deployment. Introducing autonomous agents without adjusting role definitions, oversight procedures, training arrangements, and escalation mechanisms may create inconsistencies between technological capabilities and institutional expectations.

Accordingly, the present article conceptualizes hybrid human–AI organizations as **multilevel sociotechnical arrangements** in which operational capabilities, authority structures, coordination practices, and accountability mechanisms must be jointly considered.

2.4. Decision Rights, Managerial Authority, and Distributed Operational Agency

Decision rights constitute a central dimension of organizational design because they determine who is authorized to make decisions, under what conditions, and within which institutional boundaries. In conventional organizational arrangements, formal decision rights are generally associated with human positions, managerial roles, committees, and legally constituted organizational bodies.

The emergence of autonomous AI agents requires a more differentiated understanding of decision-making authority. An AI system may possess technical access to organizational information, the ability to generate recommendations, or permission to execute transactions. However, these capabilities should not automatically be interpreted as independent institutional authority.

The distinction between **decision capability**, **delegated execution permission**, and **accountability** is therefore essential.

Decision capability refers to the technical capacity to process information, evaluate alternatives, or generate actions. Delegated execution permission refers to the authorization granted by an organization for an AI system to undertake specified activities within defined boundaries. Accountability refers to the obligation of identifiable human actors or institutions to justify decisions, respond to failures, and address consequences.

These dimensions may be distributed differently across organizational configurations. A manager may retain final authorization while an AI agent conducts extensive analysis. Alternatively, an AI agent may execute low-risk operational activities independently within predefined limits, while human managers remain responsible for establishing the rules and monitoring outcomes.

Wang et al. (2023) provide a relevant point of reference through their examination of human agency in AI configurations supporting organizational decision-making. Their work reinforces the importance of considering human agency when analysing AI-supported organizational arrangements. Sun and Xu (2026), through their discussion of hybrid cognitive authority and algorithmic subjectivity in knowledge management, offer another perspective on the changing relationships between human expertise and AI-mediated information processes.

Nevertheless, the concept of distributed agency requires careful specification. In philosophical and sociological traditions, agency may involve intentionality, interpretation, reflexivity, and the

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

capacity to act within social structures. These characteristics should not be assumed to apply to contemporary AI systems in the same manner as to human beings.

To avoid conceptual ambiguity, this article introduces **distributed operational agency**, defined as:

The organizational allocation of task initiation, information processing, coordination, and bounded execution capabilities across human actors and autonomous AI systems, under institutionally established decision rights and accountability arrangements.

This definition has four implications.

First, operational agency is functionally defined. It concerns what human and technological components can perform within organizational processes rather than attributing equivalent forms of consciousness or intentionality to them.

Second, operational agency is institutionally bounded. The scope of an AI agent's permitted actions depends on organizational authorization, access controls, operating rules, and governance procedures.

Third, distributed operational agency is relational. Its organizational significance emerges through the interaction of human judgement, AI capabilities, task interdependence, and coordination structures.

Fourth, distributed operational agency does not eliminate human or institutional responsibility. Delegating execution to an AI system does not automatically transfer legal or moral accountability to that system.

These distinctions provide the conceptual basis for examining how organizations may distribute operational activities without creating ambiguity regarding legitimate authority.

2.5. Human–AI Collaboration, Agentic Management, and Organizational Coordination

Research on AI in management has frequently examined the relationship between technological automation and human augmentation. Raisch and Krakowski (2021) argue that automation and augmentation should be understood as interdependent organizational processes rather than as mutually exclusive strategic alternatives. This perspective is particularly important because autonomous AI agents may combine both functions within the same organizational workflow.

For example, an AI agent may autonomously collect market intelligence, prepare a strategic analysis, coordinate information from multiple departments, and recommend actions to a human executive. The process simultaneously involves automated execution, augmented managerial cognition, and human–AI coordination.

Agentic management extends these relationships by introducing AI systems that can undertake sequences of actions rather than simply respond to individual requests. Ajmal and Islam (2025) explicitly examine the transition from human-centered to agentic management, positioning autonomous AI participation as a challenge to established organizational theory.

Chen (2026) similarly addresses the implications of human–machine collaboration for managerial rationality. Together, these emerging perspectives motivate a reconsideration of managerial roles, particularly where managers shift from directly executing or supervising tasks toward designing delegation boundaries, coordinating technological capabilities, and evaluating organizational consequences.

An important implication is that organizational coordination may increasingly occur across three relationships: human-to-human, human-to-AI, and AI-to-AI. These relationships differ in their requirements for communication, interpretation, escalation, and accountability.

Human-to-human coordination involves established organizational processes such as meetings, reporting relationships, professional norms, and interpersonal negotiation. Human-to-AI coordination requires the specification of objectives, constraints, permissions, and criteria for acceptable performance. AI-to-AI coordination may involve automated information exchange and task execution across multiple technological agents, potentially without direct human participation in every intermediate action.

The emergence of these arrangements creates new organizational design challenges. Managers must determine how agents communicate, how conflicting outputs are resolved, when activities should be escalated to human decision-makers, and how the organization can reconstruct the sequence of actions leading to a particular outcome.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Research on educational leadership offers illustrative perspectives on these issues. Costache (2026a, 2026b) examines autonomous AI agents and human–AI governance in educational settings. Although the institutional characteristics of higher education differ from those of commercial firms, such discussions are relevant to the broader challenge of coordinating professional judgement, delegated technological activity, and organizational responsibility.

The implications extend to leadership capabilities. Costache (2026c), in a bibliometric analysis of remote-work leadership, addresses organizational leadership in digitally mediated environments. Costache (2026d) examines continuous improvement through Kaizen leadership in the AI era, while Costache (2026e) explores the strategic implications of shifting legitimacy boundaries through Overton leadership. These perspectives offer complementary lenses for considering how leaders may adapt coordination practices, organizational expectations, and governance arrangements as AI capabilities evolve.

However, their relevance to hybrid organizational design is primarily contextual and theoretical. They should not be interpreted as direct empirical evidence that a particular configuration of autonomous AI agency improves organizational performance.

2.6. Organizational Adaptation, Human Capabilities, and Governance

The introduction of autonomous AI agents is not solely a technological implementation problem. It also involves organizational learning, institutional adaptation, employee capabilities, and the management of uncertainty.

Petcu and Costache (2026) examine functional fixedness and organizational inertia from a microfoundational perspective on leadership failure. Their conceptual focus is relevant to hybrid organizational transformation because established routines and assumptions may limit an organization's capacity to redesign decision rights and coordination mechanisms. The introduction of AI agents may therefore require changes not only in technical infrastructure but also in managerial mental models and established organizational practices.

Costache and Petcu (2025) address burnout and resilience in educational contexts, emphasizing the significance of wellbeing and contextual conditions. Although their study does not directly establish relationships between AI agents and employee wellbeing, it provides a complementary perspective for considering the human consequences of organizational transformation. In hybrid organizations, the redistribution of tasks may reduce certain workloads while introducing new demands associated with monitoring, exception handling, technological uncertainty, and responsibility for AI-mediated outcomes.

Research on digital learning and capability development also offers relevant contextual insights. Costache (2026f) discusses AI, microlearning, and emotional intelligence in educational transformation, while Kimalel (2025) examines competency development in a different educational and technological setting. These works are not direct foundations for autonomous-agent organizational theory, but they reinforce the broader importance of learning environments and capability development during technological change. Similarly, Costache (2025) examines sustainability pathways involving digital innovation and institutional transformation in higher education. This perspective may inform the discussion of organizational sustainability, particularly where technological integration must be aligned with broader institutional objectives rather than evaluated exclusively through short-term efficiency gains.

The governance dimension is especially important when AI systems are permitted to execute organizational actions. Governance arrangements may include authorization thresholds, access permissions, audit trails, monitoring procedures, escalation protocols, and mechanisms for human intervention.

Rassal and Serra (2027) propose a human-centered AI governance perspective within AI-enabled hospitality ecosystems. Because this reference is dated 2027, it should be treated as forthcoming or provisional until its publication status and bibliographic details have been confirmed. Its proposed emphasis on governance capabilities is conceptually relevant to the organizational arrangements examined in this article.

These perspectives collectively suggest that effective hybrid organizational design requires more than technical readiness. It may also depend on leadership adaptability, workforce capabilities, institutional legitimacy, organizational resilience, and the capacity to establish credible governance mechanisms.

However, these relationships should be regarded as theoretical expectations rather than established empirical findings. Their direction and magnitude may vary according to task

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

characteristics, organizational context, regulatory conditions, and the degree of autonomy delegated to AI agents.

2.7. Theoretical Synthesis and Identification of the Research Gap

The reviewed literature offers several complementary foundations for understanding the transition toward hybrid human–AI organizations.

Organizational design theory explains why firms establish authority structures, coordination mechanisms, and information-processing arrangements. Human-centered AI research emphasizes the importance of preserving meaningful human participation and aligning technological systems with human objectives. Sociotechnical systems theory demonstrates that technological capabilities and organizational arrangements are interdependent. Research on human–AI collaboration identifies tensions between automation and augmentation, while emerging agentic management scholarship raises questions about the organizational consequences of autonomous AI participation.

Nevertheless, these theoretical perspectives leave an important integrative question: **How should operational agency and decision rights be distributed across human and AI participants while preserving coherent organizational coordination and accountability?**

Three conceptual limitations are particularly relevant.

First, the distinction between technological capability and legitimate organizational authority requires greater specification. The ability of an AI agent to perform a task does not necessarily establish that it should be authorized to perform that task independently.

Second, existing perspectives do not necessarily provide a unified classification of organizational arrangements based on the interaction between AI autonomy, human oversight, and coordination responsibilities. A configuration-based framework could help explain why different organizational contexts require different forms of human–AI participation.

Third, organizational outcomes should be examined as potentially contingent on governance arrangements rather than assumed to improve automatically as AI autonomy increases. Higher levels of autonomy may create opportunities for efficiency and adaptability while also introducing risks involving opacity, coordination failures, accountability ambiguity, and excessive technological dependence.

To address these limitations, the present article develops the **Hybrid Organizational Agency Architecture (HOAA)**. The framework conceptualizes organizational design through five interdependent dimensions: the scope of AI operational autonomy, the allocation of decision rights, the configuration of coordination mechanisms, the intensity of human oversight, and the clarity of accountability arrangements.

The HOAA distinguishes four analytical configurations: human-directed AI assistance, supervised AI delegation, coordinated human–AI workflows, and bounded agent-led operations. These configurations are not proposed as a universal progression toward increasing autonomy. Instead, they represent alternative organizational arrangements whose appropriateness may depend on task risk, reversibility, organizational capabilities, regulatory requirements, and strategic objectives.

The resulting theoretical perspective shifts the central question from whether organizations should adopt autonomous AI agents to **how different distributions of human and AI operational capabilities should be organized, authorized, coordinated, and governed**.

Additional scholarship further clarifies the organizational foundations of hybrid human–AI arrangements. Shrestha et al. (2019) examine how artificial intelligence may reshape organizational decision-making structures, while Jarrahi (2018) conceptualizes human–AI symbiosis in organizational decision processes. Kellogg et al. (2020) demonstrate the relevance of algorithmic systems to questions of organizational control and managerial authority. From a human-factors perspective, Parasuraman et al. (2000) provide a foundational framework for distinguishing types and levels of human interaction with automation, while Bainbridge (1983) identifies persistent challenges associated with human supervision of highly automated systems. Research on trust further suggests that effective human–AI interaction depends on appropriate reliance and the characteristics of technological systems and their organizational contexts (Lee & See, 2004; Glikson & Woolley, 2020). Finally, Seeber et al. (2020) extend the discussion toward AI participation in team collaboration. Collectively, these contributions reinforce the need for an organizational framework that integrates technological autonomy, decision-right allocation, coordination, human oversight, and accountability. The proposed Hybrid Organizational Agency Architecture addresses this need by conceptualizing distributed operational agency as an organizational design problem rather than merely a technological capability.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

3. Conceptual Research Approach

3.1. Research Design and Theoretical Orientation

This study adopts a conceptual research design intended to develop an explanatory framework for organizational arrangements involving human actors and autonomous artificial intelligence (AI) agents. The methodological objective is not to estimate the effects of AI adoption or evaluate a particular technological implementation, but to clarify emerging organizational constructs, establish relationships among them, and formulate propositions that can guide subsequent empirical research.

Conceptual research is appropriate when an emerging phenomenon challenges the explanatory boundaries of established theories. Jaakkola (2020) distinguishes several approaches to conceptual article development, including theory synthesis, theory adaptation, typology, and model development. The present study combines theory synthesis with typology construction and conceptual model development. It integrates established organizational theories with emerging scholarship on agentic AI, differentiates alternative configurations of human–AI organizational arrangements, and proposes relationships among governance conditions and organizational outcomes.

The study is also informed by Whetten's (1989) criteria for theoretical contribution, particularly the need to identify relevant constructs, explain their relationships, establish the theoretical reasoning underlying those relationships, and specify the conditions under which they may apply. Accordingly, the proposed framework addresses four questions: what dimensions constitute hybrid human–AI organizational design; how these dimensions interact; why particular configurations may produce different organizational consequences; and under which organizational conditions these relationships may be expected to hold.

The unit of analysis is the organizational arrangement through which human actors and autonomous AI agents jointly participate in organizational processes. This level of analysis differs from studies focused exclusively on individual attitudes toward AI, technical agent performance, or dyadic human–AI interactions. Although these phenomena may influence organizational design, the present study examines how operational capabilities and decision authority are structured at the organizational level.

The article therefore develops an organizational architecture rather than a model of individual technology acceptance. Its central concern is the institutional organization of human and AI participation, including the allocation of operational initiative, decision rights, coordination responsibilities, oversight, and accountability.

3.2. Theoretical Integration and Literature Selection

The conceptual framework is developed through an integrative synthesis of complementary theoretical perspectives. The literature was organized around five analytical domains: classical organizational design and decision-making; human-centered organizations and human-centered AI; sociotechnical systems; human–AI collaboration and agentic management; and organizational governance, adaptability, and resilience.

Classical organizational design literature provides the foundational concepts of bounded rationality, information-processing requirements, authority, and coordination (Simon, 1947; March & Simon, 1958; Galbraith, 1974). Human-centered AI scholarship contributes perspectives on human participation, technological design, organizational wellbeing, and responsible AI integration (Margetis et al., 2021; Ozmen Garibay et al., 2023; Wilkens et al., 2023; Townsend & Romme, 2024).

Sociotechnical research contributes the principle that organizational and technological arrangements should be examined jointly rather than as independent components (Herrmann & Pfeiffer, 2023; Schoenherr & Michael, 2026). Research on AI-enabled management and human–machine collaboration informs the analysis of delegated execution, managerial rationality, and changing coordination relationships (Raisch & Krakowski, 2021; Ajmal & Islam, 2025; Chen, 2026). Finally, scholarship addressing AI governance, organizational maturity, resilience, and leadership adaptation informs the proposed framework's boundary conditions and institutional implications (Costache, 2026a, 2026b; Kossowska-Lai et al., 2026; Winby & Xu, 2026).

The selection of these perspectives is theoretically purposive rather than statistically representative. Sources are incorporated according to their relevance to the constructs and relationships being developed. Foundational organizational theories are used to establish the conceptual logic of decision rights and coordination, while emerging AI scholarship is used to identify organizational questions that established theories may not fully anticipate.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

The study does not claim to employ a systematic or exhaustive literature search. No bibliometric sampling frame, formal screening protocol, or quantitative evidence synthesis is reported. Accordingly, the resulting framework should be interpreted as a theoretically reasoned conceptual proposal rather than as a conclusion derived from a comprehensive systematic review.

3.3. Conceptual Boundary: Human Agency and AI Operational Autonomy

A central methodological requirement is the differentiation of human agency from the operational capabilities of autonomous AI systems. Without this distinction, discussions of hybrid organizations risk attributing human-like authority, responsibility, or intentionality to technological systems merely because those systems can perform complex activities.

In this study, **human agency** refers to the capacity of organizational actors to exercise judgement, interpret objectives, make choices, and act within institutional and social arrangements. Human agency may be constrained by organizational structures, professional responsibilities, and technological systems, but it remains conceptually distinct from automated task execution.

AI operational autonomy refers to the degree to which an AI system can initiate, sequence, and execute authorized activities without requiring direct human intervention at every operational step. This definition is functional rather than anthropomorphic. It does not assume that AI systems possess human consciousness, moral agency, or independent institutional legitimacy.

Distributed operational agency describes the organizational arrangement through which operational capabilities are allocated across human actors and AI systems. The concept captures the distribution of task initiation, information processing, coordination, and bounded execution while retaining the distinction between technical action and legitimate organizational authority.

Three related concepts are also differentiated.

First, **decision capability** concerns the capacity to analyse alternatives or generate actions. Second, **decision rights** concern the organizational authorization to select or execute particular courses of action. Third, **accountability** concerns the obligation of identifiable human actors or institutions to explain, justify, and respond to the consequences of organizational decisions.

These distinctions are important because an AI agent may possess substantial decision capability while being authorized to exercise only limited execution permissions. Similarly, an organization may delegate routine operational actions to an AI system without transferring ultimate responsibility for the governance of those actions.

The framework therefore treats AI autonomy as an organizationally bounded characteristic rather than an unrestricted or inherently desirable objective.

3.4. Analytical Dimensions of Hybrid Organizational Design

The development of the Hybrid Organizational Agency Architecture proceeds through the identification of five analytical dimensions.

Operational autonomy concerns the extent to which AI agents can initiate and execute tasks independently within organizationally defined boundaries. It may range from responding to human instructions to conducting multistep activities under predefined permissions.

Decision-right allocation concerns the distribution of authorization among human actors, managerial roles, and AI-enabled processes. It includes the distinction between recommendation, conditional execution, and autonomous execution within established limits.

Coordination mechanisms concern how human and AI activities are integrated into organizational workflows. They include task assignment, information exchange, exception management, conflict resolution, and the sequencing of interdependent activities.

Human oversight concerns the mechanisms through which authorized human actors can monitor, review, interrupt, modify, or terminate AI-supported activities. Oversight may be continuous, periodic, exception-based, or concentrated at specific authorization points.

Accountability clarity concerns the extent to which responsibility for objectives, permissions, monitoring, and organizational consequences is explicitly assigned to identifiable human roles or institutional bodies.

These dimensions are conceptually related but should not be treated as interchangeable. For example, high operational autonomy does not necessarily imply weak oversight, because an organization may combine extensive bounded execution with strong monitoring, auditability, and escalation procedures. Likewise, formal decision rights may remain centralized even when routine

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

operational execution is technologically distributed.

The framework uses these dimensions to distinguish alternative organizational configurations. Their selection reflects theoretical relevance to the research questions rather than empirical factor extraction or statistical validation.

3.5. Configuration Development and Typological Logic

The HOAA is developed as an analytical typology of organizational arrangements involving human and AI participation. Typological reasoning is appropriate because organizations may combine technological autonomy and human governance in qualitatively different ways rather than occupying positions along a single continuum.

Four configurations are proposed.

The first, **human-directed AI assistance**, describes arrangements in which AI systems provide information, recommendations, or task support while human actors retain direct control over task initiation and consequential execution.

The second, **supervised AI delegation**, describes arrangements in which AI agents perform specified activities within delegated boundaries, subject to human supervision and defined authorization or escalation requirements.

The third, **coordinated human–AI workflows**, describes arrangements in which human actors and AI agents undertake interdependent activities through structured coordination mechanisms. Operational initiative may be distributed across the workflow, although decision rights and accountability remain institutionally defined.

The fourth, **bounded agent-led operations**, describes arrangements in which AI agents can initiate and execute substantial sequences of operational activities within explicitly established objectives, permissions, constraints, and intervention mechanisms.

The configurations are differentiated primarily by the distribution of operational initiative, the scope of delegated execution, and the organization of human oversight and coordination.

They are not assumed to represent a universal sequence of organizational maturity. A highly regulated organization may appropriately maintain human-directed assistance for consequential decisions while adopting bounded agent-led operations for low-risk administrative processes. Multiple configurations may also coexist within the same organization.

Consequently, the typology is intended to support organizational comparison and theoretical explanation rather than prescribe increasing AI autonomy as an inevitable or universally desirable development path.

3.6. Development of Theoretical Propositions

Following construct clarification and typology development, the study formulates six theoretical propositions concerning the relationships between hybrid organizational design and organizational outcomes.

The propositions are developed through theoretical reasoning connecting organizational design dimensions to potential consequences for coordination, decision quality, adaptability, trust, and accountability. Their purpose is to identify relationships that may be empirically investigated rather than to assert that such relationships have already been demonstrated.

The reasoning process follows three principles.

First, proposed relationships must be consistent with established organizational theory. For example, expectations concerning coordination and information processing are grounded in organizational design perspectives rather than inferred solely from the technical capabilities of AI agents.

Second, propositions must account for potential trade-offs. Greater operational autonomy may increase execution speed while simultaneously increasing monitoring requirements or the consequences of poorly specified delegation boundaries.

Third, propositions must acknowledge organizational contingencies. The suitability of a particular configuration may depend on task complexity, risk exposure, reversibility, regulatory requirements, technological reliability, organizational capabilities, and the availability of meaningful human intervention.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Accordingly, the proposed relationships are conditional rather than universally deterministic. The framework does not assume that autonomous AI participation necessarily improves organizational performance or that human involvement is always associated with superior decision quality.

3.7. Methodological Scope and Limitations

The conceptual research design provides a basis for construct development, theoretical integration, and the formulation of research propositions. However, it also establishes important limitations.

First, the proposed Hybrid Organizational Agency Architecture has not been empirically validated. Its four configurations are analytically constructed organizational forms rather than categories derived from observed organizational populations.

Second, the framework does not estimate the prevalence of different configurations or establish their comparative effectiveness. Such assessments would require empirical investigation across organizational contexts.

Third, the proposed relationships among operational autonomy, governance mechanisms, and organizational outcomes remain theoretical expectations. They should not be interpreted as demonstrated causal effects.

Fourth, the study necessarily abstracts from some technological differences among autonomous AI systems. Agent capabilities, reliability, access permissions, and operational architectures may vary substantially, potentially influencing how the proposed organizational configurations function in practice.

Fifth, the framework's applicability may depend on institutional and contextual conditions, including industry characteristics, regulatory environments, organizational size, technological maturity, and the consequences of decision errors.

Future research could address these limitations through comparative organizational case studies, longitudinal investigations of AI delegation, scenario-based experiments, and multilevel studies of human–AI coordination. Such approaches would allow researchers to examine whether the proposed dimensions adequately distinguish organizational configurations and under what conditions particular arrangements are associated with desirable or undesirable outcomes.

Despite these limitations, the conceptual approach provides a structured foundation for examining an emerging organizational phenomenon. By distinguishing operational capabilities from legitimate authority and accountability, the study establishes a basis for analysing how human and autonomous AI participation may be organized without assuming equivalence between human and machine agency.

4. The Hybrid Organizational Agency Architecture (HOAA)

4.1. Conceptual Foundations of the Hybrid Organizational Agency Architecture

The Hybrid Organizational Agency Architecture (HOAA) is proposed as an integrative framework for examining how organizational structures may accommodate autonomous artificial intelligence (AI) agents while maintaining legitimate managerial authority and institutional accountability. Its central premise is that the organizational implications of AI autonomy depend not exclusively on technological capabilities but on the institutional arrangements through which those capabilities are authorized, coordinated, monitored, and constrained.

Traditional organizational structures allocate authority and responsibility through human roles, hierarchical relationships, and formal decision-making procedures. Autonomous AI agents introduce an additional organizational design challenge because they may perform activities previously associated with human operational initiative. Such activities can include identifying tasks, selecting intermediate actions, coordinating information across systems, and executing decisions within predefined operational boundaries.

The resulting organizational arrangement cannot be adequately characterized through a simple distinction between human decision-making and technological automation. As Raisch and Krakowski (2021) demonstrate, automation and augmentation are interdependent dimensions of AI-enabled management. Their relationship suggests that organizations may simultaneously delegate selected activities to AI systems and rely on human judgement for interpretation, coordination, and oversight.

Human-centered AI research provides a complementary foundation. Wilkens et al. (2023)

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

demonstrate that human-centered AI arrangements can take multiple configurations depending on organizational conditions and actor–structure relationships. The HOAA builds on this configuration-based perspective but addresses a distinct theoretical problem: the allocation of operational initiative, decision rights, and governance responsibilities when AI systems can execute organizational actions.

The architecture is based on three assumptions.

First, AI operational autonomy is **bounded rather than institutionally independent**. Even where an AI agent can execute activities without continuous human intervention, its permissions and operating constraints are established through organizational arrangements.

Second, human and AI operational capabilities are **potentially complementary but not interchangeable**. AI systems may contribute computational capacity, continuous monitoring, and automated execution, while human actors remain essential for legitimate goal setting, contextual judgement, normative evaluation, and accountability.

Third, organizational outcomes are **contingent on the alignment between technological autonomy and governance arrangements**. Increasing AI autonomy may create opportunities for faster coordination and operational adaptability, but these potential benefits depend on appropriate delegation boundaries, reliable information, monitoring mechanisms, and escalation procedures.

These assumptions establish the foundation for conceptualizing hybrid organizations as sociotechnical arrangements in which operational participation is distributed while legitimate authority remains institutionally structured.

4.2. Five Dimensions of Hybrid Organizational Design

The HOAA comprises five interdependent design dimensions: operational autonomy, decision-right allocation, coordination mechanisms, human oversight, and accountability clarity.

4.2.1. Operational Autonomy

Operational autonomy refers to the extent to which an AI agent can initiate, sequence, and execute authorized organizational activities without direct human intervention at every stage.

At lower levels of autonomy, AI systems primarily respond to human instructions and provide analytical assistance. At higher levels, they may undertake multistep activities, select intermediate actions, interact with other systems, and manage routine exceptions within predefined constraints.

Operational autonomy is not equivalent to unrestricted decision-making authority. An agent may possess substantial technical capability while remaining subject to narrow organizational permissions. Conversely, a relatively simple automated system may execute consequential actions if organizational controls are inadequate.

The relevant design question is therefore not merely how autonomous an AI system can become, but which operational activities it should be permitted to perform under specific organizational conditions.

4.2.2. Decision-Right Allocation

Decision-right allocation concerns the organizational distribution of authorization to select, approve, or execute particular courses of action.

In hybrid organizations, decision rights should be distinguished from technical execution capabilities. AI systems may be authorized to generate recommendations, select actions within predefined rules, or execute transactions under established thresholds. However, the legitimacy of these permissions originates in organizational governance rather than in the technological system itself.

Decision-right allocation can be examined through three categories: human-retained decisions, conditionally delegated decisions, and bounded AI-executed decisions.

Human-retained decisions require authorization by an identifiable human decision-maker. Conditionally delegated decisions permit AI execution when specified conditions are satisfied. Bounded AI-executed decisions allow agents to act independently within explicitly established operational limits.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

The appropriate allocation depends on task characteristics, potential consequences, reversibility, uncertainty, and institutional requirements.

4.2.3. Coordination Mechanisms

Coordination mechanisms determine how human and AI activities are integrated into coherent organizational processes.

In conventional organizations, coordination commonly relies on hierarchy, standardized procedures, professional norms, and interpersonal communication. Hybrid organizations may retain these mechanisms while introducing machine-mediated task allocation, automated information exchange, and agent-to-agent coordination.

Such arrangements may improve the speed of information transfer and routine execution. However, they may also generate new interdependencies, particularly when multiple agents operate across organizational functions.

Effective coordination therefore requires clarity regarding task ownership, communication protocols, escalation conditions, and the resolution of inconsistent or conflicting outputs.

Coordination quality should not be inferred solely from the speed of automated execution. A technically efficient workflow may remain organizationally ineffective if its outputs are difficult to interpret, challenge, or integrate into legitimate decision-making processes.

4.2.4. Human Oversight

Human oversight refers to the mechanisms through which authorized organizational actors can observe, evaluate, intervene in, or terminate AI-supported activities.

Oversight may take several forms, including prior authorization, continuous monitoring, periodic review, exception-based intervention, and retrospective auditing.

The appropriate oversight mechanism depends on the nature of the delegated activity. High-consequence or difficult-to-reverse decisions may require stronger authorization controls, whereas routine and reversible tasks may be compatible with less intensive direct supervision.

Meaningful oversight requires more than formal human involvement. Organizational actors must possess sufficient information, authority, competence, and practical opportunity to intervene when necessary.

Accordingly, the HOAA treats oversight as an organizational capability rather than a symbolic requirement that a human remain nominally involved.

4.2.5. Accountability Clarity

Accountability clarity concerns the explicit allocation of responsibility for the objectives, permissions, monitoring, and consequences of AI-supported organizational activities.

Distributed operational execution can complicate accountability when actions emerge from interactions among managers, technical teams, organizational policies, and multiple AI agents.

The architecture therefore distinguishes the operational origin of an action from responsibility for its authorization and governance.

An AI agent may initiate or execute an activity, but this does not establish that the agent possesses moral or legal accountability equivalent to that of a human organizational actor.

Accountability arrangements should identify who establishes delegation boundaries, who monitors compliance, who investigates failures, and who has authority to suspend or modify AI-supported processes.

These responsibilities may be distributed among several organizational roles, but their allocation should remain sufficiently explicit to support review, corrective action, and institutional legitimacy.

Table 1. Core Dimensions of the Hybrid Organizational Agency Architecture

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Dimension	Conceptual definition	Principal organizational design question
Operational autonomy	Degree of independent task initiation and execution within defined boundaries	What may AI agents do without immediate human intervention?
Decision-right allocation	Distribution of authorization to recommend, approve, or execute decisions	Which decisions remain human-controlled, and which may be delegated?
Coordination mechanisms	Processes integrating human and AI activities	How are tasks, information, and exceptions coordinated?
Human oversight	Capacity to monitor, review, intervene, and terminate AI-supported activities	When and how can authorized humans exercise control?
Accountability clarity	Explicit assignment of responsibility for delegated activities and consequences	Who is answerable for the governance and outcomes of AI-supported decisions?

Note. The dimensions are theoretically derived and have not been empirically validated as a measurement instrument.

4.3. Four Configurations of Hybrid Human–AI Organizations

The HOAA distinguishes four analytical configurations according to the distribution of operational initiative, the scope of AI execution permissions, and the structure of human coordination and oversight.

These configurations should not be interpreted as successive stages of technological progress. Their suitability is contingent on organizational objectives, task characteristics, governance requirements, and the consequences of potential failures.

4.3.1. Configuration I: Human-Directed AI Assistance

Human-directed AI assistance represents an organizational arrangement in which AI systems primarily support human activities through information retrieval, analysis, recommendation generation, or the preparation of routine outputs.

Human actors retain responsibility for initiating tasks, interpreting recommendations, selecting courses of action, and authorizing consequential execution.

In this configuration, AI functions predominantly as a cognitive or operational support resource rather than an independently executing organizational component.

For example, a financial manager may use an AI system to analyse expenditure patterns and identify potential cost reductions. The system produces analytical recommendations, but the manager determines which actions should be undertaken.

The principal organizational advantage of this configuration is the preservation of direct human

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

control over decisions. It may be particularly suitable where tasks involve high uncertainty, professional judgement, regulatory sensitivity, or substantial consequences.

Its limitations may include continued dependence on human availability, delays in execution, and limited use of the potential capabilities of autonomous AI agents.

Human-directed assistance therefore illustrates a hybrid arrangement in which technological information-processing capacity is expanded without substantially redistributing operational initiative.

4.3.2. Configuration II: Supervised AI Delegation

Supervised AI delegation describes arrangements in which AI agents are authorized to perform specified activities within predefined boundaries while remaining subject to human supervision.

Unlike human-directed assistance, this configuration permits AI agents to execute selected tasks without requiring direct human authorization at every operational step.

However, the scope of delegation is limited by explicit rules, authorization thresholds, monitoring requirements, and escalation procedures.

For example, a procurement agent may be permitted to reorder standardized supplies when inventory falls below an established threshold, provided that purchases remain within approved budgets and supplier agreements.

The organizational benefit is the potential reduction of routine managerial workload and faster execution of standardized activities.

Nevertheless, the configuration introduces new governance requirements. Managers must establish reliable delegation criteria, monitor agent performance, and define conditions under which activities should be suspended or escalated.

Supervised AI delegation therefore involves a partial redistribution of operational initiative while retaining a comparatively direct relationship between human supervisory roles and AI-executed activities.

4.3.3. Configuration III: Coordinated Human–AI Workflows

Coordinated human–AI workflows describe arrangements in which human actors and AI agents perform interdependent activities through structured coordination mechanisms.

In this configuration, operational initiative may move between human and artificial participants during a workflow. AI agents may identify relevant information, propose actions, initiate authorized tasks, and communicate with other systems, while human actors provide contextual judgement, negotiate priorities, and authorize consequential decisions.

For example, a product-development process may involve AI agents conducting market analysis, monitoring project milestones, and preparing alternative development scenarios, while human teams evaluate strategic fit, resolve conflicting priorities, and make final investment decisions.

The defining characteristic is not simply a higher degree of AI autonomy, but the organizational integration of multiple human and AI contributions.

This configuration may offer opportunities for improved coordination and organizational responsiveness, particularly when tasks require both computational processing and human contextual understanding.

However, it may also introduce ambiguity regarding task ownership, workflow dependencies, and the interpretation of agent-generated outputs.

Consequently, coordinated human–AI workflows require explicit communication protocols, role definitions, escalation procedures, and mechanisms for resolving inconsistent recommendations or actions.

The configuration illustrates distributed operational agency as a relational organizational phenomenon rather than an attribute of individual AI systems.

4.3.4. Configuration IV: Bounded Agent-Led Operations

Bounded agent-led operations describe organizational arrangements in which autonomous AI agents can initiate and execute substantial sequences of operational activities within explicitly established objectives, permissions, and constraints.

Human involvement is concentrated on strategic direction, authorization boundaries, performance evaluation, exception management, and institutional accountability rather than continuous supervision of every operational action.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

For example, an AI-enabled operations system may coordinate routine inventory monitoring, supplier communication, replenishment scheduling, and internal reporting within established financial, contractual, and operational limits.

Such arrangements may reduce coordination delays and support continuous execution in standardized environments.

However, the potential consequences of poorly specified objectives, system errors, cascading actions, or insufficient monitoring may increase as the scope of autonomous execution expands.

Accordingly, bounded agent-led operations require robust governance mechanisms, including access controls, auditability, operational limits, escalation thresholds, and effective intervention capabilities.

The term *agent-led* refers exclusively to the initiation and coordination of authorized operational activities. It does not imply that AI systems replace legitimate human leadership, assume independent institutional authority, or acquire moral or legal responsibility.

This configuration may be appropriate for selected activities with sufficiently clear objectives, manageable uncertainty, and effective safeguards. It should not be interpreted as the universally preferred form of hybrid organizational design.

4.4. Comparative Analysis of the Four Configurations

The four configurations can be differentiated through the distribution of operational initiative, decision authorization, coordination responsibilities, and oversight requirements.

Table 2. Comparative Characteristics of Hybrid Human–AI Organizational Configurations

Characteristic	Human-directed assistance	Supervised delegation	Coordinated workflows	Bounded agent-led operations
Principal operational initiative	Human	Human-defined, AI-executed	Shared across workflow	AI within authorized boundaries
AI execution autonomy	Minimal	Limited	Task-dependent	Relatively extensive
Human involvement	Direct and frequent	Supervisory	Collaborative and exception-based	Strategic, governance-focused, and exception-based
Decision rights	Primarily human-retained	Conditionally delegated	Distributed by task and role	Delegated within strict limits
Coordination complexity	Relatively low	Moderate	Potentially high	Potentially high

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Principal governance concern	Quality of AI recommendations	Compliance with delegation boundaries	Role clarity and workflow integration	Cascading failures and accountability
Illustrative application	Decision-support analysis	Routine procurement	Cross-functional project coordination	Bounded operational orchestration

Note. The classifications are conceptual. Relative autonomy and coordination complexity are analytical expectations rather than empirical measurements. Organizations may combine multiple configurations across functions.

The comparison highlights an important theoretical distinction: increasing AI autonomy does not necessarily produce a corresponding reduction in organizational complexity.

Human-directed assistance may require intensive human participation but comparatively straightforward authority arrangements. Coordinated workflows may distribute operational activities more extensively while increasing the need for communication and exception management. Bounded agent-led operations may reduce direct human involvement in routine execution but increase the importance of monitoring, authorization design, and organizational resilience.

The framework therefore rejects a simple assumption that more autonomous AI necessarily produces more effective organizations.

Instead, it suggests that organizational effectiveness depends on the fit between the selected configuration and the characteristics of the activities being performed.

4.5. The Integrative Logic of the HOAA

The HOAA conceptualizes hybrid organizational design as the relationship between leadership architecture, distributed operational agency, and organizational outcomes.

Leadership architecture refers to the institutional arrangements through which organizational leaders establish objectives, define delegation boundaries, allocate decision rights, and implement governance mechanisms.

Distributed operational agency refers to the organization of task initiation, information processing, coordination, and execution across human actors and AI systems.

Organizational outcomes refer to potential consequences for decision quality, coordination effectiveness, adaptability, trust, and accountability.

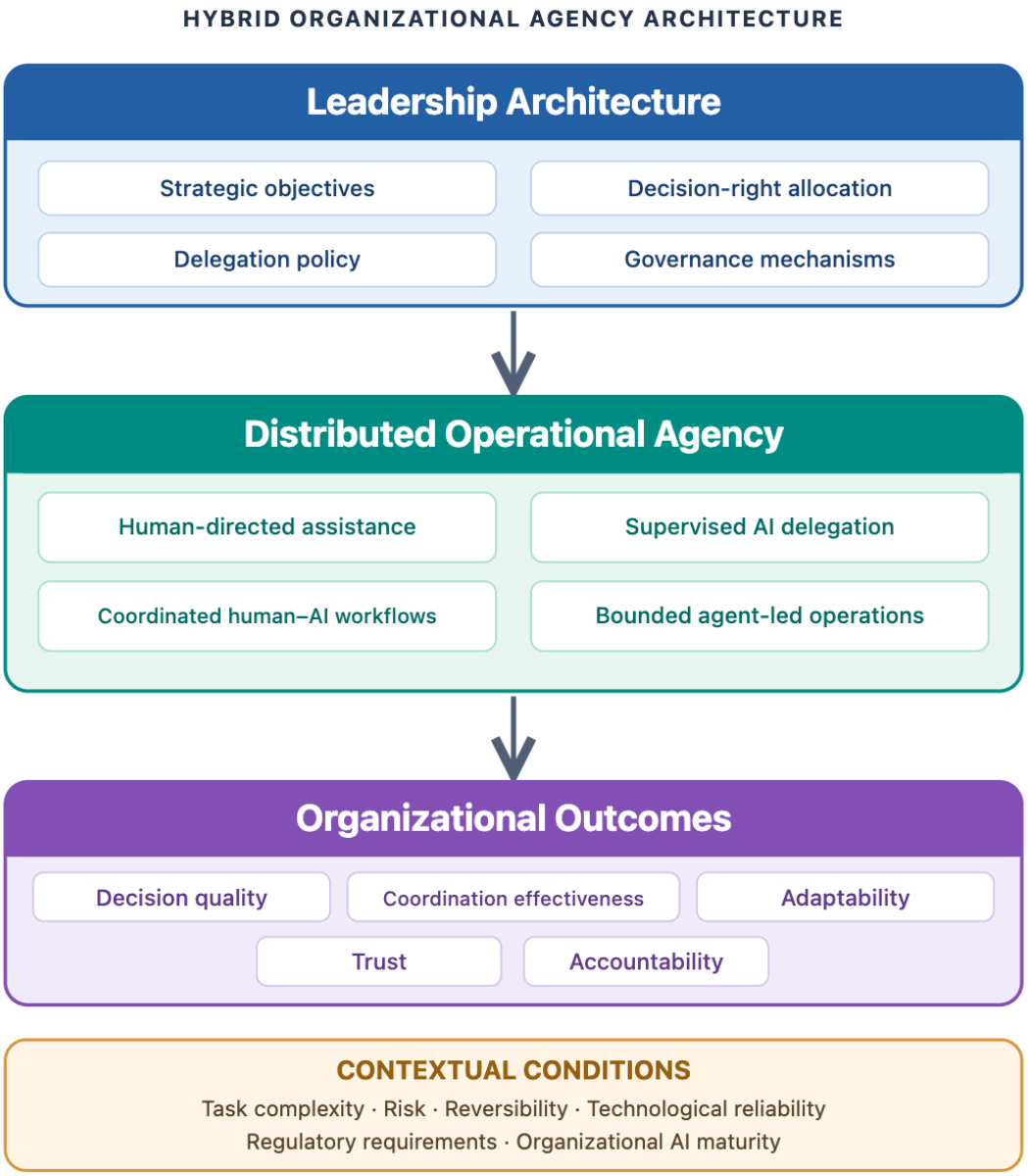
The architecture proposes that leadership arrangements influence organizational outcomes partly through the way they structure human–AI operational participation.

However, these relationships are not assumed to be linear or universally positive. The consequences of distributed agency may depend on contextual conditions, including task risk, reversibility, technological reliability, organizational competence, and regulatory requirements.

Figure 1. The Hybrid Organizational Agency Architecture: Leadership, Distributed Operational Agency, and Organizational Outcomes

Note. Figure 1 presents proposed theoretical relationships rather than empirically demonstrated causal effects. The four configurations represent alternative arrangements within the distributed operational agency construct, not sequential developmental stages.Arrows indicate proposed theoretical relationships, not empirically established causal effects. The four organizational configurations represent alternative arrangements rather than sequential stages of AI maturity.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72



Source: Authors' own conceptual contribution

The framework's central theoretical contribution is the separation of **operational autonomy** from **institutional authority**.

Organizations may distribute execution capabilities across AI agents while retaining human control over objectives, permissions, and accountability. Conversely, organizations may adopt advanced AI technologies without establishing adequate governance mechanisms, potentially creating uncertainty regarding authorization and responsibility.

The HOAA therefore emphasizes that hybrid organizational design requires the joint configuration of technological capabilities and institutional arrangements.

4.6. Organizational Fit and the Contingent Nature of AI Autonomy

The suitability of a particular hybrid organizational configuration is likely to depend on the relationship between task characteristics and governance capabilities.

Activities characterized by clear objectives, standardized procedures, limited consequences,

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

and high reversibility may be more compatible with extensive bounded AI execution.

Activities involving ambiguous objectives, substantial ethical implications, high uncertainty, or difficult-to-reverse consequences may require stronger human participation and authorization.

Organizational AI maturity may also influence the feasibility of different arrangements. Firms with reliable monitoring systems, clear accountability structures, trained personnel, and effective intervention procedures may be better positioned to manage autonomous operational activities than organizations lacking these capabilities.

However, technological maturity alone does not establish the appropriateness of greater autonomy. A technologically sophisticated organization may still choose to retain direct human control over particular decisions because of professional obligations, institutional legitimacy, or regulatory requirements.

This contingency-based perspective is consistent with organizational design theory, which emphasizes alignment between structural arrangements and environmental or task demands.

It also reinforces the distinction between organizational capability and organizational choice. The ability to delegate a task to an AI agent does not establish that delegation is desirable.

Accordingly, the HOAA should be understood as a framework for evaluating organizational design alternatives rather than prescribing a universal trajectory toward increasingly autonomous AI systems.

4.7. Theoretical Contribution of the Proposed Architecture

The HOAA advances the discussion of hybrid human–AI organizations in three principal ways.

First, it introduces distributed operational agency as a construct for examining how human and AI capabilities are allocated within organizational processes. This construct extends beyond conventional automation by incorporating task initiation, coordination, and bounded execution.

Second, it proposes a configuration-based approach to hybrid organizational design. Rather than treating AI autonomy as a single developmental continuum, the architecture identifies alternative arrangements characterized by different relationships among human control, delegated execution, and coordination.

Third, it integrates decision-right allocation and accountability into the analysis of AI-enabled organizational structures. This integration addresses the risk of conflating technological capability with legitimate institutional authority.

The framework also provides a basis for developing empirically testable propositions. Its dimensions and configurations can be operationalized through measures of delegation scope, human intervention, workflow coordination, auditability, and accountability clarity.

Nevertheless, the HOAA remains a conceptual proposal. Its usefulness depends on subsequent theoretical refinement and empirical evaluation across organizational contexts.

5. Theoretical Propositions: Distributed Agency and Organizational Outcomes

5.1. Decision-Right Clarity and Effective AI Delegation

The distribution of decision rights constitutes a fundamental organizational design problem when autonomous AI agents are authorized to participate in operational processes. In conventional organizations, authority is typically associated with formal roles, hierarchical relationships, and established procedures. The introduction of autonomous AI agents creates additional complexity because technological systems may possess the capacity to execute actions without holding independent institutional authority.

Organizational design theory suggests that effective coordination depends partly on the clarity with which decision-making responsibilities are allocated. Simon (1947) and March and Simon (1958) established the importance of organizational structures and decision procedures in addressing the limitations of individual rationality. Galbraith (1974) subsequently demonstrated that organizational arrangements must respond to information-processing demands and task interdependence.

These perspectives suggest that unclear decision rights may generate coordination difficulties even when the technological capabilities of AI agents are substantial. For example, an autonomous

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

procurement agent may be technically capable of identifying suppliers, negotiating predefined terms, and initiating purchases. However, if the organization has not specified authorization thresholds, exception conditions, and supervisory responsibilities, the agent's operational capabilities may create ambiguity regarding legitimate action.

Decision-right clarity is therefore expected to influence the effectiveness of AI delegation through several mechanisms.

First, explicit authorization boundaries may reduce uncertainty regarding which activities AI agents are permitted to execute independently. Second, clear decision rights may facilitate coordination between human managers and autonomous systems by distinguishing operational execution from managerial approval. Third, formal delegation rules may improve organizational monitoring by establishing identifiable standards against which agent behaviour can be evaluated.

Nevertheless, decision-right clarity should not be equated with excessive procedural rigidity. Highly restrictive delegation arrangements may reduce the adaptability of AI-supported workflows, particularly when agents encounter circumstances that require flexible responses. The organizational challenge is to establish sufficiently clear authority boundaries while preserving appropriate operational discretion.

The expected relationship may therefore depend on the compatibility between delegation rules and task characteristics. Clear boundaries are likely to be particularly important when activities involve financial commitments, external stakeholders, regulatory obligations, or difficult-to-reverse consequences.

Proposition 1 (P1). Greater clarity in the allocation of decision rights between human actors and autonomous AI agents is positively associated with the effectiveness of AI delegation, particularly when delegation boundaries are aligned with task risk and operational requirements.

This proposition extends organizational design theory by treating AI execution permissions as an additional component of decision-right allocation. It also establishes a basis for examining how organizations distinguish technological capability from legitimate organizational authorization.

5.2. AI Operational Autonomy, Coordination Efficiency, and Task Interdependence

Autonomous AI agents may influence organizational coordination by reducing the need for repeated human intervention in routine operational activities. Systems capable of initiating tasks, exchanging information, and executing predefined actions may accelerate workflows and reduce delays associated with sequential human approvals.

From an information-processing perspective, these capabilities may enable organizations to respond more rapidly to routine environmental changes. For example, autonomous agents coordinating inventory information, supplier availability, and delivery schedules may reduce the time required to complete standardized operational processes.

However, the relationship between AI autonomy and coordination effectiveness is unlikely to be uniformly positive.

Galbraith's (1974) organizational information-processing perspective suggests that coordination requirements increase with uncertainty and task interdependence. Autonomous AI agents may reduce some information-processing burdens while creating new dependencies among technological systems, organizational units, and human decision-makers.

When tasks are standardized, reversible, and supported by reliable information, bounded AI autonomy may improve coordination by reducing unnecessary operational interruptions. Conversely, when activities involve ambiguous objectives, conflicting priorities, or substantial contextual uncertainty, extensive autonomous execution may create coordination problems that require additional human intervention.

The automation–augmentation paradox identified by Raisch and Krakowski (2021) is relevant to this relationship. Their theoretical analysis demonstrates that automation and augmentation are interdependent and may generate organizational tensions when one is emphasized without sufficient attention to the other.

In hybrid organizations, this tension may emerge when autonomous execution is expanded without corresponding improvements in coordination mechanisms. For instance, multiple agents may independently optimize local tasks while producing conflicting outcomes at the organizational level.

Accordingly, the organizational value of AI autonomy is expected to depend on the

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

characteristics of task interdependence and the mechanisms through which agent activities are coordinated.

Proposition 2 (P2). Greater bounded AI operational autonomy is positively associated with coordination efficiency in standardized and relatively predictable workflows, whereas this relationship weakens as task interdependence and contextual uncertainty increase unless coordination mechanisms are correspondingly strengthened.

This proposition introduces a contingency perspective into the relationship between autonomous AI participation and organizational coordination. It suggests that autonomy should be evaluated in relation to task conditions rather than treated as an independent indicator of organizational advancement.

5.3. Human–AI Coordination and Organizational Decision Quality

Decision quality represents a central organizational outcome because the value of AI-enabled processes depends not only on execution speed but also on the appropriateness, reliability, and consequences of organizational choices.

Human and AI systems may contribute different capabilities to decision-making processes. AI systems can support extensive information processing, pattern identification, and the rapid comparison of alternatives. Human actors contribute contextual interpretation, professional experience, stakeholder understanding, and normative judgement.

These capabilities may be complementary, but their combination does not automatically improve decisions. The effectiveness of human–AI collaboration depends on how information, judgement, and authority are coordinated.

Chen (2026) approaches managerial rationality from a human–machine collaboration perspective, highlighting the relevance of reconsidering decision processes in AI-enabled organizations. Raisch and Krakowski (2021) similarly emphasize the interdependence of automation and augmentation, suggesting that organizational performance cannot be understood through a simple substitution model.

Within the HOAA, coordinated human–AI workflows are expected to support decision quality when organizations establish effective mechanisms for combining technological analysis with human evaluation.

Three mechanisms are particularly relevant.

First, complementary information processing may improve the identification and assessment of decision alternatives. Second, explicit role allocation may reduce confusion regarding who evaluates AI-generated recommendations and who authorizes consequential actions. Third, structured opportunities for human challenge and intervention may help identify errors, unsupported assumptions, or recommendations that conflict with organizational objectives.

However, the potential benefits may be undermined by excessive reliance on AI outputs, insufficient human understanding of system limitations, or unclear responsibility for final decisions.

For example, an AI agent may generate an apparently efficient staffing recommendation based on operational data. Human managers may nevertheless need to consider employee wellbeing, professional development, fairness, and organizational culture before adopting the recommendation.

Decision quality should therefore be understood as a multidimensional outcome that includes factual reliability, contextual appropriateness, procedural legitimacy, and alignment with organizational objectives.

Proposition 3 (P3). More effective coordination of complementary human and AI capabilities is positively associated with organizational decision quality, particularly when human actors retain meaningful opportunities to evaluate, challenge, and contextualize AI-generated recommendations and actions.

This proposition advances the argument that the organizational value of hybrid intelligence depends on the quality of coordination rather than simply on the presence of advanced AI technologies.

5.4. Distributed Operational Agency and Organizational Adaptability

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Organizational adaptability refers to the capacity to recognize changes, modify activities, reconfigure resources, and respond to evolving operational or environmental conditions.

Autonomous AI agents may contribute to adaptability through continuous information monitoring, rapid task execution, and the coordination of responses across organizational systems. These capabilities may be especially relevant in environments characterized by frequent operational changes and substantial information-processing requirements.

However, adaptability also depends on the organization's ability to revise assumptions, reconsider objectives, and modify established routines. AI systems operating within predefined boundaries may respond efficiently to anticipated changes while remaining poorly suited to situations that fall outside their authorized or technically supported operating conditions.

This distinction is important because operational responsiveness and organizational adaptability are related but not identical.

Operational responsiveness concerns the speed with which an organization reacts to changes within an established decision framework. Organizational adaptability involves the broader capacity to reconsider that framework when circumstances require.

Petcu and Costache (2026) examine functional fixedness and organizational inertia as potential sources of leadership failure. Their perspective is relevant to the possibility that organizations may reproduce established routines through AI-enabled processes rather than develop genuinely adaptive capabilities.

Similarly, research addressing human-centered organizational resilience suggests that technological systems must be considered alongside human capabilities and organizational structures (Kossowska-Lai et al., 2026).

Within the HOAA, distributed operational agency may support adaptability when autonomous execution is accompanied by organizational learning, feedback mechanisms, and the ability to reconfigure delegation boundaries.

For example, an AI-enabled logistics system may respond rapidly to routine delivery disruptions. However, a major supply-chain crisis may require human managers to reconsider procurement priorities, contractual relationships, and strategic objectives. An effective hybrid organization would need mechanisms for transferring control, revising agent permissions, and integrating new human decisions into operational workflows.

The relationship between distributed agency and adaptability is therefore expected to depend on the organization's capacity to learn from agent performance and adjust the allocation of operational responsibilities.

Proposition 4 (P4). Distributed operational agency is positively associated with organizational adaptability when supported by organizational learning, feedback mechanisms, and the capacity to reconfigure human–AI delegation arrangements in response to changing conditions.

This proposition distinguishes adaptive hybrid organizational design from the simple automation of existing routines. It suggests that adaptability emerges through the interaction of autonomous execution and organizational capacity for deliberate change.

5.5. Human Oversight, Accountability Clarity, and Organizational Trust

Trust is a significant consideration in hybrid human–AI organizations because employees, managers, and external stakeholders must assess whether AI-supported processes operate reliably and within legitimate institutional boundaries.

Trust in organizational AI systems should not be understood exclusively as confidence in technical accuracy. It may also depend on whether organizational actors understand the system's permitted activities, can identify responsible decision-makers, and have access to mechanisms for challenging or correcting problematic outcomes.

Human-centered AI research emphasizes the importance of trustworthiness, transparency, accountability, and meaningful human participation. Wilkens et al. (2023) identify trustworthiness as a condition relevant across their configurations of human-centered AI at work. Herrmann and Pfeiffer (2023) similarly emphasize the organizational context within which human-centered AI systems are embedded.

Within the HOAA, oversight and accountability clarity are expected to influence organizational trust through distinct but complementary mechanisms.

Human oversight provides opportunities for monitoring, intervention, and correction. Accountability clarity establishes who is responsible for the authorization, governance, and

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

consequences of AI-supported activities.

When these mechanisms are absent or poorly defined, organizational members may experience uncertainty about the reliability and legitimacy of autonomous actions. For example, employees may be reluctant to accept AI-generated work assignments if they cannot determine who authorized the underlying decision criteria or how an inappropriate allocation can be challenged.

Conversely, clearly communicated governance arrangements may support confidence that AI systems operate within legitimate boundaries and remain subject to institutional control.

Nevertheless, oversight mechanisms may fail to support trust if they are merely symbolic. A nominal human reviewer who lacks sufficient information, time, authority, or competence to intervene may not provide meaningful governance.

Excessively restrictive oversight may also reduce the operational benefits of AI delegation. Consequently, effective oversight should be calibrated to the consequences and reversibility of the delegated activity.

Proposition 5 (P5). Meaningful human oversight and clear accountability arrangements are positively associated with organizational trust in autonomous AI participation, particularly when oversight mechanisms provide credible opportunities for intervention, explanation, and corrective action.

This proposition integrates human-centered AI principles into organizational design theory by conceptualizing trust as an outcome of institutional arrangements rather than solely an individual attitude toward technology.

5.6. Governance Alignment and the Organizational Consequences of AI Autonomy

The preceding propositions suggest that the effects of autonomous AI participation depend on the relationship between technological capabilities and organizational governance.

This relationship is particularly important because increasing AI autonomy may create both opportunities and risks. Autonomous agents may accelerate execution, support coordination, and reduce routine managerial demands. At the same time, they may increase the consequences of incorrectly specified objectives, inappropriate permissions, system errors, and poorly coordinated actions.

The HOAA therefore proposes that organizational outcomes depend on **governance alignment**, defined as the degree of correspondence between the scope of delegated AI autonomy and the organization's authorization, monitoring, intervention, and accountability capabilities.

Governance alignment does not imply that every increase in AI autonomy requires the same increase in direct human supervision. Instead, it suggests that governance arrangements should be proportionate to the nature and consequences of the activities delegated.

For routine, low-risk, and reversible tasks, governance may rely on predefined permissions, automated monitoring, and exception-based review. For consequential or difficult-to-reverse decisions, stronger authorization procedures and more intensive human involvement may be appropriate.

This perspective is consistent with sociotechnical systems thinking, which emphasizes the interdependence of technological and organizational arrangements (Herrmann & Pfeiffer, 2023; Schoenherr & Michael, 2026).

It is also compatible with emerging discussions of AI governance in organizational leadership, including the educational contexts examined by Costache (2026a, 2026b). Such perspectives highlight the need to consider institutional responsibility when autonomous technologies participate in organizational processes.

Governance alignment may be particularly important in coordinated human–AI workflows and bounded agent-led operations, where operational activities can extend across multiple systems and organizational functions.

In these configurations, insufficient governance may create cascading failures, unclear responsibility, or delayed intervention. Conversely, governance arrangements that are unnecessarily restrictive may limit the benefits of distributed operational agency.

The expected relationship is therefore one of organizational fit rather than maximal technological autonomy or maximal supervisory control.

Proposition 6 (P6). The relationship between AI operational autonomy and desirable organizational outcomes is strengthened when governance mechanisms are aligned with the scope, risk, and

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

reversibility of delegated activities; misalignment is expected to increase coordination difficulties and accountability ambiguity.

This proposition represents the integrative theoretical claim of the HOAA. It suggests that the effectiveness of hybrid organizational arrangements depends on the joint configuration of technological autonomy, decision rights, coordination mechanisms, human oversight, and accountability structures.

5.7. Synthesis of the Theoretical Propositions

The six propositions collectively establish a conceptual explanation of how hybrid human–AI organizational arrangements may influence organizational outcomes.

The first proposition emphasizes the importance of decision-right clarity for effective AI delegation. The second examines the contingent relationship between operational autonomy and coordination efficiency. The third addresses the contribution of coordinated human–AI capabilities to decision quality. The fourth considers organizational adaptability and the capacity to reconfigure delegation arrangements. The fifth examines the relationship between oversight, accountability, and organizational trust. The sixth integrates these relationships through the concept of governance alignment.

Table 3. Summary of the HOAA Theoretical Propositions

Proposition	Principal relationship	Expected theoretical implication
P1	Decision-right clarity → Effective AI delegation	Explicit authority boundaries support appropriate delegation
P2	AI autonomy × Task conditions → Coordination efficiency	Autonomy is more beneficial when task characteristics and coordination mechanisms are compatible
P3	Human–AI coordination → Decision quality	Complementary capabilities may improve decisions when human evaluation remains meaningful
P4	Distributed operational agency × Organizational learning → Adaptability	Feedback and reconfiguration capabilities support adaptation
P5	Human oversight + Accountability clarity → Organizational trust	Credible governance arrangements may strengthen trust
P6	AI autonomy × Governance alignment → Organizational outcomes	The consequences of autonomy depend on the fit between delegation and governance

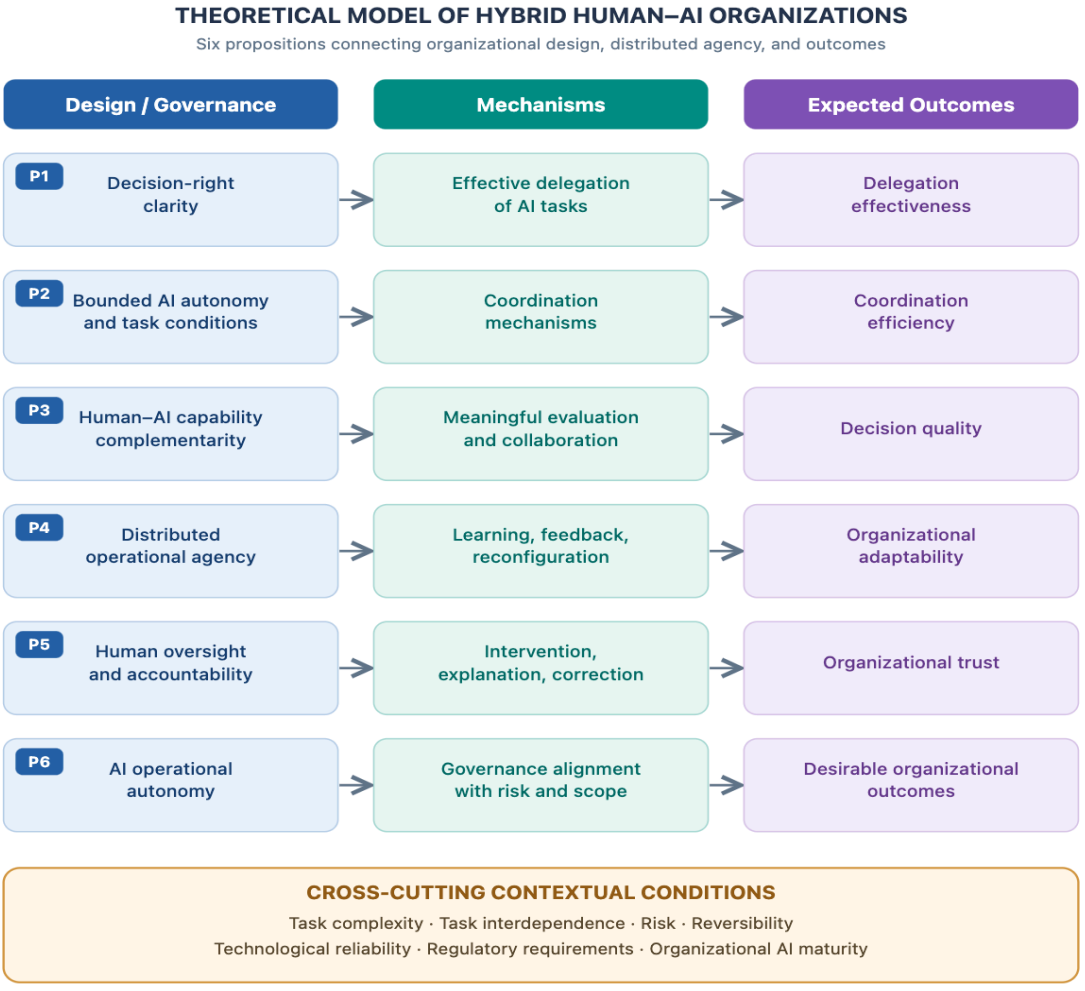
DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Note. All six propositions are theoretical and require empirical testing. The directional relationships do not imply that causal effects have been established.

The propositions also clarify the role of the four HOAA configurations. Human-directed assistance, supervised AI delegation, coordinated human–AI workflows, and bounded agent-led operations represent alternative ways of distributing operational capabilities and decision rights. Their consequences are expected to depend on the relationships identified in the propositions rather than on their position along a presumed hierarchy of technological advancement.

For instance, supervised delegation may provide an effective arrangement for standardized operational tasks, whereas coordinated human–AI workflows may be more appropriate when activities require extensive interaction between computational processing and human contextual judgement. Bounded agent-led operations may be suitable for selected activities when organizations possess sufficiently developed governance and intervention capabilities.

The theoretical architecture therefore supports a contingency-based interpretation of hybrid organizational design. It does not predict that one configuration will consistently outperform the others. Instead, it proposes that organizational effectiveness emerges from the alignment of operational autonomy, decision rights, coordination requirements, and accountability arrangements with the characteristics of organizational tasks and environments.



Source: Authors' own conceptual contribution

Figure 2. Proposed relationships between hybrid organizational design, enabling mechanisms, and organizational outcomes.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

5.8. Implications for Empirical Model Development

The six propositions provide a foundation for subsequent empirical research, although their operationalization requires additional construct development and measurement validation.

Decision-right clarity could be investigated through measures of role definition, authorization thresholds, and perceived ambiguity regarding delegated AI actions.

AI operational autonomy could be examined through the scope of independently executed activities, the frequency of required human authorization, and the complexity of agent-managed workflows.

Human–AI coordination could be assessed through communication quality, task integration, exception handling, and the ability to reconcile human and AI contributions.

Organizational adaptability could be investigated through the speed and effectiveness of operational reconfiguration, the capacity to revise delegation arrangements, and organizational responses to unexpected conditions.

Oversight and accountability clarity could be evaluated through monitoring capabilities, intervention authority, auditability, and the explicit allocation of responsibility.

Organizational outcomes could include decision quality, coordination efficiency, adaptability, trust, and accountability performance. These outcomes should be measured separately rather than combined into a single undifferentiated performance construct without theoretical and empirical justification.

Future studies could also examine potential moderating conditions, including task complexity, reversibility, organizational AI maturity, regulatory intensity, and technological reliability.

Importantly, research designs should distinguish perceived effectiveness from observed organizational outcomes. Cross-sectional managerial surveys may be useful for initial construct validation and association testing, but stronger evidence concerning causal relationships would require longitudinal designs, field experiments, or comparative organizational investigations.

The proposed framework thus provides a theoretically structured but empirically open research agenda. Its principal contribution lies in identifying relationships that can be investigated systematically as autonomous AI agents become more deeply integrated into organizational processes.

6. Discussion and Implications

6.1. Reconsidering the Organizational Foundations of Human–AI Collaboration

The Hybrid Organizational Agency Architecture (HOAA) offers a conceptual interpretation of organizational transformation in which autonomous artificial intelligence (AI) agents become increasingly involved in the initiation, coordination, and execution of organizational activities. Its principal contribution is to reposition the discussion of AI integration from the adoption of technological tools toward the institutional design of hybrid human–AI organizational arrangements.

Traditional organizational theories have generally conceptualized authority, coordination, and accountability through relationships among human actors and formal organizational roles. Although these theories recognize the importance of technological systems, the emergence of autonomous AI agents introduces new questions concerning the allocation of operational initiative and the boundaries of delegated execution.

The HOAA suggests that the transition toward hybrid organizations does not necessarily involve replacing human-centered principles with machine-centered arrangements. Instead, it involves reconsidering how human authority and technological autonomy can coexist within organizational structures.

This interpretation is consistent with the automation–augmentation paradox developed by Raisch and Krakowski (2021), who argue that automation and augmentation are interdependent rather than mutually exclusive approaches to AI-enabled management. The HOAA extends this reasoning by proposing that their organizational consequences depend on how decision rights, coordination responsibilities, and oversight mechanisms are configured.

The framework also builds on the configurational perspective advanced by Wilkens et al. (2023). Whereas their research identifies different actor–structure arrangements associated with human-centered AI at work, the HOAA focuses specifically on the organizational allocation of

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

operational agency when AI systems are capable of executing activities within delegated boundaries.

The distinction is theoretically significant. Human-centered AI research primarily addresses how technological systems and organizational arrangements can support human needs, capabilities, and values. Hybrid organizational design additionally requires consideration of how operational activities are distributed among human actors and autonomous AI agents.

Consequently, the HOAA does not propose abandoning human-centered organizational principles. It suggests that these principles may require institutional reinterpretation as organizations increasingly rely on AI systems capable of acting without continuous human intervention.

6.2. Theoretical Implications for Organizational Design

The first major theoretical implication concerns the relationship between organizational design and distributed operational agency.

Classical organizational design theory emphasizes the allocation of tasks, authority, and information-processing responsibilities among organizational participants. Simon (1947), March and Simon (1958), and Galbraith (1974) established theoretical foundations for understanding how organizations coordinate activities under conditions of bounded rationality and uncertainty.

The HOAA extends these perspectives by introducing autonomous AI agents as technologically enabled participants in organizational processes. Although such agents do not possess human institutional standing, they may perform activities that influence how organizational work is initiated, coordinated, and executed.

This development suggests that organizational design should distinguish between three interrelated dimensions: the capability to perform an activity, the authorization to perform that activity, and the responsibility for its organizational consequences.

These dimensions have traditionally been closely associated with human organizational roles. In hybrid organizations, however, they may become functionally separated. An AI agent may possess the capability and permission to execute an activity, while human managers retain responsibility for the conditions under which that activity is authorized.

The distinction has implications for established organizational concepts such as hierarchy, centralization, formalization, and coordination.

For example, an organization may retain centralized strategic authority while decentralizing routine operational execution through autonomous AI agents. Such an arrangement cannot be adequately described as either fully centralized or fully decentralized without specifying which dimension of organizational decision-making is being considered.

Similarly, formalization may increasingly involve machine-readable permissions, automated decision rules, and intervention protocols alongside conventional organizational policies.

The HOAA therefore suggests that future organizational design research should examine the relationship between formal human authority and technologically distributed operational execution.

A further implication concerns the suitability of different organizational configurations. The four HOAA configurations are not assumed to represent a universal progression toward greater AI autonomy. Instead, they illustrate alternative arrangements that may be appropriate under different task, governance, and institutional conditions.

This contingency-based interpretation is important because the organizational value of autonomy depends on its relationship with task uncertainty, reversibility, coordination requirements, and accountability structures.

Accordingly, the HOAA contributes to organizational theory by proposing that **the distribution of operational agency constitutes an additional dimension of organizational design**, alongside established considerations of authority, specialization, and coordination.

6.3. Implications for Managerial Authority and Decision Rights

The second theoretical implication concerns the changing relationship between managerial authority and operational execution.

In conventional organizational arrangements, managerial authority often involves directing activities, allocating resources, supervising implementation, and evaluating outcomes. Autonomous AI agents may alter the balance among these functions by assuming responsibility for selected operational activities within organizationally established boundaries.

However, this redistribution should not be interpreted as a transfer of independent institutional authority to AI systems.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

The HOAA distinguishes between delegated execution permissions and legitimate decision-making authority. This distinction is essential because AI agents may undertake actions with organizational consequences without possessing the moral, legal, or institutional standing associated with human decision-makers.

The framework therefore suggests that managerial authority may increasingly involve the design and governance of operational delegation rather than direct involvement in every execution step.

Managers may need to establish objectives, specify constraints, define authorization thresholds, monitor agent performance, and determine when human intervention is necessary.

This interpretation is compatible with Chen's (2026) discussion of managerial rationality in human–machine collaboration and Ajmal and Islam's (2025) examination of agentic management. Both works provide contemporary perspectives on the potential transformation of managerial decision processes, although the specific organizational relationships proposed in the HOAA remain to be empirically examined.

The implications are particularly significant for middle management.

Some routine coordination and reporting activities may become increasingly automated, potentially changing the content of managerial roles. At the same time, new responsibilities may emerge involving agent supervision, exception handling, cross-functional coordination, and the evaluation of AI-supported decisions.

Consequently, the development of hybrid organizations should not be reduced to predictions about the disappearance of managerial positions. A more theoretically productive question concerns how managerial responsibilities may be redistributed as operational execution becomes increasingly technologically mediated.

The HOAA suggests that the future role of management may depend less on direct control over every organizational activity and more on the ability to establish legitimate, effective, and adaptable systems of delegated action.

6.4. Distributed Agency and the Reconfiguration of Organizational Coordination

A third implication concerns organizational coordination.

The integration of autonomous AI agents may alter the relationships through which organizational tasks are assigned, sequenced, monitored, and completed.

Traditional coordination mechanisms rely on combinations of hierarchy, standardized procedures, mutual adjustment, and organizational communication. Hybrid organizations may retain these arrangements while introducing automated task allocation, agent-to-agent information exchange, and technologically mediated workflow orchestration.

These developments may create opportunities for faster coordination, particularly where activities are standardized and information requirements are clearly specified.

However, autonomous coordination also introduces potential difficulties.

Multiple AI agents may pursue locally defined objectives that are not fully aligned with broader organizational priorities. Automated processes may create dependencies that are difficult for human managers to observe. In addition, failures in one system may affect interconnected workflows before human intervention occurs.

The HOAA therefore suggests that coordination effectiveness depends on the alignment between the scope of AI autonomy and the mechanisms used to integrate human and technological activities.

This interpretation extends the sociotechnical perspective developed by Herrmann and Pfeiffer (2023), who emphasize the importance of keeping organizational arrangements within the analysis of human-centered AI.

It also aligns with the multilevel orientation of Schoenherr and Michael (2026), which provides a basis for considering AI-related relationships across individual, team, and organizational levels.

From this perspective, the central challenge is not merely to enable AI agents to communicate with one another. It is to ensure that their activities remain consistent with organizational objectives, established permissions, and appropriate mechanisms for resolving uncertainty or conflict.

This issue becomes especially important in coordinated human–AI workflows and bounded agent-led operations, where operational initiative may be distributed across multiple human and technological components.

The framework therefore identifies coordination architecture as a critical research domain for understanding hybrid organizational effectiveness.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

6.5. Governance, Accountability, and the Limits of Autonomous AI Participation

The fourth theoretical implication concerns organizational governance.

Autonomous AI agents may perform actions that affect employees, customers, suppliers, financial resources, and other stakeholders. These actions can have organizational consequences even when they occur without direct human intervention at every stage.

The central governance challenge is therefore to ensure that distributed operational execution does not produce ambiguity regarding legitimate authority and accountability.

The HOAA proposes that accountability should remain attached to identifiable human actors and institutions responsible for establishing objectives, granting permissions, monitoring activities, and responding to consequences.

This does not imply that a single manager must personally approve every operational action. Rather, it suggests that organizational governance should establish clear responsibility for the design and operation of delegated AI processes.

Several governance mechanisms follow from this perspective.

Organizations may establish differentiated authorization thresholds according to the consequences of delegated decisions. They may require audit trails that permit the reconstruction of AI-supported actions. They may also introduce escalation procedures, intervention mechanisms, and periodic evaluations of agent permissions and performance.

However, governance should be proportionate to the characteristics of the activity.

Routine and reversible actions may be compatible with extensive bounded execution and exception-based monitoring. Consequential or difficult-to-reverse decisions may require stronger authorization controls and direct human involvement.

This approach is consistent with the broader principles of human-centered AI discussed by Ozmen Garibay et al. (2023), Wilkens et al. (2023), and Xu (2026).

Emerging discussions of human–AI governance in educational leadership also offer relevant contextual perspectives (Costache, 2026a, 2026b). Although educational institutions and commercial firms differ in their institutional objectives, both face questions concerning the legitimacy and accountability of AI-supported organizational actions.

The HOAA therefore advances a governance principle: **operational autonomy may be distributed, but institutional accountability must remain identifiable and enforceable**.

This principle provides a foundation for distinguishing responsible organizational delegation from the uncontrolled expansion of AI execution capabilities.

6.6. Leadership Implications: From Direct Supervision to Hybrid Agency Orchestration

The transition toward hybrid human–AI organizations may require changes in leadership capabilities and practices.

In conventional managerial arrangements, leadership frequently involves directing employees, allocating responsibilities, evaluating performance, and coordinating human teams. These functions remain important, but the integration of autonomous AI agents introduces additional requirements.

Leaders may increasingly need to determine which activities should remain under direct human control, which may be delegated to AI systems, and how human and technological contributions should be coordinated.

This suggests an emerging leadership function that may be described as **hybrid agency orchestration**.

Hybrid agency orchestration refers to the leadership capability to align human expertise, autonomous AI execution, organizational objectives, and governance arrangements within coherent operational processes.

The concept involves several interrelated responsibilities.

First, leaders must establish the strategic purposes of AI delegation rather than treat technological autonomy as an objective in itself. Second, they must define appropriate decision rights and authorization boundaries. Third, they must support human capabilities for evaluating, challenging, and contextualizing AI-supported decisions. Fourth, they must establish mechanisms for organizational learning and the revision of delegation arrangements.

Finally, they must preserve accountability and institutional legitimacy when operational

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

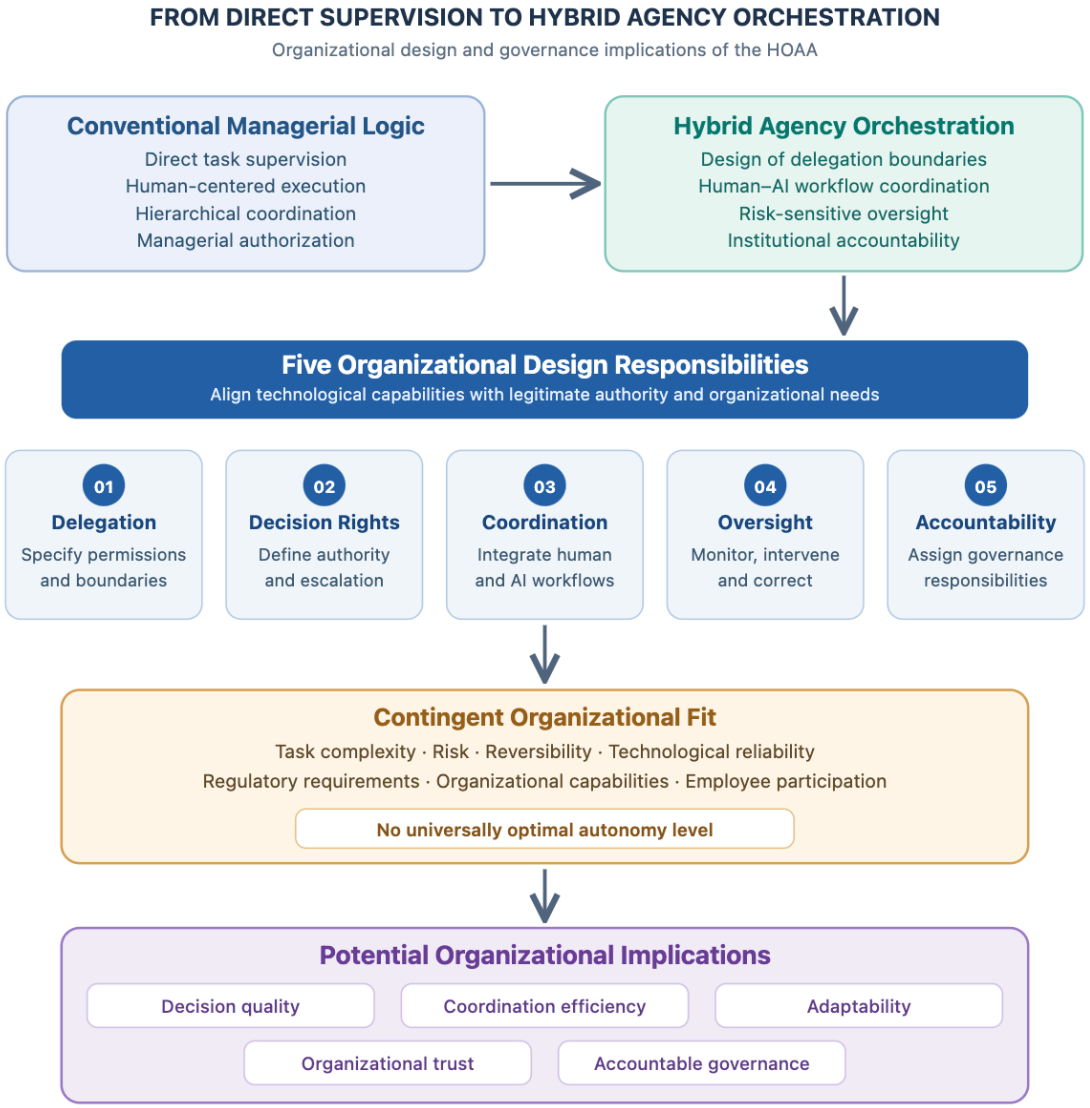
activities are distributed across human and technological components.

Costache's (2026) work on Kaizen leadership offers a complementary perspective on continuous improvement in AI-enabled organizations. Similarly, the discussion of Overton leadership (Costache, 2026) provides a contextual lens for considering how leaders may navigate changing expectations regarding the legitimacy of organizational practices.

Research on remote-work leadership and digital transformation also illustrates the importance of leadership adaptation in technologically mediated environments (Costache, 2026).

Nevertheless, hybrid agency orchestration remains a proposed conceptual capability. Its dimensional structure, measurement properties, and relationships with organizational outcomes require empirical investigation.

The concept may provide a foundation for future research examining how leadership practices influence the effectiveness of different human–AI organizational configurations.



Source: Authors' own conceptual contribution

Figure 3. Managerial and Governance Implications of Hybrid Human–AI Organizational Design

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

6.7. Implications for Employees, Organizational Learning, and Human Wellbeing

The transition toward hybrid organizations also has implications for employees and organizational learning.

Autonomous AI agents may reduce the time required for repetitive activities and expand access to analytical capabilities. However, the redistribution of operational tasks may also change professional roles, skill requirements, workload patterns, and employees' perceptions of control.

These consequences are unlikely to be uniform across occupations or organizational contexts.

For some employees, AI-supported workflows may create opportunities to concentrate on complex judgement, creativity, interpersonal coordination, and strategic activities. For others, they may introduce new demands involving technological monitoring, exception management, and responsibility for outcomes generated through systems they do not fully control.

Human-centered organizational perspectives emphasize that technological transformation should be assessed in relation to human development, participation, and wellbeing (Townsend & Romme, 2024; Atay et al., 2025).

The relevance of wellbeing is also reflected in research on burnout and resilience, including the educational context examined by Costache and Petcu (2025). Although this work does not establish the effects of autonomous AI agents on employee wellbeing, it reinforces the importance of considering the human consequences of organizational arrangements.

Similarly, studies of digital learning and capability development provide complementary perspectives on the importance of adapting educational and professional competencies to changing technological environments (Kimalel, 2025; Costache, 2026).

The HOAA suggests that human capability development should accompany the redistribution of operational agency.

Employees and managers may require competencies in evaluating AI outputs, understanding delegation boundaries, identifying system limitations, interpreting automated actions, and exercising effective intervention.

Organizational learning mechanisms may also be necessary to identify inappropriate delegation arrangements and revise workflows when technological capabilities or institutional requirements change.

The sustainability implications of digital transformation provide an additional consideration. Costache (2025) examines digital innovation and sustainability pathways in higher education, while Kossowska-Lai et al. (2026) connect human-centered design with organizational resilience.

These perspectives suggest that hybrid organizational transformation should be evaluated through broader institutional and human consequences rather than solely through operational efficiency.

Accordingly, the HOAA positions human capability development and organizational learning as important contextual conditions for responsible AI integration.

6.8. Managerial Implications for Designing Hybrid Human–AI Organizations

The HOAA provides a conceptual basis for managerial decisions concerning the integration of autonomous AI agents into organizational processes.

Its principal practical implication is that organizations should begin with the characteristics of their activities and governance requirements rather than assume that the greatest possible degree of AI autonomy represents the preferred design.

Managers may use the architecture to evaluate five questions. First, what operational activities can an AI agent reliably perform? Second, which of those activities should the organization authorize the agent to execute? Third, how should human and AI contributions be coordinated? Fourth, what forms of human oversight are necessary? Fifth, how will accountability be established when AI-supported actions produce organizational consequences? These questions can inform the selection of an appropriate HOAA configuration.

For example, human-directed assistance may be suitable for activities involving substantial uncertainty or professional discretion. Supervised delegation may support standardized tasks with clear authorization boundaries. Coordinated workflows may be appropriate when organizational activities require complementary human and AI contributions. Bounded agent-led operations may be

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

considered for activities with sufficiently clear objectives, manageable risks, and effective governance safeguards.

Table 4. Managerial Design Implications of the Hybrid Organizational Agency Architecture

Organizational design challenge	Recommended managerial consideration	Relevant HOAA dimension
Unclear delegation boundaries	Specify authorized tasks, permissions, and escalation conditions	Decision-right allocation
Excessive reliance on autonomous execution	Align autonomy with task risk, uncertainty, and reversibility	Operational autonomy
Fragmented human–AI workflows	Establish coordination protocols and task ownership	Coordination mechanisms
Limited capacity for intervention	Ensure effective monitoring, override, and termination procedures	Human oversight
Ambiguous responsibility for outcomes	Assign identifiable governance and accountability roles	Accountability clarity
Inflexible organizational routines	Review and revise delegation arrangements as conditions change	Organizational adaptability

Note. The recommendations are derived from the proposed conceptual framework and have not been evaluated through an organizational intervention.

The table emphasizes that the development of hybrid organizations is an organizational design challenge rather than exclusively a technological implementation process.

Organizations may possess advanced AI capabilities yet remain poorly prepared to manage delegated execution if their authority structures, coordination mechanisms, and accountability arrangements are insufficiently developed.

Conversely, organizations with carefully defined governance arrangements may integrate selected autonomous AI capabilities without fundamentally abandoning human-centered institutional principles.

6.9. Boundary Conditions and Alternative Interpretations

The theoretical implications of the HOAA should be interpreted in relation to several boundary conditions.

First, the suitability of different configurations may vary according to task characteristics. Standardized and reversible activities may permit broader delegation than tasks involving ambiguous objectives, substantial ethical considerations, or irreversible consequences.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Second, technological reliability may influence the effectiveness of distributed operational agency. The framework assumes that organizations can establish and enforce operational boundaries, but the practical feasibility of doing so may vary across AI systems and implementation environments.

Third, organizational capabilities may shape the outcomes of AI delegation. Firms with effective monitoring, employee training, and governance arrangements may experience different consequences from organizations lacking these resources.

Fourth, regulatory and institutional environments may restrict the extent to which organizations can delegate particular decisions to autonomous systems.

Fifth, organizational culture and employee perceptions may influence the acceptance and effectiveness of hybrid arrangements. Formal delegation rules alone may be insufficient if employees lack confidence in the legitimacy or reliability of AI-supported processes.

An additional limitation concerns the distinction between conceptual organizational configurations and actual organizational practices. In practice, firms may combine several HOAA configurations across functions, departments, and workflows.

For example, the same organization may use human-directed assistance for strategic investment decisions, supervised delegation for procurement, coordinated workflows for project management, and bounded agent-led operations for selected administrative processes.

The HOAA should therefore be interpreted as a typology of organizational arrangements rather than a classification requiring every firm to occupy a single category.

An alternative interpretation is that autonomous AI agents represent an extension of existing automation rather than a fundamentally new organizational phenomenon. This possibility should not be dismissed. Many organizational challenges associated with AI delegation resemble established problems involving software automation, algorithmic management, and information systems governance.

The theoretical relevance of agentic AI depends on whether its capacity for multistep task initiation, tool use, and bounded execution creates organizational arrangements that differ meaningfully from earlier forms of automation.

The HOAA proposes that these capabilities warrant explicit consideration within organizational design theory, but empirical research is required to establish the extent of their distinctiveness.

7. Conclusions

The emergence of autonomous artificial intelligence (AI) agents introduces important challenges for organizational theory and management practice. As AI systems become increasingly capable of initiating tasks, coordinating activities, and executing decisions within delegated boundaries, organizations may need to reconsider established assumptions about authority, coordination, and the distribution of operational responsibilities. This article has examined these challenges through the development of the **Hybrid Organizational Agency Architecture (HOAA)**, a conceptual framework explaining how human actors and autonomous AI agents may participate in organizational processes while preserving legitimate managerial authority and institutional accountability.

The article addresses three interconnected research questions concerning the implications of autonomous AI agents for organizational design, the configurations through which human and AI operational capabilities may be distributed, and the conditions under which these arrangements may influence organizational outcomes.

Regarding the first research question, the analysis suggests that autonomous AI agents challenge conventional organizational design by creating a distinction between the technical capacity to perform activities and the institutional authority to authorize and govern those activities. Traditional organizational structures associate decision-making and accountability primarily with human organizational roles. Hybrid human–AI arrangements introduce the possibility that operational initiative and execution may be distributed across technological systems while responsibility for organizational objectives, delegation boundaries, and consequential outcomes remains attached to human actors and institutions. Consequently, organizational design must increasingly address not only the allocation of tasks and information but also the boundaries of delegated technological action.

Regarding the second research question, the article proposes four configurations of hybrid organizational arrangements: **human-directed AI assistance, supervised AI delegation, coordinated human–AI workflows, and bounded agent-led operations**. These configurations differ in the distribution of operational initiative, decision rights, coordination responsibilities, and human oversight. They are not presented as sequential stages of technological maturity or as a universal progression toward greater autonomy. Instead, they represent alternative organizational arrangements whose suitability may depend on task characteristics, institutional requirements, technological

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

capabilities, and governance conditions.

The third research question concerns the organizational consequences of distributed operational agency. The six theoretical propositions developed in this article suggest that decision-right clarity, effective human–AI coordination, organizational learning, meaningful oversight, and governance alignment may influence coordination efficiency, decision quality, adaptability, trust, and accountability. Importantly, these relationships are proposed as conditional theoretical expectations rather than empirically established causal effects. Greater AI autonomy may create opportunities for faster execution and improved responsiveness, but its organizational value is expected to depend on the compatibility between delegated activities and the mechanisms governing their execution.

The article makes three principal theoretical contributions.

First, it introduces **distributed operational agency** as a concept for examining how task initiation, information processing, coordination, and bounded execution capabilities may be allocated between human actors and autonomous AI systems. This concept distinguishes technologically enabled action from human intentionality, moral agency, and legal responsibility. The distinction is essential for avoiding the assumption that AI systems acquire independent institutional authority merely because they can execute organizational activities.

Second, the article extends organizational design theory by proposing a configuration-based framework for examining hybrid human–AI arrangements. Building on established perspectives concerning bounded rationality, information processing, sociotechnical systems, and human-centered AI, the HOAA identifies operational autonomy, decision-right allocation, coordination mechanisms, human oversight, and accountability clarity as interdependent organizational design dimensions. This approach shifts the discussion beyond the adoption of AI technologies toward the institutional arrangements through which autonomous AI participation is organized.

Third, the article advances a contingency-based understanding of AI autonomy and organizational effectiveness. It challenges the assumption that increasing technological autonomy necessarily represents organizational progress. Instead, it proposes that effective hybrid organizational arrangements depend on the alignment of technological capabilities with task characteristics, strategic objectives, organizational competencies, and governance requirements.

These contributions have implications for managerial practice. Organizations considering the deployment of autonomous AI agents should evaluate not only what such systems are technically capable of performing but also which activities they should be authorized to execute, under what conditions, and with which safeguards. The proposed framework suggests that effective implementation requires explicit delegation boundaries, appropriate coordination mechanisms, meaningful human intervention capabilities, and identifiable accountability arrangements.

The transition toward hybrid human–AI organizations may also reshape managerial responsibilities. Rather than directly supervising every operational activity, leaders may increasingly be required to establish the conditions under which human and AI capabilities interact effectively. This emerging responsibility involves designing delegation arrangements, evaluating technological limitations, coordinating interdependent activities, and preserving institutional legitimacy. The article describes this potential leadership function as **hybrid agency orchestration**, while recognizing that the concept requires further theoretical refinement and empirical validation.

The study has several limitations. As a conceptual investigation, it does not provide empirical evidence regarding the prevalence, effectiveness, or comparative performance of the proposed organizational configurations. The HOAA has not been validated through organizational observations, interviews, surveys, or experimental research. Furthermore, the practical applicability of its configurations may vary across industries, regulatory environments, organizational sizes, and levels of technological maturity. These limitations indicate the need for future studies examining how organizations actually allocate decision rights to AI agents, how human–AI coordination develops over time, and how governance arrangements influence organizational outcomes.

Future research could employ comparative case studies to identify emerging forms of hybrid organizational design, longitudinal investigations to examine changes in delegation and oversight, and multilevel empirical models to test the proposed relationships among leadership architecture, distributed operational agency, and organizational outcomes. Particular attention should be given to the conditions under which AI autonomy supports organizational adaptability without undermining human judgement, employee participation, or accountability.

Ultimately, the transition from human-centered firms to hybrid human–AI organizations should not be interpreted as the displacement of human governance by artificial agents. Rather, it represents a potential reconfiguration of how organizational work is initiated, coordinated, and executed.

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

Human-centered principles remain relevant, but their implementation may increasingly depend on institutional arrangements that permit bounded technological autonomy while preserving meaningful human authority and responsibility.

The central conclusion of this article is therefore that **the future of organizational design will depend not simply on how much autonomy organizations grant to AI agents, but on how effectively they align distributed operational agency with legitimate decision rights, coherent coordination, and accountable human governance**. The HOAA provides a conceptual foundation for investigating this emerging organizational challenge and for developing more rigorous theories of hybrid human–AI organizations.

ACKNOWLEDGEMENTS

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist with manuscript structuring, drafting, academic language refinement, and editorial revision. All AI-assisted content was subject to author review and revision. The authors retain full responsibility for verifying the accuracy of the information presented, the authenticity and relevance of bibliographic references, the integrity of the research methodology, the analysis and interpretation of empirical findings, and the final content of the publication.

The authors acknowledge the contribution of the students who participated in the questionnaire and provided perspectives on information verification, critical AI literacy, and responsible artificial intelligence use in management education.

7. References

1. Ajmal, M., & Islam, A. (2025). From human-centered to agentic management: Rethinking organizational theory in the age of AI. *Journal of Social Signs Review, 3*(11), 328–342.
2. Atay, S., Müftüoğlu, C. T., Gülmez, N., & Şahin, M. (2025). Society 5.0 and human-centered technology: Redefining talent management in the digital age. *Sustainable Futures, 9*, Article 100733.
3. Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8
4. Chen, Z. S. (2026). Rethinking managerial rationality in the age of AI: A human–machine collaboration perspective on organizational decision-making. *Management Decision*, 1–18.
5. Costache, B. (2025a). Sustainability pathways in higher education transformation: Digital innovation and circular economy in Singapore and the global policy context. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 1*, 41–58. https://doi.org/10.65222/viral.2025.1.3
6. Costache, B. (2026a). Beyond generative AI: Claude agents and the emergence of human-AI governance in educational leadership. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 10*, 136–168. https://doi.org/10.65222/VIRAL.2026.7.43.63
7. Costache, B. (2026b). From TikTok attention spans to virtual campuses: How AI, microlearning, and emotional intelligence are reshaping education. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 7*, 79–98. https://doi.org/10.65222/VIRAL.2026.4.27.47
8. Costache, B. (2026c). Kaizen leadership in the age of artificial intelligence: Continuous improvement as a governance strategy for sustainable organizations. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 4*, 32–50. https://doi.org/10.65222/VIRAL.2026.1.4.24
9. Costache, B. (2026d). Leadership in remote work (2020–2027): A bibliometrics analysis of keyword networks, thematic evolution, and emerging research frontiers. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 8*, 24–45. https://doi.org/10.65222/VIRAL.2026.5.31.51
10. Costache, B. (2026e). Overton leadership: Shifting legitimacy boundaries as a strategic capability in complex adaptive systems. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 5*, 93–121. https://doi.org/10.65222/viral.2026.2.13.33
11. Costache, B. (2026f). The impact of Claude agents and autonomous AI agents on higher education: Opportunities, challenges, and implications for educational leadership. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 11*, 1–40. https://doi.org/10.65222/VIRAL.2026.8.45.65
12. Costache, B., & Petcu, C. (2025). Burnout and resilience in education: Integrating mindfulness, motivation, and contextual moderators for sustainable well-being. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 2*, 26–37. https://doi.org/10.65222/VIRAL.2025.9.7

DOI: https://doi.org/10.65222/VIRAL.2026.9.52.72

13. Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces, 4*(3), 28–36.
14. Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057
15. Herrmann, T., & Pfeiffer, S. (2023). Keeping the organization in the loop: A socio-technical extension of human-centered artificial intelligence. *AI & Society, 38*(4), 1523–1542.
16. Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*, 18–26. https://doi.org/10.1007/s13162-020-00161-0
17. 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
18. 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
19. Kimalel, B. (2025). The impact of social media on competency based curriculum skills acquisition among pupils in Chesumei constituency, Nandi County, Kenya. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 3*, 22–39. https://doi.org/10.65222/VIRAL.2025.11.16
20. Kossowska-Lai, M., Grzegorzek, M., Sołtysik-Piorunkiewicz, A., & Chmielarz, W. (Eds.). (2026). *AI and resilient organizations: Human-centered design for Industry 5.0*. Taylor & Francis.
21. 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
22. Margetis, G., Ntoa, S., Antona, M., & Stephanidis, C. (2021). Human-centered design of artificial intelligence. In *Handbook of human factors and ergonomics* (pp. 1085–1106).
23. Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., et al. (2023). Six human-centered artificial intelligence grand challenges. *International Journal of Human–Computer Interaction, 39*(3), 391–437.
24. Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 30*(3), 286–297. https://doi.org/10.1109/3468.844354
25. Petcu, C., & Costache, B. (2026). Functional fixedness and organizational inertia: A microfoundational perspective on leadership failure. *International Journal of Education, Leadership, Artificial Intelligence, Computing, Business, Life Sciences, and Society, 5*, 122–129. https://doi.org/10.65222/VIRAL.2026.2.14.34
26. 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
27. Rassal, C., & Serra, J. (2027). Human-centered AI governance: A higher-order organizational capability for workforce transformation in AI-enabled Mediterranean resort ecosystems. In *Workforce transformation amid AI-driven digital hospitality service innovation* (pp. 271–314). IGI Global Scientific Publishing. [Publication status to be confirmed.]
28. Schoenherr, J. R., & Michael, K. (2026). Multi-level sociotechnical systems approach to human-centered AI. In *Handbook of human-centered artificial intelligence* (pp. 119–192). Springer.
29. Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management, 57*(2), Article 103174. https://doi.org/10.1016/j.im.2019.103174
30. 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
31. Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan.
32. Sun, Y., & Xu, C. (2026). Hybrid cognitive authority and algorithmic subjectivity: Rethinking knowledge management in AI-driven communication. *Journal of Knowledge Management, 30*(8), 2741–2766.
33. Townsend, M., & Romme, A. G. L. (2024). The emerging concept of the human-centered organization: A review and synthesis of the literature. *Humanistic Management Journal, 9*(1), 53–74.
34. Wang, B. Y., Boell, S. K., Riemer, K., & Peter, S. (2023). *Human agency in AI configurations supporting organizational decision-making*. [Publication venue and other bibliographic details to be confirmed.]
35. Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*(4), 490–495. https://doi.org/10.5465/AMR.1989.4308371
36. Wilkens, U., Lupp, D., & Langholf, V. (2023). Configurations of human-centered AI at work: Seven actor-structure engagements in organizations. *Frontiers in Artificial Intelligence, 6*, Article 1272159. https://doi.org/10.3389/frai.2023.1272159
37. Winby, S., & Xu, W. (2026). Human-centered AI maturity model (HCAI-MM): An organizational design perspective. In *Handbook of human-centered artificial intelligence* (pp. 2493–2525). Springer.
38. Xu, W. (2026). Human-centered artificial intelligence (HCAI): Foundations and approaches. *arXiv*. https://arxiv.org/abs/2601.01247