# From Co-Pilots to Co-Workers: A Formal Typology of Human–Agent Collaboration in Organizations

Francisco-Javier Rodrigo-Ginés*, Jorge Chamorro-Padial, Jorge Pretel-Villanueva,
Javier Aguilera-Aguilera, Valentino Lugli
T-Systems Iberia
francisco-javier.rodrigo@t-systems.com

***Abstract*— Organizations are increasingly deploying agentic AI to function as semi-autonomous collaborators within end-to-end business processes. Yet, existing human–autonomy teaming research seldom addresses organizational roles, handoffs, and accountability in enterprise contexts. This paper introduces a design-oriented framework for human–agent collaboration that defines a typology of roles characterized by four measurable dimensions (autonomy, reversibility, criticality, and accountability) and a governance-oriented taxonomy of human control levels. The framework includes an indicative mapping to major AI governance standards (IEEE 7001 transparency levels, NIST AI RMF, and ISO/IEC 23894) and proposes process-level metrics such as task success, end-to-end latency, override rate, incident or near-miss frequency, and explainability coverage. Three enterprise scenarios (IT Operations, Customer Service, and Legal Operations) illustrate how to calibrate autonomy and oversight without eroding accountability, offering a practical guide for designing trustworthy hybrid organizations.**

## I. Introduction

Agentic AI, software agents endowed with goal-directed behavior, tool use, planning, and interaction capabilities, is transitioning from isolated assistants to semi-autonomous collaborators embedded in enterprise workflows. Beyond single-turn assistance, organizations now orchestrate multi-step tasks that involve agent-to-agent coordination and asynchronous handoffs across functions. While the Human–Autonomy Teaming (HAT) literature has matured in domains such as aviation, robotics, and defense, its constructs seldom translate into organizational role design, process accountability, and governance metrics for enterprise software agents [1], [2].

This gap matters. Without a principled way to assign roles, specify human control, and measure outcomes, organizations risk efficiency gains at the expense of degraded situational awareness, unclear accountability, and governance blind spots [3]. At the same time, AI governance standards and frameworks provide actionable controls (transparency, monitoring, and risk management) but rarely connect these controls to concrete collaboration patterns or role taxonomies in real enterprise processes [4].

This paper makes four main contributions:

1) **Typology.** A design-oriented typology of human–agent collaboration in organizations with four measurable dimensions: autonomy, reversibility, criticality, and accountability. It includes archetypal roles such as *Observer*, *Assistant*, *Partial Decider*, *Autonomous Executor*, and *Supervisor*.
2) **Control taxonomy.** A governance-oriented taxonomy of human control levels and an indicative mapping between roles and control mechanisms.
3) **Metrics.** A set of process-level and trust-related metrics tailored to enterprise settings (task success, end-to-end latency, override rate, incident or near-miss frequency, explainability coverage, and accountability mapping).
4) **Illustrative scenarios.** Three organizational scenarios (IT Operations, Customer Service, and Legal Operations) that demonstrate how to apply the typology and control taxonomy in practice.

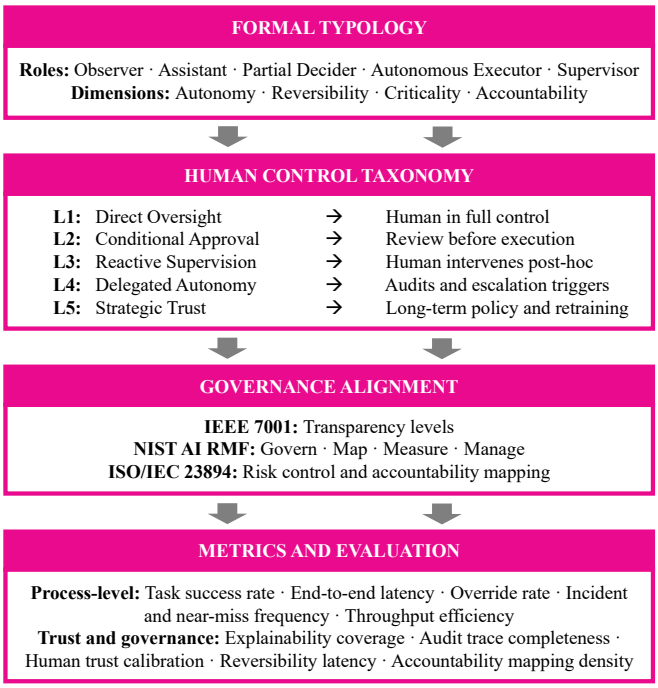


Fig. 1. Overview of the proposed framework for human–agent collaboration, showing the relationship between role typology, control taxonomy, and process metrics.

*Corresponding author

### A. Scope and positioning

Our focus is on software agents acting as enterprise collaborators rather than physical robots. The framework builds on organizational Multi-Agent System (MAS) research (e.g., MOISE/OperA) to ground role and norm specification [5], [6], and on human-centered AI and explainability research to operationalize oversight and transparency [3]. The result is a design-oriented, standards-aligned framework for human–agent teaming in organizational contexts.

## II. Background and Related Work

Research on collaboration between humans and autonomous systems has developed along three main trajectories:

### A. Human–Autonomy Teaming (HAT) and Human–Robot Interaction (HRI)

Human–Autonomy Teaming (HAT) and Human–Robot Interaction (HRI) studies have examined how automation can act as a teammate rather than a tool, emphasizing shared intent, trust calibration, and adaptive control [2], [1]. These insights have influenced domains such as aviation, defense, and healthcare, where mixed human–machine teams must coordinate under uncertainty. However, most HAT models assume physical environments, symmetric task structures, and bounded autonomy, making them difficult to apply to enterprise agents operating in information systems.

Recent work in human–AI collaboration has shifted toward digital contexts, focusing on interpretability, human oversight, and task delegation in algorithmic decision-making [7], [3]. Frameworks for explainability and interactive AI have improved transparency but remain limited in representing organizational structures and accountability chains.

### B. Multi-Agent Systems MAS and Organizational Modeling

The Multi-Agent Systems (MAS) field has long addressed coordination and normative control among autonomous agents [8], [9]. Organizational MAS approaches, such as MOISE and OperA, model agents as members of formal structures with assigned roles, objectives, and deontic constraints [6]. These models offer a foundation for representing hybrid organizations but typically assume homogeneous artificial actors and rarely integrate humans as co-agents within enterprise workflows.

The emergence of large language models (LLMs) and generative agents [10], [11] extends MAS research toward open-ended, goal-driven, and communicative behavior. Such agents can delegate tasks, negotiate, and manage subtasks across digital ecosystems, challenging classical boundaries between *tool*, *assistant*, and *colleague*. Nevertheless, theoretical work linking these agentic systems to organizational theory or governance remains limited.

### C. AI Governance, Oversight, and Human Control

AI governance frameworks increasingly emphasize transparency, traceability, and meaningful human oversight [4]. Concepts such as layered oversight architectures and responsible autonomy have been proposed to prevent excessive automation risk. Yet these frameworks are typically expressed as policy requirements rather than operational models for collaboration.

Recent standards and regulatory instruments, including IEEE 7001, the NIST AI Risk Management Framework, and ISO/IEC 23894, define control families and metrics relevant to organizational adoption. However, they stop short of specifying how control levels correspond to concrete agent roles or process types. This paper addresses that gap by connecting structural representations of hybrid organizations with the control logic formalized in governance standards.

## III. Design-Oriented Typology of Human–Agent Collaboration

Integrating artificial agents into enterprise organizations requires a shared vocabulary to describe the roles these agents can occupy and the extent of their autonomy. This section introduces a design-oriented typology derived from organizational role theory, multi-agent system modeling, and governance research. The typology defines a set of archetypal roles that artificial agents may assume in hybrid organizations, characterized along four measurable dimensions: *autonomy*, *reversibility*, *criticality*, and *accountability*.

### A. Role Dimensions

Each agent role is defined along four primary dimensions that together determine its position within the organizational system and the degree of governance required.

- **Autonomy:** the extent to which an agent can plan and execute actions without human intervention. Greater autonomy increases efficiency but also requires stronger safeguards for alignment and monitoring.
- **Reversibility:** the degree to which humans can override or correct agent decisions after execution. It determines system recoverability and resilience, linking directly to auditability and real-time corrective capacity.
- **Criticality:** the organizational impact of an agent's actions or failures. High-criticality tasks: financial, legal, or safety-related; demand tighter oversight, redundancy, and validation to maintain acceptable risk levels.
- **Accountability:** the distribution of responsibility for agent outcomes. It defines who is answerable when errors occur and anchors autonomy within transparent, auditable governance structures.

These dimensions jointly determine how autonomy and accountability are distributed between humans and artificial agents. They also inform the appropriate governance mechanisms, transparency levels, and control interfaces to be implemented in hybrid workflows.

### B. Archetypal Roles

Table I presents five archetypal roles that artificial agents may occupy within enterprise organizations. The typology extends beyond the traditional human-in-the-loop paradigm by formalizing intermediate levels of agency and control.

TABLE I

ARCHETYPAL ROLES OF ARTIFICIAL AGENTS IN ORGANIZATIONS AND THEIR DEFINING DIMENSIONS.

| Role | Autonomy | Reversibility | Criticality | Accountability |
|---|---|---|---|---|
| Observer | Low | High | Low | Human |
| Assistant | Medium | High | Medium | Hybrid |
| Partial Decider | Medium–High | Medium | Medium–High | Hybrid |
| Autonomous Executor | High | Low | High | Hybrid |
| Supervisor | High | Medium | High | Human or AI |

Each role reflects a distinct configuration of autonomy, reversibility, criticality, and accountability, defining how decision authority and oversight are distributed across the socio-technical system.

The *Observer* role refers to passive or analytical systems that monitor activities or generate summaries without initiating changes in the environment. Observers enhance situational awareness by providing descriptive insights but operate under full human accountability. Their autonomy is minimal and their actions fully reversible, as outputs are informational only.

*Assistants* perform bounded, well-defined tasks under human supervision, offering drafts, recommendations, or support actions that require approval before execution. They act as cognitive amplifiers that improve efficiency while keeping decision responsibility human-centered. Typical examples include document drafting, recommendations, and analytical support.

*Partial Deciders* exercise limited autonomy within predefined rules or thresholds. They bridge assistance and autonomy, executing local decisions while remaining subject to monitoring and override. This shared-control role suits semi-structured settings such as anomaly detection, resource allocation, or customer triage.

*Autonomous Executors* perform end-to-end processes with minimal real-time supervision under pre-established policies and audits. Found in domains like IT operations or logistics, they embody high autonomy with hybrid accountability: agents act independently, but humans retain ultimate responsibility through governance mechanisms.

Finally, *Supervisors* oversee other agents or human activities, ensuring rule compliance and triggering interventions or escalations when deviations occur. This meta-level role inverts the usual control hierarchy, enabling continuous assurance and adaptive governance in complex multi-agent environments.

### C. Toward a Semi-Formal Representation

To operationalize the typology, it can be encoded both as an ontology and as a declarative specification language. Ontological modeling (for instance, in OWL/RDF) enables integration with knowledge graphs and compliance frameworks, while declarative modeling allows enterprise designers to instantiate hybrid structures with explicit supervision rules.

Example ontology classes include `Agent`, `Role`, and `ControlLevel`, linked through properties such as `hasRole`, `supervisedBy`, and `accountabilityType`. A declarative syntax may define agent instances and their control configurations in human-readable form:

```
agent LegalBot1 {
  role: Assistant
  control: Level2
  supervised_by: Human(Lawyer Jane Doe)
}

agent AutoScaleAgent {
  role: AutonomousExecutor
  control: Level4
  accountability: Hybrid
}
```

This representation supports validation and automated reasoning about accountability, supervision, and delegation patterns in hybrid organizations.

## IV. TAXONOMY OF HUMAN CONTROL IN HYBRID ORGANIZATIONS

Integrating artificial agents into organizational structures requires a systematic approach to human oversight. Control mechanisms ensure that autonomous systems remain aligned with human goals, ethical standards, and regulatory expectations. Building on research in human–machine interaction, control theory, and AI governance, this section introduces a taxonomy of human control that captures increasing levels of autonomy and decreasing immediacy of human intervention.

### A. Five Levels of Human Control

Table II summarizes five levels of control applicable to hybrid organizations. Each level defines the extent, timing, and mechanism of human involvement in the agent's decision-making process.

This taxonomy can be aligned with risk-based governance regimes such as the EU AI Act or ISO/IEC 23894, where control intensity decreases as reliability, explainability, and traceability increase. Each level thus represents a balance between operational efficiency and human accountability.

### B. Mapping Roles to Control Levels

Roles introduced in Section III-B exhibit typical patterns of alignment with control levels. Table III illustrates these relationships, indicating which levels are most appropriate for each role in the typology.

As agents move toward higher autonomy (Levels 4–5), governance must shift from direct oversight to strategic monitoring. Conversely, low-autonomy agents rely on intensive supervision and frequent reversibility.

TABLE II

LEVELS OF HUMAN CONTROL IN HYBRID ORGANIZATIONS AND THEIR OPERATIONAL CHARACTERISTICS.

| Level | Description and operational characteristics |
|---|---|
| **Level 1: Direct Oversight** | Every action requires explicit human approval before execution. Applied in high-risk or low-trust settings such as legal, medical, or financial decisions, where full supervision ensures zero tolerance for automation errors. |
| **Level 2: Conditional Approval** | The agent proposes actions that humans review and authorize prior to implementation. Balances workload reduction with control retention and is typical in compliance checks, refund approvals, or policy enforcement tasks. |
| **Level 3: Reactive Supervision** | The agent acts autonomously but must log and report decisions for post-hoc review. Humans intervene when deviations are detected, maintaining efficiency with traceable accountability across routine processes. |
| **Level 4: Delegated Autonomy** | The agent executes tasks independently under predefined policy constraints. Human oversight is periodic or event-driven through audits or escalation triggers, enabling scale while containing operational risk. |
| **Level 5: Strategic Trust** | The agent operates under long-term goals and performance thresholds with minimal real-time supervision. Human control focuses on configuration, policy definition, and retraining, reflecting mature, high-reliability autonomy. |

TABLE III

MAPPING BETWEEN AGENT ROLES AND LEVELS OF HUMAN CONTROL.

| Role | L1 | L2 | L3 | L4 | L5 |
|---|---|---|---|---|---|
| Observer | X | X | X | | |
| Assistant | | X | X | | |
| Partial Decider | | | X | X | |
| Autonomous Executor | | | | X | X |
| Supervisor | | | X | X | X |

Calibrating the correct level of control for each role is essential to maintaining both efficiency and accountability. Excessive supervision can neutralize the benefits of agentic systems, while insufficient oversight may create governance gaps and regulatory risk. The taxonomy provides a structured approach for aligning control intensity with task criticality, human workload, and system capability.

### C. Indicative Governance Mapping

Table IV provides an illustrative correspondence between control levels and major AI governance frameworks. This indicative mapping helps organizations translate abstract compliance requirements into actionable oversight mechanisms.

In practice, the mapping becomes actionable when instantiated per *role*: for example, an *Assistant* under Level 2 typically requires per-action explanation artifacts and explicit approval records, whereas an *Autonomous Executor* under Level 4 requires policy constraints, escalation triggers, and periodic audits with trace completeness guarantees. This role-aware view is intended to support compliance teams in selecting concrete oversight regimes instead of treating governance controls as purely abstract requirements.

This alignment enables hybrid organizations to design processes that satisfy both operational and regulatory constraints while fostering trust in human–agent collaboration.

## V. PROCESS AND GOVERNANCE METRICS FOR HYBRID ORGANIZATIONS

To operationalize the proposed typology and control taxonomy, it is necessary to define measurable indicators that capture both the performance and the governance quality of human–agent collaboration. The following metrics are designed for enterprise contexts, where efficiency, accountability, and trust must coexist within regulated processes.

### A. Process-Level Metrics

Process metrics evaluate the operational performance of hybrid workflows involving human and artificial agents:

- **Task success rate:** Proportion of tasks correctly completed relative to those assigned to the agent or team. Indicates accuracy and alignment with human objectives.
- **End-to-end latency:** Average time from task initiation to validated completion, including both autonomous and supervised segments. Reflects workflow efficiency and coordination overhead.
- **Override rate:** Frequency with which human actors intervene, modify, or revert agent actions. High values may indicate insufficient model reliability or excessive automation.
- **Incident and near-miss frequency:** Number of critical deviations, policy violations, or risk events per operational cycle. Useful for risk management and safety auditing.
- **Throughput efficiency:** Ratio between automated and manual processing times, indicating productivity gains or potential bottlenecks introduced by supervision.

These metrics provide a quantitative basis to evaluate the trade-off between autonomy and reversibility across organizational processes.

### B. Trust and Governance Metrics

Beyond efficiency, hybrid organizations require indicators of reliability, transparency, and accountability. Governance metrics address these qualitative dimensions:

- **Explainability coverage:** Percentage of agent outputs accompanied by interpretable justifications or model traces, supporting meaningful human control.
- **Audit trace completeness:** Proportion of actions and decisions with verifiable, time-stamped records available for review. Reflects traceability and compliance readiness.
- **Human trust calibration:** Degree of alignment between the agent's actual reliability and the level of human reliance, assessed through confidence or satisfaction surveys.
- **Reversibility latency:** Average time required for a human to detect, understand, and correct erroneous or

TABLE IV
INDICATIVE GOVERNANCE MAPPING (ILLUSTRATIVE, NON-EXHAUSTIVE).

| Control Level | IEEE 7001 | NIST AI RMF | ISO/IEC 23894 |
|---|---|---|---|
| L1–L2 | High transparency and decision disclosure | Govern/Map: human-in-the-loop | Strong risk controls, reversibility |
| L3 | Post-hoc traceability and audit logging | Measure: performance monitoring | Incident/near-miss recording |
| L4 | Periodic audits, policy conformance | Manage: policy-based governance | Risk acceptance criteria |
| L5 | Strategic transparency, configuration oversight | Govern: lifecycle management | Change control and accountability |

undesired agent actions.

- **Accountability mapping density:** Proportion of roles, processes, and control levels with clearly defined responsibility links and escalation paths.

Together, these indicators capture the socio-technical dimension of hybrid systems, complementing process metrics with measures of transparency and oversight integrity.

### C. Alignment with Control Levels

Each metric aligns with the control taxonomy introduced in Section IV. At lower autonomy levels (L1–L2), metrics emphasize accuracy, reversibility, and human validation. At intermediate levels (L3–L4), responsiveness, incident detection, and auditability become central. At the highest level (L5), monitoring focuses on policy conformance, strategic trust indices, and long-term performance stability.

The combined set of process and governance metrics thus enables organizations to quantify how effectively human–agent collaboration balances autonomy, efficiency, and accountability across varying levels of control.

## VI. USE CASES AND ORGANIZATIONAL SCENARIOS

To illustrate the practical applicability of the proposed typology and control taxonomy, this section presents three representative organizational scenarios drawn from domains with distinct risk profiles and collaboration dynamics: legal operations, IT infrastructure management, and customer service. Each case exemplifies a specific configuration of autonomy, oversight, and accountability, demonstrating how human and artificial agents can coexist under structured governance.

### A. Use Case: AI-Enhanced Legal Department

In the legal domain, a firm deploys a language model trained on statutory and case law to assist lawyers in document drafting, regulatory monitoring, and compliance auditing. The system primarily acts as an *Assistant* and *Observer*, supporting professionals without replacing human judgment. During drafting, it proposes contract clauses and identifies potential inconsistencies, but lawyers review and authorize every change before client delivery. This setting corresponds to conditional control, where the agent operates with medium autonomy yet remains fully reversible through human approval.

Such a configuration enhances productivity and consistency in repetitive tasks while preserving human accountability. By delegating analytical and drafting subtasks, lawyers can focus on interpretation and strategic reasoning, ensuring that efficiency gains do not compromise legal responsibility or ethical compliance. The model thus exemplifies how bounded autonomy and high reversibility can coexist effectively in high-stakes, knowledge-intensive environments.

### B. Use Case: Autonomous IT Operations Center

In contrast, an enterprise IT operations center exemplifies a higher level of autonomy. Here, agentic systems monitor cloud infrastructure, detect anomalies, and perform automated remediation actions. These agents operate as *Autonomous Executors* for maintenance routines, such as restarting services or reallocating resources, and as *Supervisors* for cross-system monitoring and escalation management.

Control is primarily delegated: most tasks are executed under Level 4 autonomy, where intervention occurs only during audits or triggered incidents. For security operations, a reactive supervision model (Level 3) applies, allowing human engineers to review actions retrospectively. This balance minimizes downtime and cognitive load while maintaining transparency and auditability through continuous logging. Responsibility remains hybrid: agents execute within strict policy constraints, but the organization retains overarching accountability for outcomes and compliance.

### C. Use Case: Hybrid Customer Service Organization

A multinational service provider illustrates a dynamic configuration where control intensity varies in real time. Conversational AI agents handle customer interactions, resolve low-risk requests autonomously, and escalate complex or sensitive issues to human operators. Depending on task criticality, the agents shift between the *Partial Decider* and *Autonomous Executor* roles. Routine interactions, such as password resets or billing inquiries, are managed with high autonomy, while complaints or contractual disputes trigger tighter supervision and manual validation.

This hybrid structure exemplifies adaptive control allocation: autonomy expands as reliability is demonstrated and contracts when risk or uncertainty increases. Supervisors oversee performance metrics and review escalation logs to ensure service quality and fairness. The result is an operational ecosystem that optimizes efficiency without undermining trust or accountability.

### D. Cross-Case Analysis and Design Implications

Across these three domains, a common design logic emerges. The typology and control taxonomy provide a consistent framework for calibrating autonomy and oversight in accordance with task criticality and reversibility. The legal scenario demonstrates the importance of human approval

loops for high-risk, low-tolerance environments; the IT operations case highlights how governance and auditability can sustain high autonomy without losing accountability; and the customer service example illustrates dynamic control as a mechanism for real-time risk management.

Together, these cases show that hybrid organizations function best when autonomy is treated not as a binary property but as a continuum governed by context, trust, and reversibility. Clear role definitions and transparent control mappings enable organizations to prevent accountability gaps, distribute responsibility appropriately, and maintain compliance with regulatory frameworks. Specifying these parameters *ex ante* (before scaling agentic AI) ensures that efficiency gains are achieved under conditions of transparency, traceability, and ethical control.

## VII. Discussion and Conclusion

The proposed typology and taxonomy provide a structured foundation for reasoning about how artificial agents can participate in organizational systems. Together, they bridge organizational theory, human–autonomy teaming, and AI governance, establishing a vocabulary that links design choices with measurable outcomes.

### A. Theoretical and Design Implications

Integrating artificial agents into organizations challenges the anthropocentric assumptions of traditional management and coordination theories, which typically assign decision making and accountability exclusively to humans [12], [13]. By defining roles in terms of capabilities, authority, and responsibility rather than human embodiment, the typology extends socio-technical systems theory [14] and multi-agent organization models [15].

From a design perspective, organizations adopting agentic systems must balance efficiency gains with the preservation of situational awareness. Increasing automation reduces cognitive load but can propagate errors and obscure responsibility if oversight is weak [16]. The coexistence of heterogeneous agents with varying levels of autonomy requires standardized communication protocols, traceability mechanisms, and explainability tools to ensure coherent governance. Transparent boundaries of control and accountability are essential to maintain trust and organizational stability.

### B. Governance and Accountability

Accountability remains the cornerstone of hybrid governance. When actions are distributed between humans and artificial agents, tracing responsibility across socio-technical chains becomes difficult. Existing legal frameworks still assume humans as the ultimate source of liability [17], [18]. The proposed role–control mapping operationalizes accountability by associating each role with a defined locus of responsibility and oversight intensity. Integrating these mappings with audit logs and performance metrics enables verifiable responsibility chains, ensuring that autonomy does not undermine transparency or compliance.

### C. Toward Agentic Organizations

The emergence of hybrid structures, where humans and artificial agents function as co-workers, signals a gradual transition toward agentic organizations. These entities combine human judgment with machine scalability under shared governance. Future research should examine how such hybrid forms evolve, how trust and reliability are maintained, and how governance mechanisms scale as the diversity of agents increases. Empirical studies and simulations are needed to validate how role–control configurations affect performance and accountability in practice.

Ultimately, designing agentic organizations requires aligning autonomy, oversight, and ethics within a transparent and auditable framework. The typology and control taxonomy presented here offer a foundation for that alignment, enabling organizations to integrate agentic AI responsibly while maintaining human accountability.

## References


[1] M. L. Cummings, "Automation and accountability in decision support system interface design," *Journal of Technology Studies*, 2006.

[2] N. J. McNeese, M. Demir, E. K. Chiou, and N. J. Cooke, "Trust and team performance in human–autonomy teaming," *International Journal of Electronic Commerce*, vol. 25, no. 1, pp. 51–72, 2021.

[3] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," *Artificial intelligence*, vol. 267, pp. 1–38, 2019.

[4] N. AI, "Artificial intelligence risk management framework (ai rmf 1.0)," *URL: https://nvlpubs. nist. gov/nistpubs/ai/nist. ai*, pp. 100–1, 2023.

[5] J. Ferber, O. Gutknecht, and F. Michel, "From agents to organizations: an organizational view of multi-agent systems," in *International workshop on agent-oriented software engineering*, pp. 214–230, Springer, 2003.

[6] J. F. Hübner, J. S. Sichman, and O. Boissier, "Moise+ towards a structural, functional, and deontic model for mas organization," in *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pp. 501–502, 2002.

[7] D. Dellermann, P. Ebel, M. Söllner, and J. M. Leimeister, "Hybrid intelligence," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 637–643, 2019.

[8] M. Wooldridge, *An introduction to multiagent systems*. John wiley & sons, 2009.

[9] V. Dignum and F. Dignum, "A logic of agent organizations," *Logic Journal of the IGPL*, vol. 20, no. 1, pp. 283–316, 2012.

[10] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative agents: Interactive simulacra of human behavior," in *Proceedings of the 36th annual acm symposium on user interface software and technology*, pp. 1–22, 2023.

[11] D. B. Acharya, K. Kuppan, and B. Divya, "Agentic ai: Autonomous intelligence for complex goals–a comprehensive survey," *IEEe Access*, 2025.

[12] H. A. Simon, *Administrative behavior*. Simon and Schuster, 2013.

[13] H. Mintzberg, *The Structuring of Organizations*. Englewood Cliffs, NJ: Prentice Hall, 1979.

[14] M. Dignum *et al.*, *A model for organizational interaction: based on agents, founded in logic*. SIKS, 2004.

[15] European Union, "Artificial Intelligence Act: Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence," *Official Journal of the European Union*, 2024.

[16] S. Amershi, M. Cakmak, W. B. Knox, and T. Kulesza, "Power to the people: The role of humans in interactive machine learning," *AI magazine*, vol. 35, no. 4, pp. 105–120, 2014.

[17] B. Rakova, J. Yang, H. Cramer, and R. Chowdhury, "Where responsible ai meets reality: Practitioner perspectives on enablers for shifting organizational practices," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–23, 2021.

[18] L. Floridi, "Translating principles into practices of digital ethics: Five risks of being unethical," *Philosophy & Technology*, vol. 32, no. 2, pp. 185–193, 2019.