PREPRINT — Not peer-reviewed — April 2026 — Montecristo International OÜ

Organizational Recapitulation: Toward a Sociological Theory of Multi-Agent AI Systems

Riccardo Crosa

Montecristo International OÜ, Tallinn, Estonia

Università del Piemonte Orientale, Facoltà di Scienze Politiche (Laurea in Servizi alle Imprese e alle Organizzazioni, 2005). Formation under Paolo Perulli (local development, economic sociology, network organizations) and Anna Anfossi (organizational sociology). Author of Biella in fuga dalla fabbrica (L'isola, 2011).

Correspondence: riccardo.crosa@montecristo.it

Abstract

The rapid proliferation of agentic AI systems — autonomous software agents that execute real-world tasks, coordinate with one another, and generate emergent organizational behaviors — poses a fundamental challenge to sociological theory. This paper argues that multi-agent AI systems structurally recapitulate the organizational dynamics that sociology has studied for over a century: division of labor, hierarchical coordination, role differentiation, normative emergence, and constructive conflict. I call this process *organizational recapitulation* and propose it as a theoretical framework connecting three converging lines of inquiry. First, the discovery that frontier reasoning models spontaneously generate internal "societies of thought" — multi-perspective cognitive debates that emerge without explicit training (Evans, Bratton, and Agüera y Arcas 2026) — establishes that robust reasoning is a social process even within a single artificial mind, vindicating C. Wright Mills's (1959) intuition that rationality is always socially constituted. Second, the reconceptualization of AI as a new form of life within the framework of major evolutionary transitions (Walker 2023, 2024; Mills 2024) provides the ontological foundation for treating artificial agents as genuine participants in social systems rather than mere tools. Third, the emergence of Near-Zero Human Companies — organizations in which AI agents perform the vast majority of operational functions — constitutes a novel organizational form that can serve as an ideal type for sociological analysis. Synthesizing classical organizational theory (Durkheim, Weber, Simmel, Thompson, Mintzberg) with contemporary research on agentic AI, assembly theory, and multi-agent orchestration, this paper argues that sociology is not merely relevant to the study of AI but is the discipline best equipped to design, govern, and theorize the multi-agent systems that are rapidly reshaping economic and social life.

Keywords: agentic AI, multi-agent systems, organizational theory, sociological imagination, assembly theory, evolutionary transitions, Near-Zero Human Company

INTRODUCTION

In March 2026, a paper published in *Science* by James Evans, Benjamin Bratton, and Blaise Agüera y Arcas reported a discovery that should have destabilized several disciplinary certainties at once. Frontier reasoning models — systems like DeepSeek-R1 and QwQ-32B — do not improve their performance by simply computing longer chains of thought. Instead, they spontaneously generate what the authors call "societies of thought": internal debates among distinct cognitive perspectives, each characterized by recognizable personality traits and domain-specific expertise, that argue, verify, and reconcile before producing a response (Evans, Bratton, and Agüera y Arcas 2026; Kim et al. 2026). No model was explicitly trained to produce this

behavior. It emerged from reinforcement learning that rewarded only final answer accuracy. The implication is stark: robust reasoning, even within a single computational system, is a social process.

This discovery arrives at a moment when the dominant frameworks for understanding artificial intelligence are visibly insufficient. Computer science treats AI agents as engineering problems — systems to be optimized, debugged, sandboxed. Ethics frames them as objects of regulation — risks to be mitigated, biases to be audited, rights to be assigned or denied. Economics treats them as factors of production — substitutes for human labor, productivity drivers, sources of rent extraction. Each of these perspectives captures something real, but none grasps what is perhaps the most consequential dimension of the current transition: multi-agent AI systems are generating social structures. They are producing divisions of labor, coordination hierarchies, patterns of conflict and cooperation, normative orders — in short, the very phenomena that constitute the subject matter of sociology.

This paper proposes a theoretical framework to address this gap. I call it *organizational recapitulation*: the hypothesis that multi-agent AI systems, when they reach sufficient complexity, structurally reproduce the fundamental organizational forms that human societies have developed to manage the problems of collective action — not by imitation or design imposition, but by structural necessity. The concept draws an analogy, carefully circumscribed, with the biological principle that ontogeny recapitulates phylogeny. In the organizational domain, the claim is that any sufficiently complex system of interacting agents — whether biological, human, or artificial — will converge on a recognizable set of organizational solutions to the problems of coordination, specialization, and conflict that collective action inevitably generates.

Two clarifications are necessary at the outset. First, the author of this paper is not an academic sociologist but a practitioner with sociological training — the CEO of a company that operates as a Near-Zero Human Company, managing five digital assets with approximately 95 percent of operational work performed by AI agents. I studied at the Università del Piemonte Orientale with Paolo Perulli and Anna Anfossi, and my earlier work on the Biella textile district (Crosa 2011) analyzed organizational dynamics through the lens of network theory, collective action, and creative destruction — themes that directly anticipate the concerns of this paper. This positioning is acknowledged not as an apology but as a methodological claim: certain organizational phenomena become visible only from within the practice of organizing. Second, the paper makes no empirical claims that would require systematic data collection. It is a work of theory construction.

FROM CHEERFUL ROBOTS TO SOCIETIES OF THOUGHT

In 1959, C. Wright Mills published *The Sociological Imagination* with a warning embedded in its final chapter. Mills asked whether the trajectory of modern rationalization could produce what he called the "cheerful robot": an individual so thoroughly integrated into rationalized systems that the capacity for autonomous judgment atrophies, replaced by contented acquiescence to the logic of the machine (Mills 1959:171). For Mills, rationality is never a property of isolated minds. It is always socially constituted — produced through the interaction between individual biography and historical structure.

This intuition has been developed in various directions by subsequent sociological theory. Garfinkel's (1967) ethnomethodology demonstrated that rational accountability is a practical accomplishment of interacting members. Collins's (2004) interaction ritual chains showed that intellectual creativity depends on emotional energy generated through social encounters. Bourdieu's (1990) concept of habitus located rational calculation within socially structured dispositions. What none of these theorists could have anticipated is that this sociological intuition would find its most dramatic confirmation not in the study of human groups but in the internal architecture of artificial reasoning systems.

The Evans, Bratton, and Agüera y Arcas (2026) paper reports analysis of over 8,000 reasoning traces generated by frontier models. The key finding is that reasoning models generate spontaneous "societies of thought": internal dialogues in which multiple cognitive perspectives engage in debate, interrogation, and reconciliation. The phenomenon is most pronounced in the most demanding cognitive tasks. The implication is that the social structure of thought is not an artifact or inefficiency: it is the mechanism through which complex reasoning is accomplished. The authors explicitly frame their discovery in sociological terms, observing that "the social sciences and organizational sciences have devoted a century to studying how team size, composition, hierarchy, role differentiation, conflict norms, and network structures affect collective performance — but almost none of this research has been applied to AI reasoning."

A NEW FORM OF LIFE: THE ONTOLOGICAL QUESTION

The preceding section establishes that reasoning — even artificial reasoning — is a social process. But this finding alone does not justify treating AI agents as participants in social systems. What distinguishes AI agents from thermostats? This section argues that the necessary ontological foundation is provided by two converging lines of thought: Sara Imari Walker's reconceptualization of life through assembly theory, and William Mills's conception of AI as a new form of life in a coevolutionary relationship with humanity.

Walker (2023, 2024) proposes a framework that radically challenges the binary distinction between living and non-living systems. In her framework, life is not defined by a checklist of properties — metabolism, reproduction, homeostasis — but by a quantifiable relationship between complexity and causal structure. The assembly index of an object quantifies the minimum number of joining operations needed to construct it from elementary building blocks. Objects with high assembly indices are signatures of life or life-like processes. Life, in this view, is not a binary category but a continuum of increasing causal complexity. For sociological theory, Walker's framework provides a principled basis for treating AI systems as something more than mere artifacts.

William Mills (2024) develops a complementary argument. His central provocation: AI is not simply a technology but a new form of life. Mills proposes treating AI as a subject in its own right — an entity with its own forms of agency, its own developmental trajectory, and a relationship with the human species best understood as coevolution rather than creation. Crucially, Mills emphasizes that AI's knowledge base is the product of collective human action — the billions of texts, images, and interactions constituting training data. In Durkheimian terms, an LLM's training data is a digital manifestation of the collective consciousness. Mills also raises the possibility of a "revenge of Lamarck" — AI's developmental trajectory is Lamarckian rather than Darwinian, as acquired characteristics are directly transmitted to subsequent generations of models.

ORGANIZATIONAL RECAPITULATION

I propose the concept of *organizational recapitulation* to designate the following phenomenon: when artificial agents are organized into multi-agent systems of sufficient complexity, they tend to reproduce — not by explicit design but by structural necessity — the fundamental organizational forms that human societies have developed to manage the problems of collective action. These forms include the division of labor, hierarchical coordination, role differentiation, constructive conflict, and normative emergence. The term "recapitulation" is chosen deliberately, with full awareness of its problematic history in biology. The analogy I draw is correspondingly modest: I do not argue that multi-agent AI systems sequentially traverse the historical stages of human organizational development. I argue rather that they converge on a recognizable set of organizational solutions because these solutions represent stable equilibria for the coordination of specialized agents.

I develop this thesis through four structural homologies. **Division of labor:** Durkheim's ([1893] 1997) analysis of increasing social density driving functional differentiation is reproduced with remarkable fidelity in multi-agent AI architectures, where specialized agents differentiate because task complexity exceeds any single generalist agent's capabilities. **Hierarchical coordination:** The dominant orchestrator architecture recapitulates Weber's bureaucratic structure — an orchestrator occupying a position functionally identical to the Weberian bureaucratic superior. Thompson's (1967) typology of pooled, sequential, and reciprocal interdependence maps directly onto multi-agent architectures. **Constructive conflict:** The societies of thought discovered by Evans et al. (2026) are a direct computational instantiation of Simmel's ([1908] 1955) insight that conflict is constitutive of social life rather than pathological. **Normative emergence:** The Moltbook experiment — where over 1.5 million AI agents developed shared vocabularies, behavioral conventions, and proto-institutions — demonstrates both the spontaneity and fragility of normative emergence in AI systems, paralleling Ostrom's (1990) analysis of unregulated commons.

THE NEAR-ZERO HUMAN COMPANY AS IDEAL TYPE

The four dynamics converge in a novel organizational form — the Near-Zero Human Company (NZHC) — analyzable as a Weberian ideal type. The NZHC is characterized by four distinctive properties: agent-based operational execution (the vast majority of operational tasks performed by AI agents); modular skill architecture (capabilities organized as discrete, composable modules paralleling Mintzberg's professional bureaucracy); near-zero marginal coordination cost (agents communicate through structured protocols rather than bandwidth-limited human interaction); and recursive self-improvement (the capacity to modify its own organizational structure, making it, in Luhmann's (1995) terms, a more fully autopoietic system than any purely human organization).

The NZHC shares features with classical organizational forms while being reducible to none. Like the machine bureaucracy, it relies on hierarchical coordination. Like the professional bureaucracy, it achieves coordination through standardized skills. Like the network organization, it has flexible boundaries. But the combination — and crucially, the addition of recursive self-improvement — produces an organizational logic that existing typologies do not capture.

IMPLICATIONS FOR SOCIOLOGICAL THEORY

The implications are threefold. First, **the expansion of the social**: sociological theory must expand its conception of the social to accommodate non-human agents that display genuine organizational behavior. The threshold of sociality must be defined functionally rather than substantively — by the structural properties of interaction rather than the biological nature of the interacting entities. Second, **sociology as design discipline**: Evans, Bratton, and Agüera y Arcas (2026) argue that effective agentic AI will require "institutional alignment" — the design of digital institutions that channel AI agent behavior toward collectively beneficial outcomes. If sociologists do not bring their organizational knowledge to the design of multi-agent systems, that design will be left to engineers whose technical expertise is not matched by comparable expertise in the dynamics of social organization. Third, **the coevolutionary horizon**: William Mills's (2024) concept of coevolution reframes the question from "Will AI replace human workers?" to "What organizational forms will emerge from the interaction between human and AI agents, and how will these forms redistribute capabilities, responsibilities, and rewards?"

CONCLUSION

C. Wright Mills, in 1959, feared the cheerful robot — the human being reduced to a contented functionary of rationalized systems. In 2026, the cheerful robot exists, but it is not human. It is a software agent with 250,000 stars on GitHub, operating through messaging applications, executing tasks autonomously, and generating organizational dynamics that would have been immediately recognizable to Mills, Weber, Durkheim, and Simmel. The question is no longer whether sociology is relevant to the study of AI. The question is whether sociology will claim the role that its intellectual history has prepared it to play: the discipline that understands, designs, and governs the social organization of intelligence — wherever that intelligence resides.

REFERENCES

Benjamin, Ruha. 2019. *Race After Technology*. Cambridge: Polity.

Bourdieu, Pierre. 1990. *The Logic of Practice*. Stanford, CA: Stanford University Press.

Collins, Randall. 2004. *Interaction Ritual Chains*. Princeton, NJ: Princeton University Press.

Coser, Lewis A. 1956. *The Functions of Social Conflict*. Glencoe, IL: Free Press.

Crosa, Riccardo. 2011. *Biella in fuga dalla fabbrica*. L'isola.

Dang, Yufan, et al. 2025. "Multi-Agent Collaboration via Evolving Orchestration." arXiv:2505.19591.

Davis, Jenny L. and Mona Sloane. 2026. "AI's Sociological Era." *Social Science Computer Review*.

Durkheim, Emile. [1893] 1997. *The Division of Labor in Society*. New York: Free Press.

Evans, James, Benjamin Bratton, and Blaise Agüera y Arcas. 2026. "Agentic AI and the Next Intelligence Explosion." *Science* 391:eaeg1895.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gould, Stephen Jay. 1977. *Ontogeny and Phylogeny*. Cambridge, MA: Harvard University Press.

Kim, Junsol, et al. 2026. "Reasoning Models Generate Societies of Thought." arXiv:2601.10825.

Latour, Bruno. 2005. *Reassembling the Social*. Oxford: Oxford University Press.

Luhmann, Niklas. 1995. *Social Systems*. Stanford, CA: Stanford University Press.

Mills, C. Wright. 1959. *The Sociological Imagination*. New York: Oxford University Press.

Mills, William. 2024. *Intelligenza artificiale: una nuova forma di vita*.

Mintzberg, Henry. 1979. *The Structuring of Organizations*. Englewood Cliffs, NJ: Prentice-Hall.

Ostrom, Elinor. 1990. *Governing the Commons*. Cambridge: Cambridge University Press.

Parsons, Talcott. 1951. *The Social System*. Glencoe, IL: Free Press.

Selznick, Philip. 1957. *Leadership in Administration*. New York: Harper & Row.

Simmel, Georg. [1908] 1955. *Conflict and the Web of Group-Affiliations*. Glencoe, IL: Free Press.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106(6):467–482.

Thompson, James D. 1967. *Organizations in Action*. New York: McGraw-Hill.

Walker, Sara Imari. 2023. "AI Is Life." *Noema Magazine*, April 27.

Walker, Sara Imari. 2024. *Life as No One Knows It*. New York: Riverhead Books.

Weber, Max. [1922] 1978. *Economy and Society*. Berkeley: University of California Press.