

OnPoint Subscriber Exclusive

Longer Reads provide in-depth analysis of the ideas and forces shaping politics, economics, international affairs, and more.

Subscribe

Will AI Kill the Firm?

Oct 31, 2025 **SAMI MAHROUM**

DUBAI – Almost everywhere, debates about AI remain narrowly focused, indeed almost fixated, on its impact on employment and return on investment. Which jobs will disappear? Which skills will endure? Are current valuations justified? These questions, while important, overlook a deeper issue: whether firms – the institutions that have organized economic life for the past two centuries – will themselves survive in their current form in the wake of AI.

It's worth remembering that the firm is not a natural feature of economic life. It emerged in the 16th century, as merchants sought new ways to manage the vast distances and uncertainties of global trade. To meet those challenges, the Muscovy Company was formed in 1555 under Britain's Queen Mary, to be followed by the more famous British and Dutch East India Companies. These joint stock companies pioneered a new model: pooling capital from hundreds of investors and building bureaucracies capable of managing years-long projects. Through accounting, auditing, and hierarchy, they created an architecture of trust that made large-scale collaboration among strangers possible.

The Industrial Revolution carried this organizational logic into a new era. Steam, steel, and telegraphy demanded centralized corporate command. By the early 20th century, Henry Ford had perfected the model with the moving assembly line, fusing mechanical precision with social discipline: repetitive tasks, specialized roles, and standardized output. The factory transformed the worker into a key component of the corporate machine.

Make your inbox smarter. Select Newsletters

In the mid-20th century, Toyota's production system redefined the firm once again. Lean manufacturing and just-in-time logistics replaced rigidity with responsiveness, empowering workers on the factory floor. The firm became less like a machine and more like an information network – the last major evolution of the corporate model.

The firm thrived because it addressed four fundamental constraints: information scarcity, high transaction costs, the need for supervision, and the aggregation of capital. As the Nobel laureate economist Ronald Coase explained in his 1937 essay "The Nature of the Firm," companies exist because markets are costly to use. Firms enable people to cooperate under long-term contractual arrangements rather than through one-off transactions.

Oliver Williamson – awarded the Nobel Prize in 2009 – expanded the theory of the firm in his 1975 book *Markets and Hierarchies*. Firms, he explained, exist because humans are imperfect: we cannot foresee every contingency when making deals (bounded rationality), and we sometimes put our own interests first (opportunism).

These flaws become especially acute when investments are highly specialized. To avoid endless bargaining and renegotiation, firms rely on managers to make decisions and enforce them.

The AI Surplus Paradox

Agentic AI can now design reliable agreements that are automatically enforced by smart contracts, while cloud services make knowledge and assets instantly shareable. With the coordination problems that once justified corporate hierarchies seemingly solved by computer code, firms' traditional rationale is beginning to fade.

In his recent book *How Progress Ends*, the University of Oxford economic historian Carl Benedikt Frey identifies a deep structural tension running through economic history. Modern economies, he writes, have always balanced the search for new ideas (exploration) and the scaling and refinement of what already exists (exploitation). Exploration thrives in open, experimental environments, while exploitation depends on structure, discipline, and hierarchy. This tension, Frey argues, has shaped every phase of economic development. During the industrial age, the firm became the quintessential vehicle for exploitation. Yet those same institutions, he warns, have become barriers to progress: when structures designed for exploitation dominate, they suppress the exploratory capacities societies need to adapt. Over time, efficiency hardens into inertia, and progress stalls – not because ideas run out, but because institutions built for the old economy resist the logic of the new.

At its core, the firm is a mechanism for coordinating people, capital, and knowledge through habits that eventually solidify into routines. Management's role has always been to make those routines more efficient and scalable, and that is where most AI investment is focused today: automating what already exists.

Today, two related developments are testing firms' capacity to absorb all the surplus they generate. The first is the collapse of the boundary between exploration and exploitation. Historically, these were sequential: researchers discovered ideas in labs, and firms deployed them through established processes.

AI is dissolving that division across a growing range of domains. In drug discovery, for example, algorithms simultaneously search for new molecules and model how they can be produced at scale. In software engineering, generative models write, test, and debug code in a continuous loop. And in marketing, AI systems design, test, and optimize campaigns in real time, erasing the line between research and execution. What once required contracting out – research and development, production, operations – can now unfold within a single, integrated system.

The second development is the expansion of human capability. As AI technologies advance, they push the boundaries of what people can imagine, create, and achieve. Consider how generative AI models can assist in writing a research paper, turning a rough draft into a coherent, integrated whole by deepening the reasoning, adding nuance, and refining the prose. Similarly, an engineer can use AI tools to prototype systems that once required large teams, while a single analyst can perform the kind of work that demanded entire departments.

The combined effect of these developments is to generate more ideas, more initiatives, and more problem-solving energy bubbling up from within organizations than their hierarchical structures were built to absorb. When the same tools that generate discoveries can also act on them, and when AI-augmented employees move faster than managers can coordinate, the firm's architecture of control begins to look more like an obstacle than an advantage. As AI enhances both human and machine agency, established firms find it increasingly difficult to contain the value they create. The result of this internal overload is not greater efficiency, but entropy.

That is the AI adoption paradox: the more capable an organization becomes through AI, the harder it becomes to manage the human potential it unleashes. In a seminal 1990 paper, Wesley Cohen of Duke University's Fuqua School of Business and Daniel Levinthal of the University of Pennsylvania defined a firm's absorptive capacity as its ability to recognize, absorb, and apply new external knowledge. Building on their work, Shaker Zahra of the University of Minnesota's Carlson School and Gerard George of Georgetown University later reframed the concept as a dynamic, multi-level capability linking individual cognition to collective routines.

But when AI accelerates the pace of individual learning and decision-making beyond what organizations can handle, that balance collapses. As individuals evolve faster than the mechanisms built to coordinate them are able to adapt, absorptive capacity becomes internally misaligned. The routines that once translated learning into organizational capability are eroded, creating a crisis of coordination. Put differently, the surplus has shifted from what firms make to what individuals can do. And the same dynamic is playing out across societies, as customers, suppliers, and regulators all acquire new adaptive capacities of their own.

“Command and Control” or “Orchestrate and Empower”?

There are plenty of cautionary tales. Kodak invented digital photography but failed to foresee a world in which everyone became a photographer. While its strength lay in film technology and distribution, its governance structure remained calibrated to an era in which images were rare and costly to produce.

Nokia, once the world's largest cellphone manufacturer, dominated hardware but missed the shift to digital platforms that redefined value as coordination rather than production. The video rental chain Blockbuster collapsed when streaming emerged, its business model unable to adapt to an era in which control over time and access had moved from corporations to consumers. Each of these failures can be attributed to the same underlying weakness: the inability to reconfigure internal hierarchies once the sources of external agency had shifted.

Other firms, meanwhile, fell prey to portfolio paralysis. The Dutch electronics conglomerate Philips had become so sprawling by the 1980s that no coherent strategy could unite its numerous divisions, which ranged from light bulbs to semiconductors. Despite continuous innovation, strategic coherence broke down; each unit excelled on its own terms, but they pulled the company in different directions. Philips's German rival, Siemens, grappled with similar tensions as it tried to reconcile its industrial heritage with its expanding digital businesses. In both

cases, the failure was one of coordination: subsidiaries generated more value than the managerial hierarchies of Philips and Siemens could recognize or direct.

Public institutions are not immune to these disruptions; the questions that once haunted Kodak and Philips confront ministries, universities, and foundations struggling to govern abundance. Research funding agencies, for example, are being inundated with AI-assisted grant applications. Machine-generated proposals have surged, upending traditional evaluation processes and overwhelming committees accustomed to slow, deliberative peer review. The challenge is no longer one of administrative efficiency but of institutional cognition. Can institutions built for a slower epistemic pace remain fit for purpose at a time when ideas evolve at machine speed?

That is the essence of the AI surplus paradox: When potential and initiative outpace an organization's ability to govern them, success becomes a source of instability. Rather than facing a shortage of intelligence or skill, firms now contend with an abundance of untapped capacity. Managing that overflow requires moving decision-making closer to where insights emerge, so that those best positioned to act can do so without delay or bureaucratic friction.

Consulting partnerships provide a useful model. Partners operate with a high degree of autonomy within a shared infrastructure of trust, reputation, and financial accountability. They allocate resources directly, without waiting for managerial approval, and are rewarded for turning insight into client value.

To thrive in the AI era, firms will need to embrace a similar design: distributed authority grounded in shared infrastructure – a networked federation where alignment is maintained not through hierarchy but through transparent data flows and carefully calibrated incentives. The architecture of the firm must evolve from its current command-and-control model to one that orchestrates and empowers. Firms that master this transition will turn the AI surplus into a strategic advantage.

The Rise of the Actor-Network

As exploration and exploitation become a single process, discovery and execution no longer occur sequentially but within the same continuous cycle of sensing, learning, and acting. In this emerging order, the economy organizes itself through information flows rather than chains of command.

To make sense of this shift, we need a language that explains how agency is shared across human and machine systems. Actor-network theory offers exactly that. While traditional network theory treats people and systems as distinct entities that connect, actor-network theory argues that agency – the capacity to act and produce outcomes – emerges from the network that connects them. Together, a doctor using diagnostic software, a nurse relying on real-time patient data, and an algorithm learning from patterns constitute a single agent whose capabilities are distributed across all three participants.

AI makes such integration seamless. When people and their AI counterparts think, execute, and create together, they become a coherent economic actor whose agency

is shared between human and machine. Reproduce these relationships across thousands of hybrid participants, and the result is a true actor-network.

Actor-network theory can be traced back to the French philosopher Bruno Latour, who introduced the concept in his 1987 book *Science in Action* and expanded it in his 2005 book *Reassembling the Social*. Latour regarded agency as inherently distributed, emerging from the interplay between systems, people, and technologies.

Viewed through this lens, the AI economy appears not as a hierarchy but as an ecosystem in which humans and their agents transact with one another more efficiently than through a firm. As steady jobs and lifelong careers give way to purpose-driven projects, capital will follow problems, not companies. Governments, for their part, will regulate digital protocols and tax the exchanges where value is created rather than the profits companies report.

While such a transformation may sound promising, something vital will be lost. The firm, after all, once provided more than a paycheck; it offered community and a sense of belonging. Its decline may bring new freedoms, but also a deep longing for connection.

Agent Bosses in Charge

As agency becomes distributed between humans and AI models, a new kind of economic actor is beginning to emerge: the agent boss. According to Microsoft executive Jared Spataro, the agent boss is “someone who builds, delegates to, and manages agents to amplify their impact and take control of their career in the age of AI.”

Spataro’s observation captures a fundamental shift in the locus of economic agency. The agent boss is an individual who becomes economically coherent through augmentation. Neither a contractor offering labor nor an employee bound by corporate hierarchy, the agent boss is a micro-entrepreneur whose “startup” is a partnership between the self and a constellation of AI agents. Together, they form an economic unit that is more capable than a human or a machine alone, yet free of organizational overhead.

Unlike traditional employees, who function as interchangeable units of labor, agent bosses own their relationship with their AI agents. Able to move fluidly between clients and collaborators, they build a career made up of distinct projects, each requiring different human and machine collaborators and producing verifiable results.

These agent-boss networks may consist of only a few people and a fleet of AI agents, all working across time zones and dispersing once a certain goal is achieved. A climate analyst in Nairobi, a designer in Lisbon, and a developer in Singapore might collaborate for a month on tackling a climate-adaptation problem, leaving behind data, code, and insights for others to build on. Freelancers on platforms like Upwork are already moving in this direction, describing themselves not as contractors but as “agent orchestrators” because they manage a fleet of large language models (LLMs), toolchains, and databases.

To be sure, distributed networks already exist. Wikipedia and Linux, after all, have thrived for decades on global collaboration. But these commons-based models rely on institutional redistribution of value. By contrast, the agent-network model enables the individual to capture and own the value they create, including their code and data, rather than ceding it to organizational intermediaries.

The Hubless Economy

The Latourian economy departs from the networked society described by sociologists Manuel Castells, formerly Spain's minister of universities, and Saskia Sassen of Columbia University. In *The Rise of the Network Society* (1996), Castells traced the global flows of capital and information that reconfigured industrial hierarchies. Sassen, in *The Global City* (1991) and *Territory, Authority, Rights* (2006), showed how those same flows concentrated power in cities like New York and London. Taken together, their work highlighted the unequal effects of global connectivity: major transnational hubs have grown increasingly dominant, while peripheral countries function as consumers or low-value nodes.

In the Latourian economy, by contrast, the actor creates the network, not the other way around. Connections are loose, situational, and temporary, sustained by shared goals rather than centralized infrastructures. The Latourian economy decentralizes both flows and structures: each actor builds and maintains their own network, assembling human and machine collaborators for as long as needed.

By making digital tools lighter, cheaper, and more widely distributed, the Latourian economy enables anyone with a laptop and an internet connection to access capabilities once reserved for research labs and multinational companies. A small design studio in Accra can now train an LLM, deliver data-driven services, or collaborate with clients in São Paulo. Value creation is no longer a function of owning factories, patents, or physical distribution networks; instead, it rests on one's ability to orchestrate knowledge flows.

Even so, while the Latourian economy redistributes capability, it remains unclear whether it also redistributes power. AI certainly has the potential to narrow longstanding divides, but its impact will depend less on technology itself than on how access, credit, and governance are built into value-producing networks.

Three concerns loom large. First, why wouldn't large firms simply use AI to bolster their existing advantages? They almost certainly will. But as firms become increasingly augmented by AI, their internal dynamics may shift. Authority could be delegated, algorithms could replace layers of management, and capabilities could be redistributed across divisions. In the process, large firms may become more like holding companies for actor-networks than traditional corporations.

Second, distributed systems still struggle with coordination problems that demand hierarchy, since complex projects require both collaboration and aligned incentives. But AI helps mitigate this by reducing the need for intermediaries. Open-source projects already demonstrate that large, distributed teams can work together effectively on complex technical challenges. Linux and Kubernetes, both products of

open collaboration and decentralized governance, are prime examples. Instead of disappearing, the firm could evolve into an ecosystem of loosely affiliated networks.

Third, because fragmentation can undermine efficiency and coherence, an economy of millions of agent bosses risks devolving into a patchwork of self-contained initiatives. While the firm once solved this by enforcing shared priorities, an actor-network economy relies on autonomous actors tackling shared challenges. That model works when feedback loops are tight, but falters when timelines grow longer. As problems unfold over decades, sustained coordination becomes exceedingly difficult.

Does the Firm Have a Future?

When Coase asked in 1937 why firms exist, his answer was transaction costs. That logic defined the bureaucratic age, but in a world of AI systems and agent networks, where costs are approaching zero, it no longer holds. Bureaucracy now appears in algorithmic form, encoded into protocols and digital governance systems that automate what managers once enforced.

Against this backdrop, Coase's question takes on new urgency. For most of the 20th century, corporate hierarchy proved more effective than markets at coordinating economic behavior. As AI-augmented markets begin to handle tasks that used to require managerial oversight, the answer is less clear-cut. Still, traditional hierarchies may be better suited for certain kinds of work, especially long-term capital ventures and heavily regulated industries.

Frey warns that progress falters when the structures built to exploit innovation cannot evolve as quickly as the forces that explore it. The actor-network economy may offer a remedy, not by eliminating hierarchy altogether, but by making it the exception rather than the rule. In a world where exploration and exploitation happen within the same distributed loops, the old divide between discovery and delivery collapses. But the answer to Frey's paradox lies in a mixed economy rather than uniform actor-networks. While some problems may still require hierarchical organization, it should be applied deliberately, not assumed by default.

As the conditions that previously made the firm necessary dissolve, its survival may depend on unlearning. Research on organizational learning suggests that knowing how to let go is as important as absorptive capacity. In the age of AI, that may be the firm's last defense. Since no organization can absorb every new source of agency, it must decide what to outsource, spin off, or abandon. The future of organizational learning may rest as much on strategic forgetting as on knowledge accumulation.

The firm's replacements will be shaped by our collective choices. Can we design legal systems that empower agent bosses, or will we confine them within regulatory frameworks built for employees? Can we create tax systems capable of governing distributed value creation? And can we develop new forms of social insurance and belonging?

What's certain is that the firm is about to change, and its role in our lives will change with it. We already know actor-networks can organize economic life; the real

question is whether they can preserve the sense of equity, coherence, and shared purpose that defined the modern corporation. The 20th century perfected the art of building firms; the 21st century will test our ability to live without them.

Carl Benedikt Frey, *How Progress Ends: Technology, Innovation, and the Fate of Nations* (Princeton University Press, 2025)

SAMI MAHROUM

Sami Mahroum, Founder of Spark X, previously held posts at INSEAD, the OECD, and Nesta.

<https://prosyn.org/1vKZ1CY>

[Support Project Syndicate](#)

[Subscribe](#) [Upgrade](#) [Donate](#)

[Make your inbox smarter. Select Newsletters](#)

© Project Syndicate - 2026 [Privacy Policy](#) [Terms & Conditions](#) [Cookie Policy](#)