

✓ INSIGHT

Artificial Intelligence

Designing a Fluid Organization of Humans and AI Agents

Herman Vantrappen



Image Credit | nippon

Where humans and AI agents cohabit, agility and trust are key, but humans stay accountable for the agency of AI agents.

In ancient Greece the mythical king Sisyphus was punished and forced by the gods to roll an immense boulder up a hill, only to see it roll back, and start all over again. When redesigning their organizations, many business leaders feel like living Sisyphus' ordeal: as soon as they have settled on the best possible design, they are bound to re-start, as circumstances change continually.

RELATED ARTICLES

Vegard Kolbjørnsrud, "[Designing the Intelligent Organization: Six Principles For Human-AI Collaboration](#)."
California Management Review, 66/2 (2023): 44–64.

RELATED TOPICS

[Agentic AI](#)

[Organizational Design](#)

[Strategic Agility](#)

The entry of AI agents who cohabit, interact and collaborate with human agents may be such a change of circumstance. Their entry inevitably raises the question whether and how the organization in the presence of AI agents needs redesigning so as to keep functioning fluidly.¹ Organizational design is all about determining the organization's goals, defining the tasks required to reach those goals, distributing roles and responsibilities for each task to various agents, coordinating the execution of the tasks, motivating individual agents to continue acting in the best interest of the collective, and governing progress toward the goals. Some people doubt that the growing deployment of AI agents affects any of those design aspects: they argue that AI agents, even though increasingly human-like and

autonomous, are still a piece of software code processing data through a piece of hardware. Others are affirmative, though: they argue that the advance toward Artificial General Intelligence (AGI) makes the organizational design issue highly pertinent.

In this article we will first explain what we mean by a fluid organization. We will then describe the design features that one can manipulate to arrive at a fluid organization. Next we will discuss how the work of AI agents should be designed and governed so that their deployment enhances fluidity.

The Fluid Organization

We define a fluid organization simply as an organization that is both effective and efficient. Effective means that the organization achieves its business objectives; efficient means that it does so with minimal waste of energy. We claim that all organizational redesigns eventually seek to make an organization more fluid. The managerial question of interest is about the determinants of a fluid organization.

A number of thought leaders promote the idea that fluidity is to be achieved through the absence of organizational structure, both horizontally (“no silos”) and vertically (“no hierarchy”). They claim that an organization should be designed as a shared pool of qualified resources of diverse expertise and experience; that temporary teams of variable size should be mobilized by drawing from that pool; that each team should be given a clear mission aligned with the company’s overarching goals and boundaries; and that the teams are self-managing, i.e., they can freely decide how they go about realizing their mission.

In our experience such an approach may work for small companies (or specific small parts of a large company) in a knowledge-intensive, project-based, one-product business, say, a 100-people software development start-up. However, to achieve fluidity in general and at scale, something more is needed. Admittedly silos and hierarchy may have undesirable side effects – which have led to the negative connotation – but they exist for good reasons: they enable the consolidation of expertise, the assignment of accountability, and the cultivation of a sense of identity.²

In other words, structure does not come at the expense of fluidity. Quite to the contrary, structure enables fluidity. Structure makes the difference between liquidity and fluidity. Liquidity is like a pond of still water; fluidity is like a hydraulic system of reservoirs, canals and pumps to distribute drinking water, irrigate land, or power turbines.

That brings us back to the question that applies to the vast majority of organizations: if lateral boundaries (i.e., between silos) and hierarchical boundaries (i.e., between layers) are ineluctable, how is fluidity to be achieved? Our answer is expressed by the symbolic formula “fluidity = agility x trust”. Agility relates to the imperative that people collaborate whenever needed, in particular across silos. Trust relates to the imperative that people delegate and elevate whenever needed, in particular across layers. Let’s turn now to a more concrete description of these two notions.

Agility and Trust

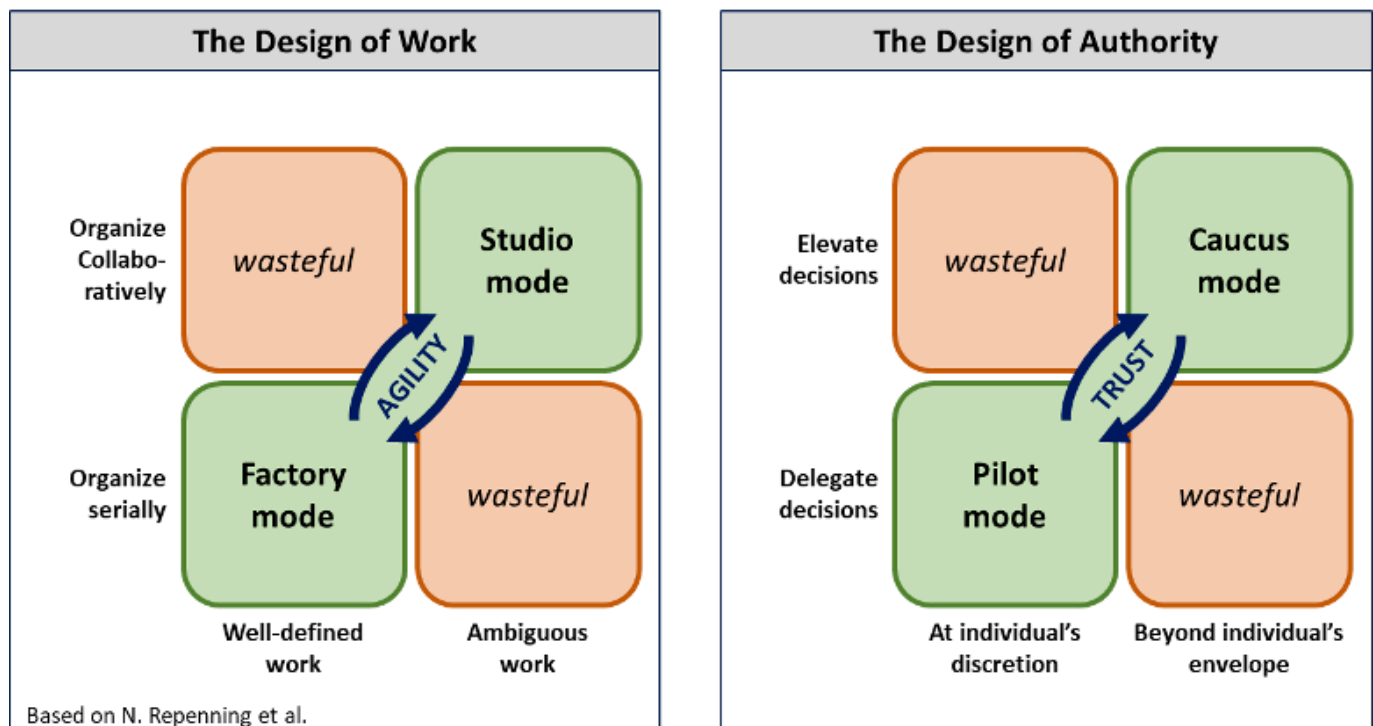
The practitioner literature is rife with starry-eyed notions about agility and trust. As far as agility is concerned, we prefer to build on Repenning et al.’s pragmatic insight that agility is the ability of an organization to cycle back and forth between “factory” mode (well-defined work organized serially) and “studio” mode (ambiguous work organized collaboratively).³ An archetypal example is the assembly line at Toyota: operators execute well-defined work serially, but when an unexpected break-down occurs, they push a button that stops the line and triggers the intervention of qualified people who collaborate and solve the problem, after which the serial work restarts.

As far as trust is concerned, we introduce the notion that trust is the ability of an organization to cycle back and forth between “pilot” mode (a person making decisions at their own discretion within the envelope of authority delegated to them) and “caucus” mode (a group of people deliberating among themselves to make a joint decision on a superordinate matter elevated to them). The person who delegates responsibility for the proper execution of a task to another person does so in the confidence that the latter will, first, execute the task autonomously and diligently, and secondly elevate emerging issues that surpass their competence or authority promptly. Conversely, the person who assumes responsibility for the proper execution of a task does so in the confidence that the person

who has delegated that responsibility will, first, respect the former's freedom of action, and secondly respond promptly to requests for support or decisions for which the former does not feel competent or authorized.

The two 2x2 grids in Exhibit 1 visualize the back-and-forth cycling that defines agility and trust, respectively. Note that the upper left and lower right quadrants represent undesirable modes, as they are wasteful. For example, the lower right quadrant of the agility grid corresponds with organizing ambiguous work serially, which tends to lead to ineffective iterations; the upper left quadrant of the trust grid corresponds with escalating decisions about delegated work, which tends to slow down the completion of the work.

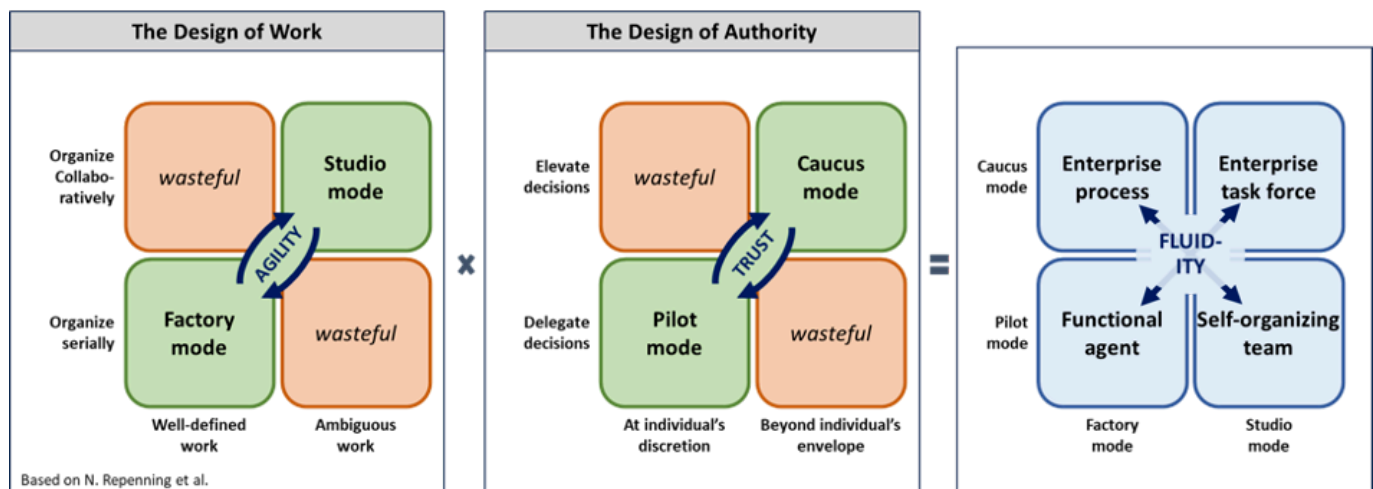
Exhibit 1. Defining agility and trust



By applying the formula "fluidity = agility x trust", we get the four organizational forms shown in Exhibit 2, each with a convenient label. "Factory x pilot" leads to "functional agent". For example, a travelling salesperson is trusted to set up and follow a routine of visits to prospects and clients as he or she sees best fit. "Studio x pilot" leads to "self-organizing team". For example, a country team is trusted to tailor and locally implement a new standard tool that has been developed at the corporate center. "Studio x caucus" leads

to “enterprise task force”. For example, the executive team in consultation with the board of directors develops an emergency response to a hostile takeover bid. “Factory x caucus” leads to “enterprise process”. For example, all of the company’s businesses and functions follow a script to iterate through the annual strategic planning and budgeting process. Fluidity, then, is the ability to cycle back and forth between any of these four organizational forms quickly and smoothly, as circumstances dictate.

Exhibit 2. Defining fluidity



Designing for Fluidity

The cycling-back-and-forth that characterizes agility and trust, hence fluidity does not come automatically – it must be designed into the organization. That can be done through five practices: subsidiarity, triggers, checks, a help chain, and transparency (the middle three of those have been introduced by Repenning et al.).

The first practice is to apply subsidiarity as the default design principle: assign responsibility for a task to the smallest possible team at the lowest possible hierarchical level, unless there are good reasons to do differently. The smallest possible team in practice may just be a single person; the lowest possible level often is the frontline, such as assembly line operators, salespeople, bedside nurses, retail store assistants, legal associates and corporate recruiters, to name a few. Subsidiarity makes sense because small teams on the frontline usually have the focus, expertise and immediacy to respond

best and fastest to emerging events, opportunities and hazards. Generally it means that functional agent – and the underlying factory and pilot modes – is the default organizational form.

The second practice is to define triggers for shifting out of the default mode when appropriate, i.e., from factory to studio mode, or from pilot to caucus mode. Practically a trigger is set up by making a list of the problems that would require the organization to shift into the other mode whenever any of these problems does occur. An example of a factory-to-studio trigger is when a bedside nurse in the course of their routine monitoring of a patient's vital signs and wounds observes an out-of-range development, thus triggering a notification to the healthcare team. An example of a pilot-to-caucus trigger is when a salesperson during a standard Know Your Customer (KYC) review discovers that the prospect may also have activities in an embargoed and sanctioned country, thus triggering a notification to the sales manager and compliance department.

Checks are the third practice. Whereas triggers are reactive, checks are proactive: they are prescheduled points at which one will expressly shift from the default mode into the other mode. The assumption is that having regular checkpoints enables the team to preemptively address issues that in all likelihood would pop up and be raised to them at some later moment. For example, retail chain stores may use daily stand-ups with their frontline people and the store manager to discuss priorities, address any issues, and ensure smooth operations during the day.

The fourth practice is to identify and design help chains. A help chain is the sequence of individuals whom a functional agent has to involve when a trigger has been activated or a check needs scheduling. For example, when a salesperson (= functional agent) in the course of their daily work (= factory + pilot) observes possibly unethical behavior by a colleague (= trigger), they are advised to report their concern through the established whistleblowing procedure (= help chain).

Finally, the triggers, checks and help chain will be effective only if the organization espouses transparency as a value, norm and attitude. For people to willingly work together across silos and layers, they need to know about each other's motivations, preferences and contributions, and observe and experience fairness and reciprocity.

AI Agents in a Fluid Organization

So far we have talked about designing a fluid organization populated by humans. What if those human agents cohabit with AI agents? What does it take to maintain, if not strengthen agility and trust if part of the work of the proverbial assembly line operator, bedside nurse, salesperson, retail assistant, legal associate or corporate recruiter is taken over or augmented by an AI agent? How to adapt the organization's design in terms of subsidiarity, triggers, checks, help chain and transparency to account for the involvement of AI agents, not only in factory and pilot mode but also in studio and caucus mode? Does any of the above turn even more complex when an AI agent interacts not only with human agents but also with other AI agents? Or when an AI agent is mandated to generate the nudges leading to a mode shift?

As an example, let's take the familiar corporate hiring process and zoom in on the candidate preselection activities for a vacancy at company ABC. Tasked with establishing a shortlist of candidates, the two recruiters and an AI agent in the HR department by default operate as functional agents, that is, they work in factory mode at start and have pilot mode authority:

- *Functional agents.* As the applications come in, the AI agent screens the resumes and preselects fitting candidates. Depending on today's workload of the two recruiters working in parallel, the AI agent directs each preselected application to one of them.
- *Trigger + help chain.* The AI agent also flags potentially interesting candidates who meet the minimum requirements but whose extraordinary profile falls outside the defined search boundaries. Each such flag triggers a shift from factory to studio mode, whereby the two recruiters sit together to discuss the candidate, possibly requesting the AI agent to search in real-time for additional information from third-party sources.
- *Check + help chain.* Once a week the AI agent prepares a report about patterns in the applications received. It flags anomalies in terms of response rates by search channel and candidate demographics. Each report constitutes a check, whereby the two recruiters shift from pilot to caucus mode, that is, they brief their HR manager, who in

turn may decide to consult with the line manager who had defined the original hiring need, before deciding on any change of course.

- *Enterprise task force.* Upon closure of the application period, the AI agent prepares a summary report with recommendations about the shortlisted candidates, as selected by the two recruiters. They validate and submit the report to the HR manager and the line manager, who jointly decide on the final list of candidates to be invited for an interview.
- *Transparency.* The role distribution between the AI agent, recruiters and managers, as described above, is clear and agreed. So are the protocols with triggers and checks. Equally importantly for the comfort and confidence of the human agents, the explainability of the AI algorithm and its outputs has been assured.

The above process looks pretty straightforward and fluid, that is, both effective and efficient. And while the involvement of the AI agent improves fluidity, it does not fundamentally change the way of working. What it does require, however, is the careful design of what Schrage and Kiron call the organization's Intelligent Choice Architecture (ICA): who decides, first, who is to define the range of potential choices and, second, who is to make the final choice?⁴

In the foregoing hiring example the HR manager and line manager must define the choice sets at every decision point where the AI agent makes a recommendation to the corporate recruiters; they must also define when the corporate recruiters must act upon the nudges generated by the AI agent, and when the recruiters can override the AI agent so that human judgment remains appropriately engaged. In other words, the HR and line manager not only continue to be accountable for the quality of the hiring decision per se but they are also becoming accountable for the quality of the decision environment in which their people operate. More generally, it is the duty of the company's upper echelons, as Schrage and Kiron explain, to "determine who has the authority and responsibility to architect, deploy, and govern choice environments where human judgment and AI capabilities intersect." Human agents remain accountable for the agency of AI agents.

References

1. Vegard Kolbjørnsrud, “**Designing the Intelligent Organization: Six Principles for Human-AI Collaboration.**” California Management Review 66, no. 2 (Winter 2024): 44-64.
 2. Herman Vantrappen and Frederic Wirtz, “**CEOs Can Make (or Break) an Organization Redesign.**” MIT Sloan Management Review 65, no. 1 (Fall 2023): 79-84.
 3. Nelson P. Repenning, Don Kieffer, and James Repenning, “**A New Approach to Designing Work.**” MIT Sloan Management Review 59, no. 2 (Winter 2018): 29-38.
 4. Michael Schrage and David Kiron, “**The Great Power Shift: How Intelligent Choice Architectures Rewrite Decision Rights.**” MIT Sloan Management Review, January 28, 2025.
-



Herman Vantrappen [Follow](#)

Herman Vantrappen is the Managing Director of Akordeon, a strategic advisory firm. He is the coauthor of The Organization Design Guide (Routledge, 2024) and publishes regularly in a variety of journals, including California Management Review, Harvard Business Review and MIT Sloan Management Review.