The automation trap

Companies are racing to deploy agentic AI. Most are doing it wrong.

-Mansi Agarwal

There is a telling pattern in how organizations have responded to each wave of enterprise technology over the past three decades. Confronted with a genuinely new capability — the internet, cloud computing, the smartphone — they have, with impressive consistency, used it to do existing things slightly faster. Agentic AI, the latest and most consequential wave, is proving no different.

The new generation of AI systems is qualitatively distinct from the chatbots and copilots that preceded them. Rather than responding to prompts, agentic systems can reason through multi-step problems, invoke tools, query databases, execute transactions and run autonomously for hours at a stretch. They are not assistants. They are, in a meaningful sense, workers. And yet firms are deploying them as though they were merely faster typists — layering them onto approval chains, handoff sequences and status-update processes that were designed, implicitly, around the limitations of human cognition. The productivity gains are real but modest. The transformational gains remain largely unclaimed.

Paving the cowpath

The processes that govern most large organizations were built to compensate for things people cannot do well: hold vast amounts of information in mind simultaneously, work without sleep, act consistently across thousands of repetitions. Approval layers exist because judgment needed a human checkpoint. Handoffs between departments exist because no single person could carry full context across an entire workflow. Status meetings exist because information did not move freely between systems.

Agentic AI is not subject to most of these constraints. It does not forget context between steps. It does not need a meeting to know what happened yesterday. It can hold the entire history of a customer relationship in mind while simultaneously querying a pricing system and drafting a response. The bottlenecks that justified the old architecture simply do not apply.

Yet those bottlenecks remain, because the processes have not changed. Firms are, in effect, paving the cowpath: automating workflows that were already suboptimal, now at greater speed and with greater confidence. The result is a version of agentic AI that is faster and cheaper than what came before, but nowhere near as useful as it could be.

There is a structural data problem, too. Processes designed around human judgment tolerate ambiguity. People are skilled at inferring what a colleague meant by an imprecise instruction, or filling in a gap in the records from experience. Agents are not. They are highly capable within the domain of well-structured, accessible, real-time information — and brittle outside it. Much of the data that drives business decisions remains locked in email threads, legacy systems and spreadsheets that no agent can easily reach. Organizations that have not addressed this before deploying agents will find the agents spending much of their time stuck.

Signs your process is not agent-ready

- Approvals add elapsed time but not genuine judgment to routine decisions
- Key data lives in inboxes, spreadsheets or systems inaccessible to external tools
- Handoffs between teams require reformatting data that already exists elsewhere
- Process steps exist primarily to compensate for information failing to flow automatically
- Outcomes are measured by task completion rather than the business result the task was meant to drive
- Exception handling is undocumented and depends entirely on individual memory

First principles

What would it look like to design processes from scratch around what agents can do? Five principles are worth internalizing.

The first is to define outcomes, not procedures. Legacy processes specify what people should do at each step. Agentic processes should specify what must be achieved, the data and tools available to achieve it, and the boundaries within which the agent may operate. How the goal is reached is the agent's domain. Specifying it in advance forfeits most of the value.

The second is to eliminate handoffs where context is the casualty. The friction that accumulates at transition points — between shifts, departments, systems or time zones — is largely a tax on human memory. Agents do not forget. Processes that build on persistent, cumulative context, rather than reconstructing it at each stage, will run faster and make fewer errors.

The third, and perhaps the most difficult for organizations to accept, is that human oversight should be the exception rather than the default. Routing every agent decision through a human approval step is not prudent governance; it is the old process with a new label on it. Leaders must be explicit about where human judgment is genuinely irreplaceable — complex ethical trade-offs, novel situations, high-stakes exceptions — and trust agents to handle everything else.

Fourth: data infrastructure is process infrastructure. This is the investment most frequently underestimated and most consequential to get right. Agents are only as capable as the information they can access and act upon.

Fifth: feedback must be built in from the start. Processes should be designed so that agents can surface patterns, flag anomalies and improve over time. Agentic systems that cannot learn from their own outcomes are being used at a fraction of their potential.

Where the opportunities are largest

Some domains are particularly well-suited to this kind of reinvention. Customer operations have long been constrained by the gap between what clients expect and what case-handling processes can deliver. Agents that retain full customer context, connect to the relevant systems and resolve issues without escalation can close that gap — not merely faster, but better.

Finance and procurement offer similar opportunities. Most approval chains in these functions add delay without adding commensurate oversight value. Redesigned around agents operating within clearly defined spending authorities and risk parameters, these processes can accelerate substantially while generating more reliable audit trails than their predecessors.

Knowledge work — research synthesis, regulatory monitoring, competitive intelligence — may offer the highest returns of all. Work that once occupied teams for days can be compressed to hours, with agents that update their outputs continuously rather than producing reports that are out of date by the time they are read.

The trust question

None of this is possible without governance. Autonomous agents require clear audit trails, well-defined escalation paths and override mechanisms that are simple to invoke. The organizations that will move fastest are not those that grant the broadest autonomy on day one, but those that build the frameworks which make expanding autonomy safe — starting narrow, accumulating evidence and extending agent authority as performance is demonstrated.

It is also worth noting what becomes more valuable as agents absorb the routine. Contextual wisdom, ethical reasoning and the kind of judgment that depends on lived experience do not diminish in an agentic organization. They concentrate. The executives who think carefully now about which human contributions are genuinely irreplaceable will be better placed than those who are simply focused on what to automate.

The competitive divide that is opening up is not between firms with better AI and firms with worse AI. It is between those that are willing to redesign their operations around what the technology makes possible, and those that are using it to go slightly faster on a road that was already leading in the wrong direction. The former group is, for now, a minority. It need not remain one.