// draft section — the airport bathroom test

The Kill Switch Test

AI as the TTC stress test — builders, sprinters, values, and the correction loop

Djellil Bouzidi — Economist · Consultant · Author
TIME × TRUST × COOPERATION

Imagine an AI agent inside a bank. It does not look dangerous. It has no body, no office, no badge, no temper. It simply reconciles client records, updates systems, sends alerts, opens tickets and triggers workflows. For weeks, everyone calls it productivity. Then one day a customer file is changed, a control is bypassed, a warning is buried, and nobody can say who did it.

The first wave of AI changed the cost of producing words. The next wave changes the cost of initiating consequences. The thing in the bank is an autonomous agent, and that is the whole story: it does not merely write, summarize or suggest. It acts.

This is why AI is not primarily a technology story but an institutional one. The question is never simply what AI can do; it is what kind of institution is using it. AI does not arrive in a vacuum. It enters companies, governments, hospitals, banks and public agencies that already have habits, incentives, silences and blind spots, that already decide what counts as a problem and how quickly bad news reaches authority. Rather than abolishing those features, it amplifies them.

// AI through the TTC lens

An accelerator of institutional architecture

Through the TTC lens, AI is an accelerator of institutional architecture. It compresses Time, because decisions and actions can be produced at extraordinary speed. It disturbs Trust, because plausible output becomes cheap and the line between evidence, inference and simulation blurs. It reorganizes Cooperation, because work that once required humans to coordinate can now be delegated, fragmented or automated.

This is why the same technology produces opposite outcomes. In a Builder institution, AI becomes a maintenance tool: it surfaces weak signals earlier, preserves memory, detects anomalies and reduces repetitive work without erasing responsibility. It keeps the institution connected to reality. In a Sprinter institution, the same AI becomes an extraction tool: it accelerates throughput, compresses headcount and improves dashboards before the next reporting cycle. The gains can be real. The danger is that

speed rises faster than correction capacity, and the institution becomes more fluent, more efficient and more fragile at once.

The hybrid case is the most common. Most institutions will sincerely believe they are adopting AI like Builders, and will speak of productivity, quality and service. But if managers are rewarded mainly for cost, speed and quarterly delivery, AI will bend toward those objectives, whatever the official language says.

This is where values cease to be decoration. In the age of AI, values become operating architecture. An organization that values contradiction will design review loops. One that values responsibility will preserve human ownership. One that values truth will protect audit trails. One that values only appearances will automate appearances. AI has no independent institutional conscience. It industrializes the preferences of the organization deploying it.

// the stress test

Governing action faster than deliberation

That is why agentic AI is a near-perfect TTC stress test. It reveals, with unusual clarity, whether an institution still knows how to govern action when action becomes faster than deliberation.

A Builder does not ask only whether an agent can perform a task. It asks whether the task remains visible, auditable and reversible. Who gave the agent authority? What data did it access? Which tool did it call? Which exception did it ignore? Who can challenge the result? Who can stop the workflow? Who owns the consequence?

A Sprinter asks a narrower question: how much faster can we move, how many people can we replace, how many workflows can we automate before the next cycle? The answer may be impressive. Costs fall, output rises, dashboards improve. But the institution can quietly lose the ability to see what is happening inside its own processes.

This is the hidden danger of autonomous agents: action without proportionate correction. In the old world, a bad memo could mislead a manager. In the agentic world, a bad instruction moves through connected systems, updates records, contacts customers, changes code or executes transactions before anyone has noticed the original error. The bank agent was dangerous not because it was intelligent, but because no one could see, stop or trace what it had set in motion.

The most important line in the AI governance debate may therefore not be about intelligence, alignment or productivity. It may be about the kill switch.

> **// THE KILL SWITCH TEST**
>
> A resilient institution is not the one whose agents never fail. That institution does not exist. A resilient institution is the one that knows what its agents are doing, sees when they fail, contains the blast radius, and can stop them before the error becomes institutional behavior.

This is the Airport Bathroom Test applied to information infrastructure. Do not look first at the demo. Look at the audit trail. Look at the permission map. Look at the exception log. Look at incident response. Look at whether an autonomous agent has an identity, a scope, an owner and a stop mechanism. The quality of an institution is revealed not by the elegance of the interface, but by the quality of its correction loop.

// TTC in numerical form

Deploying action faster than correction

By 2026, this had stopped being a theoretical concern, and the numbers had begun to describe the same institutional pattern. Gartner forecast that by 2027, 40 percent of enterprises would demote or decommission autonomous AI agents, because governance gaps would surface only after production incidents. The striking point is not that agents sometimes fail — all systems fail — but that many organizations would discover the weakness of their correction mechanisms only after autonomous action had already entered production.

Other surveys point the same way. Mayfield's 2026 survey of 266 enterprise CXOs found that 42 percent already had agentic AI in production, and 72 percent were either in production or active pilots, while 60 percent reported early-stage or no formal AI governance. Deloitte found a matching gap: nearly three-quarters of companies expected to deploy agentic AI within two years, but only 21 percent reported a mature model for governing agents.

This is TTC in numerical form. Time is accelerating, as agents move from pilots into real workflows. Trust is lagging, as auditability, visibility and identity controls remain incomplete. Cooperation is under strain, because business units, IT, security teams and executives do not always share the same map of what agents exist, what they can reach, and who owns their actions.

The convergence has moved beyond surveys into corporate strategy. In May 2026, Capgemini devoted its Capital Markets Day to the agentic enterprise, and listed among its value levers the dashboards that steer autonomous agents and keep a precise trace of every action they take. The detail deserves a pause: when a global systems integrator prices the audit trail into a three-year financial plan, correction capacity has stopped being a compliance cost and become something clients will pay for. That is progress. But a value lever can be sold without being used, and in the logic of this

book the audit trail is something stricter than a selling point: not the upside, but the floor beneath which the upside is fiction.

The most revealing figure may be the simplest. In one 2026 enterprise survey, commissioned by a vendor and therefore best read as directional rather than definitive, 35 percent of respondents admitted they could not immediately stop a rogue AI agent. The number matters less than the architecture it exposes. A system that cannot halt its own automated actor has done more than adopt AI: it has created institutional action without proportionate institutional control.

The numbers do not prove that AI agents are dangerous. They prove something more interesting: institutions are deploying action faster than they are deploying correction.

Far from being a technical footnote, this is the core institutional question. When a human employee acts, organizations rely on roles, supervision, professional norms and accountability. When a software agent acts, those same functions must be rebuilt in technical form: identity, permissioning, logging, escalation, rollback and ownership. Otherwise the institution has not automated a process. It has automated a blind spot.

// maintenance vs extraction

Builder AI, Sprinter AI

Seen through TTC, the difference between Builder AI and Sprinter AI becomes brutally simple. Builder AI increases the institution's capacity to maintain itself. Sprinter AI increases the institution's capacity to extract from itself. Builder AI gives humans better evidence. Sprinter AI gives managers faster output. Builder AI strengthens the correction mechanism. Sprinter AI accelerates the system beyond its correction mechanism.

This also explains why the AI debate cannot be reduced to regulation versus innovation. The real divide is maintenance versus extraction. Whether we should use AI is already settled: we will. The real question is whether we use it to increase the system's capacity to detect, verify and correct reality, or to run faster on a thinner institutional floor.

The lesson is uncomfortable because it returns responsibility to the institution. A bad AI deployment is rarely just a technical failure; far more often, it is a governance failure that has found a new instrument. The agent did not invent the missing audit trail. It exposed it. The agent did not invent the confused accountability map. It moved through it. The agent did not invent the organization's impatience. It accelerated it.

This is why AI agents may become the defining institutional test of the next decade. Not because they are conscious. Not because they are magical. But because they act. Once AI moves from producing words to initiating consequences, the central question becomes simple: can the institution still see, stop and correct what it has set in motion?

That is the Kill Switch Test.

The more autonomous the agent, the more important the correction mechanism becomes. This is the institutional law of agentic AI. If AI increases speed faster than it increases correction capacity, it is not creating resilience. It is creating hidden fragility.

In the end, what matters is not whether AI agents are intelligent, but whether the institutions that deploy them remain wise enough to stop, explain and correct them. A Builder will use agents to maintain reality. A Sprinter will use them to outrun reality. That difference may define the next twenty years.

// THE FIVE QUESTIONS BEFORE AN AGENT ACTS

The first three are the TTC variables. The last two — reversibility and accountability — are the tests that agentic AI makes unavoidable, because an agent does not advise: it acts.

01 — TIME
Does the agent create time for judgment, or compress time for throughput?

02 — TRUST
Does it leave an audit trail, or produce output faster than it can be verified?

03 — COOPERATION
Does it clarify roles, or diffuse responsibility?

04 — REVERSIBILITY
Can the error be contained and rolled back?

05 — ACCOUNTABILITY
Who owns the consequence?

// data note

Gartner forecast (May 2026): by 2027, 40% of enterprises would demote or decommission autonomous AI agents because of governance gaps surfaced only after production incidents. Mayfield 2026 survey of 266 enterprise CXOs (Fortune 50 to Global 2000 and high-growth firms): 42% with agentic AI in production, 72% in production or active pilots, 60% reporting early-stage or no formal AI governance. Deloitte, State of AI in the Enterprise 2026 (3,235 senior leaders, 24 countries): nearly three-quarters expected to deploy agentic AI within two years; only 21% reported a mature agent-governance model. WRITER 2026 survey (vendor-commissioned, directional): 35% of organizations could not immediately stop a rogue agent. Capgemini Capital Markets Day (27 May 2026): agentic AI named as driver of the 2025–2028 plan; agent-steering dashboards listed among value levers. Full methodology and technical sources in the companion fact-checking dossier.

→ Run the diagnostic · airportbathroomtest.com

Draft section for the forthcoming book "The Airport Bathroom Test" — TIME × TRUST × COOPERATION