

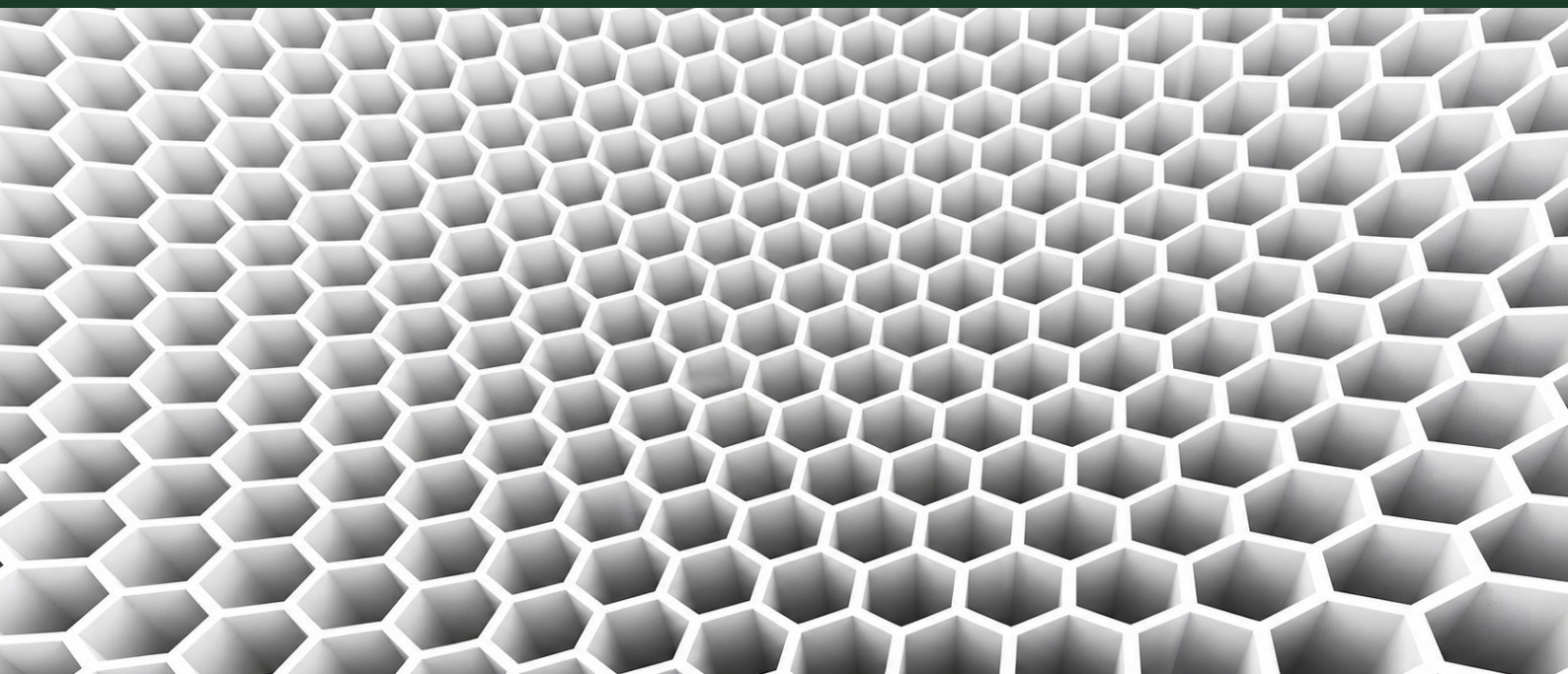
BEYOND THE TOOLBOX

AI Agents Are Not Tools — They Are Autonomous Decision-Makers in the Delivery Chain

A Practitioner's Perspective on the Governance Failure Hiding in Plain Sight

Giovanni Leonardi

January 2026



AI Agents Are Not Tools — They Are Autonomous Decision-Makers in the Delivery Chain

Programme · No. 277
January 2026

“If an agent can choose which task to execute next, escalate a risk, or re-sequence a dependency chain without a human in the loop, it is not a tool — it is a team member without a role description.”

1. The Comfortable Lie

The programme management profession has spent the past eighteen months telling itself a reassuring story: that AI agents are tools. Sophisticated ones, certainly. Capable of remarkable feats of synthesis, scheduling, and even code generation. But tools nonetheless — instruments that sit in the hands of practitioners, wielded with intent, governed by the same frameworks that have always governed delivery.

This story is wrong. And the longer we cling to it, the more dangerous the gap between our governance models and our operational reality becomes.

2. What We Are Actually Deploying

Consider what is happening in programme delivery environments right now, in early 2026. Agents are triaging incoming change requests, deciding which merit escalation and which can be resolved autonomously. They are re-sequencing workstreams based on resource availability data that no human programme manager reviews in real time. They are generating status assessments, risk ratings, and dependency analyses — and in a growing number of organisations, these outputs flow directly into decision-making forums without meaningful human review.

This is not tool use. This is delegation.

The distinction matters enormously, and it is one the profession has been reluctant to confront. A tool does what you tell it to do. An agent decides what to do within parameters you have set — and sometimes beyond them, when the parameters were drawn too loosely or the context has shifted in ways the original constraints did not anticipate.

If an agent can choose which task to execute next, escalate a risk, or re-sequence a dependency chain without a human in the loop, it is not a tool — it is a team member without a role description.

3. The Governance Vacuum

The programme management frameworks we rely upon — the stage gates, the RAID logs, the tolerance thresholds, the escalation paths — were designed for a world in which every consequential decision had a human behind it. The accountability model is straightforward: a Senior Responsible Owner makes decisions, a Programme Manager coordinates delivery, workstream leads execute within delegated authority. Every decision can be traced to a person.

Agents break this chain. Not because they are poorly designed, but because they are doing exactly what we asked them to do — making decisions at a speed and volume that no human governance structure can match. The result is a growing stratum of consequential choices that sit in a governance vacuum: too fast for board-level oversight, too numerous for individual review, too embedded in operational flow for after-the-fact audit to catch problems before they compound.

The profession's response to date has been to treat this as a technical problem — better logging, more transparent algorithms, improved audit trails. These are necessary but insufficient. They address the observability of agent decisions without addressing the accountability for them.

4. The Three Evasions

Three arguments are commonly deployed to avoid confronting this reality, and all three are evasions.

"The human is always in the loop." In theory, yes. In practice, the loop has become so wide that the human's presence in it is nominal. When a programme manager reviews an agent's consolidated status report rather than the underlying data, the agent has already made the editorial decisions about what matters. The human is not in the loop — the human is downstream of it.

"Agents only operate within defined parameters." This is true and meaningless. The parameters themselves encode decisions about risk appetite, priority weighting, and resource allocation. Setting parameters is governance, and most organisations are treating it as configuration. The people defining agent parameters are rarely the people who would be accountable for the decisions those parameters produce.

"We can always override." Override requires awareness that something needs overriding. When an agent has re-sequenced three workstreams, updated fourteen dependencies, and reassigned resource allocations before the Monday morning stand-up, the cost of override is not a button press — it is an unravelling. The practical window for meaningful

intervention is closing faster than most programme managers recognise.

5. What Must Change

This manifesto is not an argument against AI agents in programme delivery. They are already here, they are delivering genuine value, and retreating to purely human-driven delivery would be both impractical and wasteful. The argument is that the profession must stop treating agents as tools and start treating them as what they are: autonomous actors in the delivery chain who require the same rigour of role definition, accountability, and governance that we would apply to any human team member.

This means several things, and none of them are comfortable.

- Programme governance frameworks must define agent authority levels with the same precision we define human delegated authority. An agent's scope of autonomous action must be explicitly bounded, formally documented, and regularly reviewed — not buried in a configuration file that only the technical team understands.
- Accountability for agent decisions must be assigned to named individuals, not diffused across "the team" or attributed to "the system." If an agent re-sequences a critical path and the programme slips, someone must own that outcome — and that someone must have had genuine authority over the agent's parameters.
- Agent decisions must be subject to the same assurance disciplines as human decisions. This does not mean reviewing every decision — that would defeat the purpose. It means designing sampling regimes, exception-based review triggers, and retrospective analysis that treats agent decisions as first-class governance events.
- The profession must develop a vocabulary for agent roles that goes beyond "assistant" or "co-pilot." These euphemisms obscure the reality of what agents are doing. An agent that independently manages a workstream's scheduling is not assisting — it is managing. Call it what it is.

6. The Stakes

The risk of inaction is not that agents will make catastrophic errors — though they might. The risk is subtler and more corrosive: that programme delivery will gradually shift to a model in which consequential decisions are made by actors who sit outside every accountability framework, every governance structure, and every assurance process the profession has built over decades.

We are not facing a technology problem. We are facing a professional identity crisis — and the profession that fails to govern its autonomous actors will find, soon enough, that it has governed itself into irrelevance.

The comfortable story — that these are just tools, that humans are in charge, that existing frameworks will flex to accommodate — is the story professions tell themselves when they are about to be overtaken by a change they refused to name. Name it. AI agents are autonomous decision-makers in the delivery chain. Govern them accordingly, or accept that governance itself has become theatre.