DELINEA & TURNKEY X CHANGE MAKERS

WHO'S GOT THE KEYS TO THE KILL SWITCH?

AI Adoption, Non-Human Identities, and the Governance Gap Nobody's Closed

LONDON | JULY 2026

changemakers.

Delinea

TURNKEY

01 - THE STATE OF PLAY

Organisations know the risk. Almost none can control it.

Adoption maturity varies wildly by sector, but the gap isn't really about appetite for AI, it's about how much loss of control each organisation is willing to accept. One guest rated their own organisation's AI adoption a mere 2-3 out of 10, and the reason wasn't a lack of ambition: it was that nobody could yet guarantee what an agent might do with a record once it had access to one, and in that environment, an uncontrolled action isn't an inconvenience, it's a safety incident.

Set against that is a “move fast and break things” culture common elsewhere in the room. Tellingly, even this camp's favourite tactic is a control decision dressed up as a technology choice: running task-specific, open-weight models locally keeps the data, the credentials, and the compute inside a perimeter the organisation actually owns, rather than handing control of all three to a foreign-owned frontier model. One participant framed Europe's regulatory stance, the “Brussels effect”, as a way to compete on rules where it can't compete on innovation, but the rules in question are, again, mostly about who retains control of the system and its outputs.

Both ends of that spectrum share the same underlying problem, and it's not blindness. Even sanctioned, fully visible AI tools are typically approved once and then left running indefinitely, with nobody revisiting what they can still reach. Shadow AI, employees quietly using personal accounts and free services outside any approved path, is just the extreme version of a much more common failure: control is granted at the start and never checked again.

820%

The annual growth rate of non-human identities inside organisations, according to Forrester and most of that growth carries access nobody has re-examined since the day it was switched on.

> *“I have no idea how to solve this problem. I can't even see it.”*
>
> Roundtable Participant

10-SECOND TAKEAWAY

Seeing an agent tells you it exists. Control means knowing exactly what it can touch, for how long, and who can shut it off, at any point after day one, not just at approval.

02- WHERE IT BREAKS DOWN

Five patterns undermining AI control

Strip away the sector differences and the same handful of structural problems came up again and again, almost all of them failures to control an agent's access after the fact, not failures to spot it in the first place.

1 Shadow AI

Employees don't route around locked-down official tools because they want to hide. They do it because the sanctioned path offers no usable middle ground between full access and none, so an all-or-nothing control policy is exactly what pushes people outside it. The result is the same either way: unapproved tools quietly training external models on sensitive corporate data, outside any control at all.

2 The Control Vacuum

“Organisations can't see all of their AI agents,” and as one host put it, “what you can't see, you can't govern.” True, but seeing an agent doesn't control it. An engineer can spin up a bot with administrative credentials able to touch multiple systems, entirely unchecked, precisely because nothing enforces what it's actually allowed to do once it's running, discovery would only confirm a problem that was already unmanaged.

3 Regulatory Fragmentation

Every regulatory regime in the room was really asking the same question in different words: who is accountable for this system's actions, and can you prove it? NHS England pointed to the absence of a single AI act for health and care; financial services representatives described the same gap from a different angle, FCA rules, GDPR, and contractual liability clauses that were all written with a human, not an agent, as the accountable party.

4 The Board-Level Stalemate

Boards aren't only weighing savings against the risk of fines, contractual liability and reputational damage, they're being asked to approve something nobody can bound. Without a way to intervene mid-flight, narrow an agent's access, pause it, cut it off, the exposure stays open-ended no matter how good the guardrail looks on paper. That's a control gap dressed up as a financial one, and it's what stalls adoption into perpetual piloting.

5 Saying No Doesn’t Work

A blanket ‘no’ is itself the crudest possible control setting: easy to write down, impossible to enforce, and it leaves people exactly two options; comply, or go around it entirely. Employees choose the second, which leaves organisations exposed and aware of the risk, but with no real lever to manage it.

03- WHAT'S WORKING

Five moves the room is actually taking

None of these are silver bullets, but each one adds a control point somewhere that previously had none, a place where an agent's access can actually be checked, narrowed, or pulled, rather than just observed.

1 Fight AI With AI

One participant from a data driven organisation described a security strategy that has shifted toward using AI itself to detect and counter AI-driven threats. Agents act faster than a human reviewer can keep up with, so the control layer has to operate at machine speed too, otherwise it's not a control, just a report written after the damage is done.

2 Match the Model to the Task

Rather than defaulting to the biggest frontier model, choose task-specific models, like the open-weight Gemma, to balance capability against cost and geopolitical dependency. A narrower model given a narrower job also has a narrower blast radius if it's compromised or misused.

3 3 Scope Access to the Task, Not the Person

Most agents inherit the same standing access as whoever deployed them, which means an agent that oversteps can reach far more than it should. The organisations further ahead are moving away from standing credentials altogether: access is injected just for the task, scoped to what that task actually needs, and revoked automatically the moment it ends. Every action is tied to a specific identity, whether that's the person directing the agent or the service account it runs under, so there's a record to point to rather than a shared credential nobody can trace.

4 Run Models Locally Where It Matters

Open-weight models run locally ease data-sovereignty and privacy concerns, but the sharper reason is control of the perimeter itself: the data and the compute never leave the building, and the organisation isn't relying on a foreign-owned frontier model that could, in principle, be switched off by a political decision it has no say in.

5 Converge on a Shared Governance Framework

Finding an agent is only the opening move; the four questions below aren't a checklist for spotting agents, they're a lifecycle for controlling one from the moment it's approved to the moment its access is pulled. Regardless of sector or maturity level, the room converged on this as the minimum starting point.

4 QUESTIONS

Who approved its access? What does it have access to? What did it do last Tuesday? And who can hit the kill switch? The room agreed real governance starts with answering these four for every agent.

FINAL THOUGHT

The pantomime baddie isn't AI. It's not being able to stop it.

Every thread in this conversation led back to the same place. Boards aren't struggling with AI itself, they're struggling to approve something they have no way to intervene in once it's running, no dial to turn down, no lever to pause it, no guaranteed way to cut it off before the damage is done. That's what makes the risk feel unbounded, even when, on paper, it may not be. Refusing outright doesn't fix this either: a CISO's instinct to say no is understandable, but a flat block is a control with only one setting, and users will simply route around it.

What actually moves things forward is unglamorous: scope every agent's access to the task in front of it, ask the four questions below of every agent, and choose models deliberately rather than defaulting to the biggest one available. Tellingly, the room left one of those four questions deliberately open, who, exactly, holds the authority to hit the kill switch, as a live action item rather than a settled answer. Until that question has a name attached to it, control is theoretical.

> *“What you can't see, you can't govern.”*
>
> Consensus View, Delinea London Roundtable

Want to read more?

Read more of our reports on the Change Makers website and see what future events we have coming up.

Further Insights →

CHANGE MAKERS

Change Makers is a global community of senior executives driving long-lasting change within large, complex organisations. Members bench-test their thinking, share peer review, and discuss the opportunities and challenges of leading through transformation.

CHANGEMAKERSCLUB.COM

The insights in this report were drawn from a private roundtable of senior infosec, technology & risk executives and industry leaders, facilitated by James Harris, Founder of Change Makers.

FACILITATED BY

James Harris

Founder, Change Makers Club

In partnership with Delinea, Turnkey Consulting London July 2026

CONTRIBUTORS

Prabha Vijayakumar - National CIO - NHS England
Michael Stout - CISO - Need-To-Know
Victor Asanmi - Senior IAM Lead - Experian
Andreai Silva - AI Governance Lead for Data & Analytics - HSBC
Nithin Ramachandra - Global Head of Security Portfolio - Vodafone
Vitalii Chetvertakov - Senior AI & Automation Engineer - WTW
Adam Reckviczky - VP Cyber Hygiene EMEA & UKI - Deutsche Bank
Vernie Balasubramaniam - Privacy & AI Director - BDO
Marouf Malik - AI Lab Leader - RSM
Dan Goode - Regional Sales Director, UK - Delinea
Mark Lillywhite - Enterprise Sales Engineer, Delinea
Rob Lister - Associate Director, IAM - Turnkey Consulting

changemakers.

Delinea
TURNKEY