

The platform governs the agent. Who governs the work?

Databricks has built serious governance for AI agents into the platform. Permissions, guardrails, traces, cost controls. It is real infrastructure and it solves real problems. It also answers only half of the governance question, and the half it leaves open is the half that decides what ships.

What the platform now covers

In April 2026 Databricks folded its AI Gateway into Unity Catalog as Unity AI Gateway. The same permissions model that has governed enterprise data now extends to agents, models, MCP servers and coding assistants. Agents act on behalf of named users. Every model call and tool invocation is traced. Guardrails filter inputs and outputs at runtime. Cost is attributed to the token, the user and the endpoint.

The design philosophy is explicit: you cannot govern an agent by reviewing what it might do, so you govern what it can access and you monitor what it actually did. Within its layer, that logic is sound. Access governance and runtime governance are necessary conditions for safe agentic delivery, and a platform that gives you both out of the box is ahead of most of the market.

The layer it does not touch

None of this governs the delivery itself. The platform can tell you that an agent, acting under a named identity, read three tables and wrote a pipeline at 14:32 on Tuesday. It cannot tell you whether anyone asked for that pipeline, whether the work stayed inside the ask, whether a human reviewed it with substance or approved it because the checks were green, or whether the constraints your organisation claims to hold actually shaped what was built.

Access governance answers: what could the agent touch, and what did it touch. Delivery governance answers: how did intent become a shipped artefact, and who stands behind it. The first is a platform property. The second is an operating model property. Buying the first does not give you the second, and the better the first gets, the easier it becomes to believe it has.

THE GOVERNANCE THEATRE RISK

A mature platform makes documented governance nearly free. Permissions exist, traces are captured, audit logs are complete. That is Level 4 on the Cognitive Delivery Index: the control is in place and the evidence is on file. The Level 5 question is whether any of it demonstrably changes behaviour. A control that produces evidence but never changes a decision is theatre, and a platform this good at producing evidence makes the gap between the two harder to see, not easier.

Six questions the platform cannot answer for you

The Cognitive Delivery Index measures governance health across six dimensions. Each is scored on five levels, from Absent through Informal, Documented and Complete to Substantive. The table below maps each dimension against a Databricks estate: what Unity Catalog and Unity AI Gateway already give you, and what remains open regardless of how well they are configured.

CDI DIMENSION	WHAT THE PLATFORM GIVES YOU	WHAT REMAINS OPEN
Provenance Integrity	Lineage from output to source data. Audit logs naming the identity behind each action.	Whether pipelines, notebooks and models carry a record of who or what produced them, with which model, against which piece of work.
Supervisory Clarity	On-behalf-of access binds agent actions to a named user.	Whether anyone is accountable for the agent's output, who reviews it, and where escalation goes when it is wrong.
Scope Definition	Fine-grained permissions on what an agent can reach.	What the agent was asked to do, and whether the work it produced stayed inside that ask.
Review Substance	A trace of every model call and tool invocation, ready for inspection.	Whether anyone inspects it with substance. Reviewing access is not reviewing work. Green checks are not review.
Constraint Fidelity	Runtime guardrails on model inputs and outputs.	Whether your stated delivery constraints, standards, patterns and data contracts demonstrably shaped what was built.
Capability Calibration	Model routing, fallbacks and rate limits across providers.	Whether the task was deliberately matched to the capability, and reviewed in proportion to the risk of the mismatch.

Read the table column by column and a pattern appears. The middle column is evidence. The right column is judgement. The platform manufactures evidence at industrial scale. Judgement still has to be designed, assigned and practised, and no amount of evidence substitutes for it.

Absent · Informal · Documented · Complete · · · Substantive

The distance between Complete and Substantive is the governance theatre gap. On a well-configured Databricks estate, most dimensions reach Complete by default. None reach Substantive by default. That last step is organisational, not technical.

WHAT TO DO WITH THIS

Pick one agent-built artefact from the last month. A pipeline, a notebook, a model. Ask the six questions above of it. Write down which ones nobody can answer.

Find one control that has never changed a decision. A guardrail nobody has tuned, a trace nobody has read, an approval that has never been withheld. That control is at Level 4. Decide whether it earns Level 5 or gets retired.

Separate the two budgets. Platform governance is bought. Delivery governance is built. If your governance plan has a line for the first and nothing for the second, the plan is half a plan.

Jason Cruz Falzon · creator of the Cognitive Delivery Framework and the Cognitive Delivery Index
jason@cognitivedelivery.co.uk · cognitivedelivery.co.uk · cognitivedelivery.substack.com