BUSINESSGPS

THE CAPABILITY BRIEF

The Capability Brief

On 16 November Gartner reported that 95% of 298 infrastructure and operations leaders say their organisation had an agent-related incident in 12 months, and only 15% have automatic circuit breakers. Before an AI agent goes live, put a stop row on its rule sheet

Week of November 16, 2026

PUBLISHED

NOVEMBER 23, 2026

The two-minute version: decide how the agent stops before it starts

The scene. An AI agent is coming to your programme. Before it goes live, write down 3 things. What halts it. Who is told when it halts. Who decides that it runs again.

An AI agent here means software that takes actions on its own, not only answers. It might update a record, send a message or raise an order without a person pressing a button each time.

A rule sheet is the one page you keep for one such agent before go-live. It names the rule the agent takes over, what the agent does and who owns it. It holds the evidence that it works, the way back to the old process and its running cost. This edition adds a row for the stop.

In this edition. On 16 November 2026 Gartner published a survey of 298 infrastructure and operations leaders in the US and the UK. In it, 95% say their organisation faced at least one agent-related incident in the past 12 months. Only 15% have automatic circuit breakers in place.

What changed on the rule sheet:

- New. A stop row: what halts the agent, who is told and who decides it runs again.
- Escalating. The owner row. The owner now needs a named deputy, because a halt can come at any hour.
- Holds. The other rows: the rule replaced, the action, the evidence, the way back and the cost.

Bottom line: give the agent its own identity on an inventory, then write its stop row and test it once, before go-live.

Gartner finds most leaders trust their agents, and few can stop them automatically

Gartner's release of 16 November 2026 reports these findings from its survey:

- 95% of the 298 leaders say their organisation has faced at least one agent-related incident in the past 12 months.
- 70% of the 298 leaders believe their AI agents currently operate "within safe boundaries". Gartner says that belief shows "a massive operational readiness gap".
- 57% of the 298 leaders said their current governance is weak. 55% of the 298 said existing IT risk tools are "completely outdated".
- Only 15% of the 298 have "automatic circuit breakers in place to stop runaway agent errors or infinite logic loops".

The survey ran in June and July 2026. It asked leaders who run infrastructure and operations: the people who keep an organisation's systems running. It did not ask business owners. Every figure is what those leaders report about their own organisation.

The release does not define “agent-related incident”. So the 95% rests on whatever each leader counted as one.

Gartner sells research and advice to the leaders it surveys, and its full reports are for its clients. That is its interest in this finding.

A circuit breaker, in plain terms, is an automatic switch. It halts an agent when a set limit is crossed, as a breaker in a fuse box cuts the power. An infinite logic loop is an agent repeating the same step without end: changing a record, then changing it back, hundreds of times.

The gap between 70% and 15% is where the practitioner starts. Most leaders believe their agents are safe. Only 15% have an automatic way to stop one.

Put the agent on an inventory, with an identity of its own

Gartner’s release gives one action plainly. Leaders should “inventory active AI agents and classify them all as nonhuman identities in their identity platforms.”

In practice, that means the agent gets its own login, not a person’s. Ask your identity team to create one for it. Then add the agent to whatever list your organisation keeps of systems that can change data.

An agent that signs in as a person cannot be stopped without stopping that person. Its actions also read as theirs in every log. An agent with its own identity can be switched off on its own, and its trail is its own.

Gartner’s release lists 7 pillars for governing agents. Identity is the first. Guardrails, observability, escalation and containment are 4 of the others. The moves below draw on those 5.

Done when: the agent has its own identity, separate from any person’s, and appears on the inventory with its owner’s name.

Write down what halts it, automatically and by hand

The stop row starts with the triggers. List the limits that halt the agent without anyone deciding. Then name who can halt it by hand.

Good automatic limits are ones you can count:

- the number of actions in an hour, set above a normal busy hour
- the same record changed twice in a short window, which catches a loop
- an error rate above a level you set from testing
- any action outside the agent’s stated scope, such as a record type it should never touch

Ask whoever builds the agent which of these the system can enforce today. Write down the ones it can. Write down any it cannot, so the gap is on the sheet. If the system can enforce none of them, tell whoever owns the go-live decision before the date, and write their answer on the sheet.

The manual stop matters as much. Name the roles who may press it and say where the control is. A stop only the builder can find is a stop nobody will use at 2am.

Done when: the row lists each automatic limit with its number and names who can halt the agent by hand, and how.

Name who is told, and how quickly

A halt nobody hears about is a second incident waiting. Write down who is told when the agent stops, by what route and within what time.

Usually that is 3 people. The owner of the agent hears first. The owner of the process the agent supports hears next, because their work is now waiting. The service desk hears too, so that calls from users reach someone who knows.

Set the time in minutes, not “promptly”. Then check that the alert really goes to those people, and not to a shared inbox nobody reads at night.

If you work for the integrator, agree the alert route with the client’s owner. The people told should be the client’s, with your team copied in while support lasts.

Done when: the row names who is told, by what route and within how many minutes, and one test alert has reached each of them.

Name who decides the agent runs again

Stopping is the easy half. The harder half is the restart. Write down who decides that the agent runs again, and what they must see first.

That person should not be the builder alone. Choose someone who answers for the process: the business owner or their deputy. Give them a short list to check before restart:

- the log of what the agent did in the hour before it stopped
- what caused the halt, in one sentence
- what was changed, if anything, and who changed it
- any records the agent touched that need putting right

If the cause is unknown, the agent stays stopped. Write that on the sheet as the rule, so nobody has to argue it at the time.

Done when: the row names the person who decides the restart, their deputy and the 4 things they check first.

Test the stop once, before go-live

In testing, trigger each automatic limit on purpose. Then press the manual stop. Time how long the alert takes to arrive, and follow the restart list as written.

If your test environment has no alert route, send one test alert from the live route before go-live. If you cannot, write that on the sheet.

Record what happened. If a limit failed to halt the agent, fix it or remove it from the row. A limit written down but never tested gives false comfort.

Repeat the test after any change to the agent, and at least once a quarter while it runs.

Done when: each limit and the manual stop have been triggered once in testing, with the date and result written on the sheet.

In an invented example, a loop is caught in 10 minutes

Illustrative, and invented: an AI agent tidies supplier records. It merges duplicates and fills missing fields from invoices. It handles about 300 records a day.

> Halts when: more than 80 records change in one hour; the same record changes twice in 10 minutes; any bank detail is touched. Manual stop: the data team lead or the service desk, from the agent console. Told: the agent's owner, the purchasing manager and the service desk, by alert, within 10 minutes. Runs again when: the purchasing manager has read the last hour's log and the cause is named.

One morning the agent starts merging 2 suppliers with similar names, then splitting them, then merging them again. The second change to the same record within 10 minutes trips the limit. The agent halts after 2 passes, not 2,000.

The purchasing manager reads the log, sees the cause and asks for a rule that leaves those 2 suppliers alone. The agent runs again before lunch.

Set your own limits from your agent's normal day.

What to watch

Gartner's next published survey of infrastructure and operations leaders on AI agent governance, if released by 30 June 2027, will show more than 15% with automatic circuit breakers in place.

Sources

1. Gartner, press release on AI governance and infrastructure and operations, London, 16 November 2026, read on 22 November 2026. https://www.gartner.com/en/newsroom/press-releases/2026-11-16-gartner-weak-artificial-intelligence-governance-will-lead-to-business-critical-service-disruptions-for-40-percent-of-infrastructure-and-operations-organizations-by-2028

BusinessGPS publishes capability.ai.