BUSINESSGPS

THE CAPABILITY BRIEF

The Capability Brief

On 3 November CrowdStrike introduced an AI Kill Switch: an administrator can cut an AI agent off from its model. Before your agent acts, decide who stops it and what happens to its half-done work

Week of November 2, 2026

PUBLISHED

NOVEMBER 9, 2026

What this week means for you: the two-minute version

The scene. Before an AI agent acts in your programme, decide who may stop it, how fast, and what happens to the work it leaves half done. Write those answers as one row on the agent's rule sheet.

A rule sheet is one page for one AI action. An agent that takes several actions has a sheet for each, and its stop row is the same on all of them. In plain words, it says how the work now runs: the old rule, the new action, its owner, the evidence that shows it helps and the way back. It is written for the person who must run the work tomorrow, on the client's side or the integrator's.

In this edition. On 3 November 2026 CrowdStrike, the cybersecurity company, introduced an AI Kill Switch in its Falcon Guardian product. CrowdStrike says an authorised administrator triggers it, and every action is logged. Today it works at the AI gateway, where an agent reaches the AI model it runs on, and blocks the agent's access to that model.

What changed in the rule sheet:

- New. A stop row. A security vendor now offers a named control that cuts an agent off. The sheet needs a row that says who may use one, and when.
- Escalating. The way-back row. Going back to the old rule is one decision. Stopping the agent in the middle of its work is another, and it comes first.
- Holds. The other rows: old rule, new action, owner and evidence. CrowdStrike's announcement reports no measured customer result.

The bottom line: a switch answers "can it be stopped?", and your sheet has to answer "who stops it, and what do we do with what it left?"

A switch cuts the agent off, and the sheet says what happens next

CrowdStrike announced the AI Kill Switch at its Fal.Con Europe event in Barcelona. Its press release and its blog do not describe the switch in the same words. The part this Brief relies on is narrower. CrowdStrike's press release says an authorised administrator triggers it and every action is recorded in an audit trail, a log of what was done and by whom. Both pages say that today it cuts the agent off from its AI model at the gateway, where an agent reaches the model it runs on. Stopping agents directly on the computers they run on is announced as coming.

The announcement describes a control, not an outcome. The release quotes one customer, Rick Orloff, chief information security officer at Everpure, on why his team wants the switch. It reports no measured result.

A control like this settles whether an agent can be cut off. If your agent does not run through a gateway like this, the same questions still stand. They are yours to answer:

1. Who may pull it, and who must be told?
2. How quickly must it happen once someone asks?

3. What happens to the action the agent was halfway through, and to the record of what it had done?

The third question is the one that is easy to leave out. An agent that updates records does not stop between tasks. It may stop where it is.

Name who may stop the agent, and who is told

Write one role, and one deputy, who may stop this agent. Then write who must be told when it is stopped, and how.

The person with the technical access is often in IT security. The person who knows whether the agent's work is wrong is usually in the business. The row needs both: the one who decides, and the one who can act.

On the client's side, the deciding role sits with the manager whose team relies on the agent's work. If no manager will take the role, tell the person who put you in charge of this agent, and write that on the row. If you work for the integrator, draft the row for each agent you build, and let the client name the roles. Check with your delivery lead what your firm has agreed to do when an agent is stopped.

Done when: the row names a role and a deputy who may stop the agent, and who is told. Both people have read it.

Write how soon a stop must happen, and rehearse it once

Put a time on the stop, in minutes, from the moment someone asks. Then try it once before go-live, in a test system, with the people named in the row. If you have no test system, walk the named people through the stop on paper and time it.

A time nobody has tried is a guess. One rehearsal shows who has the access, who answers the phone and how long it really takes.

Done when: the row states a time in minutes, and one rehearsal has met it, or the row now says what it really takes.

Decide what happens to the work the agent left half done

For each kind of action the agent takes, write one of 3 answers. The half-done action is finished by a person. Or it is reversed. Or it is held, untouched, until a named person decides.

Then write one more line: nothing is restarted until that person has read what the agent did before the stop.

Done when: every kind of action the agent takes has one of the 3 answers, and the restart line names a person.

Keep the record of what the agent did before anyone restarts it

After a stop, the team needs one list: every item the agent touched since the last point you know was correct (a checkpoint), and what it did to each. Find out now where that list would come from. It may be the agent's own log, the system's change history or a report someone has to build.

If nobody can produce that list today, write that on the sheet. Building it comes before the agent goes live.

Done when: the sheet says where the list of touched items comes from, who produces it and how long that takes.

In an invented example, knowing which 20 invoices were half done can save about 10 hours after a stop

Illustrative, and invented: an AI agent matches supplier invoices to purchase orders and posts them, about 400 a day.

> Stop: the accounts payable manager decides; the duty IT security lead acts. Deputy: the finance systems lead. Told: the finance controller. How soon: 15 minutes from the call. Half-done work: invoices matched and not posted are held for the accounts payable manager. Record: the agent's log of every invoice it touched today, exported before any restart. Restart: only the accounts payable manager, after reading the log.

At 11am the agent is stopped. It has posted 150 invoices and matched 20 more without posting them.

Without the row, nobody knows which items are half done. A clerk checks all 170 at 4 minutes each: 680 minutes, or about 11 hours.

With the row and the log, the 150 posted invoices are on a list. If the stop was not caused by wrong postings, the clerk samples them and checks the 20 held ones in full: about 80 minutes plus the sample. If it was, the 150 must be checked too, and the row only saves the time spent finding them.

Writing the row takes about 30 minutes. Producing the log it points to, and the rehearsal, take longer, so count them before the agent goes live. Use your own figures. The sum shows the cost of not knowing what was half done. It does not show how often a stop will happen.

One question before the agent is switched on

If someone stopped this agent at 11 o'clock tomorrow, which of its actions would be half done, and who would know?

What to watch

By Thursday 31 December 2026, CrowdStrike will not have published a measured result from a named customer's use of the AI Kill Switch. That means a customer's name and a number, such as how many agents were stopped.

Sources

1. CrowdStrike, CrowdStrike Introduces the AI Kill Switch for AI Safety, press release, Fal.Con Europe, Barcelona, 3 November 2026. https://www.crowdstrike.com/en-us/press-releases/crowdstrike-introduces-ai-kill-switch-for-ai-safety/
2. CrowdStrike, blog post introducing the AI Kill Switch, 3 November 2026. https://www.crowdstrike.com/en-us/blog/crowdstrike-introduces-ai-kill-switch/

BusinessGPS publishes capability.ai.