FREE KIT · AI AGENT OPERATIONS

Stop any agent in 45 seconds.

The AI-agent kill-switch kit: one switch, breakers that trip on their own, and a monthly drill, with prompts to hand your AI and code to start from.

3:07 AM → an agent loops → the same email hits every customer, 10 times → spend spikes → you're hunting through 12 dashboards → and you're still responsible.

60% OF ORGANISATIONS HAVE NO WAY TO TERMINATE A MISBEHAVING AGENT · KITEWORKS 2026 DATA SECURITY & COMPLIANCE RISK FORECAST (225 ORGS)

How to use this: each step has the **pain** it prevents, the **fix**, a **prompt** you can hand straight to your AI coding assistant, starter code, and its honest **limit**. Then run the 10-minute audit on page 4.



Afeez Badmos · @afeez_sdbytes

SOFTWAREDEVBYTES × GLOBAL MCLIEN SOFTWARE SOLUTIONS

01 One switch for everything

PAIN At 3 AM nobody has time to find the right console among a dozen dashboards. Every minute spent hunting is another minute the agent keeps acting.

FIX One command that does three things at once: revoke every agent key, pause every schedule, and end every running agent job. Keep it in one runbook everyone on call can find and run.

> PROMPT YOUR AI “Write a one-command kill-switch runbook (a single script) that revokes every AI-agent API key, disables every scheduled agent job, and ends every running agent process. Read the list of agents from one inventory file and log what it did.”

```
#!/usr/bin/env bash
# killswitch.sh  ·  stop every AI agent in one command
set -euo pipefail
log(){ echo "$(date -u +%FT%TZ) $*" | tee -a killswitch.log; }
log "[1/3] revoking agent keys";   revoke_agent_keys   agents.yaml  # provider key API / secrets manager
log "[2/3] pausing schedules";     pause_schedules     --tag ai-agent # cron, queue, workflow tool
log "[3/3] ending running agents"; stop_running_agents --tag ai-agent # job runner, k8s, process manager
log "ALL AGENTS STOPPED"
```

LIMIT The three functions are placeholders for your own stack. Key revocation can lag for requests already in flight, so time it in a drill rather than assuming it's instant.

02 Breakers before the switch

PAIN The usual disaster isn't one bad action, it's a loop: the same email to every customer ten times overnight, or API spend climbing until the account drains, all before anyone wakes up.

FIX Add automatic breakers so the system stops itself: hard daily and monthly spend caps on every AI account, rate limits at your gateway, and an approval queue that holds anything that moves money or messages a customer until a human says yes.

> PROMPT YOUR AI "Set daily and monthly spend caps on every AI provider account, add per-agent rate limits at our API gateway, and route every payment and outbound customer message through an approval queue that a human must approve."

```
// before any agent action runs
const NEEDS_HUMAN = new Set(["payment", "refund", "outbound_message"]);
if (NEEDS_HUMAN.has(action.type))
  return approvals.enqueue({ ...action, status: "pending_human" });
if ((await spendToday(agent)) + action.estCost > CAPS[agent].daily)
  return tripBreaker(agent, "daily spend cap");   // pauses this agent, pages on-call
```

LIMIT Provider-side spend caps can be soft or delayed. Keep your own counter as well, and alert on the trip, so a breaker doesn't fail silently.

03 Clean up, review, then drill

PAIN Stopping isn't undoing. The emails are already sent, the records already changed. And a kill switch you've never pulled is a guess: you only find out it's broken during the real incident.

FIX Log every agent action with what it would take to reverse it (what it sent, changed or deleted), keep a rollback runbook, run a short postmortem after any incident, and once a month pull the switch on purpose and time how long a full stop takes.

> PROMPT YOUR AI "Log every agent action to an append-only table with a reverse_with field, write a rollback runbook from that log, and schedule a timed monthly kill-switch drill for the whole team."

```
CREATE TABLE agent_actions (
  id           bigserial PRIMARY KEY,
  agent        text NOT NULL,
  action       text NOT NULL,          -- sent_email, updated_record, deleted_file…
  target       text,
  payload      jsonb,
  reverse_with text NOT NULL,          -- how to undo it: correction email, restore
snapshot…
  created_at   timestamptz NOT NULL DEFAULT now()
);
```

LIMIT Some actions can't be reversed (a sent email, a published post). For those, the log tells you who to contact and what to correct, which is why approval for them belongs in step 2.

The 10-minute audit YES / NO

QUESTION	HOW TO CHECK	ANSWER
Can one person stop every agent with one command?	Find the runbook link from a phone. Is it one command, or a checklist of consoles?	☐ Y ☐ N
Does every AI account have a hard daily and monthly cap?	Open each provider's billing settings; note the cap and who gets the alert.	☐ Y ☐ N
Do payments and customer messages wait for a human?	Trigger a test refund and a test email from an agent. Both should land in the approval queue.	☐ Y ☐ N
Is every agent action logged with how to reverse it?	Pick yesterday's 3 most recent agent actions. Can you say how to undo each?	☐ Y ☐ N
Has the switch been pulled on purpose this month?	Check the drill log: date, who pulled it, time to full stop. Under 45 seconds?	☐ Y ☐ N

Drill log COPY MONTHLY

DATE	PULLED BY	TIME TO STOP	WHAT DIDN'T STOP

Why this can't wait

60%

of organisations have no way to terminate a misbehaving agent (Kiteworks, 2026).

Shipping

Security vendors now sell agent kill switches, e.g. CrowdStrike's AI Kill Switch (Nov 2026).

Liable

UK regulators: an agent's autonomy "does not diminish the responsibility" of the firm that deploys it.

A vendor kill switch is an emergency brake, not a rewind button: it stops further actions, but it can't un-send an email or reverse a transaction. That's why this kit pairs the switch with breakers (step 2) and an undo log (step 3).

Want it built for you?

Global McLien Software Solutions builds AI-agent controls: one-command kill switches, spend caps and approval queues, action logs and rollback runbooks, plus a first timed drill with your team.

globalmcliensoftwaresolutions.com →

Sources

Kiteworks, "Data Security and Compliance Risk: 2026 Forecast Report" (225 organisations; 60% have no way to terminate a misbehaving agent), as reported by Kiteworks, Aug 2026

CSO Online, "CrowdStrike launches AI kill switch to shut down rogue agents" (Nov 2026); Help Net Security, JetStream AI kill switch (Jul 2026); SecurityBrief, Straiker kill switch (2026)

InfoWorld, "Enterprises must take blame for their AI agents' mistakes, say regulators" (Nov 2026), on the DRCF joint statement (CMA, FCA, ICO, Ofcom)

General engineering guidance, not legal or compliance advice. Regulatory positions vary by country and change over time; check what applies to your business. Code is illustrative starter code; adapt and test it for your own stack. Product names are trademarks of their owners; this kit isn't affiliated with or endorsed by them. © 2026 Afeez Badmos / SoftwareDevBytes · Global McLien Software Solutions.