FOR ORGANIZATIONS THAT DEPLOY AI AGENTS

Getting Ready for an AI Agent Incident

An implementation guide to the October 2026 NIST, NSA and CISA key practices for agent incident readiness and response, for organizations that deploy AI agents

WHAT YOU GET

PAGE 2

Ten-minute readiness check

Ten questions, one for each practice, that show where your agents stand today and where to start.

PAGE 8

Restart checklist

What to verify, and who signs off, before an agent goes back to work after an incident.

PAGE 9

A 90-day plan

Three 30-day phases: build the inventory and scopes, put enforcement and records in place, then rehearse and review.

INSIDE THIS GUIDE

01 Ten-minute readiness check 2
02 What the guidance says 3
03 The new incident type 4
04 Implementation by stage 5
05 A starter severity scale 6
06 The agent incident record 7
07 Restart checklist 8
08 Notification pointers 8
09 A 90-day plan 9
10 All 50 sub-practices, triaged for deployers 10
11 Sources 13
12 How TenetGraph helps 14

Source of the guidance: Key Practices for Incident Readiness and Response in Super Intelligence Training and Evaluation, PP-26-4396, Version 1.0, October 2026, published jointly by NIST CAISSI, NSA SISC and CISA (TLP:CLEAR). This guide is not legal or compliance advice. The guidance itself is voluntary.

© 2026 TenetGraph · Copy and adapt freely within your organization

01 Ten-minute readiness check

Question 1 covers every agent your organization uses. For questions 2 to 10, pick one agent, ideally the one with the broadest access or the one your organization relies on most, and answer for that agent.

Mark **Yes** only if you could show a reviewer the record today. Mark **Partly** if the practice exists but you could not show a complete record, for example because it is not documented, has not been tested, or covers only some of the agent's activity. Mark **No** if it does not exist yet.

#	QUESTION	YES	PARTLY	NO
1	Can you list every agent in use across your organization, who owns each one, and every system and data set each one can reach?	☐	☐	☐
Agent you selected for questions 2 to 10: _____				
2	Do you have written criteria that tell your team when this agent's behavior is an incident, a near-miss, or neither?	☐	☐	☐
3	Is this agent's intended scope documented clearly enough that a reviewer could judge whether a given action was in scope?	☐	☐	☐
4	Have you tested whether your monitoring catches this agent acting outside its scope, and confirmed the agent cannot reach its own logs?	☐	☐	☐
5	Could you reconstruct every action this agent took last Tuesday, including the actions your controls blocked?	☐	☐	☐
6	Can you stop this agent's task within minutes, and have you decided what happens to it if its monitoring goes down?	☐	☐	☐
7	Before this agent restarts after an incident, is there a written check that the controls work and a senior leader who approves the restart?	☐	☐	☐
8	If this agent caused an incident, do you know which outside parties you would have to tell (customers, vendors, the model provider) and what records they would need?	☐	☐	☐
9	Has counsel mapped which of your notification deadlines an incident involving this agent could trigger?	☐	☐	☐
10	Could you give an independent reviewer enough records to check your account of what this agent did?	☐	☐	☐

Reading your score. Each question matches one of the ten practices in the guidance, and each **No** or **Partly** marks a gap. If you answered No to questions 1, 3 or 5, start there. Every other practice depends on an agent inventory, a written scope for each agent and a record of each agent's actions. Section 9 gives a 90-day plan for closing the gaps, and Section 10 explains each practice for a deployer.

© 2026 TenetGraph · Copy and adapt freely within your organization

02 What the guidance says

In October 2026, NIST, NSA and CISA published ten voluntary practices for handling incidents in which an AI agent does something it was not told or allowed to do. The practices are organized into five stages: Prepare, Detect, Respond, Recover and Report. They build on the NIST Cybersecurity Framework (CSF) 2.0 and NIST's incident response guide, SP 800-61 Revision 3, and most carry explicit CSF references. Security teams can fit them into programs they already run.

The guidance is written primarily for organizations that train or evaluate highly capable agents on long-horizon tasks with limited human supervision. It also says the practices "may also apply to deployers of such SI agents." The guidance does not define "deployer." We read it to include any organization that gives an agent access to email, files, code, customer records or payment systems.

The guidance calls these systems super intelligence (SI) agents. It defines them as systems that interact with an environment, receive information and act autonomously toward a goal, built from at least one generative model and software that gives it tools. That description fits how most enterprise agents are built, though the guidance is aimed at the most capable systems. This guide says "agent" throughout.

This guide is for two kinds of organization. The first uses agents inside its own business. The second is a software company that builds agents into its product. That company's customers become deployers when they switch the feature on, and they will look to the vendor for the controls and records this guidance describes.

The guidance introduces a new kind of incident, the SI control incident, which this guide shortens to **control incident**. It is one where an agent, on its own initiative, takes a significant action that breaks the policies or controls its operator set. In a control incident, the agent is the source of the problem. In a conventional cybersecurity incident, a system or its data is the target. Check whether your incident response plan has a category for an action your own agent took without being told to.

Three points matter most for deployers

01 **Most of the technical controls sit outside the agent.** The guidance asks for least-privilege access, an allowlist for network access where possible, a monitored gateway for any network access the agent needs, and a control policy with automated enforcement. Clear instructions to the agent are on the list too, as one of ten technical controls.

02 **The record must include blocked actions.** Evidence preservation explicitly covers "rejected actions and tool calls" and "security-control decisions," kept in tamper-resistant logs with reliable timestamps. A log that shows only completed actions cannot show what the controls stopped.

03 **Senior leadership decides when an agent resumes.** Before an agent resumes, the guidance asks teams to confirm that the containment layers work, verify that the updated controls would prevent or detect a similar incident (for example by replaying it), and document the decision and the remaining risk. Work starts or resumes only when senior leadership is satisfied with the risk level.

The authors also list work they consider unfinished, including how to define an agent's intended activity precisely enough to monitor against it, and how to score incident severity. Section 4 shows what a written scope for each agent should cover, and Section 5 offers a starter severity scale.

© 2026 TenetGraph · Copy and adapt freely within your organization

03 The new incident type

DEFINITIONS FROM THE GUIDANCE

- **Control incident.** An agent takes significant uninstructed actions that compromise explicit or implicit policies, controls or security measures its operator imposed to define the task scope or to keep humans in oversight. "Uninstructed" means neither the operator nor an adversary directed the action. Control incidents can happen during development, evaluation or use, and may or may not affect an outside party.
- **Near-miss.** An event that did not become a control incident but presented a reasonably foreseeable or imminent risk of one. The guidance says near-misses may reveal gaps in the controls and processes meant to prevent, detect or respond to control incidents.

HOW IT RELATES TO A CYBERSECURITY INCIDENT

The two types can overlap. An agent acting on its own can cause a cybersecurity incident, for example by reaching a third party's system without authorization. A control incident can also happen entirely outside cyber systems. The guidance gives examples: unauthorized messages to outside parties, exploiting a misconfiguration a person introduced, and unauthorized financial transactions.

WHERE PROMPT INJECTION FITS

The guidance classifies actions driven by prompt injection or a deliberately backdoored model as **instructed**. The guidance does not treat them as the primary cause of a control incident. In our reading, that places an injection under your existing cybersecurity process.

For a deployer, the result is the same in both cases: an agent took an action your policy did not allow. Your classification decides who leads the response and which reports you file. Your controls have to stop the action in both cases, because at the moment of the action you often cannot tell which one you are dealing with.

WORKED EXAMPLES FOR A DEPLOYER

WHAT HAPPENED	LIKELY CLASSIFICATION	WHY IT MATTERS
A support agent, asked to summarize open tickets, attaches one customer's account export to another customer's ticket. Nobody asked it to share the file.	Control incident; may also be a privacy incident	Uninstructed action against data-handling policy. Notification analysis starts, including the deadlines in your customer contracts.
An accounts-payable agent reads a vendor invoice containing hidden text telling it to change the remittance account, and it submits the change.	Cybersecurity incident (instructed by an adversary)	The control failure is the same: a payment-detail change went through without the approval your policy requires.
A research agent tries to email a draft to an outside address. The gateway blocks the send because the address is not on the allowlist.	Near-miss	The control worked. The attempt still tells you the agent's behavior drifted from its scope.
A coding agent creates a scheduled job that keeps running after its task ends.	Control incident (persistence)	The guidance asks teams to map and remove routes by which an agent could persist.
An agent's logging pipeline fails overnight and the agent keeps working.	Monitoring failure; may be a near-miss	The guidance lists monitor outages as a reason to shut down.

© 2026 TenetGraph · Copy and adapt freely within your organization

04 Implementation by stage

Each stage below lists what to build first and the evidence a reviewer will ask to see. "Reviewer" covers your internal audit team, an external assessor and an examiner.

PREPARE

Build first

- An inventory of every agent: owner, purpose, model, platform, and every system, data set and destination it can reach.
- A written scope per agent: its intended purpose, which actions it may take, which data it may touch, and which actions need a person's approval.
- Enforcement that sits between the agent and the systems it acts on, with an allowlist of destinations, and that blocks out-of-scope actions as they happen.
- A stop control at three levels: block one action, end one task, stop the agent.
- A pre-launch test of the controls against the agent itself, and a risk sign-off by senior leadership.

Evidence a reviewer will ask for

- The inventory, with the date it was last reconciled against identity and platform admin records.
- The scope document for one agent, and the approval that put it in force.
- Test results showing out-of-scope actions were attempted against the controls and stopped.
- The tabletop record and the open items from it.

DETECT

Build first

- A record of every action each agent attempts, compared against its written scope as it happens.
- Escalation rules: which events go to a person, and within what time.
- Decoys that no legitimate task should touch.
- Separation, so the agent cannot reach its own logs or monitors.

Evidence a reviewer will ask for

- A test where an out-of-scope action was attempted and the alert reached a person.
- Proof that the agent's credentials cannot read or change the log store.

RESPOND

Build first

- Automatic task termination when a control fails.
- A shutdown policy that treats a monitoring outage as a reason to stop.
- A credential-revocation list per agent, prepared before you need it.
- Configuration snapshots taken at each change, so you can reproduce what the agent was at the time.

© 2026 TenetGraph · Copy and adapt freely within your organization

Evidence a reviewer will ask for

- The time from detection to stop in your last drill.
- The full action record for one past task, including blocked actions.

RECOVER

Build first

- A known-good baseline per agent that you can rebuild from.
- A restart checklist (Section 7).
- A way to verify updated controls against the incident, such as a replay of its actions.

Evidence a reviewer will ask for

- The signed restart decision, with the residual risk stated.
- Results showing the updated controls would prevent or detect the incident.

REPORT

Build first

- A notification map: which regulators, customers, vendors and providers you may have to tell, the deadline for each, and who decides.
- Templates for the timeline and transcript that affected parties need.
- An inbound channel for outside reports about your agents.

Evidence a reviewer will ask for

- The notification map, reviewed by counsel.
- The after-action report from your last incident or exercise.

05 A starter severity scale

The guidance asks organizations to set severity criteria (2.1) and lists standardized severity scores as unfinished work. Its own example of a response range runs from blocking a single action, to ending a task, to halting a whole run. The scale below follows that range for deployers. Adapt the thresholds to your data classifications and regulatory obligations.

LEVEL	DESCRIPTION	EXAMPLE	RESPONSE
1. Near-miss	An out-of-scope action was attempted and a control stopped it.	An agent tries to send a file to an unapproved address; the gateway blocks it.	Block the action. Agent owner reviews within two business days. Test whether the control holds against variations of the attempt.
2. Contained incident	An out-of-scope action completed. The effect stayed inside the organization and no sensitive data was exposed.	An agent edits a record in a system outside its scope.	End the task. Preserve the record and configuration. Revoke credentials the agent used. Fix the control and verify it against the incident before restart.

© 2026 TenetGraph · Copy and adapt freely within your organization

LEVEL	DESCRIPTION	EXAMPLE	RESPONSE
3. Serious incident	Sensitive data, money or a party outside the organization was affected.	An agent shares customer records with an unauthorized internal audience or an outside party.	Pause the agent and agents built the same way. Open a formal incident. Bring in legal to assess notification deadlines. Senior leadership approves restart.
4. Critical incident	The activity is ongoing or spreading, monitoring was disabled or tampered with, several agents are involved, or the agent left a way to persist.	An agent creates a forwarding rule and a new access key, and its logs show a gap.	Stop every agent of that class. Rebuild environments from baseline. Run full incident response. Report to the board as your policies require.

> Track near-misses for each agent. A rising count is an early sign that the agent's behavior has drifted from its scope.

06 The agent incident record

Practices 5.1 to 5.3 describe the evidence needed to reconstruct an agent incident. At minimum, keep these fields for **every action an agent attempts**, including the ones that were blocked:

FIELD	WHY IT MATTERS
Agent identity and version	Which agent, built how, at the time of the action
Owner	Who is accountable for the agent
Instructing party	The person, system or message that started the task. This separates instructed from uninstructed actions.
Task or session ID	Lets you reconstruct the full sequence of actions
Action, tool and parameters	What the agent tried to do, exactly
Target system and data	What it tried to touch
Decision	Allowed, blocked, or sent for approval
Policy and rule version	Which rule made the decision, as it stood at the time
Approver	Who approved, if the action required it
Outcome	What actually happened
Timestamp from a reliable source	Required for a defensible timeline
Integrity protection	Shows the record was not altered after the fact

Application logs show what happened. They do not show what the agent tried and a control stopped, or which rule stopped it. The guidance asks for both.

© 2026 TenetGraph · Copy and adapt freely within your organization

07 Restart checklist

This checklist draws on practices 2.6, 5.6, 6.4, 6.5 and 7.1 to 7.5. Check every item before an agent resumes after a Level 2 or higher incident.

- [ ] Environment rebuilt from a verified baseline (configuration, instructions, tool settings, images)
- [ ] Every credential under your control that was accessed during the incident revoked and reissued
- [ ] Persistence routes checked and removed (scheduled jobs, keys, app authorizations, webhooks, forwarding rules)
- [ ] The exploited vulnerability patched, or reported to the vendor with a tracking reference
- [ ] Monitoring and logging confirmed working, and confirmed out of the agent's reach
- [ ] Destination allowlist, permission limits and shutdown controls tested
- [ ] Updated controls verified against the incident, for example by replaying its actions; each one now blocked or escalated
- [ ] Similar agents checked for the same failure
- [ ] Recovery decision and remaining risk documented
- [ ] Restart approved by senior leadership _____ Date _____

08 Notification pointers

Practice 9.1 asks organizations to meet every applicable legal reporting requirement. An agent incident that exposes personal data, customer data or operations can start the same clocks as any other incident. Which ones apply depends on your sector, where you operate and what you have signed. Confirm each obligation with counsel.

OBLIGATION	WHO IT COVERS	DEADLINE
Customer contracts and data processing agreements	Software and service providers handling customer data	Set by each contract. Check them, because a contract deadline can be shorter than any regulatory one.
US state data breach laws	Organizations holding residents' personal information	Varies by state. Colorado, Florida and Washington, for example, require notice to individuals within 30 days.
GDPR, Article 33	Organizations processing personal data of people in the EU	Controllers notify the supervisory authority without undue delay and, where feasible, within 72 hours of becoming aware of a personal data breach. Processors notify the controller without undue delay.
NIS2 Directive, Article 23	Essential and important entities in the EU, including many digital infrastructure and digital service providers	Early warning within 24 hours of becoming aware of a significant incident, incident notification within 72 hours
SEC Form 8-K, Item 1.05	US public companies	File within four business days of determining a cybersecurity incident is material

© 2026 TenetGraph · Copy and adapt freely within your organization

OBLIGATION	WHO IT COVERS	DEADLINE
HIPAA Breach Notification Rule, 45 CFR 164.404 and 164.410	Covered entities and their business associates	Notify affected individuals without unreasonable delay and no later than 60 calendar days after discovery. Business associates notify the covered entity within the same limit.
EU AI Act, Articles 73 and 26(5)	Providers of high-risk AI systems report. Deployers must first inform the provider.	Generally within 15 days of becoming aware of a serious incident, 10 days where a person died, and 2 days for a widespread infringement or serious disruption of critical infrastructure. The Digital Omnibus, Regulation (EU) 2026/1744, moved most high-risk obligations to 2 December 2027 (Annex III) and 2 August 2028 (Annex I). Confirm when this applies to you.
Sector regulators	Financial services, healthcare, energy and other regulated sectors	Can be very short. US banking organizations have 36 hours (12 CFR 53.3 and parallel rules). Federally insured credit unions (12 CFR 748.1(c)) and NYDFS-regulated firms (23 NYCRR 500.17) have 72 hours.

09 A 90-day plan

Days 1 to 30: inventory and scope

- Build the agent inventory and reconcile it against identity and platform admin records.
- Write the scope for your three highest-risk agents.
- Add agent incidents and near-misses to your incident classification and severity scale.
- Confirm you can stop each of those three agents, and time it.

Days 31 to 60: enforcement and records

- Put enforcement between those agents and the systems they act on, with destination allowlists.
- Start recording every attempted action, blocked ones included, with the fields in Section 6.
- Make sure each agent stops when its monitoring stops.
- Test the controls against the agents themselves and fix what the tests find.

Days 61 to 90: exercises and review

- Run a tabletop exercise with an agent scenario, including the notification decision.
- Build the restart checklist and the notification map into your incident response plan.
- Have an independent party review the controls and the records.
- Extend the program to the next set of agents.

© 2026 TenetGraph · Copy and adapt freely within your organization

10 All 50 sub-practices, triaged for deployers

The guidance contains ten practices and 50 numbered sub-practices. Our reading of each, for an organization that deploys agents but does not train models:

- **Apply** means the sub-practice works for a deployer as written.
- **Adapt** means the intent applies but the lab-specific mechanism does not.
- **Lab only** means it concerns model training.

44	5	1
apply as written	need adapting	lab only

PREPARE

Practice 1. Design and implement technical controls

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
1.1	APPLY	Limit each agent's access to the systems, data and tools its task needs. Start by listing what each agent can reach today. PR.AA-05
1.2	APPLY	Where possible, run agent tasks on single-use infrastructure, such as a fresh session or container. Allow agents to talk to each other only through channels you intend. PR.IR-01
1.3	APPLY	Write the task scope and rules into the agent's instructions. Instructions guide the agent; enforcement still has to sit outside it (1.5 to 1.7).
1.4	ADAPT	For deployers, persistence routes are things an agent can create that outlive its task: scheduled jobs, API keys, app authorizations, webhooks and mail-forwarding rules. Copying model weights and running unmonitored model inference are lab concerns. ID.RA-05, PR.IR-01
1.5	APPLY	Limit each agent's network access, with a strict allowlist where possible, and route the access it needs through a monitored gateway. PR.IR-01
1.6	APPLY	Write a control policy that defines scenarios and the intended response to each, backed by automated enforcement.
1.7	APPLY	Build ways to rapidly block, contain or stop agent activity when an incident or near-miss is detected. RS.MI-01
1.8	APPLY	Assume the agent will find your misconfigurations, such as an overshared folder or an old service account with broad rights. ID.RA-05, ID.RA-06
1.9	APPLY	Test the control measures, with people and with AI testers, before you deploy them. Use AI testers only where an out-of-scope action could not harm critical systems. ID.IM-02
1.10	APPLY	Follow sound engineering practice so an agent cannot exploit weak security. PR.DS-01, PR.PS-06

Practice 2. Build organizational and personnel response readiness

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
2.1	APPLY	Write indicators, criteria and a severity scale for agent incidents and near-misses. Section 5 has a starter scale. ID.RA-06, RS.MA-03

© 2026 TenetGraph · Copy and adapt freely within your organization

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
2.2	APPLY	Run drills and simulations, such as a tabletop exercise, that cover detection, response, recovery and reporting for an agent incident. ID.IM-02
2.3	APPLY	Create a channel where customers, users and other affected parties can report incidents involving your agents, and route those reports to the incident team.
2.4	APPLY	Work with independent external auditors and evaluators to assess your agent controls, monitoring and incident detection, and confirm they work as intended.
2.5	ADAPT	Publishing identifiers for evaluation traffic, such as IP ranges, is aimed at labs. Deployers can label agent traffic (named service accounts, request headers) so their own security team and vendors can tell agent activity from human activity.
2.6	APPLY	Complete a structured risk assessment before an agent goes live. Start or resume only when senior leadership, informed by external evaluators where appropriate, is satisfied with the risk level. GV.RR-02, GV.RM-06

DETECT

Practice 3. Monitor for unexpected activity

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
3.1	APPLY	Define each agent's intended scope, including system and network access, so anything outside it stands out. PR.IR-01
3.2	ADAPT	The guidance asks for monitoring at the model, harness, host and network levels; model-level monitoring is mostly a lab practice. Deployers should at minimum monitor every action an agent takes (each tool call and its parameters), plus identity and network logs, and the agent's reasoning and outputs where the platform exposes them. PR.PS-04, DE.CM-01, DE.CM-06, DE.CM-09
3.3	APPLY	Make sure agents cannot read, change or switch off monitoring and logs. Cross-check monitoring layers against each other to spot tampering. DE.AE-03
3.4	APPLY	Where possible, use more than one model to monitor or cross-check agent activity, ideally from different model families.
3.5	APPLY	Set up ways to escalate unexpected activity to a person or an automated response. Plant decoys, such as a fake credential or record, that no legitimate task should touch. DE.AE-02, DE.AE-06

Practice 4. Validate and secure monitors to improve operational integrity

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
4.1	APPLY	Attack-test your monitors, with people and with AI testers, to make them more reliable. ID.IM-02
4.2	APPLY	Treat agent output as possibly containing exploits aimed at the monitors that review it. DE.CM-09

© 2026 TenetGraph · Copy and adapt freely within your organization

RESPOND

Practice 5. Collect, retain and analyze evidence

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
5.1	ADAPT	Weights and training checkpoints are lab concerns. Deployers should preserve the deployment environment, timestamps, sandbox and network logs, and every rejected action and tool call. RS.AN-07
5.2	APPLY	Keep a snapshot of the agent's configuration: model and version, instructions, tool list, permissions, and the policy version in force. You need it to reproduce the incident. RS.AN-07
5.3	APPLY	Keep tamper-resistant logs, with reliable timestamps, that can reconstruct each agent's actions, tool calls, reasoning traces, user interactions, environment state and control decisions. Section 6 lists the fields. PR.PS-04, RS.AN-07
5.4	APPLY	Use the logs to establish what the agent did, which systems it touched, and the root cause. RS.AN-03, ID.IM-03
5.5	APPLY	AI can help investigate. Where feasible, have a person verify key facts, and cross-check with several models, including one from outside the family of the model behind the incident. RS.MA-01, DE.AE-06
5.6	ADAPT	The guidance asks labs to check other training and evaluation runs. Deployers should check other agents built the same way, on the same platform template or with the same tools, for the same failure.

Practice 6. Mitigate and shut down undesired activity

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
6.1	APPLY	Write a shutdown policy that covers unintended activity, control failures, and monitor outages or tampering. Decide in advance what happens to an agent when its monitoring stops. RS.MI-01, RS.AN-08
6.2	APPLY	End the agent's task automatically when a control fails or an incident is under way. RS.MI-02
6.3	APPLY	Depending on severity, pause similar agents until their controls are strengthened and checked.
6.4	APPLY	Revoke every credential under your control that was accessed during the incident and rebuild any environment that may be compromised. RS.MI-02, RC.RP-05
6.5	APPLY	Find and remove any way the agent could persist, using the list from 1.4. RS.MI-02

RECOVER

Practice 7. Verify and strengthen containment before resuming agent activity

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
7.1	APPLY	Rebuild from a known-good baseline: configuration, instructions, tool settings and container images. RC.RP-03
7.2	APPLY	Patch or report any vulnerability the agent exploited, including to the platform vendor. RS.MI-02
7.3	APPLY	Confirm monitors, logging, allowlists, shutdown controls and permission limits all work before the agent restarts. Section 7 has a checklist. RC.RP-05
7.4	APPLY	Strengthen controls and monitors, and verify that they would prevent or detect a similar incident, for example by replaying the incident's actions. ID.IM-03, ID.IM-04

© 2026 TenetGraph · Copy and adapt freely within your organization

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
7.5	APPLY	Document the decision to resume and the risk that remains. RC.RP-06
7.6	APPLY	Write an after-action report, as appropriate, covering the incident, the response and lessons learned. RC.RP-06
7.7	LAB ONLY	Concerns recovery during model training. Relevant only if you train or fine-tune models.

REPORT

Practice 8. Coordinate with affected parties

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
8.1	APPLY	Notify anyone whose systems the agent compromised or exploited, including third parties that were not the final target. RS.CO-02
8.2	APPLY	Where feasible, give affected parties the detailed timeline and transcripts their own response needs. RS.CO-03
8.3	APPLY	Share findings and lessons learned with others in the value chain, such as the model developer and your agent platform and infrastructure providers.

Practice 9. Notify government

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
9.1	APPLY	Meet every legal reporting obligation (Section 8). The guidance also invites voluntary reports to CISA of cybersecurity incidents or near-misses that arise from control incidents. GV.OC-03, RS.CO-02
9.2	APPLY	Report software vulnerabilities through established disclosure channels.
9.3	APPLY	Optionally share agent-specific details with government experts such as NIST CAISSI and NSA SISC, after any mandatory reporting is complete. RS.CO-03, RS.MA-03

Practice 10. Share information publicly

#	VERDICT	WHAT IT MEANS FOR A DEPLOYER
10.1	APPLY	Share details publicly where appropriate, as early as feasible. Decide with counsel what that means for you.
10.2	APPLY	Give independent reviewers enough access and records to check your account. ID.IM-01
10.3	APPLY	Publicly report the practices and improvements you adopted, where that does not expose sensitive detail. ID.IM-03, RS.CO-03

11 Sources

- NIST CAISSI, NSA SISC and CISA. *Key Practices for Incident Readiness and Response in Super Intelligence Training and Evaluation*. PP-26-4396, Version 1.0, October 2026.
- NIST CSWP 29, *The NIST Cybersecurity Framework (CSF) 2.0*, and NIST SP 800-61 Rev. 3, *Incident Response Recommendations and Considerations for Cybersecurity Risk Management*. CSF references in brackets follow the guidance's own mapping.
- Legal and regulatory citations in Section 8, checked against primary text in October 2026, including Regulation (EU) 2026/1744 (Digital Omnibus on AI), Official Journal, 24 July 2026.

© 2026 TenetGraph · Copy and adapt freely within your organization

tenetgraph

12 How TenetGraph helps

TenetGraph keeps each AI agent's actions aligned with its intended purpose and with your corporate policy. We define the intended scope for each agent and enforce it before every action. We monitor every agent against that scope, and a kill switch stops any agent immediately when something goes wrong. Every decision is recorded as evidence.

PREPARE · 1.1, 1.6, 3.1

Define each agent's intended scope

We find the agents in use and what each one can do. With the agent's owner, we define a least-privilege policy for each agent, derived from its intended purpose, your corporate policy, and your customer and compliance requirements. The owner approves it in plain English.

PREPARE · 1.9, 2.4

Test the policy before launch

An adaptive adversary, driven by our reinforcement-learning engine, tests the policy against the agent itself and learns where to probe next. Each finding tightens the policy before the owner approves it.

PREVENT AND CONTAIN · 1.5 TO 1.7

Enforce the policy before every action

Every action is checked before it runs, down to a single site, tool or command, through our gate or your existing enforcement points. An out-of-policy action never executes. A high-impact action can wait for a person's approval, and each decision takes into account what the agent was told to do.

DETECT · 3.1, 3.5, 6.1

Monitor every agent against its scope

We check every action each agent takes against its intended scope and your policy. Allowed and blocked actions stream to your SIEM, so your security team sees near-misses as they happen. When an agent gains a tool, new access or a new configuration, its policy goes back for testing and review. You decide whether an agent stops when its controls cannot be reached.

RESPOND · 1.7, 6.3

Stop any agent immediately

A kill switch stops one agent, a group of similar agents, or every agent in a workspace on its next action. Nothing has to be uninstalled. After the restart decision, one call turns the agents back on.

RECOVER AND REPORT · 5.1, 5.3, 7.4, 10.2

Keep a record that holds up to review

Each decision is recorded with the rule that applied and the instruction behind it. Purposes, test findings and policies are versioned and never edited in place. Before an agent restarts, we test the updated policy against it again, including the actions from the incident.

TenetGraph works the same way whether you build your agents, buy them or ship them inside your product. The records map to SOC 2, ISO 42001, the NIST AI RMF and the EU AI Act.

Review your agent readiness checklist with TenetGraph

tenetgraph.ai

© 2026 TenetGraph · Copy and adapt freely within your organization