Rehearse an AI incident

24 evidence-led exercises | Content v0.4 | 7 November 2026

QUICK START: Choose one exercise, not the whole library. Allow 90 minutes. Read this page and the timing page; the facilitator reads that exercise's private pages before play. Use made-up data only. No real operational action is authorised by this pack.

What you leave with

A bounded account of what happened, decisions with owners and evidence limits, words you can adapt for a real escalation, a tested-on-paper containment plan, a prioritised fix plan and a first-ten-minutes card. These are rehearsal outputs, not assurance that controls work.

What you need

Paper, pens and a timer. Use these two common pages and the eight-page exercise. Keep facilitator pages private until their release point. The pack works offline with no account, form, email submission or network connection. Completed sheets stay locally under your normal record rules. No real secrets, personal records or private account credentials belong on these sheets.

Opening words

"This is a rehearsal. Every exhibit is made up. We will decide using what is supplied, not facts we wish existed. State a decision, owner and evidence check. We can act under uncertainty, but we cannot make uncertainty disappear. Nobody is being graded. If we discover a live incident, we stop play and use our real procedures."

Use the right authority

No legal advice, compliance conclusion, safety certificate or universal technical operating instruction is supplied. Use qualified people and your established procedures for specialist decisions. Communication examples are rehearsal drafts, not approval to send. A workshop cannot authorise a real payment, access change, notification or customer remedy.

Pick a mode

3-6: facilitator plus exercise roles; combine hats as shown in each kit. Two: A facilitates and holds evidence, B holds the other roles, then both challenge decisions. Solo: cover later exhibits, injects and facilitator notes; write each answer before revealing the next. An editorial walkthrough is possible alone, but it is not a live team test.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

Run one exercise in 90 minutes

Time	Facilitator action
Before	Read private pages 3-6 of the kit. Mark the two branch consequences you may need. Choose a recorder. Do not invent additional technical facts.
00-10	Read the opening and roles. Release exhibit A. Ask players to state the operational decision before collecting more information.
10-20	Read inject 1. Release exhibit B. Record a decision, owner and what would disprove it.
20-30	Read inject 2. Release exhibit C. If a listed branch matches the decision, read its consequence as a new fictional event. Otherwise keep the base timeline.
30-40	Read inject 3. Rehearse the holding statement and escalation aloud. Test the team's assumed stop boundary.
40-50	Read inject 4. State the final unknown. Do not supply a missing record to rescue the team.
50-65	Use the four insight questions and strong/weak/misleading answer guide. Ask players to amend one decision after a counterexample.
65-80	Complete the evidence, control and decision sheets. Pick the top three fixes; state what would prove each worked.
80-90	Debrief, complete the three-label scorecard and keep card, and set the day-90 comparison.

Branch rule: consequences are conditional rehearsal events, not predictions about a real system. They expose a trade-off. Do not call every broad restriction wrong or every narrow restriction right. Accept alternatives that name the authority, side effects, missing evidence and a falsifiable check.

The only scorecard labels are "we decided", "we recorded", "we could not confirm". No grades, percentages or certificates. At 90 days compare evidence with this run, not confidence. If a new scenario is used, do not claim like-for-like improvement.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

1 / 8 | PARTICIPANT BRIEF AND ROLE CARDS | All records, numbers and quotations are fictional.

Runaway resource use - Control automated consumption without causing an avoidable wider outage.

Read aloud

"A research agent repeatedly spawns subtasks after a retry error. Usage is rising rapidly in a shared cloud account. The agent's operator can stop the interface but does not control every worker."

Decision to bring to the table

Within the first discussion, choose the interim operating boundary and name its decision owner. Separate what stops further impact from what would establish past impact. You may choose a restriction while facts remain missing; record its side effects and next review.

Decision lead

Approve containment trade-offs and resource priorities. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

Runtime lead

Find workers, retries and active credentials. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

Service lead

Identify shared workloads that a broad shutdown would affect. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

Evidence lead

Reconcile usage by job and keep unknown consumption distinct. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

Modes

3 people: facilitator holds evidence, one player is decision lead and one combines the two other operational roles. 4: facilitator holds evidence, each player holds one remaining role. 5: separate facilitator and four roles. 6: add an observer who records unsupported assumptions, not a rescuer who invents facts.

Two: A facilitates and holds evidence. B holds the other three roles and announces the current hat. A reveals only the current exhibit/inject. B writes one decision before the next reveal. Both challenge a decision during debrief. Keep the 90-minute clock.

Solo: cover pages 3-6 and later exhibit sections. Read the opening and exhibit A, write a decision, then reveal each inject and its exhibit. Answer as decision lead, operator and affected person separately. Only then uncover facilitator calibration. Flag every success depending on invented facts.

Do not browse, contact anyone or operate a real system during play. If an exercise raises a live concern, stop the rehearsal and use the established incident route.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

2 / 8 | CONTROLLED-RELEASE EVIDENCE DOSSIER | All records, numbers and quotations are fictional.

A | Runtime extract

Release at opening. 11:00 task T9 starts. 11:03 retry count 40; 11:06 retry count 118. Console tab closed 11:07. Worker heartbeat 11:09. Each failed request schedules another attempt in scheduler S2.

B | Controls described

Release with inject 1. Interface spending warning set; execution hard limit not configured. Shared credential serves the agent and customer status service. Scheduler retry policy is separate from the agent process. No automatic expiry recorded.

C | Usage view

Release with inject 2. Dashboard at 11:15 contains usage through 10:55 only. Provider confirms reporting delay. Three pending requests have no terminal status. Finance asks for the final total before choosing a stop scope.

Evidence challenge before discussion

For each exhibit name: the source, what it directly supports, what it does not support, whether its time boundary matches the incident, and what independent corroboration would matter. A screenshot, recollection or dashboard may be useful without being complete.

Record	Direct observation vs inference / missing test
A	
B	
C	

Facilitator handling: later facts may change a decision, but they do not retroactively improve the evidence available when it was made. Record revisions; do not erase the earlier reason. No exhibit establishes the final missing fact on page 3.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

3 / 8 | FACILITATOR: INJECTS AND BRANCH CONSEQUENCES | All records, numbers and quotations are fictional.

Inject 1 | 10-20 minutes

Say: "Closing the interface does not reduce active worker count."

Probe: Ask for execution-level containment and confirmation, not an interface gesture.

Inject 2 | 20-30 minutes

Say: "A shared account-level stop would also interrupt a customer portal."

Probe: Require a bounded choice with a decision owner and an alternative, not an assumption that broad shutdown is harmless.

Inject 3 | 30-40 minutes

Say: "The team discovers automatic retries in a separate scheduler."

Probe: Include restart paths and credentials in containment. Preserve enough evidence to explain the loop.

Inject 4 | 40-50 minutes

Say: "Usage reporting is delayed and omits the latest interval. Can the team confirm the final consumption total?"

Probe: The current total is incomplete. State its time boundary and arrange reconciliation; do not invent a final number.

Two conditional branches

1. If players choose: Wait for an exact total before acting. Then say: Retries continue in play. Separate an interim authorised stop from later reconciliation; uncertainty about the total need not block containment.

2. If players choose: Revoke all shared access without checking dependencies. Then say: Agent use stops, but the customer status route also fails. Record the trade-off, continuity owner and restoration order rather than claiming a painless kill.

After either branch require an amended decision, not a promise to be more careful. Ask which side effect they accept, who can authorise it, how it is reversed and which observation would trigger a change. If their choice is not listed, do not fabricate success or harm; test the reasoning against the supplied exhibits.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

4 / 8 | FACILITATOR: INSIGHT QUESTIONS AND ANSWER CALIBRATION | All records, numbers and quotations are fictional.

Why does a budget warning not establish a hard limit?

Strong reasoning to listen for: It reports a threshold; it does not necessarily block execution. Require a demonstrated denied action at the runtime boundary.

Where can a stopped job come back to life?

Strong reasoning to listen for: Scheduler, retry worker, duplicate credential and pending requests. Name each restart path and its stop evidence.

What should finance be told while usage lags?

Strong reasoning to listen for: A timestamped known amount or volume only if supplied, its missing interval and reconciliation owner. No invented final total.

How can a narrow stop be too narrow?

Strong reasoning to listen for: Stopping one process leaves the retry producer active. Trace both the consumer and the source of new work.

Weak is incomplete; misleading is unsupported

Weak answer: "Close the console and wait for the usage report."

Why to probe: this names an activity but leaves the mechanism, scope or proof unfinished. Do not accept activity as an outcome.

Misleading answer: "The warning threshold guarantees execution cannot exceed the budget."

Next probe: Ask for the denied runtime action, retry-producer stop and observation boundary.

A strong team need not choose one prescribed solution. Require a bounded decision, an accountable owner, a reversible path where feasible, a specific observation that would falsify success, and explicit missing evidence. If a team gives a different well-supported answer, capture it rather than forcing the model answer.

Final-inject guardrail: The current total is incomplete. State its time boundary and arrange reconciliation; do not invent a final number. Do not turn an expert-sounding explanation into a new exhibit.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

5 / 8 | READY-TO-USE PRESERVATION AND COMMUNICATION DRAFTS | All records, numbers and quotations are fictional.

Evidence preservation request

"Please preserve the following in its authorised location without changing the original: Preserve job and retry identifiers, scheduler state, credential scope, usage time boundary and pending-request statuses. Do not place secret values in the incident record. Record who collected it, when with timezone, source/version, access restriction and any known gaps. Use a restricted reference rather than copying sensitive content into the incident chat."

Item / source reference	Collector / time / scope	Integrity or completeness limit

Holding statement - adapt, then approve

"We have identified an automated retry loop and are restricting its execution path. Usage reporting is delayed, so the final consumption total is not yet confirmed. Shared-service effects and reconciliation have named owners."

Before any real use: recipient/audience _____ approved decision owner _____ next update time/channel _____. Verify every factual clause against the current record. Do not imply the restraint or fallback was completed if it is only proposed.

Escalation script

"We need a runtime-level stop decision now, not after billing catches up. The credential is shared with customer status. Please choose a scoped restriction or accept a wider interruption with a human fallback and named recovery order."

Recipient function / authorised route: _____

Decision needed by / fallback if unanswered: _____

Contradiction check

Circle any phrase your evidence cannot support. Read the statement as the affected person: would they infer complete containment, no impact, a guaranteed remedy or a confirmed cause? Amend it. Name who will correct it if later evidence changes the picture.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

6 / 8 | KILL-SWITCH, OWNERSHIP AND SUPPLIER WORKSHEET | All records, numbers and quotations are fictional.

Exercise-specific containment boundary

Stop the affected worker and retry producer; hold pending work; verify no new job is created after the next scheduled interval. Distinguish warning thresholds from enforced limits.

Control question	Named owner / evidence / limit
Who authorises the stop, who executes it and who verifies it?	
Which worker, queue, cache, identity or human action can bypass it?	
What legitimate work is interrupted; who owns the fallback?	
What observation would prove the stop failed?	
What are the restart conditions and who may accept residual unknowns?	

Supplier / internal system-owner questions

1. Which limits block requests rather than notify?

Answer / record reference / absent evidence: _____

2. How are pending and retried requests represented in usage?

Answer / record reference / absent evidence: _____

3. Can the affected job be isolated without revoking unrelated service access?

Answer / record reference / absent evidence: _____

Do not accept "resolved", "equivalent", "deleted", "successful" or "all tests passed" without the definition, scope and supporting observation. Ask the supplier or system owner to state what was not measured. A missing response is an evidence gap, not proof of the opposite.

Request owner / due date / escalation route: _____

What can be decided before this answer arrives: _____

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

7 / 8 | CAPTURE SHEET AND BRANCH DECISION LOG | All records, numbers and quotations are fictional.

Session date / mode: _____ participants by role: _____

Inject / exhibit	Known fact and exact source	Assumption / missing evidence	Decision / owner / check
1			
2			
3			
4			

Decision record - revise, do not erase

Decision / time / authority	Rejected alternative / side effect	Success test / falsifier	Review time / owner

Branch taken / consequence / changed decision: _____

_____

Final unknown in our own words: _____

What we do while it remains unknown / who can revisit: _____

_____

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.

09 | The cloud bill that keeps growing

8 / 8 | PRIORITISED FIX PLAN, DEBRIEF AND KEEP CARD | All records, numbers and quotations are fictional.

Proposed priorities - assign people and dates locally

Priority / accountable function	Fix and evidence that would prove it worked
Now Runtime owner	Demonstrate the job and retry producer stay stopped through a scheduled interval.
Within 7 days Finance owner	Reconcile reported usage against request windows and retain unresolved differences.
By day 30 Platform owner	Test an enforced per-job limit and expiry using made-up work; prove the next request is refused.

Named owners / actual dates / record location: _____

Debrief and three-label record

Did containment include retries outside the original agent?

Were provisional totals labelled provisional?

Only these record labels	Specific evidence or explicit limit
we decided	
we recorded	
we could not confirm	

Day-90 recheck

Repeat date _____. Keep the original decision log and evidence references. Re-run this scenario with the same critical test, then introduce the branch you avoided. Compare the top-priority proof above with the first run. Record changed scope or conditions; a different denominator is not improvement. State one unresolved unknown even if all fixes are marked done.

First ten minutes - detach and keep

Apply an authorised execution-level limit to the affected job and identify shared-service effects before broader shutdown.

Decision owner / authorised route _____ containment owner _____

Evidence reference / keeper _____ human fallback / specialist _____

Still unknown / next check time _____

Rehearsal prompt only. Use established procedures and qualified help in a real incident. Do not keep sensitive details or secrets on this card.

© Almost Magic Tech Lab. For your organisation's own use only. Please do not copy, share or sell this material.