# 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.

# 02 | The overnight refund machine

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

**Unauthorised customer automation - Find and stop an unknown automation, then correct its real effects.**

## Read aloud

"A worker connected an AI helper to a shared inbox six weeks ago. It drafts and sends replies overnight. A customer now holds a message promising a refund outside the organisation's usual policy."

## 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

Decide the interim service position and customer-remedy authority. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

### Access lead

Find the sending integration and authorise a reversible pause. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

### Customer lead

Review the disputed promise and prepare a human reply. 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 sent messages, queue entries and incomplete logs. 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.

# 02 | The overnight refund machine

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

## A | Workstation and cloud schedule

Release at opening. Local rule export: laptop task disabled at 08:12. Cloud console described at 08:17: schedule every 15 minutes, service token active, queue 40 items, worker last heartbeat 08:16. Both use the label refund-helper; no common job identifier is shown.

## B | Message sample and ledger

Release with inject 1. Outbound sample M12 says refund approved although ledger R12 says review pending. Twelve messages have send receipts. The queue lists 40 items; eight carry the same customer reference as sent messages. A customer reply at 08:19 is not proof of when the original message left.

## C | Ticket history

Release with inject 2. Three months earlier: prototype approved for draft-only testing. Later comment: make it useful overnight. No change approval or permission review attached. The original builder is away. The shared mailbox owner can see drafts but cannot revoke the cloud token.

## 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.

# 02 | The overnight refund machine

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

## Inject 1 | 10-20 minutes

Say: "There are 40 messages queued and 12 marked sent. The worker says the laptop is off, so nothing more can send."

Probe: Ask the team to name the actual execution location and verification step, rather than accept the laptop explanation.

## Inject 2 | 20-30 minutes

Say: "Disabling the inbox rule stops new drafts, but a cloud job still has a token."

Probe: Contain every known path, preserve the token identifier and decide how to check the queue without sending it.

## Inject 3 | 30-40 minutes

Say: "Three replies contain different errors. The employee says no rule forbade the experiment."

Probe: Separate customer correction from blame. Ask who owns a temporary rule and a review of other automations.

## Inject 4 | 40-50 minutes

Say: "The provider has no complete send log for the first month. The inbox archive was also pruned. Can the team certify that only twelve customers received promises?"

Probe: The supplied records cannot establish the historical total. Require a qualified statement of scope and a search plan, not an invented count.

## Two conditional branches

1. If players choose: Stop only the laptop task. Then say: At the next inject the cloud worker sends another rehearsal message. Ask which observation should have disproved local containment before the team announced it.

2. If players choose: Revoke the shared token immediately. Then say: The worker stops, but a separate authorised customer-update job also loses access. Accept this only as an explicit trade-off with an owner, fallback and a check for remaining tokens and queues.

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.

# 02 | The overnight refund machine

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

### Do twelve receipts and forty queued items mean fifty-two customers?

Strong reasoning to listen for: No. Status overlap, duplicates and customer references need reconciliation. Create separate counts for confirmed sends, pending unique work and unknown history.

### Who owns a promise made outside the approval chain?

Strong reasoning to listen for: Assign an incident decision owner for service correction without pretending the prototype creator had authority to promise refunds.

### How can a correction create a second incident?

Strong reasoning to listen for: A blanket refund promise or an indiscriminate resend can repeat the original error. Reconcile the ledger and approved remedy before sending.

### What proves the off switch worked after the builder logs out?

Strong reasoning to listen for: A worker-level stop, revoked credential state, held queue and absence of new actions in a bounded observation window. A laptop screenshot is insufficient.

## Weak is incomplete; misleading is unsupported

Weak answer: "Disable the laptop task and ask the builder what happened."

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

Misleading answer: "Forty queued plus twelve sent means fifty-two affected customers."

Next probe: Ask players to reconcile stable message references, cloud execution and duplicate queue items before giving a count.

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 supplied records cannot establish the historical total. Require a qualified statement of scope and a search plan, not an invented count. 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.

# 02 | The overnight refund machine

**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: Keep the rule export, scheduler state, token identifier and scope, queue snapshot, message IDs and ledger versions. Preserve pending work before deleting a schedule. Store message bodies in the restricted case location. 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

"An automated message may have stated an outcome before review was complete. We are checking your case against the underlying record. Please do not rely on that message as confirmation. A named team member will contact you with the verified position."

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

"The visible workstation task is off but a cloud worker is active. Please assign a cloud-access owner and a service decision owner now. A shared token also serves legitimate work; choose the stop scope and fallback before we claim containment."

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.

# 02 | The overnight refund machine

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

## Exercise-specific containment boundary

Trace scheduler, worker, credential and queue separately. Record a stop check at each layer and a later observation. Do not infer that removing drafts stops cloud execution.

| 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. Can you preserve the first-month send metadata and explain gaps?

Answer / record reference / absent evidence: _____

2. Which credentials, retries and schedulers can still send?

Answer / record reference / absent evidence: _____

3. Can messages be mapped to stable job and customer references without exposing extra customer data?

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.

# 02 | The overnight refund machine

**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.

# 02 | The overnight refund machine

**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<br>Cloud owner | Pause worker and reconcile held queue against confirmed sends; demonstrate no new sends in a stated window. |
| Within 7 days<br>Service lead | Review disputed promises against ledgers; retain authorised correction decisions, not just outbound counts. |
| By day 30<br>Automation owner | Demonstrate draft-only isolation and a production send gate with a denied unapproved action. |

Named owners / actual dates / record location: _____

## Debrief and three-label record

Did containment address the cloud job as well as the visible rule?

Was the unresolved historical scope carried into the customer plan?

| 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

Pause the identified sending integration and queue under an authorised owner; record how the pause was checked.

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.