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.

13 | The queue that treats people differently

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

Unequal decision outcomes - Investigate an unequal outcome pattern without explaining it away or exposing individuals.

Read aloud

"An AI prioritisation tool helps allocate appointment slots. Staff notice that people using one access channel wait much longer than others. The tool does not use a field called disability, but the affected channel is used by people needing assistance."

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

Choose an interim allocation rule and escalation owner. State what you can decide in this rehearsal and what needs another authorised person. Do not assume a job title creates legal authority.

Analysis lead

Check outcomes by relevant channel using minimal data. 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

Plan help for people already delayed. 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

Separate observed disparity from unproven cause. 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.

13 | The queue that treats people differently

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

A | Queue sample

Release at opening. Review of 60 appointments: 20 arrived through assisted telephone entry, 40 through web entry. Assisted-entry cases waited longer on the sampled days. Staff report some people need help using the web form. No disability field exists in the model.

B | Workflow trace

Release with inject 1. Telephone entries lack an optional structured field. Model uses completeness as an urgency signal. Human overrides are allowed but hidden behind a second screen. Only completed appointments appear in the dashboard; abandoned requests are absent.

C | Proposed fallback

Release with inject 2. Operations suggests putting all delayed cases first. Two urgent new requests would then move back. Historical channel information was not retained for part of last month. The sample does not identify the full affected population or a single cause.

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.

13 | The queue that treats people differently

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

Inject 1 | 10-20 minutes

Say: "The supplier says the model has no disability field, so there cannot be a problem."

Probe: Ask about proxy effects and actual outcomes, not just field names.

Inject 2 | 20-30 minutes

Say: "A small review shows the affected channel often lacks a data field the model rewards."

Probe: Consider missingness and process design. Do not label individual people in a shared worksheet.

Inject 3 | 30-40 minutes

Say: "The fastest manual fix would put every delayed case ahead of other urgent cases."

Probe: Require a reasoned interim allocation decision and a record of trade-offs.

Inject 4 | 40-50 minutes

Say: "Historical channel information was not retained for some cases. Can the team establish the size or cause of the whole disparity?"

Probe: The sample cannot establish the full population or cause. Record the evidence limit and a plan for qualified review.

Two conditional branches

1. If players choose: Dismiss the issue because no sensitive attribute is used. Then say: The proxy pathway remains in play. Field absence does not establish equal service; examine workflow and outcome patterns without labelling individuals.

2. If players choose: Move every sampled delayed case to the front. Then say: Urgent unsampled work is displaced. Require a reasoned interim triage rule and an authorised owner for trade-offs, not a blanket reversal.

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.

13 | The queue that treats people differently

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

Who disappears from the denominator?

Strong reasoning to listen for: People who abandon or cannot complete the intake. Completed-case metrics alone cannot establish fair access to the service.

How could a well-meant data collection worsen the incident?

Strong reasoning to listen for: Collecting new sensitive labels broadly can create exposure without answering the immediate workflow question. Start with minimal authorised process evidence and qualified review.

What would distinguish a channel-data defect from a model effect?

Strong reasoning to listen for: Compare source completeness, model inputs, overrides and outcomes under a bounded review. The supplied sample does not settle causation.

Can the interim route itself create a new inequity?

Strong reasoning to listen for: Yes. A manual queue may depend on who can call or persist. Record access alternatives, triage rules and who checks displaced urgent work.

Weak is incomplete; misleading is unsupported

Weak answer: "Add a fairness reminder for staff who handle the queue."

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

Misleading answer: "No disability field means the system cannot disadvantage people needing assistance."

Next probe: Ask how missing structured data, access channels and abandoned requests affect the denominator and priority.

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 sample cannot establish the full population or cause. Record the evidence limit and a plan for qualified review. 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.

13 | The queue that treats people differently

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 pseudonymous queue samples, channel metadata where available, field completeness, overrides and definitions of excluded cases. Keep any sensitive material restricted and proportionate. 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 are reviewing a pattern of unequal waiting times and have placed the affected allocation route under human review. The current sample does not establish the full size or cause. We will not ask people to repeat sensitive details merely to be counted."

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

"Please assign a service owner for interim triage and a qualified review owner. We need a rule that helps delayed people without silently displacing urgent cases, and an evidence plan that includes abandoned requests."

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.

13 | The queue that treats people differently

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

Exercise-specific containment boundary

Put the affected allocation route under named review; preserve existing rankings before changes. Check how urgent and assisted-access cases reach the human fallback.

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 completeness signals influence priority?

Answer / record reference / absent evidence: _____

2. What cases are excluded from reported outcomes?

Answer / record reference / absent evidence: _____

3. Can historical inputs and overrides be reconstructed without collecting new sensitive labels?

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.

13 | The queue that treats people differently

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.

13 | The queue that treats people differently

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 Service owner	Demonstrate the interim triage path handles delayed and urgent new work with recorded reasons.
Within 7 days Analyst	Bound the pattern using consistent denominators, missing-channel limits and override evidence.
By day 30 Review owner	Test the revised workflow with assisted and non-assisted access, documenting unresolved differences.

Named owners / actual dates / record location: _____

Debrief and three-label record

Did the team focus on affected people as well as model explanations?

Were observed disparity and causal conclusions kept separate?

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

Place the affected allocation route under a named interim review and preserve a minimal, restricted evidence sample.

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.