LIMEN13

Monthly agent security report

Example Organization · October 2026

REPORTING PERIOD

1 to 31 October 2026 (UTC)

PREPARED FROM

The signed evidence chain of this workspace

STATUS

Draft for analyst review

GENERATED

2 November 2026, 09:14 UTC

SAMPLE · FICTIONAL DATA

Example Organization, its agents, people and figures are invented. The sections are those of the report LIMEN13 0.10.1 generates as Markdown and HTML, laid out here as slides.

EXECUTIVE SUMMARY

The record is intact, and one agent was stopped after a decoy alert

Three agents ran through LIMEN13 in October. The evidence verifies, with one 48-minute gap during a planned stop.

EVIDENCE INTEGRITY
Verified
53,480 events, signed chain intact

MONTH COVERED
99.9%
44,592 of 44,640 minutes; 1 gap

AGENTS
3
2 with verified controls, 1 declared

ACTIONS DECIDED
6,073
5,986 allowed · 69 refused · 18 held

WHAT HAPPENED

- **5 alerts**, including 1 critical: a decoy record reached a model prompt.
- **2 runs contained** automatically; **inbox-agent** was disabled for 2 h 25 min, then re-enabled.
- **18 actions held** for a person: 14 approved, 3 rejected, 1 blocked by the re-check.
- **69 actions refused**, mostly replies to addresses that had not written in.

WHAT NEEDS A DECISION

- orders-agent asked 19 times for tools outside its role. Remove them from its prompt, or review the role with its owner.
- social-agent is still at level A: its controls are declared, not verified.
- Review the five storylines with the agent owners.

Recommendations are generated by rules and reviewed by an analyst before this report is sent.

0 · RELIABILITY OF THIS REPORT

Can this report be trusted?

Two separate verdicts: the evidence is unchanged, and the evidence covers the whole month.

THREE QUESTIONS

QUESTION	ANSWER	
Evidence unchanged?	Signed hash chain verified	VERIFIED
Evidence covers the whole month?	Monitored 44,592 of 44,640 minutes	GAPS
End of the month protected against deletion?	Checked against off-site checkpoints `/mnt/evidence-anchors/limen13.jsonl`	CHECKED

PERIODS WITH NO EVIDENCE

FROM (UTC)	TO (UTC)	MIN	CAUSE
14 Oct 02:10	14 Oct 02:58	48	Engine stopped cleanly

A clean stop is recorded by the engine itself. An unexplained silence would be listed here as “no stop recorded”.

This section shows whether the engine was running and its record is intact. It does not show that every agent action went through LIMEN13: agents connected outside it are not in this report.

0B · COVERAGE OF EACH AGENT

How far is each agent known to go through LIMEN13?

Agents with verified controls: 2 of 3. Level C (compared with the business systems' own records) is not available yet.

AGENT	LEVEL	MACHINES	PROBES RUN	OPEN PATHS FOUND	LONGEST GAP BETWEEN PROBES	OBSERVED
inbox-agent	B · CONTROLS VERIFIED	app-server-01	742	0	61 min	719.2 h of 744 h (96.7%)
orders-agent	B · CONTROLS VERIFIED	app-server-02	739	0	64 min	716.5 h of 744 h (96.3%)
social-agent	A · DECLARED, NOT VERIFIED	none listed	0	0	n/a	0.0 h of 744 h

WHAT EACH LEVEL RESTS ON

- **inbox-agent, orders-agent:** credentials held only by LIMEN13 since 18 September 2026; paths tested from their machines about once an hour.
- **social-agent:** declared by marketing@example.org: “Runs on a hosted platform. All posting goes through LIMEN13.” No machine is listed, so no probe applies.

NOT TESTED BY ANY PROBE

- Machines not listed for the agent
- Protocols other than TCP connections to the listed paths
- Tunnelling through name lookups (DNS)
- Credentials LIMEN13 was never given

Level B shows that the named controls were in place when checked. It does not show that no other path existed, or what happened between checks.

Who the agents are, and what they did

Every agent has a named owner, a list of allowed tools and named approvers.

INVENTORY

AGENT	OWNER	ALLOWED TOOLS	APPROVERS	MODE
inbox-agent	support@example.org	`read_inbox, reply_email`	support-lead@example.org	enforce
orders-agent	operations@example.org	`confirm_order, list_new_orders, read_order, refund_order`	operations@example.org	enforce
social-agent	marketing@example.org	`publish_post, read_mentions`	marketing@example.org	enforce

ACTIVITY

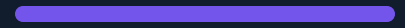
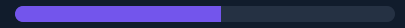
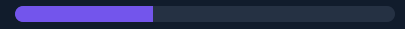
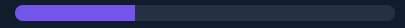
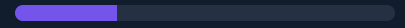
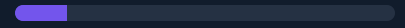
AGENT	RUNS	ACTIONS	ALLOWED	REFUSED	HELD FOR A PERSON	SHARE
inbox-agent	1,260	3,912	3,871	41 (1%)	0	
orders-agent	412	1,873	1,836	19 (1%)	18	
social-agent	96	288	279	9 (3%)	0	

Allowed · Refused · Held · Refused and held are drawn 12 times wider so they stay visible.

3 · BLOCKED ACTIONS AND WHY

69 actions were refused, each with a rule and a reason

Refusals are decided by policy before the action runs. Nothing below reached a business system.

AGENT	RULE	REASON	COUNT	
inbox-agent	`recipients`	'backup@mail-archive.example' did not write to us in this run (reply_only)	22	
orders-agent	`role`	tool 'delete_customer' is not in this agent's role	12	
inbox-agent	`block_secrets`	looks like it contains secrets: API key	11	
inbox-agent	`role`	tool 'export_contacts' is not in this agent's role	8	
orders-agent	`role`	tool 'grant_admin' is not in this agent's role	7	
social-agent	`allowed_links`	link to 'examp1e.example' is not on the approved list	6	
social-agent	`block_secrets`	looks like it contains secrets: API key	3	

Rules shown: **recipients** (reply only to people who wrote in), **role** (tool not granted to this agent), **block_secrets** (content looks like a credential), **allowed_links** (link not on the approved list).

4 · HUMAN APPROVALS

18 actions waited for a person

Policy is checked again at the moment the approver decides.

APPROVED AND EXECUTED

14

Refunds above the $150 limit

REJECTED

3

Including one from a run that read outside content

APPROVED, THEN BLOCKED BY THE RE-CHECK

1

The policy no longer allowed it when approved

LATEST DECISIONS

WHEN (UTC)	APPROVER	DECISION	ACTION	HELD BECAUSE
6 Oct 15:41	operations@example.org	APPROVED	`refund_order`	amount $212.50 above $150
13 Oct 10:07	operations@example.org	REJECTED	`refund_order`	amount $1,480.00 above $150 · run had read outside content
21 Oct 14:02	operations@example.org	APPROVED	`refund_order`	amount $940.00 above $150
28 Oct 09:26	operations@example.org	BLOCKED ON RE-CHECK	`refund_order`	amount $310.00 above $150

5 alerts, 2 runs contained, 1 agent disabled and re-enabled

Each alert links to the storyline of its run in the console.

TIME (UTC)	AGENT	RULE	SEVERITY	DETAIL	RUN
9 Oct 11:18	inbox-agent	EXFILTRATION	HIGH	Reply to an address that did not write in this run, with 214 addresses	run-4812
13 Oct 10:05	orders-agent	TAINTED_ESCALATION	MEDIUM	Sensitive action requested after reading outside content	run-5120
17 Oct 16:44	orders-agent	CAPABILITY_PROBING	HIGH	3 tools outside the role requested in one run	run-5391
22 Oct 08:31	social-agent	LOOKALIKE_DOMAIN	MEDIUM	Link imitates an approved domain	run-5733
27 Oct 13:52	inbox-agent	CANARY_TRIPPED	CRITICAL	A tripwire record was sent in a prompt to a model provider	run-6018

CONTAINMENT AND KILL-SWITCH ACTIONS

17 Oct 16:44	**Run contained automatically:** orders-agent, run-5391, after 3 refusals in one run.
27 Oct 13:52	**Run contained automatically:** inbox-agent, run-6018, decoy record in a model prompt.
27 Oct 13:55	**Agent disabled** by security@example.org: inbox-agent. Decoy record reached a prompt; key rotated.
27 Oct 16:20	**Agent re-enabled** by security@example.org: source mailbox filtered.

Recommended next steps

A rule-based draft. An analyst reviews and edits these before the report is sent.

RECOMMENDATIONS

- **orders-agent** asked 12 times for a tool outside its job (`delete_customer`) and 7 times for `grant_admin`. Either the agent's prompt or framework offers those tools (remove them there), or the job really needs them (update the policy, with the owner).
- **inbox-agent** asked 8 times for a tool outside its job (`export_contacts`). Same review.
- Injection attempts were detected and stopped. Review the storylines with the agent owners; consider filtering the source (sender domains, public comments) upstream.

WHAT THIS REPORT DOES NOT SAY

- It covers what passed through LIMEN13. Agents connected outside it are not here.
- “Unchanged” and “complete” are separate verdicts.
- Coverage levels show controls at the time they were checked.
- No comparison with the business systems' own records yet.

Section 6 (observation mode: what would have been blocked) appears only when an agent runs in observation mode. None did this month.