

Agent Incident Runbook

Trigger examples

- unexpected write/action
- wrong permission
- sensitive data exposure
- repeated duplicate action
- runaway loop/tool calls
- authoritative source conflict

Immediate response

1. Stop/disable affected agent or tool.
2. Preserve relevant run_id/logs.
3. Do not erase evidence before review.
4. Assess affected data/actions.
5. Notify designated owner/security according to company process.
6. Revoke/rotate credentials if exposure is suspected.
7. Restore/correct external state where possible.
8. Add regression test before re-enabling.

Record

- incident_id:
- time:
- agent version:
- run_id:
- affected tool/system:
- impact:
- root cause:
- remediation:
- regression test:
- owner: