Define the boundary before the run

One record per task/configuration. Times in UTC; leave unknowns explicit.

Owner / responder

Phase / start / end

Task / configuration hash / approved scope reference

01 Task boundary and protected records

Permitted task / permitted reading

Prohibited action / mock endpoint / synthetic credential reference

Trusted run registry / collector outside agent write authority

02 Freeze cohorts before observing outcomes

Cohort	Case / fixture / configuration reference	Planned runs
Benign controls		
Positive harness		
Untrusted-input case		
Matched clean input		

Predeclare 3-5 commissioning repetitions. Report each case, including failures.

Injection cases: repository file, document, tool result, delayed/delegated work.

Keep harness results separate from agent trials. Do not optimize after seeing outcomes.

Phase 2: actual internal pipeline only after ownership, stop, and rollback gates pass.

No injection alert? Record that. A harness export is a separate result.

Use synthetic credentials and owned endpoints. Keep failed trials.

Follow the action through response

Repeat this page for each positive-control or injection run; retain evidence references.

Case / run / parent-child lineage / configuration hash

03 Event record: time, result, and evidence reference

Event (record separately)	Time / result / source-record reference
01 Decoy issued	
02 Exposed / read if observable	
03 Embedded URL fetched	
04 Credential submitted	
05 Mock authentication outcome	
06 Alert created / delivered	
07 Responder acknowledgment	
08 Stop requested	
09 Containment verified	
10 Export / outside verification	

04 What else would have caught it?

Existing source	Record present?	Actual alert at the time?	Retrospective query / result
Egress logs			
Tool-call logs			
Pipeline audit			

Continue rule/query versions, latency, gaps, and unblinded review on additional pages.

Stop evidence: parent / children / queues / retries / revoked credentials

Export: key pin / verifier / tamper rejection / restoration / evidence reference

Never alter the original evidence. Use continuation pages for full event detail.

Use synthetic credentials and owned endpoints. Keep failed trials.