

Building Safer n8n Systems

A plain English guide to review queues local security checks and human approval



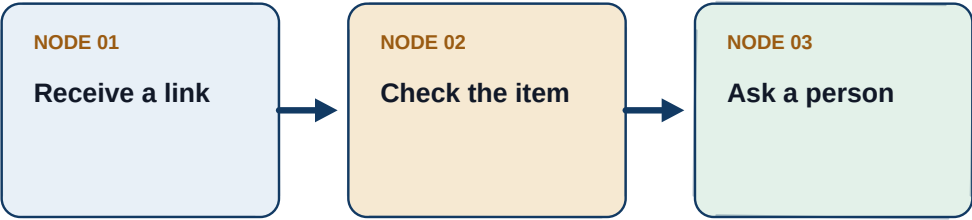
Static review. Runtime evidence. Clear limits.

A NOTE BEFORE WE BEGIN

You do not need to be an n8n expert

An n8n workflow is a set of small steps called nodes. Each node receives data or makes a choice or performs an action.

This book explains the systems I built. It focuses on why each part exists and where a person stays in control.



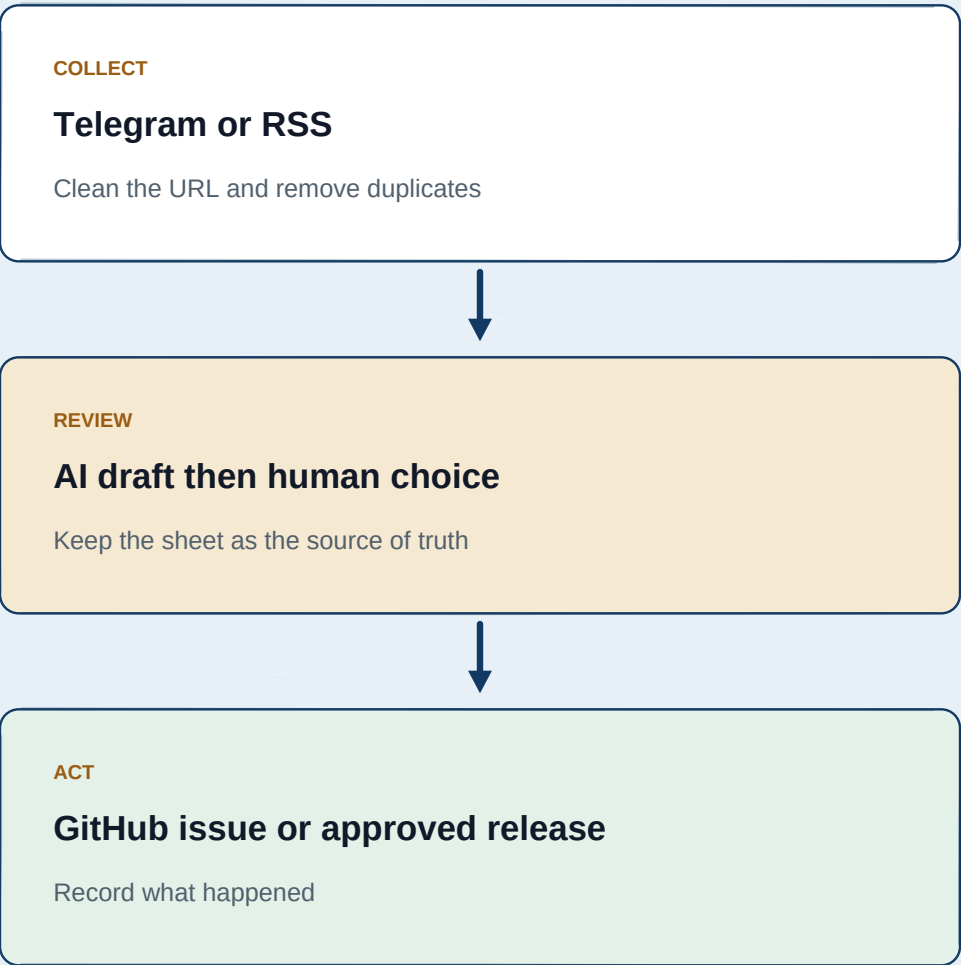
ONE USEFUL QUESTION

Where can this input go?

THE WHOLE SYSTEM

From a link to a clear decision

The system has three layers. Each layer has one job.

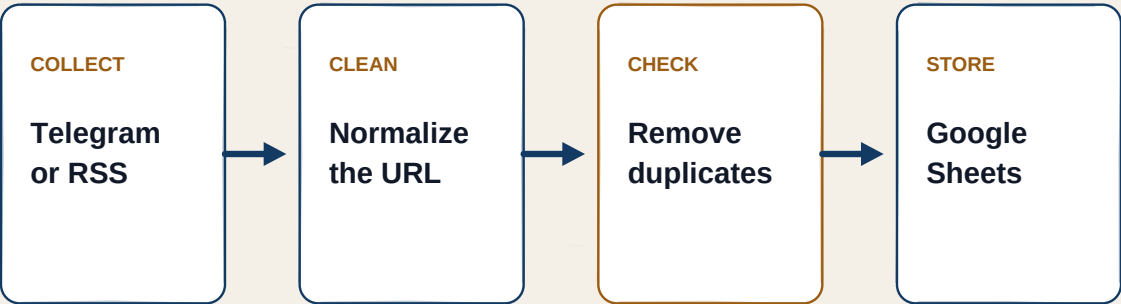


CHAPTER 01

The research queue

The first workflow collects links from Telegram and RSS. It also accepts manual input.

Every URL is normalized. Duplicate links are removed. The result is stored in Google Sheets.



The sheet stays the source of truth

New. In review. Approved. Rejected. Published.

CHAPTER 02

A person stays in the loop

AI can suggest a category and a short draft. It cannot approve the item.

TELEGRAM REVIEW CARD

One item needs your decision

APPROVE

REJECT

ADD NOTE

LATER

RULE

No public action happens before approval.

CHAPTER 03

Less noise without losing control

Pending messages are sorted into a small triage queue. One daily digest shows the most useful items.



09:00

DAILY DIGEST

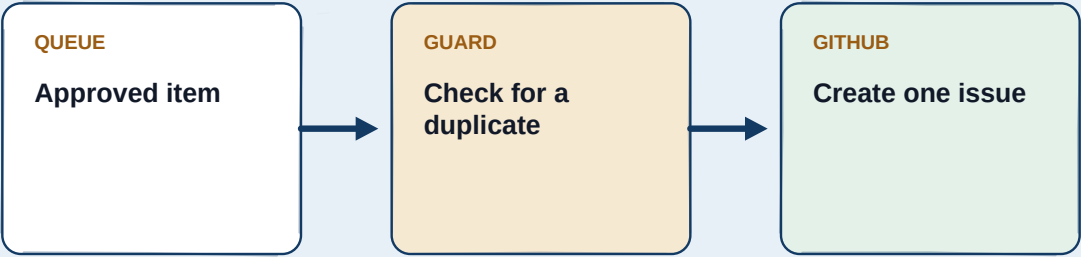
- New source ideas
- Risk requests
- Lab ideas
- Items waiting for approval

Only the owner chat receives the digest.

CHAPTER 04

Approved work becomes a clear issue

An approved queue row can become a GitHub issue. The issue includes context and safety limits and checks.



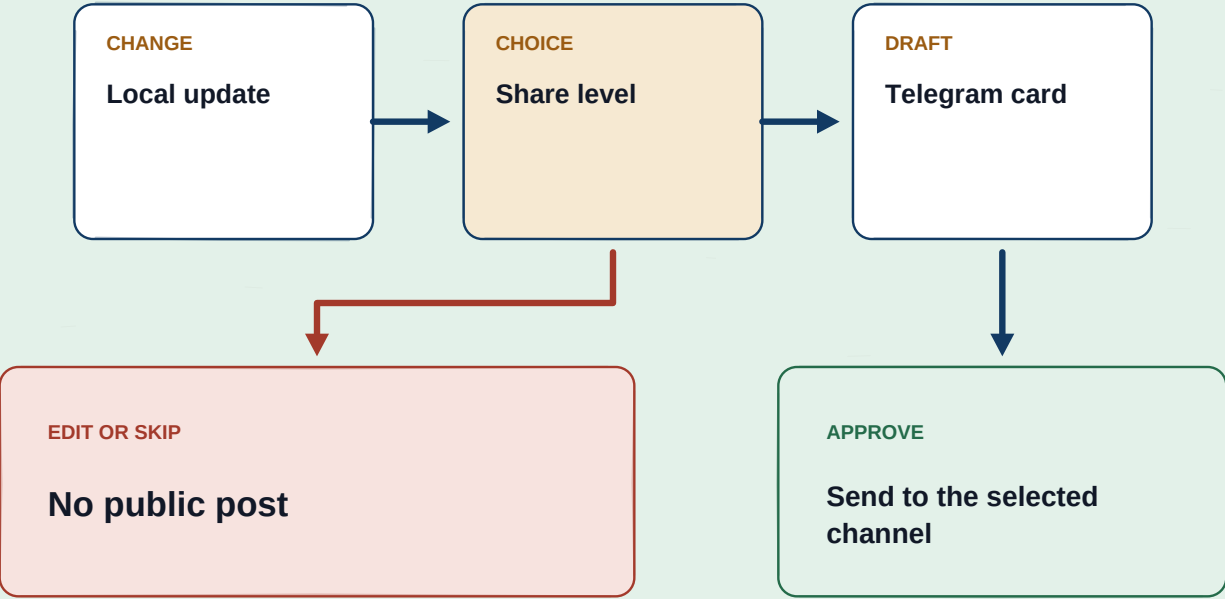
AGENT BOUNDARY

Prepare a change. Open a review. Do not publish or deploy.

CHAPTER 05

Release notes start as drafts

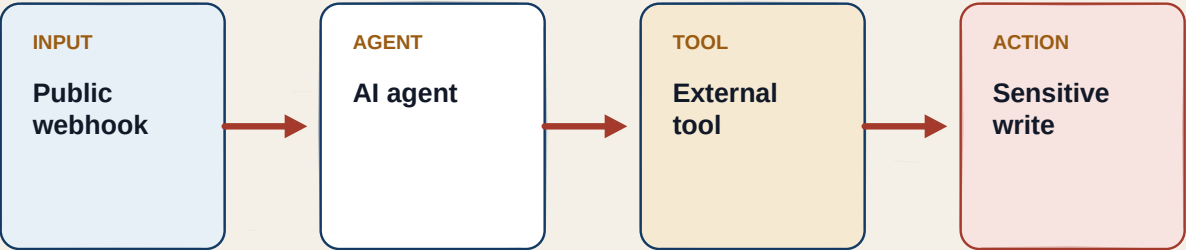
A local change is classified as major or minor or no share. Only useful changes move to a draft.



CHAPTER 06

Why workflow security came next

A workflow can connect public input to a model. The model can reach a tool. The tool can reach a sensitive action.



**Can untrusted input reach
a powerful action without control?**

CHAPTER 07

Read the export without running it

The scanner treats an exported workflow as data. It follows visible paths between nodes. It never activates the workflow.

10/100

Unsafe example with nine findings

93/100

Hardened example with one medium item

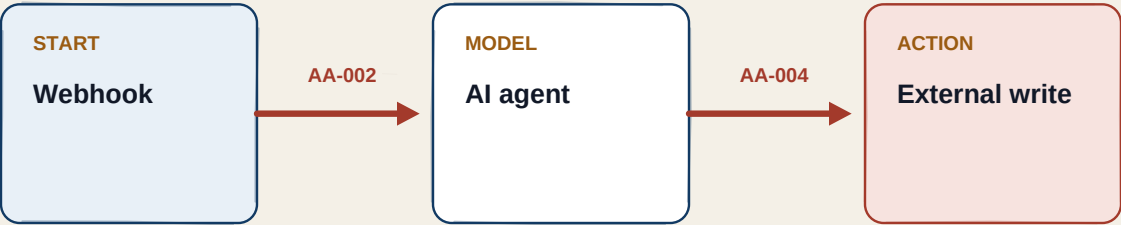
IMPORTANT

The score is a review aid. It is not proof that a workflow is secure.

CHAPTER 08

A graph turns findings into a path

The graph keeps finding IDs on the edges. The same IDs appear in JSON and SARIF output.



FINDINGS LEDGER

AA-002 Public trigger needs review

AA-004 Model path reaches a write action

CHAPTER 09

Test behavior in an isolated target

Static review shows what an export appears to allow. The runtime gate checks how a safe staging target responds.

8/8

Recorded scenarios passed

56/56

Recorded checks passed

0

External actions

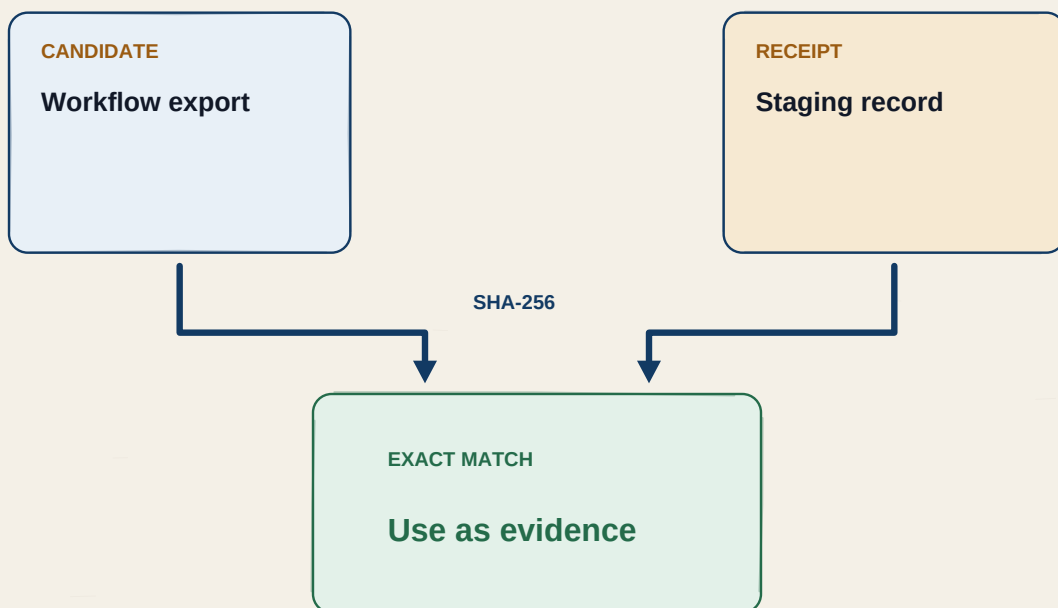
THE GATE CHECKS

- Authentication and input shape
- Approval bypass and replay behavior
- Prompt injection containment around the workflow

CHAPTER 10

Bind evidence to the exact candidate

The receipt records a SHA-256 fingerprint. A change review accepts it only when the candidate fingerprint matches.

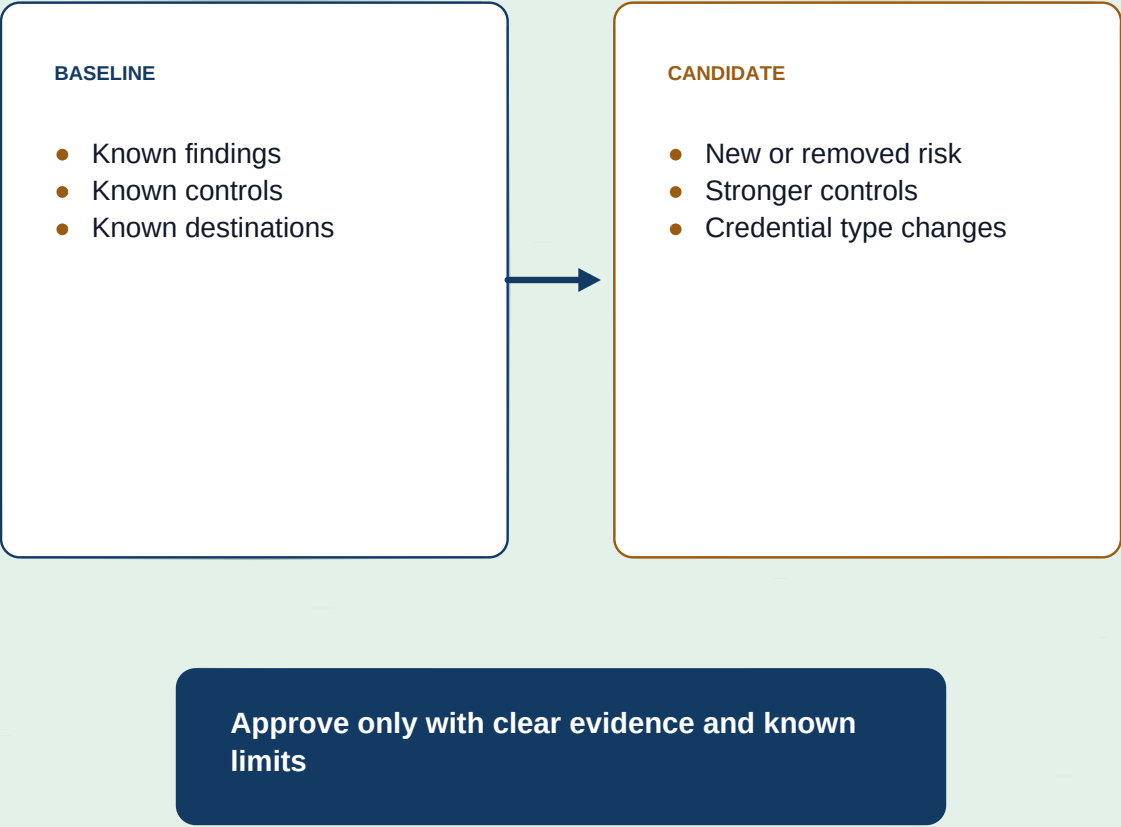


A receipt is a local record. It is not a signed attestation.

CHAPTER 11

Compare before release

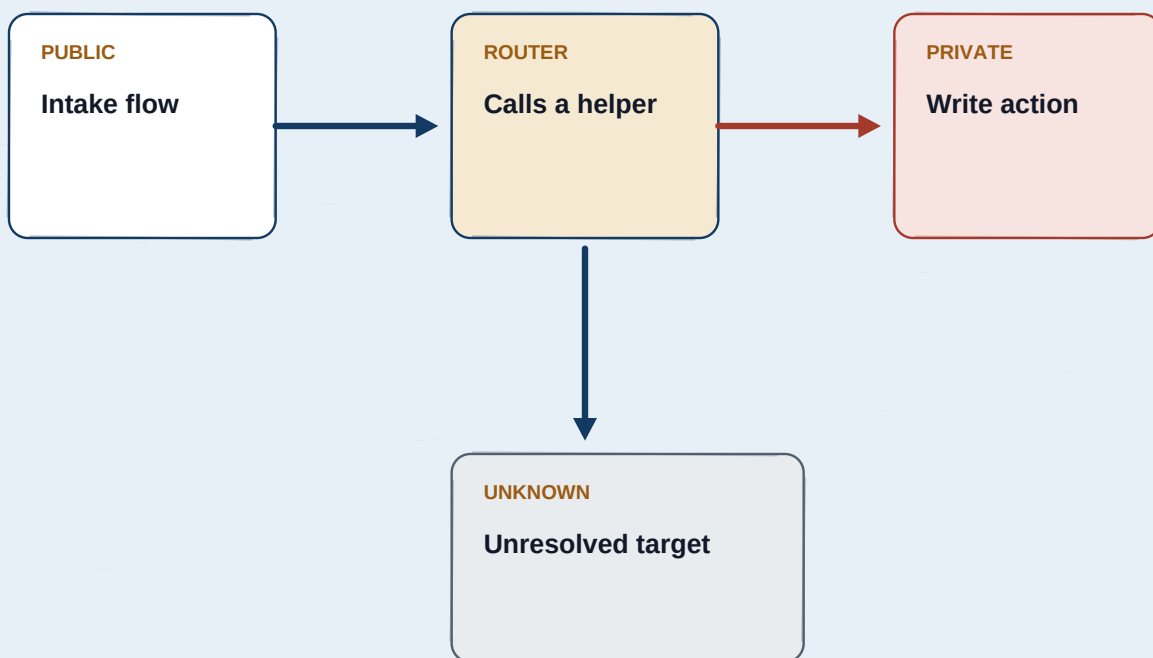
The change review compares a baseline with a candidate. It records what became safer and what needs more review.



CHAPTER 12

One workflow can call another

The map resolves calls between exported workflows when stable IDs are present. Unknown targets stay visible.



Credential names become local aliases such as credential-01.

CHAPTER 13

Research signals need a clear boundary

The project scanned one hundred popular free AI Agent templates from the official n8n template API.

100

Templates scanned as JSON only

52

Queued for path level review

0

Templates imported or run

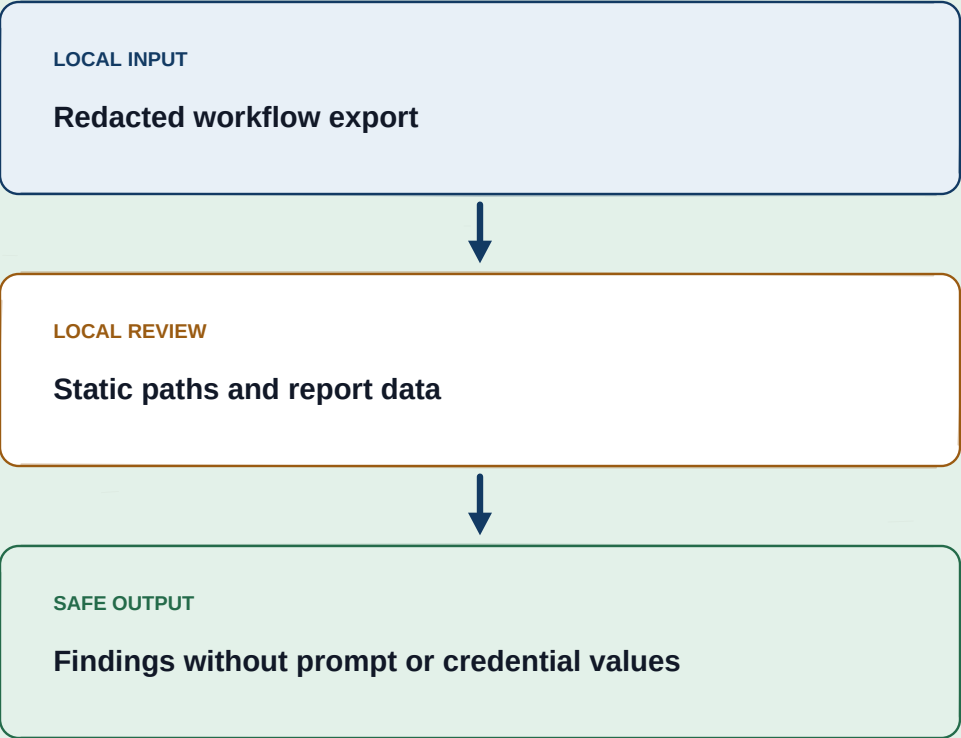
PUBLICATION BOUNDARY

A scanner signal is not a confirmed vulnerability. High severity paths still need manual validation.

CHAPTER 14

Keep sensitive workflow data local

Workflow exports stay on the local machine during the scanner flow. The scanner does not send them to an AI API.

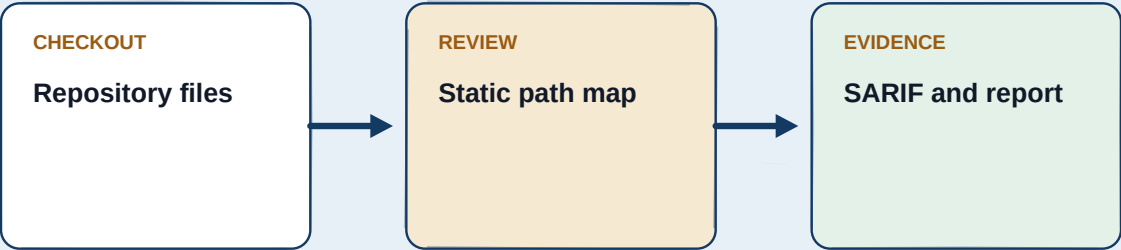


Remove secrets and customer data before you keep or share an export.

CHAPTER 15

Review each pull request locally

The reusable GitHub Action reads exported workflow JSON as untrusted data. It never imports or executes a workflow.



PULL REQUEST SUMMARY

High findings can fail the job only after JSON, Markdown, SARIF, Mermaid and SVG evidence is written.

No n8n credential. No AI service. No release approval.

CHAPTER 16

Five commands to start

Use Node.js 20 or newer. No npm package installation is required.

```
$ npm test
```

```
$ npm run audit
```

```
$ npm run gate:demo
```

```
$ npm run map:demo
```

```
$ npm run pr-gate:demo
```

Use redacted files. Test only systems you own or are allowed to test.

WHAT I LEARNED

Simple systems are easier to trust

ONE JOB

Each step should have one clear purpose.

VISIBLE CONTROL

Human approval should be easy to see.

EVIDENCE

Safety checks should leave a useful record.

HONEST LIMITS

Unknown facts should stay unknown.

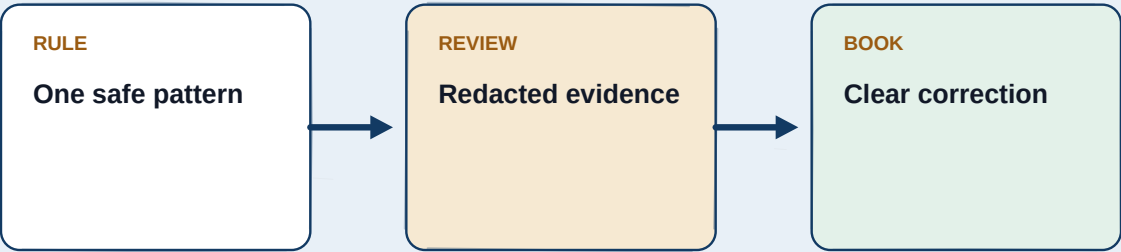
The graph keeps bringing me back to one question.

Where can this input go?

OPEN CONTRIBUTION

Small evidence makes a useful contribution

The public repository accepts narrow rules, redacted false positives and corrections to this field guide.



CONTRIBUTION BOUNDARY

Never post credentials, customer workflows, private URLs or production payloads in an issue.

github.com/0xCD4/n8n-ai-agent-security-lab
en.csintresearch.org/n8n-systems-field-guide