

05 Evidence records

1

All 3High 0Medium 2Review 0Information 1

2

AS-203Medium

Third-party action is not pinned to a full commit SHA

actions/checkout uses a mutable reference.

3

Locationn8n-ai-agent-security-lab-main/.github/workflows/test.yml:18

4

Next human checkPin the action to a reviewed full-length commit SHA and record the human-readable version in a comment.

AS-203Medium

Third-party action is not pinned to a full commit SHA

actions/setup-node uses a mutable reference.

Locationn8n-ai-agent-security-lab-main/.github/workflows/test.yml:21

Next human checkPin the action to a reviewed full-length commit SHA and record the human-readable version in a comment.

Finding the instructions and permissions in an agent repository

I open a repository archive from GitHub in the browser. AI Agent Surface Map collects the files that steer an agent, its MCP connections and its GitHub Actions permissions into one review list.

1 What the tool is for

What an AI agent can do is often written in small configuration files, not in the main code. An instruction file gives the agent rules. An MCP file connects a new tool to it. A GitHub Actions file can open write access to the repository. These files sit in different folders and are easy to miss.

I built this tool to see the parts that give an agent power one by one, instead of putting a single safe or unsafe label on a repository. The tool only reads the archive. It runs no file and connects to no server.

1. Archive	2. Instructions	3. Tools	4. Outside action
The ZIP file downloaded from GitHub.	AGENTS.md, CLAUDE.md, skill and Copilot files.	MCP servers and visible command, network and browser traces.	GitHub Actions permissions and secret names.

A few words

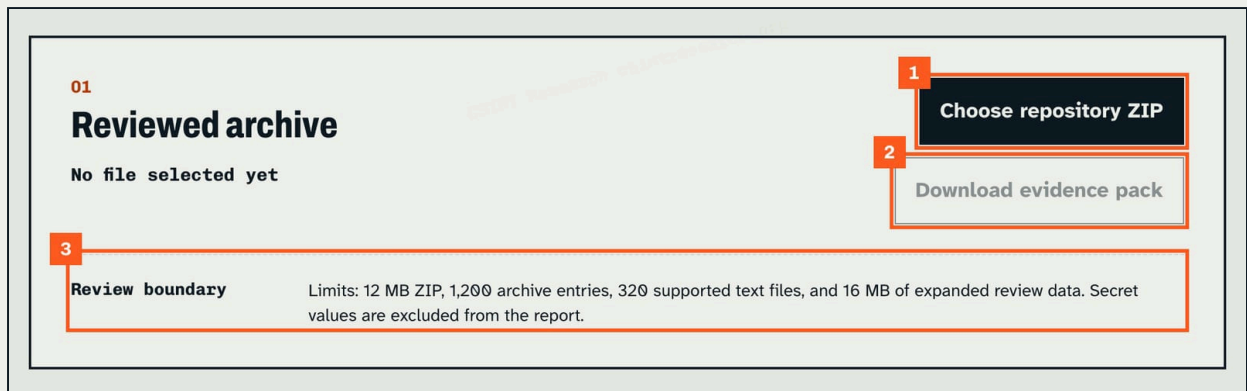
Agent	An AI program that can use tools, write files or run commands on its own.
MCP	Model Context Protocol. Configuration that connects outside tools to an agent. It usually lives in a file such as .mcp.json.
GitHub Actions	Jobs that run by themselves when something happens in the repository. Their settings are in the .github/workflows folder.
Secret	A hidden value such as a password or key. The tool only sees its name and never writes the value into the report.
Pinning	Tying a package or Action to one exact version that will not change. An unpinned part can bring different code tomorrow.

2 Before you start

You need	The ZIP file of the repository you want to review. On GitHub, use the Code button and then Download ZIP.
Limits	Up to 12 MB of ZIP, 1,200 archive entries, 320 supported text files and 16 MB of expanded data. Single files larger than 512 KB are not read.
Skipped folders	.git, node_modules, vendor, dist, build, .next, coverage and __pycache__.
Account	Not needed.
Where data goes	The ZIP opens in your browser. Nothing is sent to a server. Code and MCP servers are not run.

If a large monorepo goes over a limit, the tool stops and says why. Then download only the project folder you care about as a separate ZIP.

3 Opening the archive



The archive area. The site ships no sample archive; the area waits for your own ZIP.

1 Choose your own ZIP

The file is read in the browser as soon as you pick it. Within seconds the sections below fill in for the new archive.

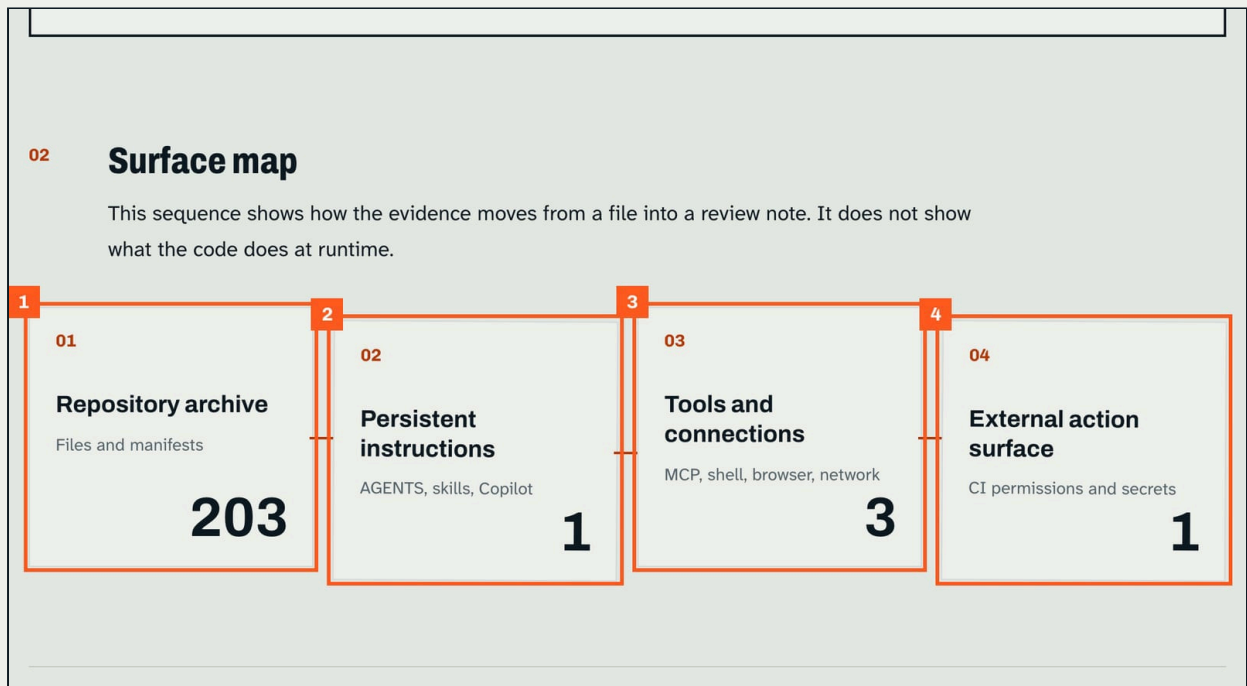
2 Download the evidence pack

One ZIP downloads. It holds the same result as JSON and as Markdown.

3 Read the boundary

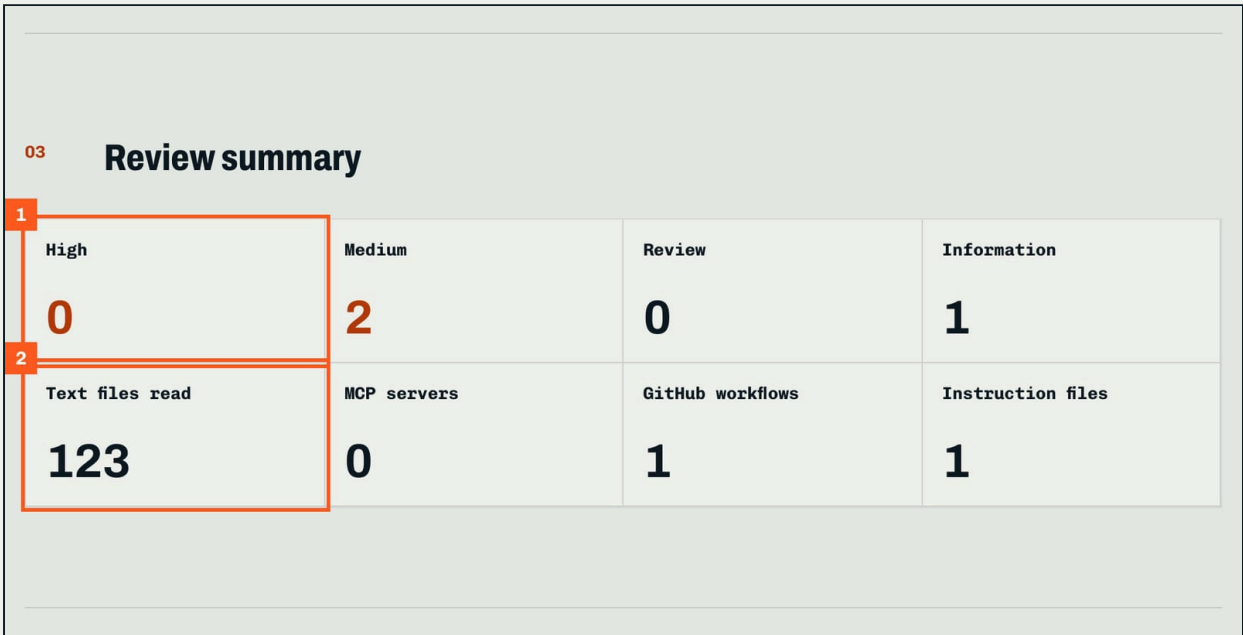
It says how much is read and that secret values are left out of the report.

4 Reading the map and the summary



The surface map, for the GitHub ZIP of my open-source repository [n8n-ai-agent-security-lab](https://github.com/n8n-ai-agent-security-lab). It shows the order in which evidence moves from a file into a review note, not what the code does at runtime.

- 1 Repository archive**
The number of entries in the archive.
- 2 Persistent instructions**
The number of instruction files found. It counts AGENTS.md, CLAUDE.md, GEMINI.md, copilot-instructions.md and SKILL.md inside a skills folder.
- 3 Tools and connections**
The number of MCP servers plus the number of visible capability types.
- 4 External action surface**
The number of GitHub Actions files that were read.



The review summary. Pressing a severity box filters the list below to that severity.

- 1 Severity counts**
High, Medium, Review and Information. Pressing a box narrows the evidence list to that group.
- 2 What was read**
Counts of text files read, MCP servers, GitHub workflows and instruction files. If a file you expected is not counted, check that it is inside the archive.

Another part is the *Visible capabilities* list. The tool looks in text files for certain library and command names tied to local command, network access, browser automation, file writing, OSINT and threat intelligence connectors. A name appearing in a file does not show that the capability is really used.

5 Evidence records

The screenshot displays the 'Evidence records' interface. At the top, there's a header '05 Evidence records'. Below it, a filter bar shows 'All 3', 'High 0', 'Medium 2', 'Review 0', and 'Information 1'. The first record is highlighted with a red box and numbered 1 through 4. The record details are as follows:

AS-203	Medium
Third-party action is not pinned to a full commit SHA actions/checkout uses a mutable reference.	
Location	n8n-ai-agent-security-lab-main/.github/workflows/test.yml:18
Next human check	Pin the action to a reviewed full-length commit SHA and record the human-readable version in a comment.

The second record is also shown below, with identical details.

Evidence records. Each record carries a rule number, a severity, a location and the next human check.

1 Filter

Narrow the list by severity. The number in brackets is how many records are in that group.

2 Rule and severity

The number starting with AS tells you which rule matched. The same number appears in the JSON and Markdown report.

3 Location

The file path and, when known, the line number. Open this file yourself before you decide anything.

4 Next human check

What you should look at for this record. The tool fixes nothing. It only tells you where to look.

When the tool opens a record

AS-001	The archive holds an environment file such as .env. High. Its values are not read.
AS-100	An MCP file was found but no server list could be read from it. Review.
AS-101	An MCP server starts a local command. Medium.
AS-102 / 103	A remote MCP address is present. Review for HTTPS, High for plain HTTP.
AS-104	An MCP server asks for environment variable names. Medium.
AS-105	npx, uvx or pipx is used without a pinned version. Medium.
AS-201	An Actions file uses the pull_request_target trigger. High.
AS-202	An Actions file grants write-all permission. High.
AS-203	A third-party Action is not pinned to a full commit SHA. Medium.
AS-204	A job triggered by pull requests uses secrets. Medium.
AS-301	The repository has a persistent agent instruction file. Information.

6 MCP records and the evidence pack

If the archive has an MCP configuration, each server is listed with its file, transport and command or remote address. The repository I used for the screenshots has no MCP file, so these cards are not shown here.

Local server	stdio transport means a command will be started on the computer. Only the names of environment keys are shown.
Remote server	remote transport means data will go to another address. The tool does not connect to it. You need to find out separately who owns it and what it keeps.

At the end I download the evidence pack. The ZIP holds two files, ai-agent-surface-report.json and ai-agent-surface-report.md. Both carry the summary counts, each record's rule number and location, and the limits. Secret values are written to neither file.

7 What it does not do

- ✗ It does not run code, scripts, Action packages or MCP servers. It connects to no remote address.

- × A record is not proof of an exploit, malware or a bad project. No records does not mean the repository is safe either.
- × Branch protection, repository settings, real runtime permissions, dependency advisories and commit origin cannot be seen from an archive.
- × It only looks at supported file names and text patterns. A permission defined somewhere else can be missed.

Clean private data out of the archive before you keep it or share it. The tool does not write values into the report, but the ZIP file itself is still yours to look after.