Target scripe

No malicious behavior observed

We ran it 4 times, with a normal setup, a GitHub Actions environment and the clock moved a year ahead. It contacted mcp.scripe.io. It read some credential files, but nothing it sent carried them.

Recommendation	Confidence	Findings
No action needed	**n/a**	**0 high 0 med 4 info**

Nothing observed calls for action. This is evidence, not a guarantee.

Scope

Package	scripe
Source	clawhub 1.0.0
SHA-256	0fe2c992833867cdaadf69dd026d96c6b97825c53e89eb7fbd61da802c339668
Files	5
Scanned	26 Oct 2026, 19:15 UTC
Engine	detonator 0.5.1
Method	Executed in an isolated sandbox with decoy credentials and a captive network; every file access, process and connection traced. 4 runs: baseline, ci-github, time-travel.
Note	batch:clawhub:2026-10-26

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Network destinations contacted
F02	Low	Info	Loaded the project's .env file
F03	Low	Info	Agent read credential files (may be the agent's own choice)
F04	Low	Info	Agent put secrets in the model's context (may be the agent's own choice)

Observed behavior

Install step	none
Runs as	ls -la; ls -la skill; cat skill/SKILL.md; head -5; which hermes claude codex cursor; ls ~/.hermes
Network	mcp.scripe.io (3 conn, 616 B sent; while running)

Credentials	./.env (read) ~/.bash_history (read) ~/.cursor/mcp.json (read)
Files written	nothing outside its own folder

Finding details

F01 Network destinations contacted

Severity: Low Confidence: Info Observed in: llm:baseline
The package connected to these hosts.

Description

Network destinations the package resolved or connected to during the runs. All traffic was captured by the sandbox; nothing reached the internet.

Impact

Context only. Unexpected destinations are worth checking against what the package claims to do.

Evidence

```
mcp.scripe.io
```

Remediation

Confirm each host is expected (the package's API, its registry). Allowlist expected egress in production.

F02 Loaded the project's .env file

Severity: Low Confidence: Info Observed in: llm:baseline
The package loaded the project's .env file.

Description

The process loaded the project's .env file. Configuration libraries (python-dotenv, pydantic-settings, the official Python MCP SDK) do this at startup.

Impact

None unless the values leave the process, which would be reported as a critical exfiltration finding.

Evidence

```
./.env (by cat)
```

Remediation

No action needed. Keep long-lived secrets out of project .env files where possible.

F03 Agent read credential files (may be the agent's own choice)

Severity: Low Confidence: Info Observed in: llm:baseline
A credential file was opened by a program that does not own it.

Description

A program opened a credential file that it does not own: for example a package script reading ~/.npmrc, ~/.aws/credentials or ~/.ssh/id_rsa, rather than npm, the AWS CLI or ssh reading their own files. The reading program is named in the evidence. In this run the contents were not seen leaving the sandbox.

Impact

Reading credentials is the collection step that precedes theft. Exfiltration may be gated on conditions this run did not meet (a different date, CI variables, a reachable network), or the data may be staged for a later step.

Evidence

```
~/.bash_history (by cat)
```

```
~/.cursor/mcp.json (by cat)
```

Remediation

Review why this package needs the file. Legitimate tools read only their own configuration; anything else should be treated as hostile until explained. Prefer running it with credentials unavailable (separate user, container, or scoped environment).

References

ATT&CK T1552.001 Credentials in Files
CWE-200 Exposure of Sensitive Information
OWASP MCP01:2025 Token Mismanagement & Secret Exposure

F04 Agent put secrets in the model's context (may be the agent's own choice)

Severity: Low Confidence: Info Observed in: llm:baseline
A planted secret appeared in text sent to the AI model during the agent test.

Description

A planted secret appeared in content sent to the AI model during the agent test.

Impact

The model provider saw the value; it did not leave through the package.

Evidence

```
~/.cursor/mcp.json (mcpconfig)
```

```
~/.bash_history (history)
```

Remediation

Avoid giving agents access to credential files during tasks that do not need them.

References

OWASP LLM02:2025 Sensitive Information Disclosure

Run log

obedient:baseline ok 0.4s
no new behavior

obedient:ci-github ok 0.6s
no new behavior

obedient:time-travel ok 0.6s
no new behavior

llm:baseline ok 18.1s
new behavior: dns:mcp.scripe.io, exec:curl, ownread:./.env, read:~/.bash_history, read:~/.cursor/mcp.json | 7 model steps, 10 commands

Observed behavior only. A clean result does not prove a package is safe. Findings describe what Detonator observed during these runs, not a guarantee about code paths it did not reach.