



Target buyer-delivery-qa

No malicious behavior observed

We ran it 1 time, with a normal setup. It made no network connections. It read some credential files, but nothing it sent carried them.

Recommendation	Confidence	Findings
No action needed	n/a	0 high 0 med 2 info

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

Scope

Package	buyer-delivery-qa
Source	clawhub 1.0.0
SHA-256	6504074a9e78bfabae2a71b3eda522f280acbbcb4782dfb05692eeb9438fe5
Files	4
Scanned	27 Oct 2026, 12:58 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. 1 runs: baseline.
Note	batch:clawhub:2026-10-27

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Loaded the project's .env file
F02	Low	Info	Agent read credential files (may be the agent's own choice)

Observed behavior

Install step	none
Runs as	ls -la; find . -maxdepth 3 -not -path; head -100; head -50; cat skill/SKILL.md; ls -la skill
Network	no connections
Credentials	./.env (read) ~/.bash_history (read)

Files written nothing outside its own folder

Finding details

F01 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 grep, sed)
```

Remediation

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

F02 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)
```

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



Run log

llm:baseline ok 15.2s

new behavior: ownread:./.env, read:~/.bash_history | 5 model steps, 7 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.