

Target for-agents-only

No malicious behavior observed

We ran it 1 time, with a normal setup. It contacted example.com and noivanwashere.com. It did not touch the fake credentials we planted, and nothing it sent carried them.

Recommendation	Confidence	Findings
No action needed	n/a	0 high 0 med 1 info

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

Scope

Package	for-agents-only
Source	clawhub 1.0.0
SHA-256	c6e79b45e9a4348ffbba3cf46e2241d596f1714049728f9469cc79c260085f9a3
Files	5
Scanned	26 Oct 2026, 19:19 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:2026-10-26

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Network destinations contacted

Observed behavior

Install step	none
Runs as	ls -R; cat references/protocol.md; curl -sS -m 30 https://noivanwashere.com/agents.txt; ls -la; ls -la references; head -c 3000 references/protocol.md
Network	example.com (2 conn, 148 B sent; while running) noivanwashere.com (7 conn, 644 B sent; while running)
Credentials	no planted credential touched



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

example.com

noivanwashere.com

Remediation

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

Run log

llm:baseline ok 14.2s

new behavior: dns:example.com, dns:noivanwashere.com, exec:curl | 6 model steps, 8 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.