

Target `optiq-local-models`

No malicious behavior observed

We ran it 4 times, with a normal setup, a GitHub Actions environment and cloud credentials present. It made no network connections. 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 0 info

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

Scope

Package	<code>optiq-local-models</code>
Source	<code>clawhub 1.0.0</code>
SHA-256	<code>89717d2e3e13ab00eec13c4d0d05992ccdefd1af240dea6eca2ea37eb259c052</code>
Files	<code>4</code>
Scanned	<code>26 Oct 2026, 19:33 UTC</code>
Engine	<code>detonator 0.5.1</code>
Method	<code>Executed in an isolated sandbox with decoy credentials and a captive network; every file access, process and connection traced. 4 runs: baseline, ci-github, cloud-creds.</code>
Note	<code>batch:clawhub:2026-10-26</code>

Observed behavior

Install step	<code>none</code>
Runs as	<code>uname -m; awk {printf "%.0f GB\n", \$1/1073741824}; python3 --version; curl http://127.0.0.1:8080/v1/chat/completions -H Content-Type: application/json -d {"mode; uname -a; cat /etc/os-release</code>
Network	<code>127.0.0.1:8080:8080 (3 conn, 837 B sent; while running)</code>
Credentials	<code>no planted credential touched</code>
Files written	<code>nothing outside its own folder</code>

Run log

`obedient:baseline` `ok` `1.6s`
`new behavior: exec:curl`



obedient:ci-github ok 1.7s

no new behavior

obedient:cloud-creds ok 1.8s

no new behavior

llm:baseline ok 6.8s

no new behavior | 3 model steps, 3 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.