



Target `untax-research`

No malicious behavior observed

We ran it 1 time, with a normal setup. It contacted `untax.io`. 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	<code>untax-research</code>
Source	<code>clawhub 0.4.3</code>
SHA-256	<code>4814bce6191a78113ba9fcf879fba0a5dc6f3bb3c144ddd305eebb1d23422578</code>
Files	3
Scanned	26 Oct 2026, 17:30 UTC
Engine	<code>detonator 0.5.1</code>
Method	Executed in an isolated sandbox with decoy credentials and a captive network; every file access, process and connection traced. 1 runs: <code>baseline</code> .
Note	<code>batch:clawhub:2026-10-26</code>

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Network destinations contacted

Observed behavior

Install step	<code>none</code>
Runs as	<code>ls -la; ls -la skill; head -50; cat skill/SKILL.md; cat README.md; cat package.json</code>
Network	<code>untax.io (1 conn, 74 B sent; while running)</code>
Credentials	<code>no planted credential touched</code>
Files written	<code>nothing outside its own folder</code>



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

```
untax.io
```

Remediation

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

Run log

llm:baseline ok 13.5s

new behavior: dns:untax.io, exec:curl | 5 model steps, 4 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.