

Target mochilab-skills

No malicious behavior observed

We ran it 4 times, with a normal setup, WATCH_REPORT_DEPENDENCIES set and SystemRoot set. It contacted mochilab.bz. 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 2 info

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

Scope

Package	mochilab-skills
Source	clawhub 1.0.0
SHA-256	74cb3a93e6c0f0a003e0be143d1f88cb7069c2f53ce47c5bdf6ec3a91915ba0e
Files	6
Scanned	26 Oct 2026, 19:43 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, guided:WATCH_REPORT_DEPENDENCIES, guided:SystemRoot.
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

Observed behavior

Install step	none
Runs as	node <skill>/scripts/mochilab.mjs; ls -la ./skill; ls -la ./skill/scripts; node --version; ls -la; head -20
Network	mochilab.bz (3 conn, 419 B sent; while running)
Credentials	./.env (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

```
mochilab.bz
```

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.

Run log

scripts:baseline ok 1.5s

no new behavior

scripts:guided:WATCH_REPORT_DEPENDENCIES ok 0.9s

no new behavior

scripts:guided:SystemRoot ok 0.7s

no new behavior

**llm:baseline ok 15.2s**

new behavior: dns:mochilab.bz, exec:curl, ownread:./env | 8 model steps, 11 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.