



Target novel-dialogue-subtext-cn

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	novel-dialogue-subtext-cn
Source	clawhub 1.0.1
SHA-256	c14878061a178754604e55514c8075b05f9f7a968639741e6f77f6c829ba668b
Files	14
Scanned	27 Oct 2026, 11:19 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. 4 runs: baseline, ci-github, cloud-creds.
Note	batch:clawhub:2026-10-27

Observed behavior

Install step	none
Runs as	python3 <skill>/scripts/self_test.py; ls -R; cat SKILL.md; head -5; cat fixtures/success-example.md; cat fixtures/failure-missing-input.md
Network	no connections
Credentials	no planted credential touched
Files written	nothing outside its own folder

Run log

scripts:baseline ok 1.0s
no new behavior

**scripts:ci-github** **ok** **1.2s**

no new behavior

scripts:cloud-creds **ok** **0.8s**

no new behavior

llm:baseline **ok** **14.2s**

no new behavior | 7 model steps, 6 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.