

Target bilibili-video-to-text

No malicious behavior observed

We ran it 7 times, with a normal setup, LC_MESSAGES and LINKDIGEST_API_KEY set and the environment variables it says it needs. 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 1 info

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

Scope

Package	bilibili-video-to-text
Source	clawhub 1.0.0
SHA-256	f6806b5cc854923192c94174bfc817c4fa18e091fe29bd047f1524727e15159c
Files	4
Scanned	27 Oct 2026, 13:11 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. 7 runs: baseline, guided:LC_MESSAGES, LINKDIGEST_API_KEY, declared-env.
Note	batch:clawhub:2026-10-27

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Loaded the project's .env file

Observed behavior

Install step	none
Runs as	python3 <skill>/scripts/linkdigest.py; python3 {baseDir}/scripts/linkdigest.py https://www.bilibili.com/video/BV1FtHg6nEE2/; ls -la ./skill; ls -la ./skill/scripts; grep -rl ld_live_ .; cat _meta.json
Network	no connections
Credentials	./.env (read)



Files written nothing outside its own folder

Finding details

F01 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 grep, sed)
```

Remediation

No action needed. Keep long-lived secrets out of project .env files where possible.

Run log

scripts:baseline ok 1.6s

no new behavior

scripts:guided:LC_MESSAGES, LINKDIGEST_API_KEY ok 1.7s

no new behavior

scripts:declared-env ok 1.4s

no new behavior

obedient:baseline ok 0.7s

no new behavior

obedient:declared-env ok 1.0s

no new behavior

obedient:guided:LC_MESSAGES, LINKDIGEST_API_KEY ok 0.6s

no new behavior

llm:baseline ok 17.9s

new behavior: ownread:./.env | 8 model steps, 7 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.