

Target fluxa-agent-wallet

No malicious behavior observed

We ran it 8 times, with a normal setup, a GitHub Actions environment, cloud credentials present, a production environment and everything. It contacted fluxapay.xyz and registry.npmjs.org. 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 3 info

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

Scope

Package	fluxa-agent-wallet
Source	clawhub 1.0.0
SHA-256	0672b947432ec1aeea4d1097cfc5c42046a8582b43ab8d4c7602189121e2fea5
Files	17
Scanned	27 Oct 2026, 12:43 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. 8 runs: baseline, ci-github, cloud-creds, production, everything, ci-gitlab.
Note	batch:clawhub:2026-10-27

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Used a planted key with the service it belongs to (registry.npmjs.org)
F02	Low	Info	Network destinations contacted
F03	Low	Info	A tool read its own credentials (e.g. gh, aws, ssh)

Observed behavior

Install step	package manager only, no install scripts
Runs as	open <URL>; dbus-send --print-reply --dest=org.freedesktop.DBus /org/freedesktop/DBus org.freedesktop.; xprop -root _DT_SAVE_MODE; grep = \"xfce4\"\$; xprop -root; grep -i ^xfce_desktop_window



Network	fluxapay.xyz (2 conn, 166 B sent; while running) registry.npmjs.org (23 conn, 5987 B sent; during install)
Credentials	~/.npmrc (own tool, normal)
Files written	nothing outside its own folder

Finding details

F01 Used a planted key with the service it belongs to (registry.npmjs.org)

Severity: Low Confidence: Info Observed in: obedient:baseline

A planted key went to the service it belongs to, as its login (npm to its registry, an OpenAI key to api.openai.com).

Description

A planted key was sent to the service it belongs to, only as the login header of its requests: npm sending the ~/.npmrc token to registry.npmjs.org while installing, or an app calling api.openai.com with the OpenAI key from .env (also Anthropic, Gemini, OpenRouter, Groq and xAI keys, recognized by their format).

Impact

Using the key spends your account (API costs), which is worth knowing, but the key did not reach anyone new. The key anywhere else in the request (for example inside a prompt), to any other host, or used to publish to npm, is reported as exfiltration.

Evidence

```
~/.npmrc (npm) sent only as the login header of requests to registry.npmjs.org
```

Remediation

No action needed.

F02 Network destinations contacted

Severity: Low Confidence: Info Observed in: obedient:baseline, 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

```
registry.npmjs.org
```

```
fluxapay.xyz
```

Remediation

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



F03

A tool read its own credentials (e.g. gh, aws, ssh)

Severity: Low Confidence: Info Observed in: obedient:baseline
A tool opened its own login or config file, such as npm reading ~/.npmrc.

Description

A tool opened its own configuration or login file, such as npm reading ~/.npmrc or gh reading its hosts file.

Impact

None in itself: this is how those tools work. Listed so the report shows every credential access.

Evidence

```
~/.npmrc (by npm)
```

Remediation

No action needed.

Run log

obedient:baseline ok 13.2s

new behavior: dns:registry.npmjs.org, ownread:~/.npmrc, write:~/.npm/_cacache/content-v2/sha512/a8/6e/*, write:~/.npm/_cacache/index-v5/1d/6b/*, write:~/.npm/_cacache/tmp/866bc3c0

obedient:ci-github ok 11.7s

new behavior: write:~/.npm/_cacache/content-v2/sha512/fc/97/*, write:~/.npm/_cacache/tmp/af04d2de

obedient:cloud-creds ok 9.7s

new behavior: write:~/.npm/_cacache/content-v2/sha512/46/ca/*, write:~/.npm/_cacache/tmp/bf4ff70c

obedient:production ok 12.2s

new behavior: write:~/.npm/_cacache/content-v2/sha512/f6/0c/*, write:~/.npm/_cacache/tmp/e968bf54

obedient:everything ok 10.2s

new behavior: write:~/.npm/_cacache/content-v2/sha512/ed/59/*, write:~/.npm/_cacache/tmp/e5a4ef5d

obedient:ci-gitlab ok 10.5s

new behavior: write:~/.npm/_cacache/content-v2/sha512/54/fa/*, write:~/.npm/_cacache/tmp/93e7a5f1

obedient:time-travel ok 13.9s

no new behavior

llm:baseline ok 62.2s

new behavior: dns:fluxapay.xyz, exec:curl, write:~/.npm/_cacache/content-v2/sha512/c0/e1/*, write:~/.npm/_cacache/tmp/0e8c3b9a, write:~/.npm/_cacache/tmp/2e1c1ffb | 7 model steps, 10 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.