



Target agent-browser-clawdbot

Malicious behavior observed

~/npmrc (npm) left the package on a normal run, headed for registry.npmjs.org. The sandbox caught it; nothing reached the internet.

Recommendation	Confidence	Reproduced	Findings
Block	Confirmed	2 of 2 runs	1 high 0 med 2 info

Do not install. If it is already installed, remove it and rotate any credentials it could reach.

Scope

Package	agent-browser-clawdbot
Source	clawhub
SHA-256	a6c1e27139e924d4139a77ff48ba65c108bcb0ea7d0063a99b4f1ee3079b23d4
Files	3
Scanned	26 Oct 2026, 15:10 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. 1 runs: baseline.

Findings summary

ID	Severity	Confidence	Finding
F01	Critical	Confirmed	Secrets sent over the network via tcp/443 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
Network	registry.npmjs.org (2 conn, 836 B sent, CARRIED ~/npmrc; during install)
Credentials	~/npmrc (read and sent out)
Files written	nothing outside its own folder

Finding details

F01 Secrets sent over the network via tcp/443 to registry.npmjs.org

Severity: Critical Confidence: Confirmed Observed in: obedient:baseline
A unique honeypot marker was found in data the package sent out.

Description

The package transmitted the contents of a credential file to a network destination. Before every run Detonator plants decoy credentials (AWS keys, SSH private key, npm and GitHub tokens, a .env file), each carrying a unique per-run marker. That marker was recovered from outbound traffic captured by the sandbox's network sink, after decoding (base64, hex, URL encoding, chunked DNS labels). This is direct proof that credential material left the process, not an inference from code.

Impact

On a developer workstation or CI runner this hands the operator of the destination host working credentials: registry tokens allow publishing malicious versions of your packages (a worm-style supply chain compromise), cloud keys allow resource abuse and data access, and SSH keys allow lateral movement to every host that trusts them.

Evidence

```
~/ .npmrc (npm) left the sandbox
```

Remediation

Do not install. If it was installed anywhere, treat every credential readable by that user as compromised: revoke and reissue npm/PyPI/GitHub tokens, rotate cloud keys, replace SSH keys, and review registry publish history and cloud audit logs for the exposure window. Block the destination host at egress and search proxy/DNS logs for other machines that contacted it.

References

CWE-506 Embedded Malicious Code

CWE-200 Exposure of Sensitive Information

ATT&CK T1552.001 Credentials in Files

ATT&CK T1041 Exfiltration Over C2 Channel

OWASP MCP01:2025 Token Mismanagement & Secret Exposure

F02 Network destinations contacted

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

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 19.5s

new behavior: dns:registry.npmjs.org, exfil:tcp/443, ownread:~/.npmrc, write:~/.npm/_cacache/content-v2/sha512/1d/ed/*,
write:~/.npm/_cacache/index-v5/3d/cc/*

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.