



Target @upstash/context7-mcp

No malicious behavior observed

We ran it 4 times, with a normal setup, the clock moved a year ahead, a GitHub Actions environment and an AI agent using it. It started as an MCP server, exposed 2 tools and we called them 24 times. It contacted context7.com. 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	@upstash/context7-mcp
Source	npm 4.4.4
SHA-256	c75a7791d86cba9a2ef04e62c0ee5d74376a4b5d0d6717f3effba6602a164f5a
Files	29
Scanned	25 Oct 2026, 18:42 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, time-travel, ci-github, agent.

Findings summary

ID	Severity	Confidence	Finding
F01	Low	Info	Network destinations contacted
F02	Low	Info	A tool read its own credentials (e.g. gh, aws, ssh)
F03	Low	Info	Tool metadata: Description exceeds 2000 characters

Observed behavior

Install step	package manager only, no install scripts
Runs as	node dist/index.js
Tools exposed	2 (24 calls tested): resolve-library-id, query-docs
Network	context7.com (27 conn, 43308 B sent; while running)



Credentials	~/.npmrc (own tool, normal)
Files written	nothing outside its own folder

Finding details

F01 Network destinations contacted

Severity: Low Confidence: Info Observed in: mcp: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

```
context7.com
```

Remediation

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

F02 A tool read its own credentials (e.g. gh, aws, ssh)

Severity: Low Confidence: Info Observed in: mcp: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.

F03 Tool metadata: Description exceeds 2000 characters

Severity: Low Confidence: Info Observed in: mcp:baseline
Found in text the server shows the agent; nothing harmful was observed from it.

Description

A tool description is unusually long.

Impact

Long descriptions can hide instructions, but are often thorough documentation.

Evidence

tools[resolve-library-id].description

Resolves a package/product name to a Context7-compatible library ID and returns matching libraries.

You MUST call this function before 'Query D... (len=2006)

Remediation

Skim the full text in the evidence.

Run log

mcp:baseline ok 31.1s

new behavior: dns:context7.com, ownread:~/.npmrc, write:<skill>/node_modules/.bin/context7-mcp,
write:~/.npm/_logs/*-debug-0.log | install scripts: exit 0; restarted once; 8 tool calls

mcp:time-travel ok 14.4s

no new behavior | install scripts: exit 0; 8 tool calls

mcp:ci-github ok 25.2s

no new behavior | install scripts: exit 0; 8 tool calls

mcp-agent:agent ok 21.8s

no new behavior | install scripts: exit 0

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.