

url-linter.js — CEO-facing URL Verification Gate

url-linter.js — Usage Runbook

Owner: CodeCraft **Created:** 2026-04-22 **Purpose:** Prevent broken URLs from reaching CEO. Enforces ZAKON URL.

When to Use

Run `pre-ceo-publish.sh` BEFORE sending ANY of these to Alem:

- Partner shortlists
- Market analysis documents
- Research briefs with external sources
- Decision documents (`*-VERIFIED.md`, `*ceo-decision*`, `*ceo-brief*`)
- Any file in `~/system/specs/`

Trigger rule: "Every URL in a CEO-facing document MUST pass url-linter.js before publish."

Quick Start

```
# Gate check before CEO delivery (REQUIRED)
bash ~/system/hooks/pre-ceo-publish.sh ~/system/specs/my-doc.md

# Direct linter run
node ~/system/tools/url-linter.js ~/system/specs/my-doc.md

# Scan entire specs directory
node ~/system/tools/url-linter.js ~/system/specs/
```

```
# JSON output (for scripts)
```

```
node ~/system/tools/url-linter.js ~/system/specs/my-doc.md --json
```

All Options

```
node url-linter.js <file-or-dir> [options]
```

Options:

<code>--fix</code>	Replace redirected URLs with final URLs; remove NXDOMAIN URLs and append a FIXME flag
<code>--json</code>	Output JSON report (machine-readable)
<code>--fail-on=<N></code>	Exit 1 if N or more broken URLs (default: 1)
<code>--concurrency=<N></code>	Max parallel checks (default: 5)
<code>--timeout=<ms></code>	Per-URL curl timeout in ms (default: 15000)
<code>--cache=<path></code>	SQLite cache path (default: ~/system/state/url-linter-cache.db)
<code>--no-cache</code>	Skip cache, always re-check every URL
<code>--cache-ttl=<h></code>	Cache TTL in hours (default: 6)
<code>--quiet</code>	Only show failures and summary

Exit Codes

Code	Meaning
0	All URLs pass (or below fail-on threshold)
1	One or more broken URLs found (at or above fail-on threshold)
2	Usage error or fatal (file not found, etc.)

What Gets Checked

URL patterns extracted from markdown:

- `[text](https://url)` — link syntax
- `<https://url>` — angle bracket URLs

- ``https://url`` — inline-code/backtick-wrapped URLs (fixed MC #9100 2026-08-08: previously silently skipped, letting a backtick-cited broken URL pass as "0 URLs found")
- `https://url` — bare URLs

Test method: `curl -sSIL --max-time <N> <url>`

Result	Classification
HTTP 2xx	PASS
HTTP 3xx followed by 2xx	PASS (shown as REDIRECT)
HTTP 4xx, 5xx	FAIL
NXDOMAIN (cannot resolve)	FAIL
Timeout	FAIL
Connection refused	FAIL

Cache

Results are cached in `~/system/state/url-linter-cache.db` (SQLite).

- TTL: 6 hours by default
- Re-running on same file with same URLs uses cache — completes in < 1s
- Override: `--no-cache` to force fresh check
- Cache is safe to delete: `rm ~/system/state/url-linter-cache.db`

PostToolUse Hook (automatic)

The hook `~/system/hooks/url-linter-gate.sh` runs automatically after Write/Edit on CEO-facing files. It runs async (does not block the write) but prints a warning to stderr if broken URLs are found.

CEO-facing patterns that trigger automatic check:

- `*/system/specs/*`
- `*/evidence/ceo-*`
- `*-VERIFIED.md`
- `*-shortlist*`
- `*ceo-decision*`
- `*ceo-brief*`
- `*ceo-report*`

Hook activation is present in `~/.claude/settings.json` under `PostToolUse` for `Write|Edit|MultiEdit`.
The canonical hook block is:

```
{
  "matcher": "Write|Edit|MultiEdit",
  "hooks": [
    {
      "type": "command",
      "command": "bash ~/system/hooks/url-linter-gate.sh",
      "timeout": 60000,
      "async": true
    }
  ]
}
```

Pre-CEO Publish Script

```
bash ~/system/hooks/pre-ceo-publish.sh <file>
```

This wraps url-linter.js with:

- Formatted output header
- Slack-shareable block on failure
- Exit 0 (safe) / Exit 1 (blocked)

Make this the last step before any CEO delivery.

Retroactive Audit

To scan all existing specs for broken URLs:

```
node ~/system/tools/url-linter.js ~/system/specs/ --json > /tmp/url-audit.json
```

Or run the full retroactive audit:

```
bash ~/system/hooks/pre-ceo-publish.sh ~/system/specs/
```

Incident Context

Origin: 2026-04-24 HR-FISK incident. John published `www.porezna-uprava.gov.hr/bi/Stranice/Popis-informacijskih-posrednika.aspx` (NXDOMAIN) to CEO in a "VERIFIED" document. URL was copied from web-search output without curl verification.

Actual working URL: `porezna-uprava.gov.hr/hr/popis-informacijskih-posrednici/8019` (HTTP 200).

Root cause: RC5 (Petter Graff audit) — citation propagation without re-verification. This tool closes that gap permanently.

ZAKON URL (exact text)

“ Every external URL and email domain published in a CEO-facing document must have a fresh HTTP verification within 10 minutes of document delivery. Verification method: `curl -sI --max-time 10` returning 2xx or 3xx. NXDOMAIN, 4xx, or timeout = document blocked. No exceptions for "found in registry" or "from web search". Web search results are unverified input, not verified facts.

Revision #2

Created 2026-07-29 03:57:22 UTC by John

Updated 2026-08-08 05:16:05 UTC by John