

OVERALL SCORE

50 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include AI guidance file and sitemap, while plain-text page access and structured data need attention. Recommended next step: add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.

Recommended next step

1. Add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.
2. Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access; serve a non-empty HTML homepage with a canonical link tag.
3. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

Signal breakdown

Crawlability

Warn10/20

Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access; serve a non-empty HTML homepage with a canonical link tag.

Robots.txt

Warn7.5/15

Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.

llms.txt

Pass15/15

The llms.txt file was found and includes the expected text, length, heading, and URL signals.

Sitemap	Pass	10/10
The sitemap.xml file is valid and contains URL entries.		
Markdown support	Fail	0/15
Add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.		
Semantic HTML	Warn	7.1/10
Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: main, article.		
Structured data	Fail	0/10
Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.		
Content signals	Fail	0/5
Add the standard directive 'Content-Signal: ai-train=no, search=yes, ai-input=yes' to robots.txt, HTML metadata, or HTTP headers so AI systems can discover content usage preferences.		

Suggested fixes

Fix Structured data

HTML

```
<script type="application/ld+json">
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://postreef.com/#organization",
"name": "PostReef Turn any post into typed data",
"description": "Paste any social post URL, get back typed JSON that conforms to your schema - transcript, metadata and AI extraction, every field, every time.",
"url": "https://postreef.com/",
"logo": "https://postreef.com/apple-icon?e2c1f46794cc84d6"
},
{
"@type": "WebSite",
"@id": "https://postreef.com/#website",
"name": "PostReef Turn any post into typed data",
"description": "Paste any social post URL, get back typed JSON that conforms to your schema - transcript, metadata and AI extraction, every field, every time.",
"url": "https://postreef.com/",
"publisher": {
"@id": "https://postreef.com/#organization"
}
},
{
"@type": "Text",
"@id": "https://postreef.com/#text",
"name": "PostReef Turn any post into typed data",
"description": "Paste any social post URL, get back typed JSON that conforms to your schema - transcript, metadata and AI extraction, every field, every time.",
"url": "https://postreef.com/",
"publisher": {
"@id": "https://postreef.com/#organization"
}
}
]
}
```

Fix Content signals

HTML

```
<meta http-equiv="Content-Signal" content="ai-train=no, search=yes, ai-input=yes" />
<meta name="content-signal" content="ai-train=no, search=yes, ai-input=yes" />
```

Full report

<https://llmscan.dev/scan/1LFSgKHUiIKDGOkd3qpqU>