

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 homepage access and crawler policy need attention. Recommended next step: remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Recommended next step

- 1. Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.
- 2. Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.
- 3. 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.

Signal breakdown

|                                                                                                                                                             |      |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-------|
| <p>Crawlability</p> <p>Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.</p>              | Fail | 0/20  |
| <p>Robots.txt</p> <p>Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.</p> | Fail | 0/15  |
| <p>llms.txt</p> <p>The llms.txt file was found and includes the expected text, length, heading, and URL signals.</p>                                        | Pass | 15/15 |

|                                                                                                                                                                                                                               |      |       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-------|
| <b>Sitemap</b>                                                                                                                                                                                                                | Pass | 10/10 |
| The sitemap index is valid and the checked child sitemap contains URL entries.                                                                                                                                                |      |       |
| <b>Markdown support</b>                                                                                                                                                                                                       | 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.                                 |      |       |
| <b>Semantic HTML</b>                                                                                                                                                                                                          | Pass | 10/10 |
| The homepage has a valid title, meta description, heading structure, semantic landmarks, substantial body copy, and descriptive anchor text.                                                                                  |      |       |
| <b>Structured data</b>                                                                                                                                                                                                        | Pass | 10/10 |
| Valid JSON-LD structured data was found with core Organization or WebSite schema.org types.                                                                                                                                   |      |       |
| <b>Content signals</b>                                                                                                                                                                                                        | Pass | 5/5   |
| Consider adding Content-Signal HTTP header, AI-specific head meta tags, robots noai/noimageai directive so AI systems can consistently discover content usage preferences across robots.txt, HTTP headers, and HTML metadata. |      |       |

# Suggested fixes

## llms.txt

### MARKDOWN

# DocxPacer - Auto Type .docx into Google Docs | Free Chrome Extension

> Auto-type any Word .docx into Google Docs at a natural human pace. Free Chrome extension; Pro keeps your bold, italics & alignment. Resume across days.

This llms.txt file summarizes the public, canonical resources that AI assistants and crawlers should use to understand this site.

### ## Site Overview

- Canonical URL: <https://docxpacer.xyz/>
- Site type: organization
- Recommended summary: Auto-type any Word .docx into Google Docs at a natural human pace. Free Chrome extension; Pro keeps your bold, italics & alignment. Resume across days.

### ## Core URLs

- [Homepage](<https://docxpacer.xyz/>): Primary public entry point and canonical site overview.
- [Install free](<https://chromewebstore.google.com/detail/bpdilfcebpecmmhneelnhdojnhheiaj>): Important public page discovered from the homepage navigation.
- [Start] Continued in the full scan report...

## robots.txt additions

### TXT

```
# robots.txt additions
# Copy these blocks into the existing robots.txt file. Keep current rules
unless a note calls out a conflicting Disallow.

# AI crawler access
# Add explicit Allow rules for blocked AI crawlers; remove or narrow
conflicting Disallow rules if your crawler target requires precedence.
User-agent: GPTBot
Allow: /

User-agent: ChatGPT-User
Allow: /

User-agent: ClaudeBot
Allow: /

User-agent: Claude-Web
Allow: /

User-agent: PerplexityBot
Allow: /

User-agent: Google-Extended
Allow: /

# Crawl-rate guidance
Continued in the full scan report...
```

## schema.json

### JSON

```
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://docxpacer.xyz/#organization",
"name": "DocxPacer - Auto Type .docx into Google Docs | Free Chrome
Extension",
"description": "Auto-type any Word .docx into Google Docs at a natural human
pace. Free Chrome extension; Pro keeps your bold, italics & alignment.
Resume across days.",
"url": "https://docxpacer.xyz/",
"logo": "https://docxpacer.xyz/icons/icon180.png"
},
{
"@type": "WebSite",
"@id": "https://docxpacer.xyz/#website",
"name": "DocxPacer - Auto Type .docx into Google Docs | Free Chrome
Extension",
"description": "Auto-type any Word .docx into Google Docs at a natural human
pace. Free Chrome extension; Pro keeps your bold, italics & alignment.
Resume across days.",
"url": "https://docxpacer.xyz/",
"publisher": {
"@id": "https://docxpacer.xyz/#organization"
}
}
]
Continued in the full scan report...
```

## Content-Signal

**TXT**

# Content-Signal recommendations

Use these directives to make AI-use preferences explicit for compliant crawlers and AI systems. They are advisory signals, so keep them aligned with robots.txt, terms, and access controls.

## Recommended values

- ai-train=no: AI model training, fine-tuning, and dataset creation.
- search=yes: AI search indexing, snippets, and discovery.
- ai-input=yes: AI answer grounding, retrieval, and generated-response context.

## HTTP response header

```
'''http
Content-Signal: ai-train=no, search=yes, ai-input=yes
Content-Usage: train-ai=n
'''
```

Best for site-wide or route-specific policies because the signal travels with every response, including pages that AI systems fetch directly.

## HTML meta tag alternatives

Add these inside the document '`<head>`' when server headers are not

continued in the full scan report...

## Full report

[https://llmscan.dev/scan/F-Qz\\_pKTKnI\\_8FXAI\\_sGY](https://llmscan.dev/scan/F-Qz_pKTKnI_8FXAI_sGY)