

OVERALL SCORE

36 /100 Poor

Executive summary

This site is difficult for AI tools to read right now. 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

Crawlability	Fail	0/20
Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.		
Robots.txt	Fail	0/15
Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.		
llms.txt	Pass	15/15
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	5.7/10
Add exactly one h1 element that describes the page topic. Add missing semantic elements: main, article, nav, footer. Add visible body copy after script and style removal so the page has at least 200 words.		
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	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

Fix Structured data

HTML

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://langpanda.com/#organization",
      "name": "Learn Languages from Netflix, YouTube & 20+ Sites - LangPanda",
      "description": "LangPanda helps you learn 36 languages by watching your favorite shows. Instant dictionary, flashcard creation, vocabulary tracking. 7-day free trial.",
      "url": "https://langpanda.com/",
      "logo": "https://langpanda.com/apple-touch-icon.png"
    },
    {
      "@type": "WebSite",
      "@id": "https://langpanda.com/#website",
      "name": "Learn Languages from Netflix, YouTube & 20+ Sites - LangPanda",
      "description": "LangPanda helps you learn 36 languages by watching your favorite shows. Instant dictionary, flashcard creation, vocabulary tracking. 7-day free trial.",
      "url": "https://langpanda.com/",
      "publisher": {
        "@id": "https://langpanda.com/#organization"
      },
      "inLanguage": "en"
    }
  ]
}
```

Continued in the full scan report...

Full report

<https://www.llmscan.dev/scan/IPjUs3K1MyEM787H7BL7z>