

OVERALL SCORE

15 /100 Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include sitemap and content signals, 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. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

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 Fail 0/15

Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

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	Fail	0/10
Add a title tag between 10 and 70 characters. Add a meta description between 50 and 160 characters. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: 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	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://www.mangaka.app/en#organization",
      "name": "Mangaka",
      "description": "Mangaka is an AI manga generator that helps anyone create manga without drawing skills. Write a script, choose a manga style, and generate beautiful manga in minutes.",
      "url": "https://www.mangaka.app/en",
      "logo": "https://www.mangaka.app/logo-with-text.png",
      "sameAs": [
        "https://www.linkedin.com/company/thinkprompt/",
        "https://x.com/mangakadotapp",
        "https://www.facebook.com/mangakadotapp",
        "https://www.youtube.com/@MangakaStories-n6y",
        "https://www.tiktok.com/@mangaka_app"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://www.mangaka.app/en#website",
      "name": "Mangaka",
      "description": "Mangaka is an AI manga generator that helps anyone create
Continued in the full scan report...
```

Full report

<https://llmscan.dev/scan/2scjnaTuFULLiHrYc6ozA>