

OVERALL SCORE

31 /100

Poor

Executive summary

This site is difficult for AI tools to read right now. The main areas needing attention are AI guidance file and plain-text page access. Recommended next step: publish an AI guidance file as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Recommended next step

1. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.
2. 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.
3. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

Signal breakdown

Crawlability Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.	Warn 10/20
Robots.txt Publish /robots.txt with explicit crawler rules and Sitemap references so search and AI crawlers can discover site policy.	Warn 15/15
llms.txt Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.	Fail 0/15

Sitemap	Fail	0/10
Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.		
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
Shorten the title tag to 70 characters or fewer. Shorten the meta description to 160 characters or fewer. 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://marpy.io/#organization",
      "name": "Browser-based IDE + AI assistant + Kubernetes deployment built for Python developers",
      "description": "marpy.io is an AI-powered development environment built specifically for the Python/Flask/FastAPI/MariaDB/Redis/Jinja/Tailwind stack. Write Python in the browser, attach a database that survives between deploys, and get a live URL.",
      "url": "https://marpy.io/",
      "logo": "https://marpy.io/static/images/bunny-logo-128.ico"
    },
    {
      "@type": "WebSite",
      "@id": "https://marpy.io/#website",
      "name": "Browser-based IDE + AI assistant + Kubernetes deployment built for Python developers",
      "description": "marpy.io is an AI-powered development environment built specifically for the Python/Flask/FastAPI/MariaDB/Redis/Jinja/Tailwind stack. Write Python in the browser, attach a database that survives between deploys, and get a live URL.",
      "url": "https://marpy.io/"
    }
  ]
}
```

Continued in the full scan report...

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://www.llmscan.dev/scan/kQf2P5CRBduSLki7lBrSo>