

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	Warn	10/20
Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.		
Robots.txt	Warn	15/15
Publish /robots.txt with explicit crawler rules and Sitemap references so search and AI crawlers can discover site policy.		
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	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
Use exactly one h1 element and move secondary section titles to h2-h6. 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	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://startupsubmit.app/#organization",
      "name": "StartupSubmit | Top Directory Submission Service for SaaS &
Startup",
      "description": "StartupSubmit is the premier directory submission service.
We manually submit to 250+ high-authority startup directories to boost your
SEO and traffic.",
      "url": "https://startupsubmit.app/",
      "logo":
        "https://t2.gstatic.com/faviconV2?client=SOCIAL&type=FAVICON&fallback_opts=TYPE,SIZE,URL&url=https://bolta.ai/&size=128",
      "sameAs": [
        "https://x.com/marclou",
        "https://x.com/DStartupsubmit",
        "https://x.com/AlexBelogubov",
        "https://x.com/arthurmuzbashew",
        "https://x.com/FlorinPop17",
        "https://x.com/levelsio",
        "https://x.com/Ziyigong_",
        "https://twitter.com/DStartupsubmit?ref_src=twsrc%5Etfw",
        "https://twitter.com/marclou/status/2018401294902579210?ref_src=twsrc%5Etfw"
      ]
    }
  ]
}
```

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/mARYSr0pj1ufpKRZa4Wqy>