

OVERALL SCORE

51 /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 plain-text page access and structured data need attention. Recommended next step: 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.

Recommended next step

- 1. 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.
- 2. Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access; serve a non-empty HTML homepage with a canonical link tag.
- 3. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

Signal breakdown

Crawlability Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access; serve a non-empty HTML homepage with a canonical link tag.	Warn 10/20
Robots.txt Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access.	Warn 7.5/15
llms.txt The llms.txt file was found and includes the expected text, length, heading, and URL signals.	Pass 15/15

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	8.6/10
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://www.webclarity.ai/#organization",
      "name": "WebClarityAI - Clear Understanding for Better Websites",
      "description": "Built for creative designers, developers, and SEO experts to inspect and analyze website reports in seconds.",
      "url": "https://www.webclarity.ai/",
      "logo": "https://www.webclarity.ai/apple-icon.png"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.webclarity.ai/#website",
      "name": "WebClarityAI - Clear Understanding for Better Websites",
      "description": "Built for creative designers, developers, and SEO experts to inspect and analyze website reports in seconds.",
      "url": "https://www.webclarity.ai/",
      "publisher": {
        "@id": "https://www.webclarity.ai/#organization"
      },
      "inLanguage": "en"
    }
  ]
}
```

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://llmscan.dev/scan/tWJsSLqi6Cujuh0JphIbl>