

OVERALL SCORE

61 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include crawler policy and plain-text page access, while AI guidance file and content signals need attention. 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. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
- 3. Prefer valid JSON-LD and include Organization or WebSite schema.org types in valid JSON-LD for clearer AI interpretation.

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 robots.txt allows crawler access and includes Sitemap references.	Pass	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 The sitemap.xml file is valid and contains URL entries.	Pass	10/10

Markdown support	Pass	15/15
The homepage returns markdown content when requested with Accept: text/markdown, giving AI systems a cleaner text representation with less navigation chrome to parse.		
Semantic HTML	Warn	5.7/10
Add exactly one h1 element that describes the page topic. Add missing semantic elements: article. Add visible body copy after script and style removal so the page has at least 200 words.		
Structured data	Warn	5/10
Prefer valid JSON-LD and include Organization or WebSite schema.org types in valid JSON-LD for clearer AI interpretation.		
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
<pre><script type="application/ld+json"> { "@context": "https://schema.org", "@graph": [{ "@type": "Organization", "@id": "https://www.thisweekinstartups.com/#organization", "name": "This Week in Startups - 5,000+ Episodes, Searchable Archive", "description": "Browse and search 5,000+ episodes of This Week in Startups, the #1 startup podcast hosted by Jason Calacanis. Practical lessons from Day 0 to Unicorn.", "url": "https://www.thisweekinstartups.com/", "logo": "https://www.thisweekinstartups.com/apple-touch-icon.png", "sameAs": ["https://www.youtube.com/@ThisWeekinAIPodcast", "https://www.youtube.com/channel/UckkhmBwFS7pILYIk0izkc3A", "https://x.com/TwiStartups", "https://www.youtube.com/@startups", "https://x.com/twistartups", "https://www.linkedin.com/company/this-week-in-startups", "https://www.tiktok.com/@thisweekinstartups"] }, { "@type": "WebSite", "@id": "https://www.thisweekinstartups.com/#website", "url": "https://www.thisweekinstartups.com/" }] }</pre>
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/ahQDW-Yv8uLoP55tfFvM3>