

OVERALL SCORE

33 /100 Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include sitemap, while AI guidance file and plain-text page access 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. 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	7.5/15
Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.		
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	Warn	5.7/10
Shorten the meta description to 160 characters or fewer. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: main, article, footer.		
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.microns.io/#organization",
      "name": "Micro Startups For Sale - Microns.io",
      "description": "Discover micro startups for sale with zero commission on Microns. Explore a wide range of online businesses and find your next opportunity without hidden fees. Join our community of entrepreneurs today.",
      "url": "https://www.microns.io/",
      "logo":
        "https://cdn.prod.website-files.com/60fa8972eef72783b828a3a6/64e513fcc13669f5f9ce7400_android-chrome-192x192.png",
      "sameAs": [
        "https://www.youtube.com/watch?v=8MRfLGluF9Y&ab_channel=IlyaNovohatskyi",
        "https://twitter.com/micronsio",
        "https://twitter.com/joshuardavis/status/1633498496757035008",
        "https://twitter.com/jspeiser/status/1631414165075120134",
        "https://twitter.com/samcraigjohnson/status/1557671659506909184",
        "https://twitter.com/inorganik/status/1438153704352325636",
        "https://twitter.com/Kamphey/status/1433548618569445405",
        "https://www.linkedin.com/company/microns-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/DrU4jg44UfM-95A5RRYvS>