

OVERALL SCORE

11 /100

Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include content signals, while homepage access and crawler policy need attention. Recommended next step: remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Recommended next step

- 1. Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.
- 2. Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.
- 3. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Signal breakdown

Crawlability

Fail 0/20

Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Robots.txt

Fail 0/15

Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.

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
Make the homepage respond successfully to requests with Accept: text/markdown, either by serving markdown directly or by adding a markdown variant URL with correct content negotiation.		
Semantic HTML	Warn	5.7/10
Expand the title tag to at least 10 characters. Use exactly one h1 element and move secondary section titles to h2-h6. Add missing semantic elements: article, nav.		
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	Pass	5/5
Consider adding Content-Signal HTTP header, AI-specific head meta tags, robots noai/noimageai directive so AI systems can consistently discover content usage preferences across robots.txt, HTTP headers, and HTML metadata.		

Suggested fixes

Fix Structured data

HTML

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://affilio.live/#organization",
      "name": "Affilio",
      "description": "Discover the best affiliate links, coupon codes and exclusive deals shared by your favorite creators. All in one place.",
      "url": "https://affilio.live/",
      "logo": "https://affilio.live/favicon.svg",
      "sameAs": [
        "https://x.com/provanshh",
        "https://www.linkedin.com/in/provanshh/"
      ],
    },
    {
      "@type": "WebSite",
      "@id": "https://affilio.live/#website",
      "name": "Affilio",
      "description": "Discover the best affiliate links, coupon codes and exclusive deals shared by your favorite creators. All in one place.",
      "url": "https://affilio.live/",
      "publisher": {
        "@id": "https://affilio.live/#organization"
      }
    }
  ]
}
```

Full report

<https://llmscan.dev/scan/9jy1Q3LU6-f1G3GNMK9vg>