

OVERALL SCORE

58 /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 content signals 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. Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.

Signal breakdown

Crawlability Warn 10/20

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.

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 Pass 15/15

The llms.txt file was found and includes the expected text, length, heading, and URL signals.

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	Pass	10/10
The homepage has a valid title, meta description, heading structure, semantic landmarks, substantial body copy, and descriptive anchor text.		
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

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://www.replyhey.com/#organization",
      "name": "ReplyHey: find your customers on Reddit",
      "description": "ReplyHey finds the Reddit threads where people are asking
for what you built, scores them daily, and drafts the reply. You decide what
to send.",
      "url": "https://www.replyhey.com/",
      "logo": "https://www.replyhey.com/icon.svg?icon.0v5f4nq0rpja9.svg"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.replyhey.com/#website",
      "name": "ReplyHey: find your customers on Reddit",
      "description": "ReplyHey finds the Reddit threads where people are asking
for what you built, scores them daily, and drafts the reply. You decide what
to send.",
      "url": "https://www.replyhey.com/",
      "publisher": {
        "@id": "https://www.replyhey.com/#organization"
      }
    },
    {
      "inLanguage": "en"
    }
  ]
}
```

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/7f5VINMXMGlb_oVPk26Q0