

OVERALL SCORE

38 /100 Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include crawler policy, while plain-text page access and sitemap 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. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
- 3. Regenerate sitemap.xml with valid XML and a sitemap <urlset> or <sitemapindex> root element.

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 Add Sitemap directives to robots.txt so crawlers can find canonical sitemap files.	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.	Warn 7.5/15

Sitemap	Fail	0/10
Regenerate sitemap.xml with valid XML and a sitemap <urlset> or <sitemapindex> root element.		
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
Add exactly one h1 element that describes the page topic. Add missing semantic elements: main, article, nav, footer. Add visible body copy after script and style removal so the page has at least 200 words.		
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.jorukagulaam.com/#organization",
      "name": "Joru Ka Gulaam - Pata Karo Kaun Hai Asli Gulaam ",
      "description": "Quiz lo. Apna relationship title pao. Partner ko tag karo. 2 minute ka quiz, lifetime ka roast.",
      "url": "https://www.jorukagulaam.com/",
      "logo": "https://www.jorukagulaam.com/heart-logo.png"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.jorukagulaam.com/#website",
      "name": "Joru Ka Gulaam - Pata Karo Kaun Hai Asli Gulaam ",
      "description": "Quiz lo. Apna relationship title pao. Partner ko tag karo. 2 minute ka quiz, lifetime ka roast.",
      "url": "https://www.jorukagulaam.com/",
      "publisher": {
        "@id": "https://www.jorukagulaam.com/#organization"
      },
      "inLanguage": "hi"
    }
  ]
}
```

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/KmSwiqFXUiWtvfjaTaxJd>