

OVERALL SCORE

53 /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 structured data 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. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

Signal breakdown

|                                                                                                                                                                                                       |                    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| <p><b>Crawlability</b></p> <p>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.</p> | <p>Warn 10/20</p>  |
| <p><b>Robots.txt</b></p> <p>Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access.</p>                                                              | <p>Warn 7.5/15</p> |
| <p><b>llms.txt</b></p> <p>The llms.txt file was found and includes the expected text, length, heading, and URL signals.</p>                                                                           | <p>Pass 15/15</p>  |

|                                                                                                                                                                                               |      |       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-------|
| <b>Sitemap</b>                                                                                                                                                                                | Pass | 10/10 |
| The sitemap index is valid and the checked child sitemap contains URL entries.                                                                                                                |      |       |
| <b>Markdown support</b>                                                                                                                                                                       | 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. |      |       |
| <b>Semantic HTML</b>                                                                                                                                                                          | Pass | 10/10 |
| The homepage has a valid title, meta description, heading structure, semantic landmarks, substantial body copy, and descriptive anchor text.                                                  |      |       |
| <b>Structured data</b>                                                                                                                                                                        | Fail | 0/10  |
| Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.                                                                  |      |       |
| <b>Content signals</b>                                                                                                                                                                        | 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.notion.com/#organization",
      "name": "The AI workspace that works for you. | Notion",
      "description": "Build Custom Agents, search across all your apps, and
      automate busywork. The AI workspace where teams get more done, faster.",
      "url": "https://www.notion.com/",
      "logo":
      "https://www.notion.com/_next/image?url=https%3A%2F%2Fimages.ctfassets.net%2
      Fspqsf9291f%2FjWkCPMAAj0Av69704DQoB%2F84499b606028b276e0bf94c238f07097%2F0
      penAI-black-wordmark-cropped.png&w=3840&q=75",
      "sameAs": [
        "https://www.instagram.com/notionhq/",
        "https://twitter.com/NotionHQ",
        "https://www.linkedin.com/company/notionhq/",
        "https://www.facebook.com/NotionHQ/",
        "https://www.youtube.com/channel/UCoSvlWS5XcwaSzIcbuJ-Ysg"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://www.notion.com/#website",
      "url": "https://www.notion.com/"
    }
  ]
}
```

## 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/1srluKXdhoehORqzEsetd>