

OVERALL SCORE

43 /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 page structure, 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. 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. Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.

Signal breakdown

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

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
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	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://notecove.io/#organization",
      "name": "NoteCove - Your notes. Your devices. No one else’s servers.",
      "description": "A safe harbor for what matters. Local-first notes and tasks that sync through your own cloud storage.",
      "url": "https://notecove.io/",
      "logo": "https://notecove.io/images/favicons/apple-touch-icon.png",
      "sameAs": [
        "https://www.youtube.com/watch?v=HqrYdpEfPQs",
        "https://www.youtube.com/watch?v=tFyVPaX97MI",
        "https://www.youtube.com/watch?v=BimxhrMzk9U",
        "https://www.youtube.com/@LossyBrain"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://notecove.io/#website",
      "name": "NoteCove - Your notes. Your devices. No one else’s servers.",
      "description": "A safe harbor for what matters. Local-first notes and tasks that sync through your own cloud storage.",
      "url": "https://notecove.io/"
    }
  ]
}
```

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/AUQBD5yc1ywUXsKGkCuXh>