

OVERALL SCORE

41 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include 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

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 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	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	Warn	5.7/10
Shorten the meta description to 160 characters or fewer. Add missing semantic elements: article. 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.teardowntheory.com/#organization",
      "name": "Teardown Theory - engineering, explained from the inside out",
      "description": "Teardown Theory is an engineering explainer channel produced by an autonomous AI system. Watch how the built world actually works - and how the machine that makes the videos works, too.",
      "url": "https://www.teardowntheory.com/",
      "logo": "data:image/svg+xml,<svg xmlns='http://www.w3.org/2000/svg' viewBox='0 0 100 100'><circle cx='50' cy='50' r='38' fill='%23f2a93b'/></svg>",
      "sameAs": [
        "https://www.youtube.com/channel/UCLd-9jpwj6dogbWTi1EqqNg",
        "https://www.youtube.com/channel/UCLd-9jpwj6dogbWTi1EqqNg?sub_confirmation=1"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://www.teardowntheory.com/#website",
      "name": "Teardown Theory - engineering, explained from the inside out",
      "description": "Teardown Theory is an engineering explainer channel produced 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/CEdkceKH2LpGG2GgKPriE>