

OVERALL SCORE

70 /100

AI-Readable

Executive summary

This site is mostly readable by AI tools, with a few improvements still recommended. Key strengths include AI guidance file and plain-text page access, while structured data and homepage access need attention. Recommended next step: add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

Recommended next step

1. 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.
2. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.
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 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 Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.	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	Pass	10/10
The sitemap.xml file is valid and contains URL entries.		
Markdown support	Pass	15/15
The homepage returns markdown content when requested with Accept: text/markdown, giving AI systems a cleaner text representation with less navigation chrome to parse.		
Semantic HTML	Warn	7.1/10
Shorten the meta description to 160 characters or fewer. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps.		
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	Pass	5/5
Consider adding AI-specific head meta tags, robots noai/noimageai directive so AI systems can consistently discover content usage preferences across robots.txt, HTTP headers, and HTML metadata.		

Suggested fixes

Fix Structured data

HTML

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://veedcrawl.com/#organization",
      "name": "Veedcrawl - Video intelligence API for AI agents",
      "description": "Give agents and applications the ability to understand
social videos. Search, transcripts, metadata, and structured extraction from
YouTube, TikTok, Instagram, X, and Facebook.",
      "url": "https://veedcrawl.com/",
      "logo":
"https://veedcrawl.com/_next/image?url=%2Fveedcrawl_icon.png&w=96&q=75",
      "sameAs": [
        "https://github.com/veedcrawl/mcp"
      ],
    },
    {
      "@type": "WebSite",
      "@id": "https://veedcrawl.com/#website",
      "name": "Veedcrawl - Video intelligence API for AI agents",
      "description": "Give agents and applications the ability to understand
social videos. Search, transcripts, metadata, and structured extraction from
YouTube, TikTok, Instagram, X, and Facebook.",
      "url": "https://veedcrawl.com/"
    }
  ]
}
```

Continued in the Full Scan Report...

Full report

<https://www.llmscan.dev/scan/OcQSdqHX2jawaHI-XjCJI>