

OVERALL SCORE

34 /100

Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include sitemap and structured data, while homepage access and crawler policy need attention. Recommended next step: remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Recommended next step

1. Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.
2. Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.
3. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Signal breakdown

Crawlability

Fail0/20

Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Robots.txt

Fail0/15

Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.

llms.txt

Fail0/15

Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Sitemap	Pass	10/10
The sitemap.xml file is valid and contains URL entries.		
Markdown support	Fail	0/15
Make the homepage respond successfully to requests with Accept: text/markdown, either by serving markdown directly or by adding a markdown variant URL with correct content negotiation.		
Semantic HTML	Warn	8.6/10
Add missing semantic elements: article.		
Structured data	Pass	10/10
Valid JSON-LD structured data was found with core Organization or WebSite schema.org types.		
Content signals	Pass	5/5
Consider adding Content-Signal HTTP header, 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

llms.txt

MARKDOWN

Bulkmark - AI X/Twitter Bookmark Manager + Weekly Digest

> Stop losing saved tweets. Bulkmark auto-tags your X bookmarks with AI, builds shareable lists, and delivers a weekly email digest. Free 7-day trial.

This llms.txt file summarizes the public, canonical resources that AI assistants and crawlers should use to understand this site.

Site Overview

- Canonical URL: <https://bulkmark.io/>
- Site type: software application
- Recommended summary: Stop losing saved tweets. Bulkmark auto-tags your X bookmarks with AI, builds shareable lists, and delivers a weekly email digest. Free 7-day trial.

Core URLs

- [Homepage](<https://bulkmark.io/>): Primary public entry point and canonical site overview.
- [Sign in](<https://bulkmark.io/login>): Important public page discovered from the homepage navigation.
- [Prafull Sharma@prafull_sharma2 via Product Hunt"This is kinda sick. Like the only kinda email newsletter worth looking at cuz it's directly curated by me lol. I like this"](https://www.producthunt.com/@prafull_sharma2): Continued in the full scan report...

robots.txt additions

TXT

robots.txt additions

Copy these blocks into the existing robots.txt file. Keep current rules unless a note calls out a conflicting Disallow.

AI crawler access

Add explicit Allow rules for blocked AI crawlers; remove or narrow conflicting Disallow rules if your crawler target requires precedence.

User-agent: GPTBot

Allow: /

User-agent: ClaudeBot

Allow: /

User-agent: Google-Extended

Allow: /

Crawl-rate guidance

Add Crawl-delay under a User-agent block when the site needs gentler crawler pacing.

User-agent: *

Crawl-delay: 10

schema.json

JSON

```
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://bulkmark.io/#organization",
"name": "Bulkmark - AI X/Twitter Bookmark Manager + Weekly Digest",
"description": "Stop losing saved tweets. Bulkmark auto-tags your X bookmarks with AI, builds shareable lists, and delivers a weekly email digest. Free 7-day trial.",
"url": "https://bulkmark.io/",
"logo": "https://bulkmark.io/favicon.ico",
"sameAs": [
"https://x.com/MShaiz9/status/2057862185125032195"
],
},
{
"@type": "WebSite",
"@id": "https://bulkmark.io/#website",
"name": "Bulkmark - AI X/Twitter Bookmark Manager + Weekly Digest",
"description": "Stop losing saved tweets. Bulkmark auto-tags your X bookmarks with AI, builds shareable lists, and delivers a weekly email digest. Free 7-day trial.",
"url": "https://bulkmark.io/",
"publisher": {
"@id": "https://bulkmark.io/#organization"
}
}
]
```

Content-Signal

TXT

Content-Signal recommendations

Use these directives to make AI-use preferences explicit for compliant crawlers and AI systems. They are advisory signals, so keep them aligned with robots.txt, terms, and access controls.

Recommended values

- ai-train=no: AI model training, fine-tuning, and dataset creation.
- search=yes: AI search indexing, snippets, and discovery.
- ai-input=yes: AI answer grounding, retrieval, and generated-response context.

HTTP response header

```
'''http
Content-Signal: ai-train=no, search=yes, ai-input=yes
Content-Usage: train-ai=n
'''
```

Best for site-wide or route-specific policies because the signal travels with every response, including pages that AI systems fetch directly.

HTML meta tag alternatives

Add these inside the document '<head>' when server headers are not

Full report

https://llmscan.dev/scan/aF2upcbAWkK0dbRp0t_3N