

LLM Scan

PUBLIC AI VISIBILITY REPORT

surfmind.ai

Scanned May 24, 2026, 09:07 UTC

OVERALL SCORE

61

/100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include homepage access and sitemap, while AI guidance file and plain-text page access need attention. Recommended next step: publish an AI guidance file as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Recommended next step

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

|                                                                                                                                                                                   |      |        |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|--------|
| <div>Crawlability</div> <div>The homepage resolves, connects, returns HTTP 200 OK, exposes a canonical URL, and is not blocked by robots.txt.</div>                               | Pass | 20/20  |
| <div>Robots.txt</div> <div>Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.</div> | Warn | 7.5/15 |
| <div>llms.txt</div> <div>Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.</div>                             | Fail | 0/15   |

|                                                                                                                                                                                                                               |      |        |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|--------|
| <b>Sitemap</b>                                                                                                                                                                                                                | Pass | 10/10  |
| The sitemap.xml file is valid and 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>                                                                                                                                                                                                          | Warn | 8.6/10 |
| Add missing semantic elements: article.                                                                                                                                                                                       |      |        |
| <b>Structured data</b>                                                                                                                                                                                                        | Pass | 10/10  |
| Valid JSON-LD structured data was found with core Organization or WebSite schema.org types.                                                                                                                                   |      |        |
| <b>Content signals</b>                                                                                                                                                                                                        | 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

# SurfMind: Free AI Browser Extension (GPT, Claude, Gemini, OpenRouter)

> SurfMind is an AI-powered browser extension that understands the context of any web page and provides instant intelligence to help you work smarter and faster.

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://surfmind.ai/>
- Site type: organization
- Recommended summary: SurfMind is an AI-powered browser extension that understands the context of any web page and provides instant intelligence to help you work smarter and faster.

### ## Core URLs

- [Homepage](<https://surfmind.ai/>): Primary public entry point and canonical site overview.
- [Blog](<https://surfmind.ai/blog>): Editorial updates, announcements, and current context.
- [What's new](<https://surfmind.ai/whats-new>): Important public page discovered from the homepage navigation.

[Add to] 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 AI crawlers so public content permissions are
clear.
User-agent: GPTBot
Allow: /

User-agent: ChatGPT-User
Allow: /

User-agent: ClaudeBot
Allow: /

User-agent: Claude-Web
Allow: /

User-agent: PerplexityBot
Allow: /

User-agent: Google-Extended
Allow: /

# Crawl-rate guidance
Continued in the full scan report...
```

## schema.json

### JSON

```
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://surfmind.ai/#organization",
"name": "SurfMind: Free AI Browser Extension (GPT, Claude, Gemini,
OpenRouter)",
"description": "SurfMind is an AI-powered browser extension that understands
the context of any web page and provides instant intelligence to help you
work smarter and faster.",
"url": "https://surfmind.ai/",
"logo": "https://surfmind.ai/images/icon48.png",
"sameAs": [
"https://www.linkedin.com/feed/update/urn:li:activity:7423955362074079232",
"https://www.linkedin.com/feed/update/urn:li:activity:7424110800811880448",
"https://x.com/surfmind_",
"https://www.youtube.com/@trysurfmind",
"https://www.tiktok.com/@surfmind",
"https://www.instagram.com/trysurfmind/"
]
},
{
"@type": "WebSite",
"@id": "https://surfmind.ai/#website",
"name": "SurfMind: Free AI Browser Extension (GPT, Claude, Gemini,
Continued in the full scan report..."
}
```

## 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

continued in the full scan report...

## Full report

<https://www.llmscan.dev/scan/2cCDJZRqA1sJWyr4VQuV6>