

# LLM Scan

PUBLIC AI VISIBILITY REPORT

# www.llmscan.dev

Scanned Jun 27, 2026, 10:00 UTC

OVERALL SCORE

# 90 /100

AI-Ready

## Executive summary

This site is in strong shape for AI discovery and readability. Key strengths include crawler policy and AI guidance file, while homepage access needs attention. Recommended next step: serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

## Recommended next step

1. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

## Signal breakdown

### Crawlability

Warn 10/20

Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

### Robots.txt

Pass 15/15

robots.txt allows crawler access and includes Sitemap references.

### llms.txt

Pass 15/15

The llms.txt file was found and includes the expected text, length, heading, and URL signals.

### 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

Pass 10/10

The homepage has a valid title, meta description, heading structure, semantic landmarks, substantial body copy, and descriptive anchor text.

## 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 robots noai/noimageai directive, Content-Signal robots.txt directive so AI systems can consistently discover content usage preferences across robots.txt, HTTP headers, and HTML metadata.

## Suggested fixes

### llms.txt

#### MARKDOWN

```
# AI visibility scanner - LLM Scan | LLM Scan
```

```
> Run a free scan to see whether your website is readable, crawlable, and structured for AI answer engines.
```

```
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://www.llmscan.dev/>
- Site type: organization
- Recommended summary: Run a free scan to see whether your website is readable, crawlable, and structured for AI answer engines.

```
## Core URLs
```

- [Homepage](<https://www.llmscan.dev/>): Primary public entry point and canonical site overview.
- [Leaderboard](<https://www.llmscan.dev/leaderboard>): Important public page discovered from the homepage navigation.
- [Docs](<https://www.llmscan.dev/docs>): Documentation and implementation guidance for the site.
- [Sign in](<https://www.llmscan.dev/login>): Important public page discovered from the homepage navigation.
- [Unlock monitoring](<https://www.llmscan.dev/pricing>): Pricing, plans, and 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.

# 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://www.llmscan.dev/#organization",
      "name": "AI visibility scanner - LLM Scan | LLM Scan",
      "description": "Run a free scan to see whether your website is readable,
crawlable, and structured for AI answer engines.",
      "url": "https://www.llmscan.dev/",
      "logo": "https://www.llmscan.dev/icon.svg",
      "sameAs": [
        "https://x.com/gorillliteam"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://www.llmscan.dev/#website",
      "name": "AI visibility scanner - LLM Scan | LLM Scan",
      "description": "Run a free scan to see whether your website is readable,
crawlable, and structured for AI answer engines.",
      "url": "https://www.llmscan.dev/",
      "publisher": {
        "@id": "https://www.llmscan.dev/#organization"
      }
    }
  ],
  "inLanguage": "en"
}
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/H\\_4UWaXIToNC6aZw5JZ4a](https://www.llmscan.dev/scan/H_4UWaXIToNC6aZw5JZ4a)