

# LLM Scan

PUBLIC AI VISIBILITY REPORT

# app.siteroast.cc

Scanned Jun 29, 2026, 11:01 UTC

OVERALL SCORE

# 35 /100

Poor

## Executive summary

This site is difficult for AI tools to read right now. Key strengths include homepage access, 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

1. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.
2. 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.
3. Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.

## Signal breakdown

### Crawlability

Pass 20/20

The homepage resolves, connects, returns HTTP 200 OK, exposes a canonical URL, and is not blocked by robots.txt.

### Robots.txt

Warn 7.5/15

Add a User-agent: \* section with Allow or Disallow directives and include Sitemap references for crawler discovery.

### llms.txt

Fail 0/15

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

## Sitemap

Fail 0/10

Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.

## 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 7.1/10

Shorten the meta description to 160 characters or fewer. Add missing semantic elements: article.

## 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://app.siteroast.cc/#organization",
      "name": "SiteRoast - AI Conversion Audit for SaaS Landing Pages",
      "description": "Paste your URL to get a scored conversion audit with quick wins, trust gaps, copy issues, UX friction, and a step-by-step action plan. First pass in under 60 seconds.",
      "url": "https://app.siteroast.cc/",
      "logo": "https://app.siteroast.cc/favicon.ico?favicon.0x3dzn~oxb6tn.ico"
    },
    {
      "@type": "WebSite",
      "@id": "https://app.siteroast.cc/#website",
      "name": "SiteRoast - AI Conversion Audit for SaaS Landing Pages",
      "description": "Paste your URL to get a scored conversion audit with quick wins, trust gaps, copy issues, UX friction, and a step-by-step action plan. First pass in under 60 seconds.",
      "url": "https://app.siteroast.cc/",
      "publisher": {
        "@id": "https://app.siteroast.cc/#organization"
      },
      "inLanguage": "en"
    }
  ]
}
```

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://www.llmscan.dev/scan/yVIPYba6ibJtL591YVaWK>