

LLM Scan

PUBLIC AI VISIBILITY REPORT

www.cuppalist.com

Scanned Jun 23, 2026, 08:32 UTC

OVERALL SCORE

33 /100

Poor

Executive summary

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

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. 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.

Signal breakdown

Crawlability

Warn 10/20

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.

Robots.txt

Warn 7.5/15

Review AI crawler Disallow rules and keep only the paths that should be excluded from AI crawler access.

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

Pass 10/10

The sitemap.xml file is valid and contains URL entries.

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 5.7/10

Add exactly one h1 element that describes the page topic. Add missing semantic elements: main, article, nav, footer. Add visible body copy after script and style removal so the page has at least 200 words.

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://www.cuppalist.com/#organization",
      "name": "Cuppalist: a todo app, gently brewed",
      "description": "A calm todo app. A chronological calendar for your day, tidy
project lists for the long stuff, and recurring chores that show up on their
own.",
      "url": "https://www.cuppalist.com/",
      "logo": "https://www.cuppalist.com/favicon.svg"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.cuppalist.com/#website",
      "name": "Cuppalist: a todo app, gently brewed",
      "description": "A calm todo app. A chronological calendar for your day, tidy
project lists for the long stuff, and recurring chores that show up on their
own.",
      "url": "https://www.cuppalist.com/",
      "publisher": {
        "@id": "https://www.cuppalist.com/#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/Ax53g1aUSRoJ8DIK70u0X>