

LLM Scan

PUBLIC AI VISIBILITY REPORT

openai.com

Scanned Jun 23, 2026, 08:33 UTC

OVERALL SCORE

35 /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. 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.
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 index is valid and the checked child sitemap 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 7.1/10

Shorten the meta description to 160 characters or fewer. Add exactly one h1 element that describes the page topic.

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://openai.com/#organization",
"name": "OpenAI | Research & Deployment",
"description": "We believe our research will eventually lead to artificial
general intelligence, a system that can solve human-level problems. Building
safe and beneficial AGI is our mission.",
"url": "https://openai.com/",
"logo": "https://openai.com/apple-icon.png?apple-icon.02be~wu.bus9e.png",
"sameAs": [
"https://x.com/OpenAI",
"https://www.youtube.com/OpenAI",
"https://www.linkedin.com/company/openai",
"https://github.com/openai",
"https://www.instagram.com/openai/",
"https://www.tiktok.com/@openai"
]
},
{
"@type": "WebSite",
"@id": "https://openai.com/#website",
"name": "OpenAI | Research & Deployment",

```

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/bCCEkhd1i2toZIAN8zVM4>