

OVERALL SCORE

89 /100

AI -Ready

Executive summary

This site is technically ready for AI crawlers to read. Key strengths include crawler policy and AI guidance file, while homepage access and page structure need attention. Recommended next step: point the homepage canonical link tag at the URL the homepage is served from.

Recommended next step

- 1. Point the homepage canonical link tag at the URL the homepage is served from.
- 2. Add missing semantic elements: article.

Signal breakdown

Crawlability	Warn	10/20
Point the homepage canonical link tag at the URL the homepage is served from.		
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	Warn	8.6/10
Add missing semantic elements: article.		

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 AI-specific head meta tags, 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 # KARPON - Automate Your ESG Reporting > Karpon automates ESG data collection, calculation, and audit-ready reporting for growing companies. 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://karpon-landing-version2-dev.azurewebsites.net/ - Site type: organization - Recommended summary: Karpon automates ESG data collection, calculation, and audit-ready reporting for growing companies. ## Core URLs - [Homepage](https://karpon-landing-version2-dev.azurewebsites.net/): Primary public entry point and canonical site overview. - [KARPON](https://karpon-landing-version2-dev.azurewebsites.net/index.html): Important public page discovered from the homepage navigation. - [How It Works](https://karpon-landing-version2-dev.azurewebsites.net/how-it-works/index.html): Important public page discovered from the homepage navigation. 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://karpon-landing-version2-dev.azurewebsites.net/#organization",
"name": "KARPON - Automate Your ESG Reporting",
"description": "Karpon automates ESG data collection, calculation, and
audit-ready reporting for growing companies.",
"url": "https://karpon-landing-version2-dev.azurewebsites.net/",
"logo":
"https://karpon-landing-version2-dev.azurewebsites.net/Content/icons/RedFavi
con_apple-icon-180x180.png",
"sameAs": [
"https://linkedin.com/showcase/karponfi",
"https://instagram.com/karponfi",
"https://www.facebook.com/people/Karpon/61575445229220/",
"https://www.youtube.com/@karponfi"
]
},
{
"@type": "WebSite",
"@id": "https://karpon-landing-version2-dev.azurewebsites.net/#website",
"name": "KARPON - Automate Your ESG Reporting",
"description": "Karpon automates ESG data collection, calculation, and
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://llmscan.dev/scan/TCHXu59ejkjqEtYkSkKkB>