

OVERALL SCORE

44 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include crawler policy and 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. 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

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	8.6/10
Shorten the meta description to 160 characters or fewer.		
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://agentscan.baldbee.me/#organization",
      "name": "AgentScan - A local scanner for AI agent skills",
      "description": "AgentScan is a local scanner for AI agent skills. It reads a skill directory and reports what the code does: permissions, network calls, secrets, and license. It never runs the skill and never sends data.",
      "url": "https://agentscan.baldbee.me/",
      "logo": "https://agentscan.baldbee.me/icon.png?icon.146i34da7wx3z.png",
      "sameAs": [
        "https://github.com/thesfb/scanaskill",
        "https://github.com/thesfb/scanaskill/blob/main/CORPUS-REPORT.md",
        "https://github.com/thesfb/scanaskill/blob/main/LICENSE"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://agentscan.baldbee.me/#website",
      "name": "AgentScan - A local scanner for AI agent skills",
      "description": "AgentScan is a local scanner for AI agent skills. It reads a skill directory and reports what the code does: permissions, network calls, secrets, and license. It never runs the skill and never sends data.",
      "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://llmscan.dev/scan/zZLLNcG7Y3qoMobiojW9b>