

LLM Scan

PUBLIC AI VISIBILITY REPORT

autark.sh

Scanned Jun 30, 2026, 11:00 UTC

OVERALL SCORE

76 /100

AI-Readable

Executive summary

This site is mostly readable by AI tools, with a few improvements still recommended. Key strengths include homepage access and AI guidance file, while plain-text page access and crawler policy need attention. Recommended next step: 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.

Recommended next step

1. 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.
2. Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.
3. Add missing semantic elements: article.

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

Fix robots.txt syntax issues so each rule uses Field: value format, directives appear under a User-agent, and Sitemap entries use absolute URLs.

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

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

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 Content-Signal HTTP header, AI-specific head meta tags, robots noai/noimageai directive so AI systems can consistently discover content usage preferences across robots.txt, HTTP headers, and HTML metadata.

Suggested fixes

llms.txt

MARKDOWN

Autark | The Revenue system your agent operates

> An all-in-one revenue platform. Autark builds your TAM, runs outbound, captures every interaction, and manages pipeline. Your agent, your data.

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://autark.sh/>
- Site type: organization
- Recommended summary: An all-in-one revenue platform. Autark builds your TAM, runs outbound, captures every interaction, and manages pipeline. Your agent, your data.

Core URLs

- [Homepage](<https://autark.sh/>): Primary public entry point and canonical site overview.
 - [Sign up](<https://autark.sh/?login=1>): Important public page discovered from the homepage navigation.
 - [Watch](<https://autark.sh/watch>): Important public page discovered from the homepage navigation.
 - [Blog](<https://autark.sh/blogs/>): Editorial updates, announcements, and current context.
- 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://autark.sh/#organization",
"name": "Autark | The Revenue system your agent operates",
"description": "An all-in-one revenue platform. Autark builds your TAM, runs
outbound, captures every interaction, and manages pipeline. Your agent, your
data.",
"url": "https://autark.sh/",
"logo": "https://autark.sh/apple-touch-icon.png",
"sameAs": [
"https://x.com/KushalSM5"
]
},
{
"@type": "WebSite",
"@id": "https://autark.sh/#website",
"name": "Autark | The Revenue system your agent operates",
"description": "An all-in-one revenue platform. Autark builds your TAM, runs
outbound, captures every interaction, and manages pipeline. Your agent, your
data.",
"url": "https://autark.sh/",
"publisher": {
"@id": "https://autark.sh/#organization"
}
}
]
}
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://www.llmscan.dev/scan/oVTZE3J\\_QNxM5gDfUxNU](https://www.llmscan.dev/scan/oVTZE3J_QNxM5gDfUxNU)