

LLM Scan

PUBLIC AI VISIBILITY REPORT

growthhog.com

Scanned Sep 1, 2026, 11:02 UTC

OVERALL SCORE

47

/100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include AI guidance file and sitemap, while homepage access and crawler policy need attention. Recommended next step: remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.

Recommended next step

1. Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.
2. Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.
3. 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.

Signal breakdown

Crawlability Remove AI crawler Disallow: / rules or replace them with narrower path-level restrictions for private content only.	Fail	0/20
Robots.txt Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.	Fail	0/15
llms.txt The llms.txt file was found and includes the expected text, length, heading, and URL signals.	Pass	15/15

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	7.1/10
Add missing semantic elements: article. Add visible body copy after script and style removal so the page has at least 200 words.		
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

```
# growth0G - Story-led B2B ads, done for you

> Story-led B2B video ads, done for you - the creative your buyers actually watch.

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://growthog.com/
- Site type: organization
- Recommended summary: Story-led B2B video ads, done for you - the creative your buyers actually watch.

## Core URLs

- [Homepage](https://growthog.com/): Primary public entry point and canonical site overview.
- [Work](https://growthog.com/work/): Important public page discovered from the homepage navigation.
- [Brands](https://growthog.com/brands/): Important public page discovered from the homepage navigation.
- [Team](https://growthog.com/team/): Important public page discovered from the homepage navigation.
- [Resources](https://growthog.com/resources/): Important public page discovered 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 blocked AI crawlers; remove or narrow conflicting Disallow rules if your crawler target requires precedence.
User-agent: GPTBot
Allow: /

User-agent: ClaudeBot
Allow: /

User-agent: Google-Extended
Allow: /

# Crawl-rate guidance
# Add Crawl-delay under a User-agent block when the site needs gentler crawler pacing.
User-agent: *
Crawl-delay: 10
```

schema.json

JSON

```
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://growthog.com/#organization",
"name": "growthOG - Story-led B2B ads, done for you",
"description": "Story-led B2B video ads, done for you - the creative your buyers actually watch.",
"url": "https://growthog.com/",
"logo": "https://growthog.com/apple-touch-icon.png"
},
{
"@type": "WebSite",
"@id": "https://growthog.com/#website",
"name": "growthOG - Story-led B2B ads, done for you",
"description": "Story-led B2B video ads, done for you - the creative your buyers actually watch.",
"url": "https://growthog.com/",
"publisher": {
"@id": "https://growthog.com/#organization"
},
"inLanguage": "en"
}
]
}
```

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

Full report

<https://llmscan.dev/scan/Y9UgK3lNne6Hmc3ltlYLG>