

OVERALL SCORE

37 /100

Poor

Executive summary

This site is difficult for AI tools to read right now. Key strengths include AI guidance file and structured data, 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. Replace the global Disallow: / policy with narrower path-level rules if public content should be discoverable by crawlers.
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 Replace the global Disallow: / policy with narrower path-level rules if public content should be discoverable by crawlers.	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	Fail	0/10
Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.		
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
Shorten the meta description to 160 characters or fewer. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps.		
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

Nomoads - On-Device Ad Blocker for iOS and Android

> Nomoads is a mobile ad blocker for iPhone and Android that filters ads, trackers, and malware across every app using on-device DNS blocking. No account required. Free Basic plan; Premium from \$2.99/month.

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://nomoads.com/>
- Site type: organization
- Recommended summary: Nomoads is a mobile ad blocker for iPhone and Android that filters ads, trackers, and malware across every app using on-device DNS blocking. No account required. Free Basic plan; Premium from \$2.99/month.

Core URLs

- [Homepage](<https://nomoads.com/>): Primary public entry point and canonical site overview.
- [App Store](<https://apps.apple.com/us/app/nomoads-ad-blocker/id6759786049>): Important public page discovered from the homepage navigation.
- [Google Play](<https://play.google.com/store/apps/details?id=com.crowstar.nomoads>):

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.

# Sitemap discovery
# Add this top-level directive so search and AI crawlers can discover the
canonical sitemap.
Sitemap: https://nomoads.com/sitemap.xml

# 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: ChatGPT-User
Allow: /

User-agent: ClaudeBot
Allow: /

User-agent: Claude-Web
Allow: /

User-agent: PerplexityBot
Allow: /
Continued in the full scan report...
```

schema.json

JSON

```
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://nomoads.com/#organization",
"name": "Nomoads - On-Device Ad Blocker for iOS and Android",
"description": "Nomoads is a mobile ad blocker for iPhone and Android that
filters ads, trackers, and malware across every app using on-device DNS
blocking. No account required. Free Basic plan; Premium from $2.99/month.",
"url": "https://nomoads.com/",
"logo": "https://nomoads.com/favicon.svg"
},
{
"@type": "WebSite",
"@id": "https://nomoads.com/#website",
"name": "Nomoads - On-Device Ad Blocker for iOS and Android",
"description": "Nomoads is a mobile ad blocker for iPhone and Android that
filters ads, trackers, and malware across every app using on-device DNS
blocking. No account required. Free Basic plan; Premium from $2.99/month.",
"url": "https://nomoads.com/",
"publisher": {
"@id": "https://nomoads.com/#organization"
},
"inLanguage": "en"
}
]
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/Mas1F6sWRc4_1GJ0M0_3X