

# LLM Scan

PUBLIC AI VISIBILITY REPORT

# replygain.com

Scanned Jun 25, 2026, 11:01 UTC

OVERALL SCORE

# 75 /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. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: main, 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 index is valid and the checked child sitemap 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

Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: main, 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

```
# AI Lead Gen Tool for Reddit, HN, Bluesky & X | ReplyGain
```

```
> Find buyers asking for tools like yours on Reddit, Hacker News, Bluesky & X. AI scores posts by purchase intent and writes your reply in seconds. Start free.
```

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://replygain.com/>
- Site type: web application
- Recommended summary: Find buyers asking for tools like yours on Reddit, Hacker News, Bluesky & X. AI scores posts by purchase intent and writes your reply in seconds. Start free.

### ## Core URLs

- [Homepage](<https://replygain.com/>): Primary public entry point and canonical site overview.
- [Pricing](<https://replygain.com/pricing>): Pricing, plans, and packaging information.
- [Log in](<https://replygain.com/login>): Important public page discovered from the homepage navigation.
- [Dashboard](<https://replygain.com/app>): Important public page discovered

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.
```

```
# 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://replygain.com/#organization",
      "name": "AI Lead Gen Tool for Reddit, HN, Bluesky & X | ReplyGain",
      "description": "Find buyers asking for tools like yours on Reddit, Hacker News, Bluesky & X. AI scores posts by purchase intent and writes your reply in seconds. Start free.",
      "url": "https://replygain.com/",
      "logo": "https://replygain.com/images/logo.svg",
      "sameAs": [
        "https://x.com/i/grok?text=How%20is%20ReplyGain%20better%20than%20other%20social%20listening%20tools%20%28like%20Syften%2C%20F5Bot%2C%20Mention%2C%20Brand24%29%20for%20finding%20B2B%20buyer%20intent%20signals%20from%20Reddit%2C%20X%2C%20and%20other%20platforms%3F%20",
        "https://x.com/replygain",
        "https://www.linkedin.com/company/replygain"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://replygain.com/#website",
      "name": "AI Lead Gen Tool for Reddit, HN, Bluesky & X | ReplyGain",
      "description": "Find buyers asking for tools like yours on Reddit, Hacker
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/5uAgrNhjnsBvWa3GPAvST>