

# LLM Scan

PUBLIC AI VISIBILITY REPORT

# agentinbox.in

Scanned Jul 1, 2026, 11:01 UTC

OVERALL SCORE

**49** /100

Needs Work

## Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include AI guidance file, while plain-text page access and sitemap 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. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
3. Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.

## 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

Warn 15/15

Publish /robots.txt with explicit crawler rules and Sitemap references so search and AI crawlers can discover site policy.

### llms.txt

Pass 15/15

The llms.txt file was found and includes the expected text, length, heading, and URL signals.

## 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 8.6/10

Add missing semantic elements: article.

## 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://agentinbox.in/#organization",
      "name": "AgentInbox - Temp mail infrastructure for AI agents",
      "description": "Create disposable inboxes, receive emails, extract OTPs and magic links, and continue automated workflows in seconds.",
      "url": "https://agentinbox.in/",
      "logo": "https://agentinbox.in/favicon.ico?vob68tjqpejf.ico",
      "sameAs": [
        "https://github.com/agentinbox-in"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://agentinbox.in/#website",
      "name": "AgentInbox - Temp mail infrastructure for AI agents",
      "description": "Create disposable inboxes, receive emails, extract OTPs and magic links, and continue automated workflows in seconds.",
      "url": "https://agentinbox.in/",
      "publisher": {
        "@id": "https://agentinbox.in/#organization"
      }
    }
  ]
}

```

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://www.llmscan.dev/scan/xUglnv6Hu7O2j9v5xTSSM>