

OVERALL SCORE

45 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include homepage access and page structure, while AI guidance file and plain-text page access need attention. Recommended next step: publish an AI guidance file as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Recommended next step

- 1. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.
- 2. 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.
- 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 The homepage resolves, connects, returns HTTP 200 OK, exposes a canonical URL, and is not blocked by robots.txt.	Pass 20/20
Robots.txt Publish /robots.txt with explicit crawler rules and Sitemap references so search and AI crawlers can discover site policy.	Warn 15/15
llms.txt Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.	Fail 0/15

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	Pass	10/10
The homepage has a valid title, meta description, heading structure, semantic landmarks, substantial body copy, and descriptive anchor text.		
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://www.copywatch.io/#organization",
      "name": "Shopify Copycat Monitoring for Artists & Brands | Copywatch",
      "description": "Monitor the web for products that look like yours. Review possible matches, organise evidence and draft takedown notices with Copywatch for Shopify brands.",
      "url": "https://www.copywatch.io/",
      "logo": "https://www.copywatch.io/icon.svg?icon.05ir_pd3akl91.svg"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.copywatch.io/#website",
      "name": "Shopify Copycat Monitoring for Artists & Brands | Copywatch",
      "description": "Monitor the web for products that look like yours. Review possible matches, organise evidence and draft takedown notices with Copywatch for Shopify brands.",
      "url": "https://www.copywatch.io/",
      "publisher": {
        "@id": "https://www.copywatch.io/#organization"
      }
    },
    {
      "@type": "Text",
      "@id": "https://www.copywatch.io/#full-scan-report",
      "text": "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://llmscan.dev/scan/ekhLSfna7iO_21pZxmhr4