

OVERALL SCORE

56 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include crawler policy and plain-text page access, while AI guidance file and structured data 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. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
- 3. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

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 Pass 15/15

robots.txt allows crawler access and includes Sitemap references.

llms.txt Fail 0/15

Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Sitemap Pass 10/10

The sitemap.xml file is valid and contains URL entries.

Markdown support	Pass	15/15
The homepage returns markdown content when requested with Accept: text/markdown, giving AI systems a cleaner text representation with less navigation chrome to parse.		
Semantic HTML	Warn	5.7/10
Use exactly one h1 element and move secondary section titles to h2-h6. Avoid skipped heading levels so sections progress from h1 to h2 to h3 without gaps. Add missing semantic elements: main, 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
<pre><script type="application/ld+json"> { "@context": "https://schema.org", "@graph": [{ "@type": "Organization", "@id": "https://www.scapin.co.uk/#organization", "name": "SaaS Penetration Testing Web & API Pentests Scapin", "description": "Fixed-price penetration testing for SaaS companies. Web & API security by an OSCP/OSEP-certified ex-software-engineer. SOC2-ready reports and a free retest.", "url": "https://www.scapin.co.uk/", "logo": "https://framerusercontent.com/images/FMsAWJjvICC41pgqcgo5Trrc.png" }, { "@type": "WebSite", "@id": "https://www.scapin.co.uk/#website", "name": "SaaS Penetration Testing Web & API Pentests Scapin", "description": "Fixed-price penetration testing for SaaS companies. Web & API security by an OSCP/OSEP-certified ex-software-engineer. SOC2-ready reports and a free retest.", "url": "https://www.scapin.co.uk/", "publisher": { "@id": "https://www.scapin.co.uk/#organization" } }], "inLanguage": "en" }</pre>
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/LEMO03BpWnE0l5nC_tf1y