

OVERALL SCORE

47 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include AI guidance file and sitemap, 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. Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.
- 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 Remove AI crawler Disallow: / rules or add narrower Allow/Disallow rules if AI crawlers should be able to discover public content.	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	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.		
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

Total Project & Turnaround Management - Cleopatra Enterprise

> Cleopatra Enterprise, the #1 Total Project & Turnaround Management Platform. Improve the performance of your next project.

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://cleopatraenterprise.com/>
- Site type: web page
- Recommended summary: Cleopatra Enterprise, the #1 Total Project & Turnaround Management Platform. Improve the performance of your next project.

Core URLs

- [Homepage](<https://cleopatraenterprise.com/>): Primary public entry point and canonical site overview.
 - [Cleopatraenterprise](https://www.linkedin.com/company/cleopatraenterprise/?utm_source=website&utm_medium=header&utm_campaign=linkedin_header_button): Background information about the organization or site owner.
 - [Support](<https://support.costengineering.eu/support/login>): Support, help and contact information.
- 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.

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: ClaudeBot

Allow: /

User-agent: Google-Extended

Allow: /

schema.json

```
JSON
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://cleopatraenterprise.com/#organization",
      "name": "Total Project & Turnaround Management - Cleopatra Enterprise",
      "description": "Cleopatra Enterprise, the #1 Total Project & Turnaround Management Platform. Improve the performance of your next project.",
      "url": "https://cleopatraenterprise.com/",
      "logo":
        "https://cleopatraenterprise.com/wp-content/uploads/2025/09/Enbridge-company-logo.webp",
      "sameAs": [
        "https://www.linkedin.com/company/cleopatraenterprise/?utm_source=website&utm_medium=header&utm_campaign=linkedin_header_button",
        "https://www.linkedin.com/company/cleopatraenterprise/",
        "https://www.youtube.com/@cleopatraenterprise",
        "https://www.facebook.com/CostEngineering/",
        "https://x.com/cleopatraent"
      ]
    },
    {
      "@type": "WebSite",
      "@id": "https://cleopatraenterprise.com/#website",
      "name": "Total Project & Turnaround Management - Cleopatra Enterprise",
      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/zBkGO8f4PN_52roTDueQG