

```html

In the evolving landscape of **AI crawlability** and search engine optimization, new file formats and tools are emerging to help websites optimize for both AI-driven engines and traditional search. Among these, the `llms.txt` file is gaining traction as a critical asset in controlling how large language models (LLMs) such as ChatGPT, Google AI Overviews, and Perplexity interact with your website content.

This post will walk you through the Semrush AI Site Audit capabilities, what exactly `llms.txt` is, and how to properly configure it. We'll also explore complementary toolsets like Otterly.AI and Peec AI, and key SEO themes like AI citation tracking, competitor benchmarking, and share of voice in AI engines. Expect concrete pricing info, tool features, and strategic advice tailored for both SMBs and enterprise teams.

## What Is `llms.txt` and Why Does It Matter?

The `llms.txt` file is a new standard aimed at guiding AI crawlers and generative engines on how to access, interpret, and attribute your site data properly. Similar to the familiar `robots.txt` file that prevents or allows bot crawling, `llms.txt` focuses on AI-specific crawlability and content attribution rules.

This file is paramount in the age of **Generative Engine Optimization (GEO)**, where content visibility isn't solely about rankings but how AI models understand and present your content to users.

## Why You Can't Rely on Traditional `Robots.txt` Alone

Traditional SEO tactics have long depended on `robots.txt` for crawl control and sitemap files for indexing instructions. While these remain essential, AI crawlers often retrieve content differently, sometimes via API queries or multiple crawl paths. This is where `llms.txt` steps in, offering more granular control tailored to AI's unique indexing and citation needs.

## Semrush AI Site Audit: A Game-Changer for AI Crawlability

Semrush, a leader in SEO technology, has incorporated AI-driven insight mechanisms into its auditing toolkit. Their **AI Site Audit** extends beyond traditional technical SEO checks by analysing how well your site is optimized for AI indexing, including compliance and effectiveness of your `llms.txt` setup.

## Pricing Overview (Context for Small and Enterprise Teams)

| Plan          | Price                       | Key Features                             |
|---------------|-----------------------------|------------------------------------------|
| Basic         | \$99/mo                     | Entry-level SEO toolkit                  |
| SEO Plan      | \$117.33/mo billed annually | Full SEO audit + reports                 |
| Pro+          | \$248.17/mo billed annually | AI Visibility Toolkit, advanced analysis |
| AI Visibility | From \$199/mo               | AI Visibility + SEO                      |
| Combined      | From \$199/mo               | Combined AI and SEO optimizations        |

With access to a 28B+ keyword database, 142 geographic databases, and 808M domain profiles, Semrush AI Site Audit allows you to benchmark your site against up to 9 named competitors — a solid [iplocation.net](https://www.iplocation.net) foundation for competitor benchmarking and share of voice analysis on AI engines.

## What to Put in `llms.txt`: Best Practices

Creating an effective `llms.txt` involves defining directives that guide AI models on your content's:

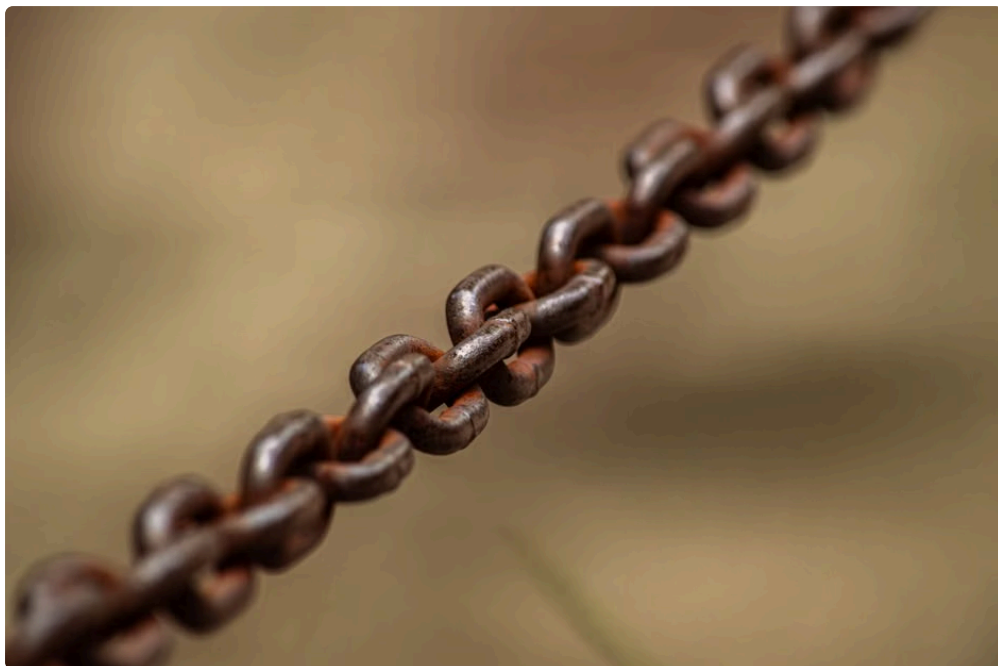
- **Indexing permissions:** Specify which sections or URLs AI crawlers can or cannot process.

- **Citation guidelines:** Encourage models to attribute content correctly, prioritizing linked URLs over plain mentions.
- **Content freshness preferences:** Indicate update frequencies or priority content sections for AI engines.

## Typical Structure of llms.txt

User-agent: ChatGPT Disallow: /private/ Allow: /blog/ Citations: must include links to original articles Update-frequency: daily User-agent: Perplexity Disallow: /docs/ Allow: /faq/ Citations: linked URLs preferred, no unlinked mentions

By tailoring the file for specific AI engines—such as ChatGPT, Perplexity, and Google AI Overviews—you improve the chances your content is used and cited properly across these platforms.



## AI Citation Tracking vs. Traditional SEO Rank Tracking

Unlike traditional SEO metrics focused on ranking positions and organic click-through rates, AI citation tracking extends into how AI engines:

1. **Use your content as source material**, which can appear in AI responses beyond direct search results.
2. **Attribute information**, ensuring your brand or website is cited with proper URLs rather than just mere mentions.
3. **Understand content context** for better prompt responses, which impacts your indirect traffic and brand authority.

This kind of citation tracking is where Semrush's AI Visibility Toolkit provided in their Pro+ plan (\$248.17/month billed annually) shines, offering in-depth analysis of how AI references your content, beyond classic keyword rankings and share of voice.

## Competitor Benchmarking and Share of Voice with AI Engines

One of the overlooked advantages of Semrush's AI Site Audit is benchmarking competitor performance within the AI landscape. You can monitor your:

- **Share of voice** in AI-driven search results and generative engines.
- **Keyword visibility** across both traditional and AI-focused databases.
- **Domain authority** as understood by AI, which may weigh citations and content differently from standard SEO.

This multi-pronged approach helps businesses identify gaps not just in Google rankings, but also in how their content performs in emerging AI channels—a key consideration when integrating tools like Otterly.AI and Peec AI, which specialize in AI content generation and AI citation monitoring.

## Engine Coverage: ChatGPT, Perplexity, and Google AI Overviews

Understanding how to configure AI crawlability in your `llms.txt` is impossible without knowing what kinds of AI engines access your content:

- **ChatGPT (OpenAI):** Primarily trained on broad datasets but increasingly integrating real-time content via plugins and API crawls.
- **Perplexity AI:** Focuses on blending search and generative answers with live web citations, making citation attribution crucial.
- **Google AI Overviews:** Google's own generative search features that combine traditional search with AI summaries, and are sensitive to SEO signals and crawl directives.

Each requires different crawl permissions and citation instructions in your `llms.txt`, which, if misconfigured, could result in AI engines ignoring critical parts of your site or misattributing your content.

## The Importance of Citation Attribution Depth: Linked URLs vs. Mentions

One pet peeve shared by experienced SEO and GEO marketers is tools and companies counting unlinked “mentions” as valid citations. However, AI models rely heavily on *attributed* and *linked* source references to reduce misinformation and improve content credibility.

Here's why structured citation is vital:

- **Linked URLs** provide verifiable sources that AI engines can check, increasing trust.
- **Mentions without links** are ambiguous and may not carry the same weight in AI response generation or SEO algorithms.
- **Unlinked citations** can dilute your authority and lead to improper attribution or content scraping risks.

When setting up `llms.txt`, you should explicitly instruct AI crawlers to prioritize linked citations, seamless URL attributions, and discourage reliance on non-linked mentions.

## Complementary Tools to Boost Your AI SEO Strategy

Besides Semrush, consider leveraging these AI-focused SEO tools:

- **Otterly.AI:** Specializes in automatic citation tracking across AI content and helps ensure your linked URLs are referenced properly in large language model outputs.
- **Peec AI:** Focuses on GEO with advanced share of voice monitoring in generative engines, providing granular insights for AI-specific keyword datasets and competitor performance.

Remember, integrating these tools is not standalone — they typically require a base SEO monitoring subscription like Semrush's SEO Plan or AI Visibility Toolkit.



## Summary and Best Practices

1. **Create a well-structured llms.txt** file that defines AI engine-specific crawling and citation rules.
2. **Leverage Semrush AI Site Audit** to identify AI crawlability issues and benchmark competitors on AI engines.
3. **Focus on citation attribution depth**, emphasizing linked URLs over generic mentions to improve AI trust and content visibility.
4. **Understand that AI citation tracking** complements, but does not replace, traditional SEO rank tracking.
5. **Use complementary tools like Otterly.AI and Peec AI** to monitor AI citations and generative engine share of voice effectively.
6. **Budget wisely:** Semrush plans start at \$99/month for basics but AI-focused plans cost around \$199-\$248/month billed annually; watch out for required base subscriptions.

With these strategies, your site will be positioned to thrive not only in Google's SERPs but also within the rapidly growing AI-driven content discovery and attribution ecosystem.

...