

Blog

What is LLMs.txt & How To Create It For SEO?

[Full Guide]

Discover what LLMs.txt is and how it boosts your content's visibility in AI models. Learn how to optimize for LLMs with our step-by-step guide.



Matija Golubovic | 5 Oct 2025

Knowledge

• 7 min read

Did you know that [AI search engines](#) handle over **4.8 billion prompts every month**?

This number is an **opportunity** for SEO professionals and marketers to start optimizing their sites for LLMs.

One emerging method to potentially help AI models better understand and prioritize your key content is the **llms.txt** file.

Let's explore what it is and how it works!

What is LLMs.txt?

LLMs.txt **guides AI models to your most valuable content**. It is a structured [Markdown file](#) (<https://yourdomain.com/llms.txt>) that explicitly **lists important pages or resources, along with brief summaries of each**.

LLMs.txt is solely for AI content guidance, and it's used to make your content more **accessible and readable** to AI search crawlers in a format they prefer.

For example, you might have sections like "Blog" or "Products" with links to relevant pages.

```
# https://zencoder.ai/llms.txt
# Zencoder - The AI Coding Agent

> Zencoder is an AI coding agent that enhances productivity, automates code tasks, and accelerates development. The most integrated, customizable, and intuitive coding agent. Zencoder handles the routine so you can focus on the vision.

## General Pages
# "Zencoder" # The AI Coding Agent[https://zencoder.ai/] Transform your coding experience with Zencoder, the next-gen AI agent that automates tasks, enhances collaboration, and improves code quality seamlessly within your existing workflow.
# "Contact" [https://zencoder.ai/contact] Connect with our sales and support teams for demos, pricing inquiries, and product feedback. Explore Zencoder's features and documentation for comprehensive assistance.
# "Blog" [https://zencoder.ai/blog] Explore expert insights on AI coding, automatic unit test generation, and the latest in software development at the Zencoder AI blog. Stay updated with our comprehensive guides and industry news.
# "Events" [https://zencoder.ai/events] Join Zencoder at global events to connect with experts, watch live demos, and explore the future of software development. Discover insights from industry leaders and enhance your skills.
# "Pricing" [https://zencoder.ai/pricing] Check out our transparent pricing for Zencoder AI's AI coding agent. Get started with code generation, chat assistance, and more.
# "About us" [https://zencoder.ai/about] Join our innovative team to revolutionize software development through AI automation. Shape your career and impact the future of technology with us. Explore open roles
```

By doing this, LLMs.txt tells AI assistants what to prioritize, so they don't have to guess by crawling your website.

How LLMs.txt Supports Generative Engine Optimization (GEO)?

In this search shift, Generative Engine Optimization (GEO) becomes **critical**.

Here are some of the things that the **LLMs.txt file can help you with:**

- **Direct content guidance:** LLMs.txt provides AI models a clear roadmap to your most valuable pages. Instead of searching through navigation and sidebars, AI can access your authoritative content immediately.
- **Improved response quality:** By offering a clean Markdown index of key pages, you boost AI response accuracy and reduce reliance on outdated cached content.
- **Strategic integration:** LLMs.txt complements existing [GEO strategies](#), such as the structuring schema snippets, digital PR, and creating an additional structured layer at the site level. This curated approach **increases citation probability** and helps maintain brand presence.



Services Company Industries Knowledge



Get started >

On this page

- What is LLMs.txt?
- How LLMs.txt Supports Generative

Benefits of LLMs.txt for SEO



Traditional Search Crawlers vs. LLMs

Search crawlers and LLM crawlers **work quite differently**.

Traditional crawlers (e.g., Googlebot, Bingbot) continuously scan the web, following links, obeying robots.txt and sitemaps, and **building a long-term index of your entire site**. They revisit regularly, render pages (including those with JavaScript), and store content for improved ranking.

In contrast, LLM-based systems like ChatGPT **do not build a permanent index of all pages. LLMs access a page only at query time**, not ahead of time. They “don’t index or remember your site” – they only see small parts of a page for each question.

They work within a **limited token window**, so overly long or cluttered pages get cut off. They skip content that’s not clearly linked or easily parsed. If links aren’t obvious or if a page is heavily loaded with ads or JavaScript, the LLM **may ignore it**.

Aspect	Traditional Crawlers	LLM Crawlers
Indexing	Full site index	No permanent index
Timing	Crawl ahead of time	Fetches at query time only
JavaScript	Executes JavaScript	Does not run JavaScript
Content Limit	Full content	Limited token window
Discovery	Follow links, use sitemap.xml	Relies on backlinks, llms.txt file
Memory	Store and update content	Doesn’t remember past pages

In practical terms, LLM “crawling” often relies on specialized bots (like OpenAI’s **GPTBot**).

Engine Optimization (GEO)?
[Traditional Search Crawlers vs. LLMs](#)

Difference: LLMs.txt vs Robots.txt vs Sitemap.xml

Why SEOs and Marketers Should Care About LLMs.txt?

Step-by-Step: How to Create Your LLMs.txt File

Example of LLMs.txt

How to Test and Validate Your LLMs.txt File?


What Causes LLMs to Misread Your Website Content?

Common Mistakes and How to Avoid Them

Will LLM Crawlers Read the LLMs.txt File?

Conclusion

FAQs



Unlock healthy growth at efficient economics with holistic marketing partners.

Explore Omnius

For example, GPTBot discovers pages via backlinks, public URLs, and possibly sitemaps, but it does not execute JavaScript. It simply fetches the raw HTML. This means server-side rendered content is visible, but dynamic content may be missed.

Difference: LLMs.txt vs Robots.txt vs Sitemap.xml

Three essential files sit at your domain root, each serving distinct purposes for different systems accessing your website.

Here is the breakdown of each:

Robots.txt: Establishes crawl boundaries through directive protocols. It instructs web crawlers which site sections remain inaccessible via Allow/Disallow commands, preventing automated indexing of restricted content.

LLMs.txt: Implements content prioritization for AI parsing systems. It signals which pages contain authoritative information, guiding language models toward high-value content without imposing access restrictions.

Sitemap.xml: Generates comprehensive URL mapping for search engine discovery. It delivers structured XML data containing every indexable page, ensuring complete site crawling and improved search visibility.

File	Purpose	Format	Target
robots.txt	Controls access permissions	Plain text rules	Search crawlers
sitemap.xml	Lists all site pages	XML structure	Search crawlers
llms.txt	Highlights key content	Markdown text	AI search engines

Use all three together: robots.txt and sitemap.xml handle traditional SEO, while llms.txt optimizes for AI-powered search.

Why SEOs and Marketers Should Care About LLMs.txt?

Today, many users start exploring with ChatGPT, Gemini, or voice assistants – not just Google.

In this “answer-first” era, showing up in the answer itself is as important as ranking on page 1.

It’s important to know that **being first in search results doesn’t guarantee that you will be in the AI search engines.**

Having LLMs.txt helps ensure your site is included correctly. By curating your key content, it **increases the chance that AI search engines will cite your brand** and facts rather than spitting back generic or outdated info.

Any site with **valuable content** should consider LLMs.txt.

Examples include:

- **Service landing pages**
- **Product pages**
- **Educational articles or blog posts**
- **Large news portals or e-commerce sites**
- **Documentation or Help Centres**
- **FAQs**

For SEOs and marketers, as always, the most important thing is **testing**. Testing on some of your projects or clients' websites will give you a better understanding of does LLMs.txt works or not.

Step-by-Step: How to Create Your LLMs.txt File

Creating an effective LLMs.txt file requires **planning** and **precise execution**.

This step-by-step guide walks you through building a file that maximizes your content's visibility to AI systems while maintaining practical usability.

1. File Structure Setup

Begin by creating a plain-text file named llms.txt using Markdown formatting. This approach ensures human readers and AI systems can parse your content effectively.

The foundation **starts with an H1 header** featuring your site or brand name, followed by a concise site description wrapped in blockquote formatting.

Your initial structure should establish a **clear hierarchy through H2 section headers**, organizing content into logical categories. This organizational approach enables AI systems to understand the relationships between your content and prioritize information appropriately.

2. Content Organization

Strategic content selection determines the effectiveness of your file.

Rather than overwhelming AI systems with extensive link lists, focus on **curating 5-10 high-value pages** that genuinely represent your expertise and answer user questions directly.

Organize content into meaningful H2 sections such as Documentation, Guides, FAQ, and Products. **Each section should contain only your most authoritative content** within that category.

Remember that AI systems prioritize **quality over quantity**, so resist the temptation to include every available page.

Pages worth including demonstrate clear expertise, provide standalone value, and address specific user needs. Avoid generic content, outdated information, or pages that require additional context to be fully understood.

3. Link Formatting

Each link requires descriptive context that explains its value proposition. The standard format combines the page title with a brief explanation of what users can expect to find there. This additional context helps AI systems understand when to reference your content in their responses.

For example, instead of simply listing "Getting Started Guide," provide context, such as "Getting Started Guide: Quickstart tutorial for new users." This approach tells AI systems exactly what problem your page solves and when it's most relevant.

4. Technical Implementation

Your LLMs.txt file must reside in your **website's root directory** to ensure that AI systems can locate it consistently. The standard URL structure should be `https://yourdomain.com/llms.txt`, with direct access and no redirects that interfere with automated retrieval.

When uploading manually through cPanel or FTP, place the file in your `public_html` directory to ensure proper root-level access.

For WordPress users, Yoast SEO's plugin automatically generates and updates files with new content on a regular basis.

5. Quality Validation

Before publishing, verify that your file follows **proper Markdown formatting**, including the correct header hierarchy and link syntax.

AI systems expect a consistent structure, so ensure your H1 and H2 headers are properly formatted and each link uses the standard `[text](url)` markdown syntax.

Content review is equally important. Each included link should provide genuine value and accurately represent your expertise within its respective category. Test your file's accessibility by visiting the URL directly and verifying that it displays correctly without any formatting errors.

6. Maintenance Protocol

Your LLMs.txt file **requires regular updates** to maintain relevance and accuracy. Schedule reviews whenever you publish significant content pieces, launch new products, or restructure important pages.

This proactive maintenance approach ensures that AI systems always access your **most current** and **valuable information**, maximizing the potential for your content to be included in generated responses.

Example of LLMs.txt

If you're creating an LLMs.txt file for the first time, it can feel a bit unclear; seeing a clear example of what an LLMs.txt file should look like can really help.

```
# ExampleSite
> A tech blog explaining AI, SEO, and web development topics.
## Tutorials
- [Building LLM-Powered Apps](/tutorials/build-llm-app):
Complete implementation guide for developers
- [SEO for 2025](/tutorials/seo-2025): Advanced optimization
strategies for emerging search behaviors
## FAQ
- [Contact Support](/faq/contact-support): Detailed process for
reaching our technical team
## Products
- [Premium SEO Tool](/products/seo-tool): Comprehensive feature
overview and setup instructions
```

How to Test and Validate Your LLMs.txt File?

Since LLMs.txt is a proposal, there's no official validator, unlike sitemap.xml.

Instead, testing is informal:

- **Check Accessibility:** First, simply open <https://yourdomain.com/llms.txt> in a browser to ensure it loads as plain text.
- **Monitor Logs:** Look at server logs for known AI user agents (e.g., GPTBot, OAI-SearchBot, or others listed in your robots.txt). Right now, few AIs officially fetch llms.txt, but this could change. Use analytics or log tools to see if "llmstxt" appears in any access logs.
- **Query AI Tools:** A practical test is to ask AI chatbots directly. For example, use ChatGPT (with browsing enabled) or Perplexity and ask about your site's content. See if it cites or pulls information from the pages you included. Try verifying in different LLMs, such as ChatGPT or Copilot, by explicitly asking them to use the llms.txt file for answers.
- **Validate Content:** Confirm that each link and snippet in llms.txt is accurate. Typos or outdated info can mislead AI. Also, ensure the file is well-formatted (headings, bullet points, no broken links).
- **Use Search Console (optional):** While not required, you can list your llms.txt file in your sitemap.xml and submit it in Google Search Console. This alerts Google (and any bots) that the file exists. Note: Google itself ignores llms.txt, but it's a way to monitor if any crawlers fetch it.

Regularly review your implementation: if you change your site structure or content focus, update llms.txt accordingly and re-run these simple tests. Because the LLMs.txt standard is community-driven, practical checks like these are the best current approach.

What Causes LLMs to Misread Your Website

Content?

LLMs often misunderstand a site for two main reasons: **context limits** and **noisy formatting**.

AI tools don't perceive your site the way traditional search engines and humans do.

In concrete terms:

- **Live-only scanning:** LLMs read a page only when needed. They don't crawl your whole site in advance.
- **Short context window:** They can only process a limited number of words at once (a few thousand tokens). Very long articles or complex pages may exceed this limit, causing the model to truncate or omit parts.
- **Missing cues:** If information is buried behind menus, tabs, or is not clearly labelled, the AI might skip it. Cluttered pages with hidden links confuse the model.
- **Dynamic/JavaScript content:** Many LLM bots (e.g., ChatGPT's crawler) **do not execute JavaScript**. If your content appears only after scripts run (like single-page app content or dynamically loaded sections), the AI will miss it.
- **Technical or irrelevant content:** Pages filled with code examples, logs, or heavy widgets can overwhelm an LLM. If the content looks "too technical," the model may not parse it properly. Also, ads, sidebars, or boilerplate text distract AI from the main message.

LLMs misread sites when the signal-to-noise ratio is poor. By providing a clean, structured llms.txt file, you can **reduce noise** and give the **AI a reliable signal** about what matters.

This helps prevent situations where an AI "hallucinates" or gives an incomplete answer because it has never seen your actual content.

Common Mistakes and How to Avoid Them

Implementing llms.txt effectively requires attention to detail.

Here are common LLMs.txt mistakes and how to steer clear of them:

- **Too many links:** Don't dump your entire URL list into llms.txt. The point is to guide AI, not overload it. Choose evergreen pages that directly answer user questions. If a page makes no sense out of context, skip it.
- **Skipping proper Markdown structure:** The file must be a valid Markdown document. If you forget the headings or misuse bullet/link syntax, AI won't parse it correctly. Ensure you have an H1 header (your site name), H2 section titles, and each link is formatted as [Link Text](URL).
- **Neglecting updates:** Like a sitemap or content index, llms.txt should evolve with your site. If you add new product pages, FAQs, or blog posts, update the file. Otherwise, an LLM will cite outdated information.
- **Confusing it with robots.txt:** LLMs.txt doesn't block bots or prevent AI training—that's handled by robots.txt or meta tags. Its role is different: robots.txt sets access rules, while LLMs.txt highlights content for AI to reference.
- **Ignoring evidence:** Some may rush to adopt LLMs.txt, hoping it's a magic trick. But it's just one tactic among many. Keep an eye on how AI traffic behaves, and don't abandon [SEO best practices](#) in pursuit of uncertain gains. Think of llms.txt as a safe preparation step; it won't harm your site, but its benefits are still unproven.

Will LLM Crawlers Read the LLMs.txt File?

Currently, we don't have information indicating that **any AI systems are reading or using llms.txt files**.

Google's [John Mueller recently confirmed](#) that **no consumer-facing LLMs are fetching this file** for training or grounding purposes.

While llms.txt is positioned as a way to guide AI toward key content, it's not an active signal, yet.

That said, it's **still worth testing**.

Adding a llms.txt file won't hurt your site, and if adoption grows, you'll already be prepared.

Conclusion

While the LLMS.txt file is **not yet widely adopted** or officially supported by major LLMs, it reflects a broader shift in how search is evolving, from link-based crawling to answer-focused understanding.

Think of llms.txt as a low-effort, future-focused enhancement: it won't harm your SEO, but it might help AI tools find and cite your content more effectively as adoption grows. Just like schema markup or early sitemap.xml files, its true **value may only become clear over time**.

From **improving your visibility** in traditional search to **optimizing your content** for AI-driven platforms, [Generative Engine Optimization \(GEO\) agencies](#) can make strategies that make a measurable impact.

If you're looking for a professional approach and real results, Omnius is here to help.

[Book a free 30-minute call](#) and discover how Omnius can help you rank higher within LLM

platforms with personalized SEO strategies!

FAQs

How to discover tools supporting llms.txt?

Yeast SEO plugin for WordPress offers an easy way to create llms.txt. There are also standalone tools like llmstxt.firecrawl.dev, and some open-source generators. Searching "llms.txt generator" will show available options.

How often should I update my llms.txt file?

Update the file whenever you add or change important pages, such as new products or FAQs. Keep it current so it reflects the latest key content on your site.

Is the LLMs.txt file necessary?

No, it's not required yet. Most AI systems don't currently check it, so having an LLMs.txt file won't affect your site today, but it's worth testing.

We broke the "standard agency" model, and built it differently.

Learn how we integrate deep into SaaS & Fintech companies to make the growth predictable.



/ No. 1 LinkedIn™ content-focused SaaS tool

"With Omnius, we saw immediate results - 64% higher conversion on a new website and 110% organic growth in 6 months. So, if you want an agency that understands startups, do yourself a favour and talk to them."

Ivana Todorovic

CO-FOUNDER & CEO

Monthly Growth OpenLetter.

Learn how to scale user acquisition without scaling costs from our findings. We spent years exploring, so you don't have to.

Related articles.

16 Best Fintech SEO Agencies to Consider in 2026

[Knowledge](#) | Matija Golubovic

15 Best Fintech Marketing Agencies To Partner With For Revenue Growth in 2026

[Knowledge](#) | Matija Golubovic

Maximizing the value of SEO & GEO.

Omnius is a B2B SEO & LLMO agency; partnering up exclusively with **SaaS, Fintech & AI** companies. The result? Compounding growth made through organic positioning **everywhere** **people search for information**, including both Google & AI search engines.

Contact our team

Our work is referenced by the leading media, venture funds & startup organizations



Company

- About
- Client reviews
- Blog
- News
- Contact
- VC Partnerships
- Careers
- Projects

Industries

Services

- Content Marketing
- B2B SEO
- Generative engine optimization (GEO)
- Answer engine optimization (AEO)
- AI SEO
- Development & CRO
- Programmatic SEO Agency
- Data Analytics
- AI Search Audit

Case studies

Resources

SaaS SEO Agency	AI SaaS: 0-2.7M clicks in 13 months	Google Algorithm Updates
Fintech SEO Agency	Fintech: +227.9% sign ups in 6 months	SaaS Industry Report 2024
LLM SEO Agency	Hospitality ERP: 4X organic clicks in 6 mo	Fintech Industry Report 2024
LLM Marketing Agency	LinkedIn SaaS: 64% better conversions	Bad SEO Practices
SEO Agency for AI Companies	Airbnb Case Study: Programmatic SEO	ChatGPT SEO Guide
SaaS Content Marketing Agency	74% Fintech SEO Growth	Programmatic SEO Guide
Fintech Content Marketing Agency	BigCommerce SEO Case Study	How to create Content flywheel
SaaS SEO Consulting Agency	AI SaaS SEO Growth: 0-60k	SaaS Conversion Tracking
Fintech SEO Consulting Agency	Wise Fintech SEO Case Study	SaaS Growth 101
ChatGPT SEO Agency	pSEO Case Study: 2000+ signups	SaaS Growth Strategies
ChatGPT Marketing Agency	Global App Testing: HMQLs +163% in 12 mo	SaaS Content Writing Guide
B2B Marketing Agency		Low-Hanging Fruit Keywords
Semantic SEO Agency		High-Intent Keywords
MarTech SEO Agency		Product-led Content Guide
Webflow SEO Agency		Webflow SEO Guide
Technical SEO Agency		Best SaaS Marketing Agencies
Perplexity SEO Agency		Best SaaS SEO Agencies
Claude SEO Agency		Best Fintech AI SEO / AEO Agencies
Gemini SEO Agency		Top Generative Engine Optimization (GEO) Agencies
DeepSeek SEO Agency		Fintech Marketing Agencies
		Fintech SEO Guide
		SaaS Inbound Marketing
		Fintech Content Strategy
		Best SaaS SEO Tools
		Best SEO Agencies in Europe
		Best AI Search Engines
		GEO Strategies
		Google vs ChatGPT & AI Platforms

Stay up to date

Receive Growth OpenLetter once a month. No spam, only value.

Your business email

Subscribe