



Scott Tischler

FOUNDER & CHAIRMAN · AIRECOMMEND.AI

TECHNICAL

What Is llms.txt? The New Standard for Talking to AI Crawlers

By Scott Tischler · July 28, 2026 · 8 min read

llms.txt is a proposed standard: a single Markdown file at the root of your website that gives AI systems a clean, curated guide to your most important content. It lives at `yourdomain.com/llms.txt`, and its job is to help large language models and the tools built on them find and understand the pages you most want them to use — without wading through your navigation, ads, cookie banners, and boilerplate.

That's the whole idea in one sentence. The rest of this article is the honest detail: where it came from, how it's different from the files you already know, what actually goes inside it, how AI systems may or may not use it today, and a template you can adapt. I'll be clear about what's settled and what isn't, because a lot of what's written about llms.txt overstates how established it is.

What is llms.txt, exactly?

llms.txt is a **community proposal**, not an official web standard ratified by a standards body. It was introduced to solve a specific, real problem: AI systems consuming web content have limited context windows and have to work hard to separate the substance of a page from the surrounding clutter. HTML is built for browsers and humans, not for a model trying to extract the core information efficiently.

The proposal's answer is a plain-Markdown file at your site root that acts as a curated table of contents. It typically contains a short description of what your site or project is, followed by lists of links to your most important pages — ideally links to clean, Markdown-friendly versions of those pages. Markdown is chosen deliberately: it's simple, it's readable by both humans and machines, and it strips away the presentation noise.

A closely related convention is **llms-full.txt**, a larger file that inlines the actual content rather than just linking to it — useful when you want to hand a model the full text in one place. Whether you offer one, both, or neither depends on your site and your goals.

The key mental model: **robots.txt tells crawlers where they may not go; a sitemap tells them everything that exists; llms.txt tells them what matters most and helps them read it cleanly.** Those are three different jobs.

How is llms.txt different from robots.txt and sitemaps?

This is the question I get most, because on the surface all three are files at your site root that talk to automated visitors. They do genuinely different things.

File	Primary job	Audience	Format	Nature
robots.txt	Permission — where crawlers may and may not go	Search and AI crawlers	Directives	Long-established standard
sitemap.xml	Inventory — a machine list of all URLs	Search crawlers	XML	Long-established standard
llms.txt	Curation — what's most important and how to read it cleanly	AI systems / LLMs	Markdown	Emerging proposal

A few distinctions worth making explicit:

- **robots.txt is about access control**, not curation. It's a set of allow and disallow rules. It doesn't tell a crawler which of the allowed pages are your best; it just draws boundaries. Importantly, it's also not a security mechanism — it's a request that well-behaved crawlers honor.
- **A sitemap is exhaustive, not selective.** It aims to list every URL so nothing gets missed. That's the opposite of llms.txt's purpose, which is to highlight the few pages that matter most and present them in a clean, digestible form.
- **llms.txt is curation and comprehension.** It answers "if an AI only reads a handful of things about us, what should they be, and how do we make them easy to parse?" Neither robots.txt nor sitemaps were designed for that.

They're complementary. A well-run site can have all three, each doing its own job. llms.txt doesn't replace the other two.

What goes in an llms.txt file?

The proposed format is intentionally simple and human-readable. A typical file follows a loose structure:

- **An H1 with the name** of the site, project, or organization.
- **A short blockquote summary** — a sentence or two describing what you are and what the content covers.

- **Optional context paragraphs** giving background a model would find useful.
- **Sections with H2 headings**, each containing a Markdown list of links. Each link points to an important page, ideally a clean Markdown version, with a brief description after it.
- **An optional "Optional" section** for links that can be skipped if the model is short on context — a way to signal priority.

The guiding principles are: keep it concise, link to clean content, describe each link briefly, and order things by importance. You're writing a curated guide, not dumping your whole sitemap into Markdown.

A practical template

Here's a straightforward template you can adapt. Replace the placeholders with your own site's details.

```
# Example Company

> Example Company helps small businesses manage inventory with simple, affordable software. This file guides AI systems to our

We're a SaaS company focused on inventory management for retail and e-commerce. The links below cover our core product, docume

## Documentation

- [Getting Started Guide](https://example.com/docs/getting-started.md): How to set up your account and add your first products
- [Inventory Sync Reference](https://example.com/docs/sync.md): How real-time sync works across sales channels.
- [API Overview](https://example.com/docs/api.md): Endpoints, authentication, and rate limits.

## Guides and Explainers

- [Inventory Basics](https://example.com/learn/inventory-basics.md): Core concepts for anyone new to inventory management.
- [Reducing Stockouts](https://example.com/learn/stockouts.md): Practical methods to avoid running out of popular items.

## Optional

- [Company History](https://example.com/about/history.md): Background on the company, safe to skip for most queries.
```

Keep the descriptions genuinely informative — they help a model decide what to pull. And point the links at clean pages: if you can serve Markdown versions of your key content, do it, because that's the whole efficiency win the format is chasing.

How do AI systems actually use llms.txt?

Here's where I'll be blunt, because this is where the hype tends to run ahead of reality. **Adoption is still early, and support is not universal or guaranteed.** llms.txt is a proposal that some tools and platforms have adopted and many have not. Publishing one does not mean every AI system will read it, honor it, or change its behavior because of it.

What llms.txt is designed to enable, where it's supported:

- **Efficient retrieval.** A tool that reads your llms.txt can jump straight to your curated, clean content instead of parsing cluttered HTML — saving context and reducing the chance it grabs the wrong thing.
- **Better comprehension.** Clean Markdown is easier for a model to interpret accurately, which can mean fewer misreadings of what you actually offer.
- **Prioritization.** Your ordering and your "Optional" section signal what matters most, helping a system spend its limited attention on your best material.

And here's what it is **not**:

- It is **not** a ranking mechanism. There's no evidence that simply having an llms.txt file makes you rank higher or get cited more in AI answers on its own. It's a comprehension and retrieval aid, not a magic visibility lever.
- It is **not** enforcement. Like robots.txt, it's a convention that cooperative systems may choose to follow. It can't compel any AI system to do anything.
- It is **not** a replacement for good content, structured data, or the rest of your AEO work. It's one small, sensible piece of a larger picture.

My honest take: llms.txt is **low-cost, low-risk, and directionally smart**. The effort to publish a good one is modest, it can only help comprehension where it's supported, and it costs you little if support stays partial. That's a reasonable bet. Just don't publish one and expect your AI visibility to transform — that's not what it does, and anyone promising otherwise is overselling.

Should you add an llms.txt file?

For most content-rich sites, yes — with realistic expectations. It's cheap insurance that your most important content is easy for AI systems to find and read cleanly, and adoption is trending upward, not down. The downside is minimal: a little upfront work to curate and maintain the file.

The one caveat is maintenance. A stale llms.txt that points to dead links or outdated pages is worse than none, because it actively misdirects. Treat it like any other piece of your public documentation — keep it current, keep it curated, and update it when your important pages change. Do that, and you've done a small, sensible thing that helps AI systems understand you on your terms.

Key takeaways

- 🕒 llms.txt is a proposed standard — a Markdown file at your site root that gives AI systems a curated, clean guide to your most important content.
- 🕒 It does a different job than robots.txt (access control) and sitemaps (full inventory): llms.txt is about curation and comprehension.
- 🕒 The file contains a summary and prioritized lists of links, ideally pointing to clean Markdown versions of your key pages.
- 🕒 Adoption is still early and support isn't universal; publishing one doesn't guarantee any AI system will read or honor it.
- 🕒 llms.txt is not a ranking lever and not enforceable — it's a low-cost comprehension aid, not a magic visibility boost.
- 🕒 It's a reasonable bet for most content-rich sites as long as you keep it current; a stale file that points to dead links is worse than none.

Frequently asked questions

What is llms.txt?

It's a proposed standard: a single Markdown file placed at the root of your website (yourdomain.com/llms.txt) that gives AI systems a curated guide to your most important content. It typically includes a short description of your site and prioritized lists of links to clean, easy-to-read versions of your key pages. The goal is to help large language models find and understand your best content without parsing cluttered HTML.

How is llms.txt different from robots.txt?

robots.txt is about access control — it tells crawlers where they may and may not go using allow and disallow rules. llms.txt is about curation — it highlights which of your pages matter most and presents them cleanly for AI systems to read. They solve different problems and can coexist on the same site.

How is llms.txt different from a sitemap?

A sitemap aims to list every URL on your site so nothing gets missed — it's exhaustive by design. llms.txt is the opposite: it's selective, highlighting only the handful of pages that matter most and presenting them in clean Markdown. A sitemap is inventory; llms.txt is curation.

What goes inside an llms.txt file?

A typical file has an H1 with your name, a short blockquote summary of what you do, optional context, and sections of H2 headings each containing a Markdown list of important links with brief descriptions. Links should ideally point to clean Markdown versions of your key pages. An optional section can flag lower-priority links that can be skipped.

Do AI systems actually read llms.txt?

Some tools and platforms support it and many do not — adoption is still early and not universal. Publishing an llms.txt file does not guarantee that any given AI system will read it, honor it, or change its behavior. Where it is supported, it can make retrieval more efficient and comprehension more accurate.

Will adding llms.txt improve my rankings in AI answers?

There's no evidence that simply having an llms.txt file makes you rank higher or get cited more on its own. It's a comprehension and retrieval aid, not a ranking mechanism. It can help AI systems read your content cleanly where the file is supported, but it's not a substitute for good content, structured data, and the rest of your AEO work.

Is llms.txt an official standard?

No. It's a community proposal, not a standard ratified by an official web standards body like robots.txt effectively is. It has gained meaningful traction and adoption, but it remains an emerging convention that cooperative systems may choose to follow rather than a formally mandated requirement.

Should I add an llms.txt file to my site?

For most content-rich sites, yes — it's low-cost, low-risk, and directionally sensible, with adoption trending upward. Just keep your expectations realistic: it aids comprehension where supported rather than transforming your visibility. The main caveat is maintenance, since a stale file pointing to dead or outdated links is worse than having none at all.

About the author. Scott Tischler is the Founder & Chairman of Alrecommmend.ai and a practitioner-authority on AI search and Answer Engine Optimization. With 20+ years in marketing technology — including American Express, MetLife, and UBS — and executive study at Wharton, Harvard, Yale, and Oxford, he helps businesses become the ones

AI recommends.