

THE NEW RULES OF AI SEARCH

**SEO FOR VISIBILITY IN CHATGPT,
GOOGLE GEMINI AND GENERATIVE SEARCH**

CRYSTAL CARTER

THE NEW RULES OF AI SEARCH

THE NEW RULES OF AI SEARCH

**SEO FOR VISIBILITY IN CHATGPT,
GOOGLE GEMINI AND GENERATIVE SEARCH**

CRYSTAL CARTER

WILEY

This edition first published 2026

© 2026 Crystal Carter

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Crystal Carter to be identified as the author of this work and has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, New Era House, 8 Oldlands Way, Bognor Regis, West Sussex, PO22 9NQ, UK

John Wiley & Sons Singapore Pte. Ltd, 134 Jurong Gateway Road, #04-307H, Singapore 600134

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

The manufacturer's authorized representative according to the EU General Product Safety Regulation is Wiley-VCH GmbH, Boschstr. 12, 69469 Weinheim, Germany, e-mail: Product_Safety@wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty:

While the publisher and the authors have used their best efforts in preparing this work, including a review of the content of the work, neither the publisher nor the authors make any representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data is Available:

ISBN 9781394399697 (Cloth)

ISBN 9781394399703 (ePUB)

ISBN 9781394399710 (ePDF)

Cover Design: Wiley

Cover Image: [ibragimovN/stock.adobe.com](https://www.gettyimages.com/detail/stock-photo/ibragimovN/stock.adobe.com)

*Dedicated to Cass, Sam, my Mom, my Dad, my Sister,
Annie, and all the entire group chat*

Contents

Chapter 1	The New Rules	1
Chapter 2	Be Real About How Your Customers Use AI Search	9
Chapter 3	Benchmark Before You Start	21
Chapter 4	Stay Ready So You Don't Have to Get Ready—for Data Ingestion	33
Chapter 5	One Strategy Doesn't Fit All—Content Strategies for Every Type of LLM Response	49
Chapter 6	Optimizations for Training-Data Responses	59
Chapter 7	Optimizations for RAG Responses	71
Chapter 8	New Brains, Same Internet: Reasoning Models for Marketers	83
Chapter 9	Optimizations for All LLM Response Types	95
Chapter 10	Embrace Experimentation	115
Chapter 11	Build Content Partnerships with AI Search Tools	131
Chapter 12	Be Prepared for the Agentic Era of AI Search	137
Chapter 13	Measure with the Right Stick	147

Chapter 14	“Rank” in AI Responses	157
Chapter 15	Track Your Performance: Making Prompt Tracking Work on Any Budget	165
Chapter 16	Tracking LLM Traffic: Finding Your Brand in the AI Sandbox	177
Chapter 17	Really, Really Know Your Customers	185
Chapter 18	Be the Answer They Can’t Ignore	193
	<i>Resources</i>	197
	<i>Acknowledgments</i>	201
	<i>Index</i>	203

1

The New Rules

AI search optimization is an emerging field that has burned through a handful of names in record time. You might know it as large language model (LLM) optimization, generative engine optimization (GEO), or even answer engine optimization.

But underneath the buzzwords, it's simply the process of making your brand more visible in AI chat responses and their underlying ecosystems. This includes optimizing directly for LLM responses through content strategy and technical integration, but it also means creating accessible pathways for AI tools to interact with your brand across the web.

In this book, we will explore GEO from the perspective of optimizing for LLM responses and integrating your business into these new AI-driven ecosystems.

Rule 1: Conversational Search Is the Focus

Before we get started, I should clarify right out of the gate: we are not covering brand visibility in Google's AI Overviews extensively in this book. That might seem counterintuitive since they are the hot topic *du jour*, but there is a fundamental difference between AI Overviews and LLMs when it comes to the user journey.

People actively use LLMs for search and information discovery. While they aren’t traditional search engines, users purposefully opt into them to find answers. In contrast, AI Overviews just appear in standard search results, like a featured snippet or a “People Also Ask” box.

LLMs for search—user journey	AI Overviews in Google search—user journey
Users actively choose LLMs for search	Users happen upon AI Overviews passively
Users control which LLM content they wish to explore	AI Overviews generate according to system needs

Because users never opt in, give consent, or even know for certain that they will see an AI Overview for a specific query, the user journey is largely passive. With LLMs, users lean in: they give feedback, ask follow-up questions, and go deeper.

This conversational experience is the playground we will explore in this book.

Why the Shift Matters for Brands

As marketers, we already know how to optimize for people who choose to visit Google or Bing. What we are tasked with now is adjusting our framework for users who seek information directly from LLMs.

These users are actively opting into having their questions and answers mediated by an AI. They build relationships with AI assistants that they simply do not have with a standard search engine. Classic search engines don’t laugh at your jokes, help you brainstorm, or keep you company with conversational dialogue on a long drive. Additionally, chatbots have the ability to generate assets, media, and dynamic tools on the fly in a way that search engines cannot.

The data backs this up. In a landmark 2025 study, OpenAI analyzed 1.5 million conversations and found that while users turn to chatbots simply to “seek information” roughly 24% of the time, their top use case is actually “practical guidance” at nearly 29%. Furthermore, about half of all messages sent to ChatGPT fall into the category of “Asking”—meaning users value the AI primarily as an advisor and decision-support tool rather than just a task executor.

These conversational intent areas—practical guidance, advising, and technical help—are exactly where your brand can step in to contribute expertise, shape the narrative, and provide solutions.

Rule 2: Follow the Direction of Travel

Why should brands care about GEO? It essentially comes down to the direction of travel. Specifically, LLM use is growing rapidly.

In 2025, ChatGPT brought in a staggering 63.3 billion visits over the course of the year. The year before, it brought in just 24 billion. It was the fastest app to ever reach a million users, and the name itself has become synonymous with the technology. Combine this with the fact that Google has jumped feet-first into the AI fray—integrating Gemini into their phones, Fitbits, and search experience—and the case is clear: if brands want to be in front of users, they must take these channels seriously.

Users are spending roughly seven minutes per session and visiting around four pages per visit on ChatGPT. That is a lot of time that a lot of users aren’t spending in traditional search engines. And ChatGPT isn’t the only LLM commanding this attention.

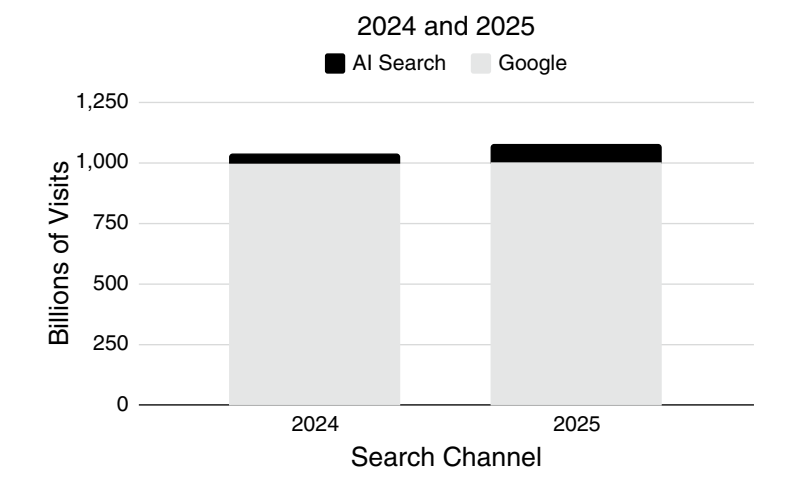
The 82-Billion Visit Reality

Across the web, there is a whole host of LLM-based chatbot channels capturing users’ time. Data from SimilarWeb in 2025 shows over 82 billion combined visits across the top players.

LLM platform	Global monthly visits 2025
ChatGPT	63.3bn
Gemini	9.2bn
Deepseek	4.7bn
Grok	2bn
Perplexity	1.7bn
Claude	1.5bn
Total AI search	82.5bn

ChatGPT leads the charge, with an overwhelming market share of 76.7%, but a diverse mix of players continues to rapidly grow their user bases.

The momentum is decidedly directional. When we compare aggregate traffic from the top LLM platforms with traffic to Google in 2025, we see that 7.7% of traffic went to LLMs. If this seems small, consider that in 2024, LLMs only accounted for 2.8% of that share.



So, the growth of AI channels is happening at an incredibly fast rate, giving way to a massive new opportunity that even Google itself is embracing.

Rule 3: Google Is Shifting Too

Google is a behemoth. It averages about 83 billion visits a month, and even on a slow month, they pull in vastly more visits than any single competitor.

Current research suggests that GEO is not taking massive chunks of customers away from Google.com just yet. In 2025, visits to Google were down by about half a percent compared to 2024—dropping from 999 billion to 993 billion. That minor drop falls within a margin of error, but it still translates to roughly 6 billion fewer visits than the year before.

Google traffic		
2025	2024	Difference YoY
993bn	999bn	−0.5%

However, the Google search results page has become incredibly AI-forward since the emergence of ChatGPT. Google’s own Gemini pulled in 9.2 billion visits in 2025, more than making up for any potential loss in classic search. They are hedging their bets. With Google rapidly integrating Gemini into its tooling for Google Fit, Google Assistant, and AI mode, plain blue links may become less and less visible. When that happens, the lines between AI search and Google itself will blur entirely.

Rule 4: Prepare for Frictionless Search

People often ask me when we should expect the total shift to AI search, and I always think about music.

Back in high school, a friend showed me an MP3 player. I was perfectly happy with my Sony Discman, but he told me this clunky new machine could hold dozens of songs—way more than my favorite mix CD. I looked at the device with all its weird buttons and thought, “*Sure, but I’ll stick with my Discman.*”

Fast forward a few years, and Apple drops the iPod. Suddenly, there were marketing campaigns with dancing silhouettes, a sleek click-wheel, and the promise of a thousand songs in your pocket. We saw a massive, rapid transition from CDs to digital downloads. It wasn't because people suddenly became obsessed with technological advances; it was because the new tool made getting what they already wanted totally frictionless. CD sales plummeted from over 95% of market share in 2002, to 31% of sales in 2012, to just 2% by 2022.

I share this to say that the transition from classic search to AI search will be slow—until it's fast. Just because traditional search currently has an enormous market share doesn't mean user habits won't flip overnight. In fact, major platforms are taking the iPod route right now, integrating generative AI into everyday smart devices to make opting into AI search completely effortless. We are already seeing examples of this from Baidu in China, as well as Microsoft's Bing.

The takeaway: The shift isn't about replacing search; it's about removing friction. Brands must be ready when their audience trades their "Discman" for an "iPod."

Rule 5: Expand Your Digital Footprint

What AI search represents is a shift from users seeking pre-packaged information to users connecting with platforms that dynamically reshape information into whatever format is most useful to them at that exact moment.

Take Google's NotebookLM, for example. It can take a collection of content in any format and allow you to review it as an AI-generated podcast, flashcards, or even an infographic. This makes information accessible for all kinds of learners and needs.

In this new paradigm, your website is your Web HQ, but the entirety of your digital presence is up for grabs as AI tools shape assets and brand information for their end users. Rather than

optimizing for single keywords or a single linear user journey, we need to consider the full scope of our digital footprint.

Your New Rules Playbook

In this book, we will talk about how you can assess where you are in this AI moment by benchmarking your activity and analyzing your audiences. We will explore how LLMs work, how to tailor your strategy for different types of LLM responses, and drill down into techniques that I've proven particularly effective.

This book is designed for practitioners who need high-level oversight but also need to deliver day-to-day. We will move quickly through straightforward tactics, but for strategies that need more fleshing out, we'll get into the weeds.

Throughout the book, you will see prompts designed to help you execute these strategies.

Try This Prompt

My business serves customers in [insert list of areas of operation]. You are an AI search strategist. Create a table ranking these locations by AI adoption. Create columns for the region, AI adoption rate, and population size.

These prompts can be used with any AI assistant and have been tested so that if you change the variables to match your needs, you will get a highly useful response.

Where possible, we've also added checklists to help you make sense of the opportunities. Paired with AI, you can use these lists to create guidance for your own custom GPTs, Gemini Gems, or Perplexity Spaces.

With these tools and strategies, you should be ready to take on the new era of search.

2

Be Real About How Your Customers Use AI Search

You may be fully bought into the AI revolution, but are your customers? If so, which platforms are they using? Because what marketers do every day is very different from what teachers or nurses or tree surgeons do.

So, before you commit time and resources to your AI search campaign, take a beat and find out where and how much your customers are using these platforms. Once you have a good idea of your customer AI adoption, you can:

- Determine the investment levels for AI channels
- Identify your best AI channels
- Decide on the implementation pace (fast-track versus gradual uptake)
- Gauge the estimated return on your investment

Identify Your Target Market with Existing Customer Data

The first place to start in the process of qualifying your AI audience is by reviewing your existing customer data. I know this is not shiny and new, but AI technology does not change the fact that businesses are solving real-life challenges for humans.

So, as the rock-star marketing team that you are, you should already have a good idea of who your target market is, what their pains are, and where they tend to gravitate online. As well as tapping into broader market statistics, chat with sales for common questions, dig into customer service requests, keyword research, internal search, social analytics, customer surveys, reviews, and website data to make sure you know who you're talking to.

And if you are working on a new brand or product, you can research competitor reviews, user-generated content like Reddit threads, and customer communities to identify who is looking for the solutions that your business offers.

For instance, if you were opening a pet-sitting business in Chicago, you might search for communities of pet owners, then filter for folks asking for pet sitting. Equally, you might review Google Business Profiles of similar businesses.

Try This Prompt

My business offers solutions for customers that need [describe your solution] in [location] with the added USP of [USP]. You are a market research analyst. Create a brief report of this target market. Create three personas that exemplify the most significant demographics for this market, what their pains are, and the top three websites each persona uses online.

This kind of long-tail query research can pair well with broader statistics to uncover real needs before chasing AI trends. And the pain points you identify here can be used to create example prompts for tracking later.

Once you know who your customers are, you can assess if and how much they have adopted AI.

Gauge the AI Adoption Rate for Your Target Market

Adoption of AI technology varies across demographics. So, understanding the potential of the AI search market for your audience is critical to deciding if, how much, and how quickly you should invest in your LLM optimizations.

Don't get me wrong, the rate of adoption for AI tools has been incredible. Folks have taken to this faster than any other technology in history—more than electricity, cellphones, and the internet itself. Within less than three years from the release of ChatGPT, 1.2 billion people had used AI tools.¹ That's around one in eight people on Earth tapped into AI.

If your audience is using AI frequently, then you will be more likely to get a valuable return on your AI search investment. If your audience is more cautious, then you can build foundations for the AI search shift in your own time.

To assess your audience's likeliness to turn to AI for information and brand discovery, consider

- Geography
- Language
- Gender
- Age
- Technology affinity
- Niche

¹ Microsoft (2025). *AI Diffusion Report: Where AI is most used, developed, and built*. AI Economy Institute. <https://www.microsoft.com/en-us/research/wp-content/uploads/2025/10/Microsoft-AI-Diffusion-Report.pdf>

Geography

Global uptake of AI varies and is concentrated in wealthier nations. At the end of 2025, nations like the United Arab Emirates and Singapore had more than 60% AI user rate, while Costa Rica came in around 27%. In a list of the top 30 countries whose working-age population was using AI, the United States came in 24th with 28%.

Rank	Country	AI adoption rate
1	United Arab Emirates	64.0%
2	Singapore	60.9%
3	Norway	46.4%
4	Ireland	44.6%
5	France	44.0%
6	Spain	41.8%
7	New Zealand	40.5%
8	Netherlands	38.9%
9	United Kingdom	38.9%
10	Qatar	38.3%
11	Australia	36.9%
12	Israel	36.1%
13	Belgium	36.0%
14	Canada	35.0%
15	Switzerland	34.8%
16	Sweden	33.3%
17	Austria	31.4%
18	South Korea	30.7%
19	Hungary	29.8%
20	Denmark	28.7%
21	Germany	28.6%

Rank	Country	AI adoption rate
22	Poland	28.5%
23	Taiwan	28.4%
24	United States	28.3%
25	Czech Republic	27.8%
26	Italy	27.8%
27	Bulgaria	27.3%
28	Finland	27.3%
29	Jordan	27.0%
30	Costa Rica	26.5%

Adapted from <https://www.microsoft.com/en-us/research/group/aici/ai-diffusion/>

Brands with international audiences should prioritize AI optimization for markets with high AI saturation over regions with less.

Try This Prompt

My business serves customers in [insert list of areas of operation]. You are an AI search strategist. Create a table ranking these locations by AI adoption. Create columns for the region, AI adoption rate, and population size.

Language

Language is also a consideration for adoption with English speakers benefiting from models which have been trained on English-heavy websites and data. For other languages, model improvements can cause significant uptake of AI. South Korea,

for example, had good AI adoption rates, but became the second largest paid market for ChatGPT, outpaced only by the United States,² after the GPT 5 update led to improved Korean language understanding. Customers flocked to the tool because it served them better than other chatbots.

If you plan to serve customers in non-English regions, identify the model with the best understanding of the language. Major AI players like ChatGPT and Gemini will certainly be in contention, but government initiatives³ and regional preferences may mean that more niche models are also being used locally.

Try This Prompt

My business serves customers who speak [insert language—e.g., Tagalog, Bengali, Portuguese]. You are an AI search strategist helping to increase AI visibility with my potential audience. Which AI assistant best understands and is most used by customers who speak this language?

Gender

AI adoption follows distinct demographic patterns that savvy marketers can't ignore. While men currently represent 59% of AI users globally in 2025, women are closing the gap at an accelerating pace. In the United States alone, women have tripled their AI adoption rates, fundamentally reshaping platform demographics. ChatGPT's user base exemplifies this shift, as the platform launched with an 80% male audience, but was over 50% female in less than two years.

²<https://www.koreaherald.com/article/10500190>

³<https://www.weforum.org/stories/2025/10/generative-ai-languages-llm/>

Platform choice matters strategically. ChatGPT commands the largest female user base at 46%—nearly gender balanced—positioning it as the optimal choice for reaching women audiences. Grok, by contrast, captures only 34% female users, making it a weaker investment if female visibility is your goal.

Age

Age segmentation reveals similarly dramatic differences. Gen Z and younger millennials dominate adoption, with 70% of 18–24-year-olds using generative AI regularly. This drops sharply to just 34% for women over 55. Millennials have emerged as the daily-use power users across all demographics.

Operating Systems

ChatGPT held on to the vast majority of AI tools' market share for much of 2024 but in 2025, Google Gemini started gaining ground. In fact, the platform doubled its monthly visits within six months. This didn't happen because their tech was suddenly much better or because users suddenly decided they couldn't live without Gemini.

The shift happened because Google started building Gemini into their technology. First, Gemini was added into Google Pixel watches, then into Google Pixel 10 phones and Android, and not soon after, Google Chrome started rolling out Gemini in Chrome agentic browsing.

In 2025, Google also began deprecating Assistant in favor of Gemini. Originally launched in 2016, Google Assistant had over 1 billion users as of 2024.⁴ In 2025, Gemini became the default assistant in Google Pixels, Google Home, and other products. If that wasn't enough, Google announced a partnership with Apple to power Siri with Gemini as well.

⁴ <https://developers.google.com/assistant/why-build>

For brands, this means that the barrier for users to access these tools will become lower as the technology becomes more integrated into our day to day. In this environment, users won't need to make a distinct choice to use AI for search or information discovery, because Gemini is increasingly becoming a default service for users of Google products.

Similarly, in recent years, PC laptops are being sold not only with Microsoft Copilot built in, but with a button to open Copilot on the physical keyboard.

Speaking of laptops, it's also worth noting that Apple users are more likely to use agentic web browsers (more on that in Chapter 12) since ChatGPT Atlas. Many other agentic web browsers are also available exclusively on MacOS.

Brands should thus take the operating systems into consideration when assessing the potential long-term customer reach for a particular AI channel.

Once you've done the macro work, it's time to get into specifics.

Calculate Adoption Market Share and Total Serviceable Market

Whether you're the business owner weighing up your precious time, the CMO getting budget sign-off, or the CEO answering to shareholders, at some point you're likely to need to validate why you're investing in AI search. And you will need some market data for that. So, let's get it.

Calculating AI Preference

Based on our research above, you might know that your audience has some general interest in AI. But where can we find out if users are actively choosing AI over more classic search surfaces?

Market research tools like Sparktoro, created by Rand Fishkin, can help you to see if *your* users are turning to ChatGPT, Claude, Gemini, and the rest more than they are using other

channels. And this information is critical for how you proceed with your GEO strategy.

As Fishkin explains:

Don't ask your audience via surveys or interviews whether they use AI tools. Don't try to back into it via your analytics. And don't make assumptions based on stereotypes or preconceived notions (or worse, the highest paid person's opinion). Get real, panel-observed data from a clickstream source that sees what millions of devices do on the web, and narrow to the group of people you care about reaching.⁵

With Sparktoro, you can use the demographic topical data to:

- Use a dedicated panel designed for AI and search tools to understand your audience's AI-adoption patterns.
- Target specific geographic markets by applying region filters.
- Identify audiences based on their search behavior and interests through keyword filters.
- Segment your audience by the AI tools and platforms they use.

All of which is incredibly valuable for marketers and can clearly illustrate whether your audience is ahead of or behind the curve with regard to AI adoption.

My favorite example of this is *programming* versus *potatoes*. These are two very different audiences and data shows that they use search in very different ways.

According to data from Sparktoro, users who are searching for *programming* are 112% more likely than the average US user

⁵ Fishkin, Rand. "Is Your Target Market Using AI? How to Find Out." Wix Studio AI Search Lab, 19 May 2026, www.wix.com/studio/ai-search-lab/is-your-target-market-using-ai

to use ChatGPT for search. But when you look at searches for *potatoes*, ChatGPT isn’t even a top-seven discovery channel, and in each case, use of Google is roughly the same.

For a brand targeting programmers, being visible in ChatGPT is absolutely essential. But if you were planning an event for potato enthusiasts, other channels could yield much better results because the potato market still has a way to go.

Defining Your Total Serviceable Market

With this information, we can really get into the weeds to create average search volume estimates and define your total serviceable market.

For instance, you could use historic keyword search volume data to calculate how many monthly searches are likely to originate in AI tools for specific terms by weighting traditional keyword volume against your audience’s specific habits.

Variable	Name	Definition
ASV	Keyword search volume	The total monthly searches for a specific term on traditional search engines (e.g., Google).
CGPT	ChatGPT base usage	The percentage of the general population (control group) that uses ChatGPT for common tasks.
Rel	Relative usage difference	How much more or less your specific target audience uses AI compared to the general population.
AIU	Average market use of AI	The overall percentage of the market that has adopted AI tools as part of their daily routine.

This is the kind of data that gets sign-off from clients and stakeholders:

$$\text{Estimated AI monthly search volume} = \text{ASV} \times \frac{\text{CGPT} \times (1 + \text{Rel})}{100} \times \text{AIU}$$

3

Benchmark Before You Start

You’ve sized up the market, now it’s time to turn the lens on *your* brand. Are your existing customers already finding you through AI tools? This baseline audit isn’t just due diligence—it’s your secret weapon for sustainable scaling.

Without knowing what AI already “likes” about your site, you’re flying blind. Are LLMs quietly sending you traffic? Is your best content already resonating with AI-driven searchers? The answers will justify your budget, win you executive buy-in, and show you exactly where to double down.

We’ll tackle this in four steps, from dead simple (minutes to implement) to more technical (but massively revealing):

- “How did you find us?” form tweaks—immediate results, zero development time.
- Analytics deep dive—spot historic AI referrals and growth patterns.

- Bot log reports—see exactly which AI crawlers visit, what they hit, and even hallucinated URLs screaming for fixes.
- Benchmark creation—turn raw data into your AI traffic growth roadmap.

By the end, you'll have hard numbers: current AI interaction levels, quick wins, and realistic growth potential. No more "AI is the future" platitudes—this is *your* before picture for proving the return on investment after.

The “How Did You Find Us?” Question Is Your Sneaky Superpower

If you feel like LLM traffic is hiding in your analytics, you're not wrong. A lot of AI-driven visits show up as “direct,” “other,” or just . . . vibes. The good news is you already own one of the simplest, most powerful attribution tools on the planet: that tiny “How did you find us?” question on your contact and demo forms.

This field is doing more work than most dashboards, because it lets the user narrate their own journey in their own words. When someone tells you “ChatGPT recommended you” or “I asked Perplexity who the best tools were for X,” they're giving you something Google Analytics 4 (GA4) can't: intent, context, and the exact channel that got them over the line.

Update the Question for the AI Era

Most teams still treat this like it's 2014: “Google,” “Social,” “Friend,” “Other.” In an AI search world, that's not enough. You need to explicitly make space for LLMs to show up as their own channel.

For example, you might change your options to:

- Search engine (Google, Bing, etc.).
- AI assistant (ChatGPT, Perplexity, Gemini, Copilot).

- Social media (LinkedIn, X, TikTok, etc.).
- Referral from a person or company.
- Other (please specify).

The wording matters. When you say “AI assistant,” you’re jogging people’s memory—“Oh yes, I did ask ChatGPT about this”—and that nudge dramatically increases how often LLMs show up as a credited source. Make it a required field at the point of *high intent* (contact, quote, demo, checkout), not at the top of the form where it feels like friction.

Why This Beats Waiting for Better Tools

Right now, AI attribution is murky. Referrers can break, URLs get copied and pasted, and some AI flows never pass a neat UTM tag along. That’s why this old-school question is so valuable: it sits at the moment of commitment. The person who just filled in a high-intent form will usually tell you the truth about what sent them there, because they want you to understand them and serve them properly.

Think of it as your LLM lie detector. Your analytics might swear most of your conversions are “direct,” but your form field quietly reveals: “ChatGPT,” “Perplexity,” “Gemini AI Overview,” “Copilot answer.” Once you see that pattern, it changes the conversation in the boardroom from “We think AI is important” to “X% of closed deals started in LLM search; here’s why we’re investing in GEO.”

Some people will argue that self-reported data is unreliable, but in today’s incredibly messy middle, it’s one of the few ways to check whether your attribution actually matches the real customer journey. There is a clear emerging pattern where users do most of their decision-making with an AI assistant, then hop over to Google for a quick branded search and convert. In that scenario, it’s not classic search engine optimization (SEO) that truly closed the deal—it’s the AI assistant that did the heavy lifting.

By adding “How did you find us?” to your forms, and even scripting it into your sales calls, you get a window into a touch-point that is otherwise almost impossible to quantify. Someone who has spent 10 minutes with a chatbot walking them through options is far more likely to remember, and happily tell you, that it was “ChatGPT” that led them to the finish line—even if the last visible click in your analytics was a simple Google brand search.

The screenshot shows the Seer Interactive website. On the left, a red sidebar contains the text: "[NEED HELP SOLVING A PROBLEM?]", "How Can We Help?", "What are the digital marketing questions that keep you up at night?", "Tell us more and a member of our team will be in touch within one business day to schedule a discovery call.", and "Searching for career opportunities? [View Open Roles](#)". On the right, a white contact form titled "Our Partnership Together Starts Here" is displayed. The form includes fields for "Email*", "First Name*", and "Last Name*", each with a red asterisk indicating it is required. Below the "First Name*" field is a dropdown menu with the following options: "Please Select", "Google Search", "Conference / Webinar", "Seer Team Member", "Agency Referral", "Referral", "Chat GPT/AI" (highlighted in blue), "Seer Email/Social", "Seer Event/Blog", "Podcast", "Outreach", and "Other". Below the dropdown is another "Please Select" field. A red error message "Please complete this required field." is visible below the "Please Select" field. At the bottom of the form is a red "Submit" button.

How to Track Your Progress

But what about the data? It's true that this won't feel as neat as standard analytics, but it's still absolutely worth tracking. When you connect “How did you find us?” responses into your existing

systems, you turn a simple form field into a proper key performance indicator (KPI) input.

If you're using customer relationship management (CRM) tools to manage customer data, you can get very granular by tying the origin of the lead to the eventual conversion value. The easiest way to start is to ask your AI assistant how to wire this up for your specific stack.

Try This Prompt

I use [CRM name] and I want to collect data from a “How did you find us?” form for each lead. What is the best way to do this for my CRM?

If you're using a standalone form system, you can still track this at the reporting layer by piping form submissions into a Looker Studio dashboard. That lets you see, over time, how many leads mention “AI assistant” or specific tools like ChatGPT or Perplexity.

Try This Prompt

I have a “How did you find us?” form on my website for each lead. The form is built using [form tool]. Can you give me a step-by-step process for how to add data from this form to a Looker Studio dashboard?

Getting eyes on this data from day one gives you a baseline. From there, you can measure whether your AI search optimization is actually moving the numbers, not just generating screenshots of LLM answers.

Review Your Analytics to Find Hidden LLM Traffic

Measuring website traffic from LLMs is new, but you can repurpose the same analytics muscles you already use for organic search and crawl bots. The key shift is to stop looking *only* at Google Search Console and Bing Webmaster Tools for “organic” traffic, and start looking at every place that records where a session came from—your third-party analytics, and your website platform’s built-in reports.

Traditional search tools only show you what came from classic organic search. They will not show the full picture of people who clicked from ChatGPT, Perplexity, Gemini, or Copilot. To understand your real AI footprint, you need to see where traffic is being attributed—referral, direct, “other,” or a custom channel—and what share of total traffic is coming from AI-driven sources versus everything else.

Track LLM Traffic in GA4

Analytics tools like GA4 can be configured with “Explorations” and custom reports to identify and analyze sessions from LLMs. Include this in your regular reporting so you’re not just tracking blue links, but also the chatbots that send people your way. This gives you a more holistic view of your traffic mix and a proof-of-concept data trail before you put serious budget behind generative search optimization.

A practical way to do this is to isolate page referrer traffic using a regex filter that matches known LLM domains, for example:

```
^https://\/(www\.meta\.ai|www\.perplexity\.ai|  
chat\.openai\.com|claude\.ai|chat\.mistral\.ai|  
gemini\.google\.com|bard\.google\.com|  
chatgpt\.com|copilot\.microsoft\.com|  
chat\.deepseek\.com)(\/.*)?$
```

You'll notice I've included legacy domains like Bard in this list. That matters when you're analyzing historic data, because older LLM brand names and hosts can still appear in past referrals. The same is true for ChatGPT, which changed domains about a year after launch—this pattern is designed to catch both the older and newer URLs.

If there's an AI assistant that's especially important to your audience and it isn't listed here, you can ask your AI co-pilot to extend the pattern for you. For example:

Try This Prompt

Here's my current GA4 regex for LLM referrers:

```
^https://(www\.meta\.ai|www\.perplexity\.ai|  
chat\.openai\.com|claude\.ai|chat\.mistral\.ai|  
gemini\.google\.com|bard\.google\.com|  
chatgpt\.com|copilot\.microsoft\.com|  
chat\.deepseek\.com)(\/.*)?$
```

Please update it to also include traffic from [AI tool name and domain].

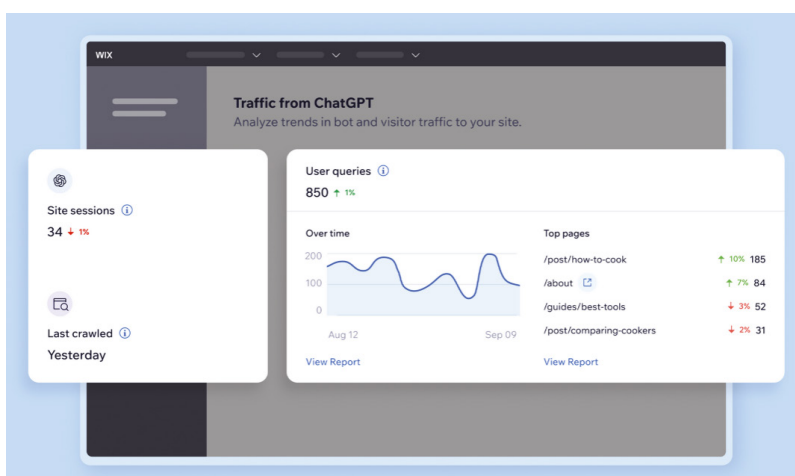
With this pattern, you can build explorations or reports in GA4 that surface sessions where the referrer is an AI tool, so you can see volume, landing pages, and behavior in near real time.

From there, you can:

- Use traffic reports to estimate potential for growth.
- Pipe the data into Looker Studio for easier trend-line dashboards.
- Compare LLM traffic against other channels to understand relative value.

Use Your Website Platform's Analytics

Your website platform may offer better visibility into referrers than external tools, so don't ignore its built-in analytics. Your website analytics will have first-party data that isn't throttled or sampled, making it an important source of truth. Within your platform's traffic reports, go to your referral view and filter for domains that include AI tools—for example, entries that contain “chatGPT,” “perplexity,” “gemini.google,” or “claude.”



This is precisely why my team built the Wix AI Visibility Overview tool into the Wix website builder. Our AI Visibility Overview tool is valuable because you can see the traffic from AI, the queries, and the pages all in the same portal, without connecting any additional APIs. Bringing all of that first-party data together on a single interface can be incredibly effective for helping you make valuable decisions, so don't overlook website analytics.

When you report all this data to stakeholders, set expectations—not all LLM traffic will be neatly labeled as “referral.” In some cases, clicks from ChatGPT and other tools arrive without a referrer and show up as “direct” traffic. That means you

should treat these numbers as a trend line rather than a perfect count—look for direction and growth, not surgical precision.

What You're Looking For

As you review your analytics across tools, focus on three questions:

- Where is LLM-driven traffic being attributed today—referral, direct, or a custom channel?
- What proportion of total sessions and conversions are coming from AI sources versus other channels?
- Are you seeing that share grow over time, especially on key landing pages and commercial intents?

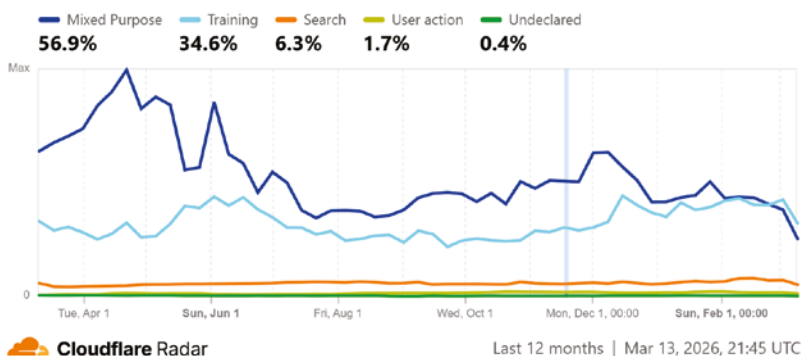
Answering these will tell you whether AI is already influencing your funnel, which content resonates most with LLM users, and how big the upside is if you invest in GEO now instead of waiting.

Use Bot Logs to See Which LLMs Actually Visit Your Site

One of the most powerful, and underrated, ways to measure your AI visibility is to look at your bot logs. On the surface, log files sound deeply technical—something only a developer would touch—but they are one of the few places you can see, in black and white, which LLM crawlers are visiting which pages and what they encountered when they got there. Bot logs record every request to your server, including user agents like OAI-SearchBot, ClaudeBot, PerplexityBot, BingBot, or Google-Extended, so you can distinguish classic search crawlers from AI-specific bots. For instance, globally, around 35% of AI bots traffic to websites is from training data bots. Is that what you're seeing on your website?

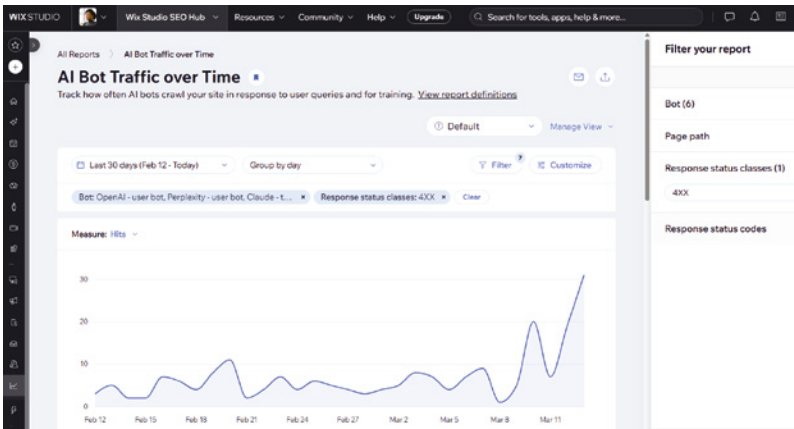
Crawl Purpose Worldwide

Traffic volume by crawl purpose



This level of detail goes far beyond what standard analytics can show. While analytics tools are great for understanding user traffic and referrals, they often obscure or aggregate bot activity and can miss cases where AI tools fetch content without a visible referrer or where traffic is sampled. Bot logs, by contrast, give you the definitive record: timestamps, user agents, URLs requested, and status codes. From that, you can see which LLMs are actually crawling your homepage versus your product pages, which sections of your site are being ignored, and where bots are hitting 404s or other errors.

This is exactly why, as AI search started moving from novelty to real channel, one of the very first GEO-focused requests I made to my team was to isolate LLM bot traffic inside our website analytics. I wanted marketers—not just engineers—to be able to open a report and immediately see: “Here’s how often OpenAI, Perplexity, Claude, and others are crawling our site, here are the pages they’re focusing on, and here’s how that’s changing over time.” The result was a dedicated AI bot traffic report within the Wix analytics environment, making log-level insights accessible through a familiar dashboard.



If you're evaluating this from scratch without a dashboard, you may need to speak to your developers to get the logs. Yes, it can feel a bit technical at first glance. But it is absolutely worth the effort.

Once you can see LLM bot traffic clearly, you can answer questions that no other dataset can touch:

- Are AI crawlers actually reaching the pages we expect to power LLM answers?
- Are they over-indexing on legacy content while skipping newer, better assets?
- Are they repeatedly requesting broken or hallucinated URLs?

With those answers, you can prioritize technical fixes, content updates, and `robots.txt` adjustments with confidence, rather than guessing why your brand is or isn't surfacing in AI-driven results.

Your AI Baseline: From Gut to Numbers

You’ve now got the full audit toolkit—from form nudges that surface hidden LLM journeys to bot logs revealing what crawlers actually see. This isn’t busywork; it’s your proof-of-concept for GEO as a channel.

Picture it like prepping for a road trip: you wouldn’t just gas up and go without checking the map, tires, and oil. Here, your “before” numbers—AI-referral share, form responses naming ChatGPT, bot crawl patterns—become the hard evidence that justifies budget and wins C-suite alignment.

Start small: tweak one form today, filter analytics tomorrow, ping your development team for logs by week’s end. In days, not weeks, you’ll know if AI’s already driving decisions (spoiler: for most brands, it is) and exactly where to amplify.

The beauty? This baseline endures. Six months from now, when stakeholders ask: “Is GEO working?” you’ll flip to fresh data showing traffic lifts, conversion shifts, and crawls deepening on optimized pages.

No more faith-based marketing—just results that scale. Benchmark ruthlessly, then build with precision. Your brand’s AI era starts with knowing where you stand.

4

Stay Ready So You Don't Have to Get Ready—for Data Ingestion

AI tools are constantly crawling the web, gathering data, and training their models. Are you actually ready for them? The way LLMs collect and process information about your brand can come down to a handful of pages or even a few tokens—which means one sloppy FAQ or half-baked product page can end up speaking for your entire brand. As you create and update content, making sure your information is always ready for AI data ingestion can be the difference between being the default recommendation and being invisible.

Before an AI model can say anything smart about your brand, it has to consume or “ingest” your content as data. It

gathers information from all over—from public web pages to structured feeds and curated datasets. These systems pull content at scale—but how do they sort through it all?

When AI tools visit your website, your brand storyline enters a data pipeline where your text and media are ingested, broken into tokens, converted into numeric IDs, and turned into vector embeddings that the model can search and reason over.

This all takes place in four stages:

Ingestion → Tokenization → Token IDs → Embeddings

This process may occur for every piece of content you publish well before a model can recognize it, reference it, or recommend it in an answer.

Are you a passenger princess in all of this? Absolutely not. Even at this early stage, you can take steps to make your content cleaner, clearer, and more effective.

Increase Your Reach to Ingestion

Ingestion is the “first-contact” moment between your brand and an AI model: it’s the phase where the system goes out into the world, collects raw content, and pulls it into a form it can later tokenize and learn from.

During the ingestion phase, bots will crawl and scrape your website for two main reasons:

- To update the model’s core training sets during pre-training or fine-tuning.
- To temporarily pull in fresh web or local data “just in time” to answer a specific query during a retrieval-augmented generation (RAG) response.

As a website owner, you can see when LLMs are processing your content by checking your bot or server logs. Each crawler has a

declared user agent string, and many of the major models publish theirs publicly.

LLM	User agent
ChatGPT	OAI-SearchBot, ChatGPT-User, GPTBot
Copilot	BingBot
Gemini	Google-Extended
Claude	ClaudeBot (plus Claude-User, Claude-SearchBot)
Perplexity	PerplexityBot, Perplexity-User

Each model has a unique process for how they gather information. Google's Gemini builds on Google's existing crawlers, and you can use `robots.txt` directives to tell Google-Extended how much of your site can be used in training and answers. Other AI players have their own crawlers and policies, and not all of them are equally transparent or consistently respectful of `robots.txt`, which means real-world bot behavior can vary by model.

For the latest models, ingestion is increasingly multimodal. All models can process plain text, but newer generations can also handle images, video, audio, PDFs, and other file formats, and multimodal RAG setups let them combine text and visuals into a single, richer data point.

For marketers, this leaves little to the imagination. These systems vacuum up millions of documents across the web, including news coverage, social posts, reviews, documentation, forums, and more, not just the polished blog posts you wish they'd see. In practice, that means your entire digital footprint is fair game for how a model learns to describe and compare your brand, from old press mentions to that long-abandoned help article nobody has touched in five years.

Established brands usually have a structural advantage here because they simply have more historical touchpoints and citations in the wild, which gives models more data to ingest and cross-reference. This means a single contradiction or opinion carries less weight. But for a smaller brand, managing online reputation has never been more important.

In either case, if you're trying to grow visibility in AI search, expanding your brand's legitimate digital reach—more high-quality, on-brand content in more credible places—gives these systems more content to ingest and, ultimately, increases the likelihood you will be included in their answers.

Optimizing for Tokenization

There are lots of ways models can structure and store information, but for website owners, tokenization is one of the most important to understand. It's the moment your content stops being a wall of words and becomes machine-readable building blocks that LLMs use to decide whether, when, and how your brand shows up in AI answers.

When an AI system ingests content—a web page, a PDF, a support doc—it breaks that text into tiny chunks called tokens. Tokens can be full words, sub-words, or even single characters, depending on how often those patterns show up in the model's training data. Using OpenAI's tokenizer, you can see this in action as the string you type is split into a sequence of tokens the model actually counts and bills against.

OpenAI Platform

Tokens	Characters
9	53

The screenshot shows the text 'introducing crystal carter' on two lines. The first line is 'introducing crystal carter' and the second line is 'Introducing Crystal Carter'. The text is color-coded to show individual tokens: 'introducing' (blue), 'crystal' (orange), 'carter' (red), 'Introducing' (blue), 'Crystal' (orange), and 'Carter' (red).

How Case and Context Change Tokens

Tokenization isn't just about splitting on spaces, it's deeply context-aware. For instance, the phrase “introducing crystal carter” has a different tokenization from “Introducing Crystal Carter.” With sentence case and capitalized proper nouns, the tokenizer can more confidently treat “Introducing,” “Crystal,” and “Carter” as whole units.

At a high level, you get something like this:

Phrase	Token split (conceptual)	Token count
introducing crystal carter	introdu / cing / crystal / car / ter	5
Introducing Crystal Carter	Introducing / Crystal / Carter	3

In the first version, the tokenizer is hedging its bets: it isn't sure whether “crystal carter” is a person, a material, or a car brand, so it breaks the phrase into smaller, statistically safe pieces. In the second version, capitalization gives it a clearer signal that these are proper nouns, so it can use fewer, more efficient tokens.

Across a whole website, that kind of difference adds up. A phrase that costs five tokens instead of three might not matter once, but multiplied across thousands of lines of copy and millions of model calls, you're paying for every extra fragment.

Context Windows and Why Efficiency Matters

During ingestion, this tokenization process happens at scale. When people talk about a model's “context window,” they're talking about how many tokens it can process at once—not how many characters or words. OpenAI's GPT-4 model, released in March 2023, had a context window of just 8,192 tokens; newer

models like GPT-5.4, released in March 2026, have a context window of over a million tokens.

As models mature, their vocabularies grow and their pattern recognition improves, so tokenization tends to get a bit less granular over time—but you still pay, in both performance and money, for every token you send.

Why Marketers Should Care About Tokenization

Understanding tokenization helps you manage both your AI budget and your brand. Token efficiency can improve the return on investment (ROI) when you build with LLMs, and choosing more token-efficient formats can help with that as well.

Better AI ROI when you build with LLMs. When you use APIs, agents, or customer-facing AI tools, you are charged by the token. Prompts, grounding docs, CRM exports, FAQs—they all get tokenized. Cleaner, more efficient text means fewer tokens per call, which means lower running costs over thousands of requests.

Using token-efficient formats. At scale, it often makes sense to store prompts, standard operating procedures (SOPs), and grounding documents in plain formats like .txt, .md, or CSV rather than noisy HTML. Because LLMs have to chew through all the markup in raw HTML, they waste tokens on CSS and JavaScript you don't actually care about. One test comparing a markdown .txt file to the same homepage HTML showed around 95× fewer tokens for the markdown version. Over hundreds of calls a day, those “wasted” tokens turn into real money.

For your brand, tokenization quirks also matter for brand names. Some names are typically treated as a single token, while others are often split by byte pair encoding (BPE), a common tokenization method that breaks words into frequent sub-word pieces. When a brand name is split, there's a higher risk the model will misinterpret it during vectorization, which can affect how reliably it recognizes and surfaces your brand.

Can you do anything about it? Yes, you can.

Let's start with car brands as an easy example. In the table below, brands like Toyota, Ford, and Mercedes are processed as single tokens, but others are often split.

Single-token brand names	Split-token brand names (typical)
Toyota	Suzu / ki
Ford	Ch / rys / ler
Mercedes	Must / ang
Audi	K / ia
Honda	Che / vy

For instance, “Kia” at the very beginning of a sentence is often split into “K” / “ia.” But when “Kia” is preceded by a space—as it is when it's the second word in a sentence (e.g., “New Kia”)—it's much more likely to be recognized as a single token. This is due to the space-prefix factor: many tokenizers treat “space + word” as a different, more predictable token than the same word at the very start of a string.

You can check how this works for your own brand using OpenAI's tokenizer or similar tools. If you find that your brand is often split at the start of a sentence, you have two practical options:

- Add a common word before your brand so it becomes the second word (“Why Kia. . .,” “With Kia. . .,” “Discover Kia. . .”).
- In internal prompts or RAG documents, add a leading space or short lead-in before your brand name so it benefits from space-prefix tokenization.

And you can absolutely have fun with this. A brand like Kia could update some of its copy with sentence starters such as:

- Why Kia stands out on every road.
- With Kia, every journey feels effortless.
- When Kia leads, adventure follows.
- Because Kia redefines comfort.
- Discover Kia and drive your story.
- Imagine Kia in your driveway.
- Feel Kia in every curve.
- See Kia and see the difference.
- Drive Kia to elevate your day.
- Experience Kia like never before.

This is the kind of detail that feels tiny but is worth experimenting with, especially on high-value pages.

Try This Prompt

Generate 20 short sentence starters where a common word comes first (e.g., “Why,” “With,” “When,” “Because,” “Discover”) and your brand name [brand] is the second word, so it benefits from space-prefix tokenization but still reads like natural, on-brand marketing copy.

Writing for Token Efficiency

When you’re creating or refreshing content, you can also optimize for token efficiency more broadly—not just by reducing character count, but by emphasizing clarity. In practice, that means:

- writing in active voice and trimming filler phrases
- writing numbers as digits instead of words where it fits your style

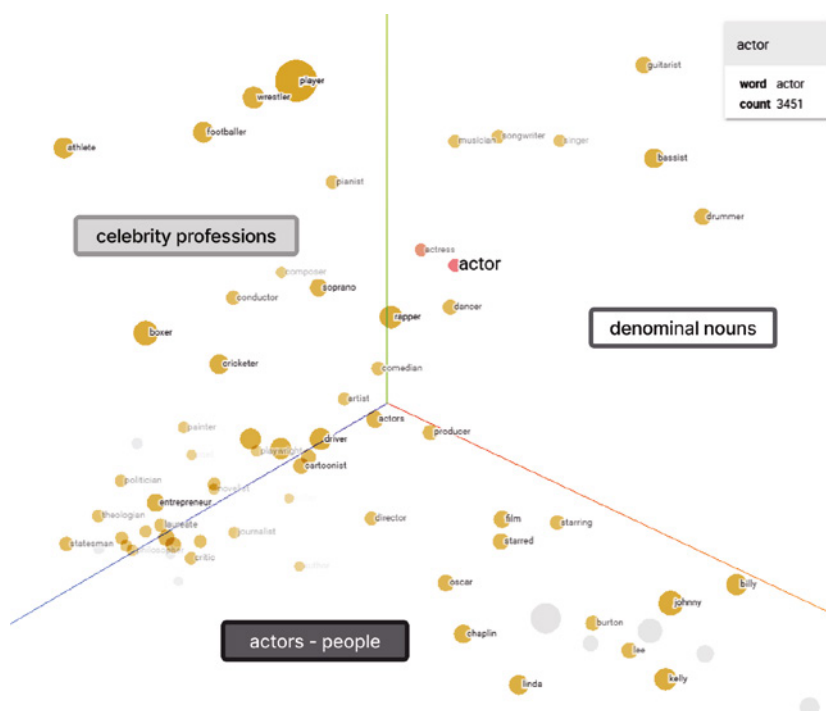
- using abbreviations or acronyms your audience already knows
- reducing repeated nouns when the subject is obvious
- stripping out empty filler words that add no meaning
- merging closely related sentences and using commas or semicolons instead of long “then. . . then. . . then. . .” chains
- choosing clear, unambiguous words over vague, flowery phrasing

Is this absolutely necessary for every piece of content? No. And it's certainly not worth obsessing over for a 600-word blog post. But for big-beast content—skyscraper posts, “ultimate guides,” long-running SOPs, or core grounding docs that can easily hit 10,000+ words—designing for token efficiency, especially in the early sections, can make a real difference to both cost and how much of your content the model can comfortably hold in its context window.

Token IDs and Vector Embeddings

Once your content is tokenized, each token gets an ID—a numeric index in the model's vocabulary. Think of token IDs as the model's internal dictionary lookup: every token you send is mapped to a specific ID before anything more interesting can happen.

Vectorization is where it gets fun. In this step, each token (and often, each chunk of text) is turned into a vector—a long list of numbers that captures its meaning and its relationship to other words and phrases. This is the embedding process: texts with similar meaning end up with vectors that point in similar directions in a shared “semantic space,” so the model can tell that “actor” is closer to “dancer” than to “motorway.”



Vector search itself is not new. For years, semantic and vector search systems have used embeddings to match queries and documents by meaning rather than exact keywords—think of it as mapping the relationships between words instead of just counting overlaps. Today’s LLMs and RAG systems are basically building on that same idea: convert everything to vectors, then look for what sits closest together.

Cosine Similarity: How “Closeness” Becomes Relevance

Under the hood, the model treats phrases that are tightly clustered in vector space as having high cosine similarity—essentially,

a small angle between their vectors. Cosine similarity is just a way of measuring how similar two vectors are in direction:

- close to 1 → the content is very similar or highly relevant
- close to 0 → the content is largely unrelated

This is how vector-based systems calculate relevance: they compare the embedding of your document to the embedding of the user's query and turn that into a relevance score. The closer the vectors, the more relevant your content looks—regardless of whether you used the exact same keyword phrasing.

When a user enters a prompt into an LLM-powered system that uses vector search (e.g., RAG-enhanced ChatGPT or Gemini), the workflow looks roughly like this:

1. The query is converted into an embedding vector.
2. The system compares that query vector against stored vectors for pages, passages, or documents.
3. It selects the chunks with the highest cosine similarity as “most relevant.”
4. Those chunks become the research base and context the LLM uses to generate a natural-language reply.

For your brand, that means well-structured, semantically clear content—especially content that uses explicit relationships and semantic triples—is more likely to sit “in the right neighborhood” in vector space, and more likely to be pulled in when an AI system goes looking for an answer.

Semantic Triples: Helping Vectors “Lock On” to Meaning

One practical way to make your content easier to embed and retrieve is to use semantic triples—clear subject–predicate–object statements. This approach has been championed by

Mike King, founder of iPullRank, as a way to make relationships in your copy painfully obvious to both search engines and LLMs.

Examples of simple semantic triples:

- Beyoncé → is a → singer
- Chocolate → tastes → sweet

In each case, you're explicitly stating "this thing is/does this other thing." When the model turns these sentences into vectors, the relationships are easier to discern and more likely to cluster correctly around the right entities and topics.

Disambiguate Your Brand

Semantic triples are especially powerful for brands with ambiguous names—or brands that live near ambiguous concepts. They help LLMs tell Amazon the company from the Amazon rainforest, Apple the tech brand from apple the fruit, and Ice Cube the artist from the ice cube in your drink.

A triple like

- Amazon provides cloud hosting
- Apple makes smartphones
- Ice Cube is a rapper

leaves very little room for misunderstanding. Each statement pins a specific entity to a specific role or offering, in a way that's hard for a model to reinterpret as "trees," "fruit," or "frozen water."

In the case of Ice Cube, a sentence like "Ice Cube is cold" actually needs more clarification, because it could apply to both the object that chills a drink and the rapper using slang. "Ice Cube is a rapper" or "Ice Cube performs hip-hop" is a much stronger semantic triple because it clearly defines who and what we're talking about.

Define Your Offering

You can use the same technique to define your own offering. These sentences are often almost brazen in their simplicity and confidence—and that's the point. “Amazon provides cloud hosting” is super clear, and that clarity is exactly what helps LLMs understand, store, and retrieve you correctly.

If you have a new product or service, using at least one semantic triple in your launch copy helps confirm to LLMs what it does and why it matters. LLMs recognize patterns, so the more often they see “brand → does → thing,” the easier it is for them to recognize you for what you are.

Specify Your Audience

You can also use semantic triples to connect your product to a specific audience or use case. For example:

- Little League is for kids
- DJs use SoundCloud
- Etsy is for handmade products

Statements like these give models a crisp signal about who something is for and how it's used, which increases the chances your brand shows up in the right AI answers, not just any AI answers.

Try This Prompt

Generate 20 semantic triples for the brand [brand]. Each triple should clearly state what [brand] is, what it does, or who it is for. Focus on simple, confident sentences that remove ambiguity about category, core products, and primary audience.

Build Topic Clusters (And Your Own “Vector Neighborhood”)

Topic clusters are a practical way to reinforce those relationships at the page level. Instead of one catch-all article, you create multiple supporting pages that all point into a single hub focused on a specific entity or theme.

For example, say you want to own AI visibility around “how to make a sandwich,” not just rank for one recipe. A stronger cluster might look like the following.

- Hub page: “How to make a great sandwich: bread, fillings, tools, and techniques.”
- Supporting pages:
 - “Best bread for sandwiches (and when to use each one).”
 - “Vegan fillings for sandwiches that actually taste good.”
 - “Sandwich-making tools every home cook should own.”
 - “25 world-famous sandwiches and what makes them iconic.”

Each supporting page links back to the hub and uses clear semantic triples like “Ciabatta is good for toasted sandwiches” or “Vegans use hummus as a sandwich filling.” Together, that cluster helps LLMs see all of those pages as sitting in the same “sandwich-making” neighborhood in vector space—and just like with a semantic triple, you’re actively pointing the model to where your brand should sit on that map.

Use Named Entities to Lock In Relevance

Vector embeddings also map relationships between specific people or companies and the things around them—products, awards, universities, investors, locations, industries, and more. That means you can use named entities to help LLMs laser-focus on why your business is relevant to a topic, not just that your name appears nearby.

If I write:

“Steve McQueen, famous for his films”

that doesn’t give enough information about who “Steve McQueen” is, or even confirm I’m talking about one specific person. There are at least two well-known Steve McQueens who fit that description, so search engines and LLMs need more detail to work out which one I mean.

If instead I write:

“Steve McQueen, Academy Award winner, film director, and Turner Prize-winning artist”

then I’ve added multiple named entities (Academy Award, Turner Prize, artist, director) that already exist in knowledge graphs and entity databases. Those extra anchors make it much easier for models to lock onto the right Steve McQueen and to understand his relevance to film, art, awards, and culture.

You can apply the same idea to your business. Compare:

- “Acme is famous for its products.”
- “Acme Analytics is a Google Partner based in Toronto, providing marketing attribution software for ecommerce brands with Wix websites.”

The second version layers in concrete named entities and attributes—Google Partner, Toronto, Wix, ecommerce—plus a clear product category (marketing attribution software). That gives LLMs far more to work with when they place you in vector space and decide whether you belong next to “enterprise analytics tools,” “local Toronto agencies,” or “ecommerce solutions for Wix sites.” By deliberately weaving in clear named

entities—your full brand name, product category, key partners or platforms, industry, and audience—you’re effectively telling LLMs: “This is who we are, this is what we do, and this is who we do it for.”

Try This Prompt

Generate 20 named entities related to the brand [brand]. Prioritize affiliations, institutions, notable individuals, and locations that are genuinely connected to [brand] (e.g., investors, partner platforms, headquarters city, key executives, and flagship events).

Bringing It All Together

If ingestion is how AI models find you, then tokenization, token IDs, and vector embeddings are how they remember, file, and retrieve you. Together, they decide whether your brand sits in the right neighborhoods in vector space—and whether you get pulled into the answers that matter.

You can’t control every part of that process, but you’re far from powerless. By writing with tokenization in mind, shaping your content around semantic triples, building smart topic clusters, and deliberately weaving in strong named entities, you’re effectively leaving a trail of neon arrows that say “put us here” in the model’s internal map.

The brands that win in AI search won’t just be the ones with the biggest budgets; they’ll be the ones whose content is easiest for machines to parse, place, and trust.

5

One Strategy Doesn't Fit All— Content Strategies for Every Type of LLM Response

Here's something a lot of marketers get wrong about AI search: they treat every LLM like it works the same way. It doesn't. And if you're applying one blanket strategy across all of them, you're probably leaving visibility on the table—or worse, optimizing for the wrong thing entirely.

Once you understand how each type of LLM generates its responses, the right strategy becomes a lot clearer. In this chapter, we're going to look at the three main types of LLM responses and, more importantly, the kind of content each one rewards.

The next three chapters will then dive into specific tactics for each response type.

How LLMs Generate Responses

Do you answer every question in exactly the same way? Of course not—and neither does AI. Some models generate responses using only information they already know, some reach for fresh information, and some pull in additional “experts” for deeper research and reasoning.

In this chapter we’re talking about response types, not obsessing over individual models or brand names. That distinction matters. Consumer-facing platforms switch between models for almost every query, and most of them are running at least three different models in parallel:

- A flagship, high-performance general model with a large context window and lots of skills, running on the latest tech.
- A “lite” or “flash” model for quick, lightweight answers—less complex, but faster and cheaper.
- A “deep research” or “thinking” model that’s reserved for complex, high-stakes queries where quality matters more than speed.

When a user types a query, the platform decides which model is best for that job. Sometimes that’s automatic; in many tools, the user can also explicitly choose—“fast answer” versus “deep thinking.” If I ask for a quick guacamole recipe, I’ll almost certainly get an answer from the lite model. If I ask for help drafting a dissertation on the history of avocado agriculture, the platform will escalate me to a research-heavy model.

This is why trying to “optimize for GPT-X” or “optimize for Gemini-Y” is a treadmill you’ll never get off. Models update constantly. Names change. Capabilities shift. You end up chasing patch notes instead of building durable visibility. Users don’t show up for a model name; they show up for answers and solutions.

So, instead of optimizing for specific models, we want to optimize for the types of responses your audience is most likely to receive. Across platforms, those fall into three broad categories.

■ **Training-data responses**

The model answers using a fixed training set and does not run an additional search outside its existing corpus of knowledge.

■ **RAG responses**

The model supplements its training data by retrieving information from an additional data source that may be updated in real time—for web-search LLMs, that usually means live search results; inside a company, it might be a product inventory, CRM, or internal knowledge base.

■ **Reasoning responses (RAG + mixture of experts)**

The model taps into a mixture of specialist “expert” networks—organized by topic or by task—and uses that alongside retrieved data to answer complex questions and go deep on research.

Each response type runs on a different data configuration and handles citations differently. That has very real consequences for when, where, and how your brand can appear. The same platform might quietly route each of these prompts to a different underlying model—but from a marketing perspective, the response type is what you can actually optimize for.

Three response types at a glance

Response type	Training data	RAG	RAG + MoE reasoning
Data sources	Fixed training set	Fixed training set augmented by live search or local data	Mixture of experts (MoE)—niche topic or task networks + RAG
Links and crawling	Citation links may not be included	Includes citations and updates based on web crawling	Often includes many citations and frequent updates based on web crawling
Last data update	A few months to over a year behind real time	Typically close to real time (depending on the search index)	Typically real time, with added reasoning steps
Example query	“Generate a simple guacamole recipe using common ingredients.”	“Has the average retail price of avocados increased in the last 12 months?”	“Design a TikTok campaign that uses the cultural meaning of avocado toast for millennial professionals.”

Let’s look at the kinds of queries that tend to trigger each response type—and what that means for your brand’s content strategy.

Think Evergreen Information for Training-Data Responses

Training-data responses are often the ones where the LLM sounds most confident. It's drawing on the large corpus of information it ingested, organized, and was trained on—no extra search required.

In the example above, the query “Generate a simple guacamole recipe using common ingredients” is unlikely to trigger a web search, because ChatGPT-style models have effectively read thousands of guacamole recipes. A basic guacamole recipe has become public information—the model is recombining knowledge it already has, not hunting for a specific blog post.

Information is the key word here. Training-data responses are most likely to surface core facts about your brand, not your latest campaign. Ask an LLM, “Why is Dolly Parton considered to be a great musician?” and it won't need to search the web. It can explain that she's been at the top of her field for decades, has won major awards, sold millions of albums, and written some of music's most iconic songs—facts supported by her own site, biographies, social channels, and organizations like the Grammys.

You don't have to be a music icon to see the pattern. For your brand, the same kind of evergreen facts come into play:

- Brand values and mission
- Flagship products and services
- Core audience and positioning
- Main USPs and top competitors
- Origin story, founder story, and where/how you operate

LLMs have to get this information from somewhere. The question is: Is it accurate, and is it available on the open web in places the model trusts? For SEOs, this is the kind of material that

usually lives in evergreen content. For PR folks, it's the content that would sit in that corporate brochure you update once a year.

Like Dolly Parton's accolades, your brand story shouldn't live only on your own domain. It should also appear on well-established knowledge hubs and organizations that act as pillars in your industry—whether that's awards bodies, trade associations, review platforms, or major directories. The more robust and consistent this public information is, the more confident LLMs will be in mentioning your brand in relevant training-data responses.

Stay Up to Date for RAG Responses

RAG responses kick in when a search is needed—when the model decides that its training data alone isn't enough. RAG simply means the model combines what it already knows with information it fetches from an external source, such as the web or a private database.

Perplexity's documentation explains that its search classifier initiates retrieval “for questions requiring current information, facts, or research.” In our avocado example—“Has the average retail price of avocados increased in the last 12 months?”—there's a clear time constraint. The model can't rely on a training snapshot that might be months out of date; it has to augment that snapshot with fresher data.

I've seen this play out in my own work. When my team and I published content around AI visibility, `LLMs.txt`, and the agentic web, we saw clear spikes in traffic from AI search tools. Those pieces weren't just timely—they were fact-rich, cited original research, and spoke directly to emerging questions users were asking. That combination made them attractive targets for RAG-style citations.

For your brand, this means time really matters. You might not be in a fast-moving industry, but keeping a finger on the pulse can help you edge ahead of competitors in RAG-style visibility. Timely, fact-rich content across your site, documentation, and key off-site channels is what earns you those citations when the model goes looking for up-to-date answers.

For RAG-driven responses, the content that pulls its weight tends to be:

- Recent product, feature, or pricing changes
- News, releases, and campaign launches
- Fresh performance data, case studies, and benchmarks
- Policy, eligibility, or availability updates
- Time-stamped explainers on “what’s changed” and “what it means”

At a strategic level, if training-data responses are about who you *are*, RAG responses are about what’s happening *right now*—and whether you’ve published anything worth citing when the model goes out to look.

Go Deep for Reasoning Models

Reasoning models build on everything that came before them. If accessing training data is like reading the encyclopedia, and RAG responses are like asking someone who’s read the encyclopedia and also skimmed today’s newspaper, then reasoning models are like talking to someone who has done both—and also has a trusted panel of experts on speed-dial.

In my keynotes, I often use Beyoncé to explain this. She’s a skilled musician who can write her own songs and direct plenty of her shows. But for her country album, she leaned on Willie Nelson and Dolly Parton. For *Renaissance*, she brought in Grace Jones. For Destiny’s Child, she collaborates with Kelly and Michelle, and Blue Ivy joins as a dancer on tour. She treats each of these people as specialists who bring something specific—but she doesn’t need everyone for everything. She picks the right combination of experts for the job at hand.

Reasoning models work similarly. When a user makes a complex query, the model passes it through a “gating” layer that decides which experts to consult. An “expert” might be a specialized skill (math, coding, translation, legal reasoning) or a topic-focused

corpus (medicine, sports, film, finance). For harder questions, the model uses a mixture of experts, then applies its own reasoning to synthesize what they return before generating the final answer.

You can see this clearly in tools like Gemini's Deep Research or Perplexity's advanced modes, where the model breaks your question into sub-tasks, researches each one, and then stitches the findings together. For the avocado TikTok brief—"Design a TikTok campaign that uses the cultural meaning of avocado toast for millennial professionals"—Perplexity split the work into tracks around marketing strategy and cultural symbolism, then recombined the results into a coherent campaign outline.

The screenshot displays the Perplexity AI interface in its 'Deep Research' mode. At the top, there are tabs for 'Answer', 'Links', and 'Images', along with a 'Share' button. The main query is: "Design a TikTok campaign that uses the cultural meaning of avocado toast for millennial professionals". Below the query, it indicates "2 steps completed".

- Step 1:** "Designing a TikTok campaign leveraging avocado toast's cultural symbolism for millennial professionals." This step includes a "Searching" phase with three sub-queries: "user marketing expertise", "user interest social media campaigns", and "user tone preference concise strategic".
- Step 2:** "Researching avocado toast's cultural origins, TikTok memes, and millennial content trends for the campaign." This step also includes a "Searching" phase with three sub-queries: "avocado toast millennials housing meme origin", "tiktok avocado toast meme millennials work burnout", and "millennial professionals tiktok content preferences 2025".

Below the search steps, there is a "Reviewing sources" section showing a list of relevant links and their sources:

- What's the deal with avocado toast?: r/OutOfTheLoop (reddit)
- Avocado Toast (knowyourmeme)
- Recipe of the Month: Avocado Toast (1967) - Blogs (blogs.lib.umich)
- A Grand Unified Theory of Avocado Toast (news.ycombinator)
- Where Did the Avocado Toast Millennial Stereotype Come ... (jeffjbutler)
- actually, i am very stressed about my avocado toast - TikTok (tiktok)
- Millennials favour connection over scrolling on Social Media in 2025 (belongcreative.com)

At the bottom left, the status "Finished" is shown.

For brands, reasoning-driven responses tend to pull from content that showcases depth and expertise, such as:

- In-depth guides and detailed explainers
- Reports, white papers, and research studies
- Thought-leadership pieces that unpack your point of view
- Technical and support documentation that shows how things really work
- Specialist microsites, subdomains, or long-form video for niche topics

These queries are especially important in categories with long consideration cycles, where buyers ask multi-step, nuanced questions before they convert. In those situations, models will lean on whoever looks most like the genuine expert in the room.

We'll come back to exactly how to build that depth on MoE in Chapter 8. For now, think of the three response types like this:

- Training-data responses tap into *what defines your brand*.
- RAG responses tap into *why you're relevant right now*.
- Reasoning responses tap into *what makes you expert and distinctive*.

In the next three chapters, we'll turn each of these into a practical playbook—starting with training-data responses and how to shape what the models “remember” about your brand.

6

Optimizations for Training-Data Responses

Many of the first generation of consumer AI chatbots, like ChatGPT-3, generated responses exclusively from training data. As the industry has matured, very few consumer tools still work that way—but every model still has core training data at the heart of its foundation. That means some responses, especially those that don't show links or citations, are still running in this training-data-only mode.

In these moments, the model answers without surfacing any new external information. It draws on a fixed training set, and visibility shifts as that data is updated. You might see links. You might not.

When optimizing for brand visibility in training-data responses, four rules matter most:

- Mentions matter more than links
- Visibility shifts as training data updates

- Training sources shape what the model knows about you
- Feedback influences future responses

Mentions Matter More than Links

When you're thinking about visibility within training-data responses, the goal is to be mentioned frequently, accurately, and positively. In this mode, the model is recalling what it already “knows” about you, not deciding which link deserves the click.

In training-data responses, links are shown very infrequently—if at all. In many cases, these systems are trained on so many sources that it becomes very difficult to attribute a single piece of content to a single source. Add to that the general messiness of the web, where content is syndicated across multiple sites, duplicated, and repurposed across YouTube, social, and other destinations, and attribution becomes even murkier.

For folks who come from an SEO background, focusing on mentions feels counterintuitive, because we've always worked for a click. But users have relationships with AI tools in a way they've never had with search engines. At this stage, a mention is almost like a brand recommendation from a knowledgeable acquaintance—you might not take it as gospel, but it plants a flag.

When links do appear, they typically point to high-level or evergreen content—often your homepage or major reference pages—rather than the most recent blog post or campaign landing page.

To measure progress towards more brand mentions:

- Run regular queries for relevant entities and brand terms
- Use bulk querying (spreadsheet integrations or purpose-built tools) to track mentions over time
- Explore services that report on how a brand appears in large samples of general AI queries

For SEO-invested teams used to traffic as the main key performance indicator, this can feel wooly. But in an LLM context, mentions act as a form of brand equity and can influence conversions and drive direct traffic.

How to drive mentions

- Align PR with how you want to be described. Seed consistent language about your category, audience, and unique selling point into press releases, interviews, and conference bios so that third-party descriptions reinforce your core positioning.
- Collaborate with credible partners. Guest posts, joint webinars, research reports, and co-branded campaigns with recognized experts or brands give models extra evidence that you belong in that conversation.
- Show up on relevant authority hubs. Prioritize mentions and profiles on awards sites, trade bodies, review platforms, and media that already rank or get cited a lot in your space.

Think of it this way: every accurate, on-topic mention is another brushstroke in the portrait models are painting of your brand. Your job is to make sure they're not painting in the dark.

Visibility Shifts as Training Data Updates

When tools are not web-search-enabled—or choose not to search the web—they generate answers from the pre-trained datasets inside their core models. One way to visualize this is to think about asking a question to someone whose knowledge is limited to the latest edition of an encyclopedia: they'll know what happened up to the date it was published, but they have to wait for the next edition for newly discovered facts. In LLM terms, the time between today and that last update is the knowledge cutoff.

Each tool manages its information and datasets differently—and that directly impacts how visible a brand is on that platform. Knowledge cutoff dates can vary by model; for example, there follows a list of models which are available in March 2026.

LLM platform	Knowledge cutoff
GPT-5	September 2024
Gemini 3 (consumer models)	January 2025
Claude Sonnet 4.5	July 2025

To find the current knowledge cutoff for any model your team is using, look for its model card or documentation page—providers typically publish the cutoff date, capabilities, and main training sources there. And for teams using third-party visibility tools on top of ChatGPT-style APIs, it’s also worth confirming which model version and knowledge cutoff those tools rely on before making major strategic changes based on their reports.

LLM providers now announce training-set updates much like product releases. So, if brand visibility inside these tools suddenly shifts, the first step is to check whether there has been a training-data update.

A Personal Example: “I Don’t Know Her”

When ChatGPT-3 first rolled out, I did what lots of people do and asked about myself. I typed, “Who is Crystal Carter?” and the model replied like that Mariah Carey meme—essentially saying, “I don’t know her.” Ouch. At that point I already had some professional acclaim and was reasonably well known within my industry, but much of that visibility had developed after I joined the team at Wix in 2022. ChatGPT-3’s training data only ran up to around 2022, so it simply didn’t have much information about me to work with.

Fast-forward about eight months and I asked the same question again. This time, ChatGPT knew exactly who I was. I hadn't suddenly done a huge visibility push just for the model; they had updated their training data to be closer to real time, and I saw a jump in my visibility as a result. This kind of delayed recognition is normal for both established and newer businesses.

This isn't something to panic about, but it is something to be mindful of. If a model is using older data and isn't tapped into the web, it's more likely to confidently hallucinate based on the information it already has—and it won't know what it doesn't know. Your job is to manage internal expectations, invest in the right foundations, and be patient while those updates catch up.

Waiting for the Next “Phone Book”

For new businesses or products, waiting for training-data updates can feel a bit like going back to the days of the phone book. If the phone book came out in January and you opened your business in February, you wouldn't appear until the next edition dropped. Pre-trained LLMs work in a similar way: until your brand makes it into the next training cycle, you're effectively invisible in those static responses.

That doesn't mean you sit still. It means you build the strongest possible foundations for training-data ingestion while you wait for the next update.

Site of Sites: Building Foundations Before the Model Catches Up

As another example, I was working on a project called Site of Sites. My team at Wix built it as a showcase for web design projects and different things you can do with our website builder, Wix Studio. The site launched in the spring, and in the summer I asked lots of different LLMs, “What is Site of Sites?” I got all kinds of responses: one said it was an architecture term, others made educated guesses, but none of them recognized the project.

So, we kept building. We spun out the site with additional pages, worked on our indexing, and built backlinks on relevant websites—bread-and-butter SEO activity. When I checked a few months later, some of the older pre-trained models were still unaware of the project, but we had made it into Gemini, which meant that our efforts were taking hold in at least one major model.

New Businesses: What to Do While You Wait to Be Recognized

If I had to give a quick checklist for new businesses waiting to be recognized and included in training data, I'd start with these rules:

- **Don't skip classic SEO foundations.** Many of the tactics that make your website machine-readable for search-engine crawlers apply to LLMs as well, and most on-page content guidelines still hold true.
- **Explicitly define your brand in your copy.** Use semantic triples (e.g., [Brand] is a [category] that serves [audience]) to illustrate the core of your brand and proliferate that messaging across your website, PR, YouTube content, and social media profiles. This gives LLMs context about you and helps them understand which queries you should surface for.
- **Map your key entities and their relationships.** Identify the named entities connected to your brand, then illustrate the relationships between them with (ideally mutual) links and content. For example, if you open a new restaurant in Los Angeles, your key entities might include the city, the neighborhood, the restaurant's name, the owner, the culinary school where the chef trained, and notable restaurants where the chef has worked.

We'll cover these tactics in more depth later, but if you follow them consistently, you'll be in a much stronger position when the next training-data update rolls around.

Training Sources Shape What the Model Knows About You

Most foundation models were originally trained on large open-web datasets such as Common Crawl, plus high-authority sources like Wikipedia and other long-standing domains. For established brands, that means training-data responses may reflect older versions of messaging, positioning, or product lines—because that is what the open web looked like when the model was trained.

If a brand has gone through a pivot, rebrand, or major strategy shift in the last few years, the model may still be pulling from legacy information. Understanding the brand's own history helps interpret what these responses surface and identify where statements no longer match reality.

Take a Moment to See What AI Tools Say About Your Brand

A practical exercise is to log into a basic, no-web-search assistant and ask foundational questions such as:

- What does [brand name] do?
- How would you describe [brand name]?
- What is [brand name] known for?
- What products, services, or experiences does [brand name] offer?
- Who is the primary audience for [brand name]?
- How do consumers typically perceive [brand name] (e.g., premium, affordable, innovative, sustainable)?
- What sets [brand name] apart from competitors in its industry?
- Who are the main competitors of [brand name]?
- Why might a customer choose [brand name] over a competitor—and vice versa?
- Which websites or publications are most likely to discuss [brand name]?

Whatever comes back is being pulled from somewhere in the training data. If the answers feel off-brand, outdated, or biased towards legacy narratives, that's a signal to modernize what exists on the open web and to use feedback and fresh content to push models towards the current reality.

When the Training Data Gets It Wrong

One small business I worked with sold custom handmade goods. They had a good website, strong brand presence, and a solid digital footprint, but when you asked an LLM what that brand did, the model became incredibly confused and started mentioning services that had nothing to do with them. The recommendations seemed completely out of left field, so I had to investigate.

Turns out, at one point the brand had spearheaded a campaign for other local creators and even put a creator showcase on their website. The campaign was really successful and got them coverage on TV news and support in their community. But *now*, LLMs were reporting that the services they showcased were the brand's own. To fix this conflicting narrative, they needed to add clarity and distinction everywhere the confusion might originate.

So, what do you do if the training data is mentioning incorrect information?

- Try asking the LLM for its source. This may or may not be maintained in their training process, but if you're there anyway, it can't hurt. Follow the trail to see if you can verify the source.
- Trigger the conflicting response across multiple chatbots to identify the common phrases that are being surfaced, then carry out a verbatim search on Google with the term in quotation marks. If you're seeing the same phrasing repeated, it's likely been cited a few times and will have an origin.

- Clean house. Update or delete out-of-date content on your websites and social. LLMs will train on the data from your website and will consider that to be a highly authoritative source of facts about your brand. If you can't keep your facts straight, they can be forgiven for being confused.
- Create new content that specifies the change—why it happened, what it means, where you are now.
- Create a unique domain for content that conflicts with your brand story. If old campaign microsites or specialty projects are creating noise, consider moving them off your primary domain.
- Get third-party content updated. This may seem daunting, but it can be worth it. Third-party content creators who have recently and inaccurately mentioned your brand have probably done so in good faith. Get in touch, give them resources to make the update really simple.

This is a challenge that many established brands face when high-profile but out-of-date information lingers on the web. It can happen when you've had a big campaign in the past or a service that you offered extensively and then discontinued, merged, or renamed. LLMs may look at old press releases, blogs, and YouTube videos, and surface outdated statements.

The correction process is not immediate—it's more like changing your brand name than updating a meta description. It takes time for the whole ecosystem to get the memo. (If I'm completely honest, I still call it Twitter most of the time.) So, aim for steady, compounding improvement rather than expecting LLM responses to flip overnight like a light switch.

One tactic that can accelerate this slow-but-steady process, though, is giving direct feedback to the models themselves.

Feedback Influences Future Responses

Even when an LLM isn't pulling fresh data from the web, it still learns from its users—which makes feedback on responses a meaningful optimization lever. From ChatGPT to Grok, Gemini, Claude, and everything in between, you'll usually see thumbs-up and thumbs-down buttons underneath each response. These are not for show; they're there so you can give feedback that helps train the model.

In its documentation, Google explains that feedback can “help make Gemini better” because “one important part of developing responsibly is expanding participation. . . You can rate responses as good or bad and send feedback each time Gemini responds.” This sits high in the hierarchy of how LLMs learn, alongside other forms of reinforcement from human feedback.

Whenever a response about a brand is incorrect or out of date, it's worth:

- Using the built-in feedback tools (thumbs up/down, rating dialogs)
- Supplying the correct information in clear language
- Pointing, where possible, to a trustworthy source such as the official site or a recognized publication

This isn't for subjective disagreements—if Claude thinks Beyoncé and I don't look like twins, it's difficult to argue—but if Claude said my business HQ was in New York when it's actually in New Mexico, I can provide evidence to support the correction and submit that through the feedback tools.

Site of Sites Strikes Again

Remember Site of Sites? When we finally got some visibility on Gemini, the first thing I did was ask it for a link.

Me:

“What is the URL for Site of Sites?”

Gemini:

“Site of Sites is actually a section within the Behance platform. It is not a standalone website. You can access it through the Behance search function.”

I was floored. Behance! Really? Now I could have just sat with that feeling, but instead I took action. I clicked the thumbs-down button and wrote:

“<https://www.siteofsites.co/> is the URL for the website Site of Sites. It is not part of Behance.”

When I checked a few weeks later and asked the same question, Gemini replied with the correct URL. You might reasonably say, “Well, lots of things could have happened in those weeks,” and that’s fair—but I’ve followed the same process for other inaccuracies and seen updates appear within minutes.

Of course, this is tricky to scale across every possible query, but for your core brand questions—where you see clear, factual inaccuracies—don’t let them slide. Let the model know. The feedback loop is real. Like any machine-learning system, LLMs depend on good signals to improve, so updating them with accurate information about your brand is good for both your visibility and the tools themselves.

Bringing Training-Data Optimization Together

Training-data responses are the part of AI search you can’t see updating day-to-day, but they quietly set the baseline story

models tell about your brand. Your job is to make that story as accurate, consistent, and flattering as possible.

If you focus on three things—being mentioned clearly and often in the right places, keeping your evergreen signals aligned with your current reality, and feeding back when the model gets it wrong—you’ll already be ahead of most brands in your category.

In the next chapter, we’ll shift from “what the model remembers” to “what it can find right now,” and look at how to optimize for RAG responses that pull in fresh web data, citations, and live context around your brand.

Training-Data Quick Checklist

- **Know the cutoff**

Confirm which models you care about and note their knowledge cutoff and training sources from the model docs.

- **Track brand mentions**

Run regular queries for your brand and entities in no-web modes, and log how often you’re mentioned and how you’re described.

- **Build strong foundations between data updates**

Keep SEO basics tight, define “[Brand] is a [category] for [audience]” across web, PR, YouTube, and social, and link key entities to each other.

- **Clean up bad or old info**

Find the sources of wrong or legacy claims, update or remove them on your own channels, and request fixes from third-party sites where needed.

- **Use feedback loops**

When LLMs get facts wrong, use thumbs-down tools with a short correction plus a trusted source, focusing on your core brand queries.

7

Optimizations for RAG Responses

RAG responses kick in when the model decides it can't rely on training data alone and needs a live fact-check from the web. In this mode, your visibility depends not just on what the model *remembers* about you, but on what it can *find* and *trust* in search results right now.

Think of RAG as an AI answer that has phoned a friend—the search engine—before replying. Your job is to make sure that when it makes that call, your content is the one it wants to quote.

Keep Content Fresh Enough to Deserve a Citation

RAG systems are built to go and get the latest information, not just whatever was true at training time. That means your “ever-green” pages need more than a nice intro—they need a pulse.

You've seen this first-hand with your Google-versus-ChatGPT piece. You refreshed the article, asked ChatGPT about the data shortly afterwards, and saw it cite the updated stats within about 30 minutes. In other words: search picked up the

update, the retriever saw it, and your new version became the default reference.

Prioritize regular, meaningful updates on:

- Time-bound explainers (“What changed in [X] this year?”)
- Market trend and performance posts
- Comparison content (you versus competitors, platform A versus platform B)
- Policy, eligibility, and pricing pages

When you make a significant change, spot-check the same query in ChatGPT, Gemini, Perplexity, and Copilot. If they don’t pick you up within a reasonable window, that’s a signal to look at crawling, indexing, or how clearly you’ve surfaced the update.

Make Discovery Faster with IndexNow and Solid Crawling

RAG can’t cite what search hasn’t found yet. One of the lowest-lift ways to speed up that discovery loop is IndexNow.

IndexNow is an open protocol that lets your site tell participating search engines—originally led by Bing—“this URL has changed” the moment you add, update, or delete a page. You host a small key file on your domain, then send a simple ping listing the URLs that changed. Those notifications are shared across all search engines that support IndexNow.

Why this matters for AI search:

- Bing and ChatGPT have a long-standing partnership, and Bing is a major grounding source for Copilot and some ChatGPT modes.
- Anecdotally, sites configured with IndexNow see new or updated content appear in Bing and ChatGPT-style citations very quickly—sometimes within minutes.

- Configuring IndexNow will not hurt you in the short run, and it can only help shorten the “we updated this” → “the model is quoting this” lag.

In practical terms:

- Use your CMS, CDN, or a plugin to send IndexNow pings automatically for key changes.
- Prioritize pings for pages that answer newsy, data-driven, or comparison queries that are likely to trigger RAG.
- Keep your basic crawl hygiene in order (sitemaps, internal links, no accidental noindex), so those pings land on pages that are easy for bots to parse.

Optimize for the Search Engine Behind the LLM

Every RAG-powered LLM is essentially standing on someone else’s index. Your strategy needs to account for both the LLM *and* the search engine it leans on.

At a high level:

LLM/model	Primary search/index	What that means for you
Copilot (incl. early “Bing Chat”)	Bing	Prometheus framework grounds GPT answers in Bing rankings.
Gemini	Google	Strong coupling to Google Search and its signals.
ChatGPT (search modes)	Mixed Bing/Google/own index	Evidence of both Bing and scraped Google SERPs (search engine results pages) feeding citations.
Perplexity	Own crawlers + partner indexes	Heavy emphasis on its own crawl plus RAG stack.

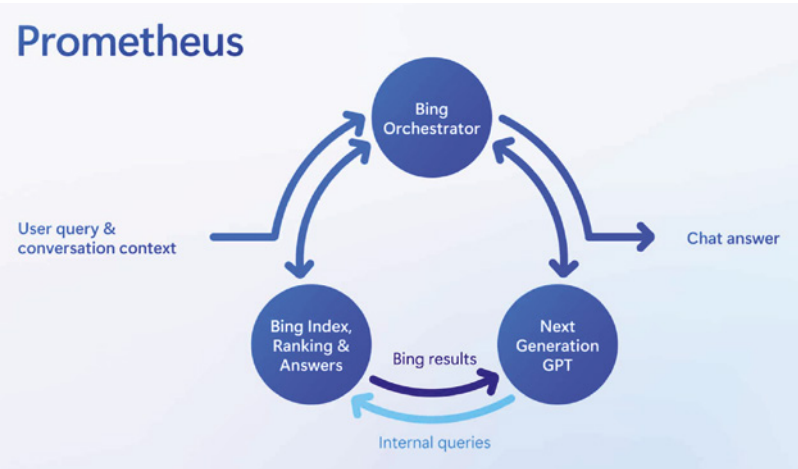
(continued)

(continued)

LLM/model	Primary search/index	What that means for you
Grok (on X)	X posts + web + Grokipedia	Deep integration with X’s live post stream and Grokipedia’s AI-generated encyclopedia; RAG answers often lean on X content and Grokipedia entries.

When Microsoft launched Bing Chat, they described the Prometheus model as a way to “ground” Copilot with Bing’s search results and reduce hallucinations.

It was groundbreaking and it meant that, if you want to rank in Copilot, it’s your Bing performance that carries the most weight—not just your Google rankings.



A lot of marketers still assume that ranking on Google is enough to get cited in ChatGPT. Research doesn’t back that up.

An Ahrefs study found that only around 10–12% of URLs cited by ChatGPT overlap with Google’s top-10 results for the same queries, and broader analyses across AI tools show similarly low overlap between AI citations and traditional SERP winners.

Practically, that means:

- Track your presence in Bing if Copilot and ChatGPT are important touchpoints.
- Treat Bing and Google as separate—overlapping, but not interchangeable—levers for RAG.
- Don’t rely solely on “we rank #1 in Google” as proof you’ll be the canonical answer in AI search.

For Grok specifically, what happens on X doesn’t stay on X. Public posts can be pulled into Grok’s real-time retrieval layer, and over time those signals feed into Grokipedia—the AI-generated encyclopedia that Grok uses as a persistent knowledge base. In practice, that means consistent, high-quality posting and being cited or discussed by credible accounts on X can increase the odds that your brand shows up both in Grok answers and in Grokipedia articles.

Check Your Core Queries on a Regular Cadence

Just as rankings fluctuate in search, your RAG visibility will move around as indices and models update.

Build a simple RAG audit loop:

- **Pick your core queries.** Start with 10–30 questions where you’d *want* to be cited: category queries, “best tools for . . .,” problem-solution searches, “[brand] versus [competitor],” and key branded questions.
- **Track them systematically.** You can do this with something as simple as a Google Sheet—log the query, the tool,

the date, and which sources are cited—or with dedicated prompt-tracking tools that automate runs and screenshots.

- **Run them through key tools.** Check ChatGPT search, Gemini, Copilot, and Perplexity on a set cadence (monthly or quarterly is enough for most teams).
- **Log what changes.** Who appears, who disappears, which of your own URLs show up, and which third-party sources keep recurring.

Treat prompt tracking like the conversion rate optimization tactics of testing and iterating in response to live data: you should only be tracking queries you're actively trying to optimize. Monitoring random, general-interest prompts will burn time (and credits) without teaching you much. Focus on the questions tied to revenue, reputation, or strategic positioning, then iterate your content and technical work against that short, meaningful list.

Prioritize Pages That Are Already Cited in LLMs

In classic SEO, getting indexed is table stakes; in RAG, being cited is the equivalent. Once a search-augmented LLM has used one of your URLs as a reference, that page is effectively “on file” as part of its working knowledge about that topic.

First, find those pages:

- Check your bot log reports for visits from AI and search-assistant user agents that hit specific URLs.
- Review referral traffic in your web analytics for sources like ChatGPT, Perplexity, Gemini, Copilot, and similar.
- Scan your prompt-tracking sheet or tool for any prompts where the LLM showed your domain in its citations or link cards.

Next, treat the cited URLs you uncover as the key to help surface other content and send stronger relevance signals across your site:

- Strengthen internal links from cited pages to other strategic content (product pages, deep guides, comparisons).
- Keep them freshly updated so they remain attractive candidates for future RAG responses.
- Align anchor text with the queries you care about, so it's obvious what each linked page helps a user (and model) answer.

If you want more of your site pulled into RAG answers, one of the fastest routes is to radiate outwards from the content the model has already decided it trusts.

Publish in Topic Clusters and Waves

RAG systems don't just want *an* answer; they're trying to triangulate on the *best* answer, across multiple angles and sources. Topic clusters help you look like the kind of domain that understands a subject in the round.

On your own site, a topic cluster behaves a bit like a mini knowledge graph: interlinked articles, explainers, FAQs, data pieces, and opinion content that all orbit the same theme. That structure makes it easier for both search engines and LLMs to see you as an expert on that topic.

To make clusters work harder for RAG:

- Cover multiple facets of a topic (definitions, how-tos, comparisons, use cases, objections, data).
- Use consistent language across the cluster so your core terms and entities repeat in natural, readable ways.

- Roll content out in waves—publishing or updating several related pieces in a short window—so you send a strong, time-bound signal that “we’re active and authoritative here *now*,” not just historically.

Done well, a single cluster can help you surface across dozens of adjacent queries in both search and AI responses.

Use Structured Data and Rich-Results-Friendly Markup

Structured data sits at an interesting crossroads for RAG: it doesn’t directly change what non-search-enhanced LLMs remember, but it does influence which pages rank in Google and Bing—where many RAG systems pull their citations from. In most web-derived LLM training datasets, preprocessing pipelines extract the visible text and strip out script tags, including those that contain JSON-LD structured data, so that information rarely makes it into the final training corpus. Which means that for pure training-data responses, your JSON-LD won’t suddenly catapult you into more answers—the model is still learning from the story you tell in the body copy, not from the markup tucked into your HTML.

Where structured data really starts to earn its keep is in search-enhanced LLMs like Gemini, ChatGPT search, Copilot, and Perplexity, which ground their answers in live search results. In these classic searches, structured data helps determine which URLs qualify for rich results, how clearly a page’s purpose is understood, and how prominently it should show up for specific intents. In my own tests, when I asked questions like “How do I make a vegan cake?” most of Perplexity’s sources matched Google’s rich-result recipe cards, and several overlapped with ChatGPT search citations as well—strong anecdotal evidence that rich-results-eligible pages are disproportionately likely to be pulled into AI answers.

Early studies and testing also suggest that structured data can help with Google’s AI Mode and AI Overview-style results, which are closely tied to traditional search signals and rich results. The space is evolving quickly: emerging standards like NLWeb aim to give LLMs and agents a more consistent way to understand structured information on the web, which should make it easier over time for AI systems to consume and act on your JSON-LD.

So, while there’s no magic “add schema, get ChatGPT ranking” switch, structured data still quietly does the work of making your content easier for search engines and AI-enhanced SERPs to parse, classify, and elevate—putting you in the pool of URLs RAG systems are most likely to reach for when they go hunting for evidence.

RAG responses are where classic SEO and AI search shake hands. When the model steps outside its training data and starts phoning the web for backup, the brands it calls first are the ones with fresh, well-structured, easily discoverable content in the right indexes.

RAG Optimization Checklist

Keep content fresh enough to deserve a citation

- Flag time-sensitive pages (trends, comparisons, pricing, policies).
- Set review cadences (monthly, quarterly) and add “last updated” notes.
- After big edits, ask related questions in LLMs and check which version they cite.

Encourage faster discovery with IndexNow and crawling

- Confirm IndexNow support in your CMS and set it up (API key and verification).

- Ensure pings fire on add/update/delete of important URLs.
- Fix crawl blockers in XML sitemaps, internal links, and robots rules.

Optimize for the search engine behind each LLM

- Map which LLMs your audience actually uses.
- Note the underlying index (Google, Bing, own crawlers, Grok's stack, etc.).
- Watch Bing if Copilot/ChatGPT matter; watch Google if Gemini/AI Overviews matter.
- Track AI citations separately—don't assume "Google #1 = cited in AI."

Check your core queries regularly

- Pick 10–30 "core RAG queries" tied to revenue, reputation, or positioning.
- Log query, LLM, date, and cited sources in a sheet or prompt-tracker.
- Test across ChatGPT search, Gemini, Copilot, Perplexity (and any priority tools).
- Repeat on a schedule and track shifts in who gets cited.
- Only track queries you plan to influence—treat this like CRO, not vanity.

Prioritize pages that are already cited

- Collect URLs already cited by AI tools (via bot logs, referrals, prompt trackers).
- Refresh those pages first for accuracy, depth, and clarity.
- Strengthen internal links from these URLs to other strategic content.
- Tune anchor text to mirror the questions and topics you want to own.

Publish in topic clusters and waves

- Choose three to five themes where you want to look definitively expert.
- For each, build a cluster: definitions, how-tos, comparisons, FAQs, data, opinion.
- Link aggressively within each cluster.
- Publish in waves (multiple related pieces in a tight time window).
- Update older pieces when something meaningful shifts in that topic.

Use structured data and rich-results-friendly markup

- Add rich-results-eligible schema markup (Article, Product, Recipe, etc.) to relevant pages.
- Validate and fix schema issues on existing pages, then target rich results where relevant.
- Monitor which pages earn rich results and whether those URLs show up in AI citations.
- Keep an eye on emerging standards (e.g., NLWeb) and platform updates that make structured data more valuable to LLMs.

8

New Brains, Same Internet: Reasoning Models for Marketers

Reasoning-model responses are where AI stops behaving like autocomplete and starts acting like a strategist. Instead of just recalling what it already knows (training data) or phoning the web for a quick fact-check (RAG), the model breaks your question into parts, calls on different experts, and then explains how it reached its answer.

Under the hood, many of these answers run on an MoE setup: a routing layer decides which specialist networks and external sources to consult for this specific job. That's true whether you're in DeepSeek R1, Gemini's "deep" models, or ChatGPT's reasoning models. Since 2025, reasoning-style

models have been available from every major AI platform, turning “deep thinking” from a party trick into a standard feature of AI search.

For marketers, these responses show up on the highest-value queries—the ones where users are making decisions, weighing trade-offs, or asking “why” and “how,” not just “what.” Your goal is to become one of the experts these systems lean on when they have to think out loud.

How Reasoning Models Use Your Content

If training-data responses are about who you are, and RAG is about what’s happening right now, reasoning responses are about what makes you expert and distinctive.

When a user makes a query to a reasoning model, it goes through a gating network. Here, the model chooses from a selection of “experts” to answer the question in the most efficient way.

What’s an expert? In LLM terms, an expert is a niche dataset, tool, or skill. These can be arranged around *tasks* (like translation, coding, math, or writing), or they can be arranged around *topics* (like medicine, sports, law, film, or music). When a model answers a question using a combination of these specialists, this is MoE.

1. They unpack your question

The model starts by restating or clarifying what you asked: “Okay, so the user is asking why the Renaissance film hasn’t been released for streaming yet.” This is the “let me break this down” moment.

2. They fan your query out into sub-questions

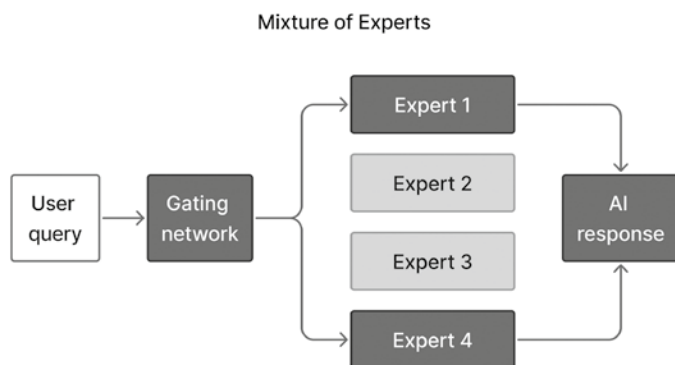
One question becomes several: Which Renaissance film? What do I know about movie release windows? What has this artist done with streaming in the past? What’s been reported in the press?

3. They activate experts

Inside the model, a routing or “gating” network decides which internal experts to call and, in MoE setups, which specialized subnetworks to light up. In search-enabled models, these experts also include external websites and products.

4. They reason, then respond

The model does a chain-of-thought pass, weighing different factors: “it’s possible. . . another factor to consider. . . based on industry norms. . .”, then produces a narrative answer with a justification baked in. In models like Gemini Deep Research or Perplexity’s “copilot”-style experiences, you often see citations for each major point.



Though an MoE model doesn’t necessarily *need* to trigger a RAG (web search) response, it often does. When this happens, the LLMs essentially treat websites as experts. And with good reason. If you think about it, what is a more expert publication on film than IMDb? Or what has a more nuanced discussion about coding than Stack Overflow?

For brands, you should be the definitive expert on everything about you, and ideally a top expert on the solutions that you offer. The strategic question is simple: When the model goes

looking for someone to “phone a friend” on your niche, does it call you or someone else?

- If the question is about your brand and how you work, does the model reach your implementation guide, or does it end up in a subreddit?
- If the question is about a core service that you provide, does it find thought leadership, guides, and research from you, or from another website?

This may seem daunting, but creating niche topic forums, specialist blogs, and dedicated web properties can go a long way to shaping your expert profile. Your job is to become the most convincing, well-structured expert source in the room, so when the model assembles its expert panel, your site is the obvious one to invite.

What Reasoning Models Reward

Think about where you can add relevance to your brand with timely, fact-rich information that warrants a citation. Essentially, give them a reason to link to you.

I ran research on over 1,000 MoE queries, and Levi’s emerged as a brand where LLMs were surfacing content from their whole digital profile—investor documentation, sustainability statements, targeted campaign microsites, blogs, product pages, and international websites—in order to answer deep queries.

For the other brands that I researched, reasoning queries relied heavily on third-party sources for answers. What set Levi’s apart was the depth of expertise. They didn’t leave a vacuum for others to fill; they were the definitive expert on themselves.

The good news is that to achieve this, you’re not throwing out the work you did for training data and RAG. You’re building

on it. The difference is that now you're optimizing for multi-step thinking, not just recall or freshness.

When we analyzed citations across reasoning-heavy models like DeepSeek, ChatGPT, and Perplexity for non-brand questions, three patterns showed up again and again.

1. Current events and live topics

News and trending-topic content is much more likely to be cited in reasoning models than in static training-data answers. DeepSeek, for example, pulled 40–50 citations for a single current-events-style query, drawing from news sites, streaming platforms, Reddit, and fan communities. These models need fresh sources to ground time-sensitive claims, so they reward brands who talk about what's happening now.

2. Opinionated, multi-perspective content

When the question is squishy—"Which tour was better, Beyoncé's Renaissance or Taylor Swift's Eras?", "Are this brand's sustainability claims credible?"—reasoning models lean hard on commentary and analysis. They want multiple perspectives, including brand sources, critics, explainers, and user-generated content (UGC). If your content never takes a stance, you're less likely to make it into those "why" and "how come" answers.

3. Rich, structured, machine-readable pages

Across reasoning tools, content that uses headings, lists, tables, and clear semantic structure is easier to quote and cite. Think of it as pre-chunked for the model: each subsection can become its own expert snippet. In Google's Deep Research, you literally see how the system turns subheadings into separate research tasks, then stitches them back into a single report.

Put simply: reasoning models love content that is fresh, opinionated, and easy to slice into evidence.

Optimizing for Reasoning Models and MoE

So, how do you actually optimize for this new class of models, without spinning up a whole new team?

Here are the highest-impact moves, building on what you already do for training data and RAG.

Embrace Current Events in Your Lane

Reasoning models lean heavily on timely sources when they're asked "why now?" questions. A static training snapshot can explain what a brand does in general; it can't justify why a film isn't streaming yet, why a platform changed its policy, or how a new regulation will affect a specific business this quarter.

In practice, I've seen reasoning-style answers pull heavily from recent news coverage, long-form analysis, and community discussions with strong, repeated themes.

To optimize:

- Publish robust explainers and recaps tied to live news in your vertical: regulation changes, platform updates, product recalls, new standards.
- Carry out "state of the market" research and publish to a research hub.
- For big launches, ship *analysis*, not just announcements—why this matters, what changed, how to adapt.

In our research, newsy and timely content drove significantly higher citation counts in DeepSeek and Gemini than static evergreen pages, especially on entertainment and culture queries.

Show Your Working on Complex Questions

When the question is simple (“What is this?”, “Who owns that?”), models don’t need much reasoning. When the question is messy (“Is this sustainability claim credible?”, “Which solution is better for this trade-off?”), they go hunting for sources that actually show their working.

Reasoning models are drawn to in-depth guides, technical documentation, methodology pages, and thought-leadership pieces that walk through a point of view rather than just announcing it.

To optimize:

- Identify the “hard questions” in your category—the ones your sales team hears on long calls, or that keep popping up in forums and reviews.
- Build content that answers those questions with evidence and original data: comparative tables, diagrams, metrics, quotes from credible third parties, and frank discussion of trade-offs.
- Use headings and lists to break that content into modular chunks; in both RAG and reasoning models, those chunks are more likely to be quoted than a wall of text.

When we looked at Levi’s sustainability-related queries for denim, the reasoning models didn’t just grab homepages—they pulled deep cuts from brand labs, long-form sustainability reports, and detailed FAQs, then stitched them into a narrative. That’s the bar.

Design for “Query Fan-Out”

Reasoning models rarely treat your question as a single task. They turn it into a research plan. In Gemini’s Deep Research, you can see this explicitly: the query is split into sub-tasks (global

context, user behavior, narrative hooks, recommendations), each with its own mini-investigation. In other reasoning tools, the same pattern plays out behind the scenes.

That means your content strategy needs to be fan-out friendly:

- Treat each likely sub-question as its own target: “who is. . .”, “how does. . . work?”, “what are the risks?”, “how does X compare to Y?”, “what changed recently?”
- Create content clusters that cover those angles around your core topics, then link them tightly together so they look like a coherent body of work.
- Write headings that mirror the sub-questions models are likely to generate, so each section can slot neatly into their chain of thought.

When the model needs a deep dive on your niche, your goal is to look like the natural place to send that sub-question: clearly scoped, well-structured, and strongly associated with the entities and topics in play.

Optimize for Citations, Not Just Rankings

In classic SEO, rankings come first and citations are a side effect. In reasoning models, citations are the main event. These answers often include a long list of sources, and only a fraction of them belong to brands. In my analysis of 1,000 branded queries, brands typically captured only the 20% minority of citations in a reasoning-heavy answer.

To improve your odds:

- **Make key facts quotable.** Put short, clear statements near the top of the page, and back them up with deeper detail below.

- **Use entity-rich language.** Consistently use exact brand names, product names, locations, and categories so the model can easily associate snippets with you.
- **Keep credentials prominent.** Models lean on clearly authoritative sources when the stakes are high, so make sure your author and brand credentials are front and center.
- **Script your stance on video and social.** Assume transcripts and captions are part of your evidence base. If you want reasoning models to explain your product or stance a particular way, say it clearly and consistently in your scripts.
- **Have an opinion.** Where there are controversial topics or ongoing debates, AI has to turn to web content to find an answer—it can't sort out opinions on its own. If you've staked your claim on whether pineapple belongs on pizza, you're far more likely to be cited in a debate than if you just wrote a generic article about pizza.

Be Deliberate with Social and UGC

Reasoning models don't stop at official sources. For opinion-heavy, experience-driven, or culture-driven questions, they pull in posts, comments, and reviews to capture sentiment and lived experience. That means your brand's social presence and UGC are not just "nice to have"—they're part of the corpus that explains who you are and how you're perceived.

To optimize:

- Be intentional on the channels that matter most in your vertical—whether that's Reddit, TikTok, YouTube, X, or specialist forums.
- Seed accurate, easy-to-quote explanations of key concepts and positions into high-visibility posts and videos.

- Monitor recurring misconceptions in UGC, then address them both on your own site (with clear correction content) and, where appropriate, in the threads themselves.

In tests around entertainment topics, reasoning responses from multiple platforms pulled citations not only from news outlets but from Reddit, comments, and video descriptions, especially when explaining fan debates or perceived sentiment. If those are the only sources with a strong stance, they'll dominate the narrative.

If Chapter 6 is about shaping what models remember, and Chapter 7 is about shaping what they can find, this chapter is about shaping how they argue on your behalf. Reasoning models are effectively building a case every time a user asks a hard question. Your job is to make sure your brand is the expert witness they call first.

Reasoning Model Optimization Checklist

Show your working on complex topics

- Identify the “hard questions” in your category where buyers hesitate or debate.
- Answer these questions using hard evidence (comparative tables, diagrams, metrics, third-party quotes).
- Take a clear stance on controversial topics in your industry so models have a definitive opinion to cite.

Build expert-level content clusters

- Create content clusters that cover every angle of a topic (definitions, comparisons, risks, data, opinions).
- Use clear headings, lists, and tables so models can easily map your sections to their research plans.

- Keep technical docs, FAQs, and support content up to date and easily discoverable.
- Ensure your site acts as the definitive expert on your brand and your specific solutions.

Optimize for citations and expert status

- Make key facts quotable using short, clear statements near the top of the page.
- Use consistent, entity-rich language (exact brand names, products, locations) across all channels.
- Keep author and brand credentials prominent, especially on high-stakes content.
- Identify which URLs are already cited by AI tools and strengthen internal links from them to other strategic pages.

Leverage current events and social evidence

- Publish timely explainers connecting your expertise to live industry topics, policy changes, or news.
- Add “what changed this year” or “what this means” sections to key evergreen pages.
- Ensure video scripts and social captions clearly state your brand’s facts, as transcripts act as evidence.
- Monitor recurring misconceptions in UGC (Reddit, forums, TikTok) and address them directly on your site.

9

Optimizations for All LLM Response Types

Over the last few chapters, we've been looking under the hood of specific AI response types. We've covered the tactical plays you need to surface your brand in RAG, MoE, and static training-data models. But AI search isn't a series of isolated algorithms—it's a living, breathing ecosystem. Now, we are going to zoom out. In this chapter, we are going to look at the foundational optimizations you can make to improve your brand's AI search visibility *regardless* of the model or response type.

All Roads Lead to Relevance

Why do we need a universal approach? Because optimizing for AI search is a multidisciplinary sport. You can't simply tweak the text on your web page, cross your fingers, and expect incredible results. Generative engines aren't just reading your site; they're

reading the whole web. They look at your entire digital footprint to understand who you are, what you do, and the authority you hold. Whether you're aiming to be cited by a lightning-fast RAG model or synthesized deeply by an advanced reasoning model, a robust, wide-reaching semantic footprint is what makes you the definitive answer.

In this chapter, we'll set the foundation. We will cover how you can do the following.

- 1. Work smarter, not harder**

Focus your efforts where AI reacts the fastest by testing changes on search-augmented LLMs first, and ensure your global content stays aligned across every language they can read.

- 2. Create AI-ready content**

Make your site effortless for AI systems to read by optimizing crawler access, structuring content into clean semantic chunks, and ensuring your visuals communicate meaning.

- 3. Strengthen your brand signals**

Strengthen how AI understands and trusts your brand by solidifying your entity signals, managing community-driven UGC, and proactively correcting misinformation wherever it appears.

- 4. Improve native platform presence**

Show up cleanly inside AI ecosystems by actively engaging with platform programs and optimizing your brand cues for AI-generated result pages.

Strategy and Workflow

Set your AI strategy on the right footing by starting where change is visible fastest. Search-augmented LLMs give you real-time feedback, making them the perfect proving ground for

testing optimizations before scaling them across slower-moving closed models. And because LLMs read every language version of your site interchangeably, you must treat your global content as one interconnected system. Clean, synchronized international pages are no longer optional—they're foundational.

In this section, we'll address why you should:

- **Phase your optimizations.** Start with search-augmented LLMs for real-time feedback and rapid iteration.
- **Think global.** Keep all international versions aligned because LLMs read everything, in every language.

Phase Your Optimizations

When it comes to LLM optimization, timing isn't just a factor—it's a tactical advantage. As we discussed in previous chapters, static training-data models come with a built-in delay. They rely on knowledge cutoffs, meaning it can take months for your latest website changes to be absorbed into their core understanding. RAG (search-augmented) models, on the other hand, are immediate. Because they are grounded in live web searches, they process new information in real time. This means that as you work on your AI visibility, you should expect to see RAG updates first—and you should plan your workflow accordingly.

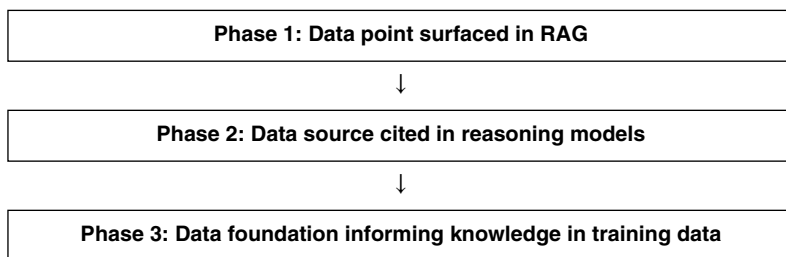
Think of RAG as your live dress rehearsal. Between the moment you publish your website updates and the time the static training data finally catches up, you have a golden window to fine-tune how your brand is perceived. If you make a tweak and it doesn't quite move the needle in a tool like Perplexity or ChatGPT's search mode, you don't have to panic. You have the opportunity to go back, adjust your phrasing, and try again. You can iterate in this live feedback loop until you are happy with the results.

But to make this phased approach work, you must follow a golden rule: start with your own domain and work your way

out. It is incredibly tempting to jump straight into off-site tactics—launching link-building campaigns or driving mentions via digital PR to boost your entity authority. But if those external publications are going to talk about you, they need a destination to point to. They need something concrete to mention. If that foundational content—the clear definitions, the chunky FAQs, the authoritative stance—does not exist on your own domain first, it is very difficult to sell your brand as something worth citing.

Create Your Own Information Pipeline By phasing your information and optimization rollout, you essentially create your own information pipeline for AI search visibility.

Phase your optimizations to create
your own information pipeline



You can move from being a simple data point that pops up in a quick RAG search, to being a trusted data source vetted by reasoning models, and finally, to becoming part of the foundational information that the model implicitly knows.

This progression means that when someone asks, “What is a good family car?” the Toyota Camry will show up no matter what the response type is. It has completely saturated the pipeline.

This kind of optimization means you can spend less time chasing your tail and more time actually building your brand. But remember: build the house before you invite the guests.

Once your domain clearly tells the LLMs who you are, the rest of your digital footprint will naturally follow suit.

Think Global

In classic SEO, we rely heavily on strict regional borders—like hreflang tags and localized domains—to ensure users see the right content for their specific location. But in the world of LLM, those borders are incredibly porous.

While AI tools will generally try to show users content from their local region, absolutely nothing is off the table. If you have inaccurate or out-of-date information lingering on one of your international domains, there is a very real chance it will surface as a confidently translated fact in an AI response.

Remember, tools like Gemini and ChatGPT can read and synthesize over 100 languages in a matter of milliseconds. If a user asks a highly specific question and your primary English site doesn't have the answer, but your localized German or Japanese site from three years ago *does*, the LLM will simply grab that older information, translate it on the fly, and serve it to the user as current fact.

If you see this happening, the fix is straightforward: you need to localize the content properly and update the information across *both* pages to ensure consistency.

The takeaway: You can no longer afford to treat your international domains as an afterthought. Because LLMs are the ultimate passport-free travelers, an outdated fact in one language is a liability in every language. Keep your global footprint clean, consistent, and up to date.

Crawlability and Content Structure

AI visibility starts with being easy for machines to read. That means giving LLM crawlers exactly the structure they prefer: clean markup, clear headings, strong internal signals, and

well-chunked content. Layer in multimodal optimization—ensuring your images, videos, and on-page visuals communicate meaning directly—and you create a site that is effortless for AI systems to parse, index, and cite.

In this section, we'll look at ways to:

- **Optimize for LLM crawlers.** Give the bots exactly what they need to understand you.
- **Keep it chunky.** Format for machines using FAQs, listicles, headings, and clear structure.
- **Optimize for multimodal.** Ensure your images, videos, and visuals are LLM-friendly.

Optimize for LLM Crawlers

LLMs use crawlbots to read your website, much in the same way traditional search engines do. This means you can track them in your bot log reports and instruct them via your `robots.txt` file. The major players are actually quite transparent about their user agents: ChatGPT uses OAI-SearchBot and GPTBot, Copilot leans on BingBot, Gemini uses Google Extended, Claude sends ClaudeBot, and Perplexity dispatches PerplexityBot.

But here's where everything changes: intent.

Traditional search engine bots treat your website like a “best-of” album. They focus on the hits—the content that resonates with the widest audience. Their job is to index your stand-out tracks and skip anything that feels like a deep cut, a draft, or a niche experiment. They're curators, picking only the songs they think deserve airtime.

LLM crawlers, on the other hand, treat your website like the newest release from their absolute favorite artist. They aren't here to judge which tracks should be popular. They want *everything*: every deep cut, every demo, every forgotten landing page, every tag folder. Their mission is to absorb your tone, themes,

vocabulary, and worldview so the model can “sample” the perfect riff later in a conversation.

Search engines skim to judge. LLMs binge to understand.

And when a user asks a question in chat? The LLM drops the exact beat it needs, sampling your content at the perfect moment to create a useful, context-rich response.

However, because these bots binge, they can get easily distracted. If you leave your site completely unrestricted, they will happily waste their time hanging out in your tag folders or deep pagination pages instead of learning the core facts about your brand.

This means you need to manage your crawl with an “all killer, no filler” approach. While some SEOs recommend blocking AI bots entirely, that is generally a mistake. Instead, use bot log reports to see exactly where these “superfans” are wasting time. Then, use your `robots.txt` to roll out the red carpet directly to the pages with your core brand narrative and value propositions, intentionally keeping them out of the weeds.

Keep It Chunky

There is a lot of ongoing discussion in the SEO world about “chunking.” Some folks have gone a little overboard, breaking everything down into microscopic, disconnected sections. But the core tenets of chunking have solid foundations in both machine learning and good content optimization.

We will dig into the technical details later, but at a high level: chunking means structuring your content into distinct, semantically meaningful sections that machines can index, embed, and retrieve cleanly.

This concept is not exclusive to AI. In web development, developers talk about breaking up the DOM (Document Object Model). In data science, when analysts classify data into different sets, they are also chunking it. In machine-learning pipelines,

information is broken into digestible pieces so that each “unit of meaning” can be processed and recalled independently.

Why? Because embedding windows are limited, and AI retrieval models return smaller units than they consume. Chunks that are efficient in terms of message and meaning can reduce noise and drastically increase semantic precision.

A perfect example of this is YouTube transcripts.

If you open the transcript on any YouTube video, you will see timestamps for every few seconds of the video. These are chunks. This is done so that it’s incredibly easy to find the exact piece of information you need if you want to surface a specific quote. Incidentally, YouTube videos are heavily cited by LLMs, and have been consistently since the beginning. The chunking of their transcripts, combined with the use of video chapters (which is just another form of chunking), is almost certainly a massive contributor to that visibility.

So, when we think about making our own content more citable and LLM-friendly, chunking is about creating natural retrieval points—places where an AI can grab a clean, self-contained piece of information without wading through a wall of text. The side benefit? It also makes your content dramatically more readable for human users.

In fact, good on-page content optimization has always included breaking up walls of text with heading tags, lists, and tables. Here are some chunking tactics that you’ve probably been using for years:

- FAQs
- H1s and H2s
- Numbered lists
- Bullet points
- Tables
- Timestamps
- Anchor links

Research on AI citations shows that listicles perform incredibly well in generative search. Why? Because listicles are automatically chunked. They require multiple headings, multiple data points, and inherently produce clean semantic blocks. FAQs excel for the same reason: they package information into predictable, digestible, machine-friendly units.

Can chunking be abused? Absolutely—like any SEO or GEO tactic.

But the real danger is over-chunking. When chunks become too short to carry contextual meaning, you create orphaned fragments. Done right, chunking is a highly valuable contributor to your overall AI strategy; done poorly, it leaves the machine—and your users—confused.

Optimize for Multimodal

I’ve mentioned YouTube a few times already, and I’m going to do it again. YouTube is a massive destination for AI-driven traffic, and its content is surfaced systematically across almost every major chatbot platform.

AI systems from Google, OpenAI, and Microsoft are rapidly expanding their multimodal capabilities—meaning their ability to process and interpret images, video, and screen layouts. As a result, your visual content carries far more influence than it used to. The images and videos associated with your brand now play a measurable role in how AI systems perceive and describe your business.

When AI tools evaluate your visual assets—images, logos, screenshots, videos—they aren’t just “looking” at pixels. They run object detection, identify entities, and often anchor those entities to a broader knowledge graph. For example, if a model sees a picture of the Statue of Liberty, it doesn’t merely detect a statue. It resolves that image to a specific location on a map and links it to associated historical, cultural, and contextual information.

Your business may not manage an iconic landmark, but you can still optimize visuals to align with entities relevant to your brand. If you run a local business in New York, prominently featuring a clear image of the Brooklyn Bridge (or other recognizable local markers) helps AI systems confirm where you are and what geographic entities your business is associated with.

Brand identity matters here too. Managing your logos, color schemes, and visual guidelines strengthens your multimodal profile. Google's multimodal assessment tools draw heavily from the same underlying technologies that power Google Vision AI and Google Lens. A consistent, clear visual identity increases the likelihood that your brand is recognized and surfaced in platforms like Gemini.

It's also worth noting that some AI systems take literal visual snapshots of your website during the crawl process to understand layout, hierarchy, and on-screen meaning. Ensuring that your core value propositions and key content are visible without excessive scrolling or visual clutter can significantly improve your multimodal crawlability and how quickly your information is retrieved.

The takeaway: AI doesn't just read your website—it sees it.

If your visual assets aren't clearly communicating who you are, where you are, and what you offer, you're leaving a major gap in your entity profile.

Entity and Reputation Management

LLMs lean heavily on entity understanding and real-world validation. Strengthening your brand's presence on authoritative sources, managing how your business appears in UGC ecosystems, and proactively correcting hallucinations ensures that AI tools form an accurate, trustworthy picture of who you are. The cleaner your entity footprint, the cleaner your AI outputs.

Let's get into how you can:

- **Manage your brand entities.** Strengthen your presence through Wikipedia, named entity recognition (NER), and authoritative sources.
- **Manage your UGC.** Let your community validate your expertise across third-party platforms.
- **Handle hallucinations.** Correct and reinforce misinformation by publishing consistent, canonical facts.

Manage Your Brand Entities

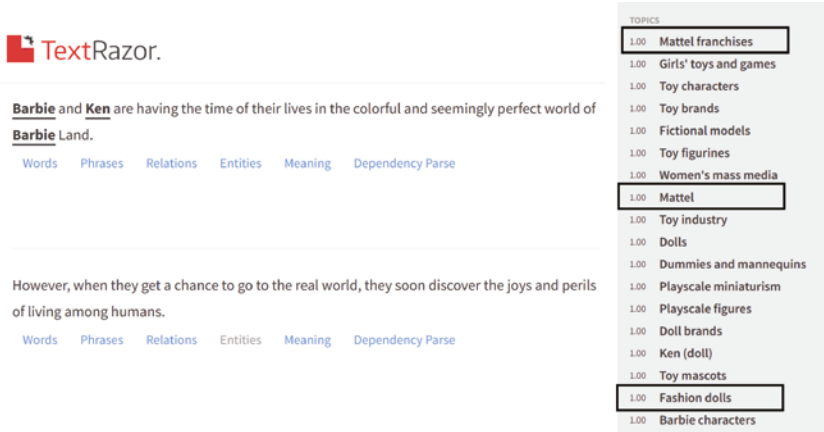
Both search-augmented LLMs and static LLMs are heavily trained on Wikipedia's collection of over 65 million web pages. This information forms an essential part of Google's knowledge graph and the semantic web as a whole. This means that entities form a core component of LLM training, and brands with robust entities will inherently have high visibility within LLM responses.

Case in point? Let's look at Barbie.

Consider the official plot summary for the 2023 blockbuster movie *Barbie*. It reads:

“Barbie and Ken are having the time of their lives in the colorful and seemingly perfect world of Barbie Land. However, when they get a chance to go to the real world, they soon discover the joys and perils of living among humans.”

Notice what's missing? It says nothing explicit about toys, plastic, or the corporate brand. But when you analyze that exact text with a natural language processor, entities like “fashion doll,” “Mattel,” and “toy industry” emerge instantly—despite the fact that the summary does not explicitly say them.



Why does this happen? Because language processing models can identify Barbie as an entity, particularly when she is mentioned in close proximity to Ken. More importantly, there are thousands of data points across the semantic web that tell these tools exactly what Barbie is.

This means that when you ask an LLM—whether it’s Claude, Perplexity, ChatGPT, Copilot, or Gemini—to “name a fashion doll,” they all answer Barbie. When this is reinforced by the opening line of the brand’s Wikipedia page literally reading, “Barbie is a fashion doll,” the entity association becomes unbreakable.

While your business may not have the colossal intellectual property of an iconic toy brand, you can absolutely replicate this strategy on a smaller scale. Working on your structured data, your Wikipedia presence, your NER, and your wider semantic footprint can go a long way in improving the accuracy and visibility of your products and services in LLMs.

Different business types will have different ways of managing their entities. A local service business might focus heavily on their Google Business Profile, while a B2B SaaS company might focus on niche databases like G2 or Crunchbase. But the goal is not to be everywhere; the goal is

to focus on what makes you unique and well-connected within your specific niche.

Manage Your UGC

When you look closely at how LLMs ingest and process data, it becomes immediately clear that they absolutely *love* UGC. Because LLMs rely heavily on this data for training, managing your UGC must be a foundational part of your AI search optimization strategy.

Look at the deals being made right now: both Google and OpenAI have direct, high-value partnerships with Reddit to use their conversational data to train their models. As a result, these engines regularly surface Reddit discussions in their RAG responses.

Why? Because Reddit threads are incredibly strong at providing real-world context for product reviews, non-commercial advice, and nuanced, ongoing discussions. Furthermore, these threads natively mimic the interface of a chatbot. A user asks a question, someone answers, and follow-up questions ensue. This structure makes them the perfect training ground for LLMs to learn natural language and anticipate the exact types of questions users will ask about particular topics.

Beyond Reddit, it is crucial to remember that general reviews, user comments, and niche forums also come heavily into play. As I mentioned earlier, LLMs are looking at your *entire* digital footprint. When an AI generates a response about your brand, it isn't just looking at what you say on your website; it's synthesizing the collective consensus of the internet.

And it doesn't stop at text. While YouTube videos have been consistently cited in AI responses for a while, we are now seeing LLMs regularly pull from Facebook videos, and increasingly, content from Instagram.

When managing your UGC, you must recognize that your social profiles themselves are part of your brand's core entity and

knowledge graph. Platforms like LinkedIn are cited regularly. These profiles are prime real estate—a great place to plant high-priority, definitive statements about who you are as a brand.

Additionally, almost every major social platform is now integrating AI directly into their systems. Meta has Meta AI, X has Grok, and TikTok is developing its own AI systems to support users and surface information.

My advice here is the same as it has always been: connect with the social media platforms that matter most to your users. It is exhausting—and usually counterproductive—to try to excel on every single platform. For instance, if your target demographic leans heavily towards TikTok, it is absolutely worth your while to optimize the content you create there. You need to ensure your videos are ready for AI responses, not just by choosing the right topics, but by being highly intentional with the actual words you say.

But the ultimate takeaway is this: you can no longer afford to treat social media and review platforms as completely separate from your search strategy. In an AI-driven landscape, your community *is* your content. Taking steps to influence what your users say about you in forums and comment sections is actively training the machine on how to answer the next person who asks about your brand.

Handle Hallucinations

AI hallucinations get a lot of dramatic press, but let's take a step back. When an AI “hallucinates” a URL or a piece of information, it is essentially just making a logical assumption and confidently stating it as fact.

Honestly, humans do this all the time. I've worked with plenty of clients, and I am not too proud to admit that I have, on occasion, published a link that was slightly off. You tell the

client the URL will be `example.com/halloween-night` and, at the last minute, they decide it should be `example.com/halloween-nite`. For all intents and purposes, that is exactly what an AI chatbot is doing. It's guessing the path based on patterns it has seen before, and then publishing it with absolute, unwarranted confidence.

What *is* new here is that classic search engines almost never did this. Google might occasionally send a user to a page that had been deleted or changed, but it didn't just invent URLs out of thin air. That's why I like to describe AI hallucinations as a “newish” problem.

Fortunately, there are a few straightforward ways to handle them and turn a potential crisis into an opportunity:

1. **Spot them in your bot logs.** You can identify hallucinated URLs by reviewing your 404 error reports and filtering for AI bot traffic. This will show you exactly which non-existent pages the AI is trying to access or send users to.
2. **Redirect (or build) the page.** Once you find a hallucinated link, set up a 301 redirect to point that traffic to the most relevant, actual page on your site. In some cases, the AI's guess might actually highlight a gap in your content—if the hallucinated page sounds like something your users actually need, just go ahead and build it!
3. **Give direct feedback to the LLM.** Go into the LLM, prompt the hallucination, and use the platform's native feedback mechanisms to correct the error. Just like with pre-trained responses, giving direct feedback tells the engineers and the model that the output is inaccurate.

The takeaway: by tackling hallucinations this way, you are handling the issue on all fronts. You are sending a signal to the

LLM that it needs to course-correct, you are ensuring that any user who clicks that link from their historic chat logs lands on a working page, and you are taking proactive action to align your content with exactly what the machine thinks your audience wants.

Native Platform Presence

Visibility isn't just about your website—it's about how you show up inside the AI ecosystem itself. Engaging directly with platform programs and optimizing your brand signals for AI-generated result pages helps ensure you appear clearly, consistently, and credibly. From ecosystem participation to your favicon in the AI SERP, platform presence is now a core part of your search strategy.

Finally, we'll cover how to:

- **Get involved with the platforms.** Partner with AI providers and contribute to their ecosystems.
- **Prepare for the AI SERP.** Don't underestimate the power of your favicon, metadata, and visual brand cues.

Get Involved with the Platforms

Don't just sit around waiting to be found in the answers of LLM chatbots—get actively involved within their ecosystems. Think of AI platforms right now the way we thought about social media features a decade ago. Whenever a platform rolls out a shiny new toy, their algorithm inherently favors the early adopters who play with it. Being a first mover in these AI ecosystems gives you a significant advantage and massive additional publicity. By positioning your brand natively inside these tools, you open yourself up to millions of active users and drastically reduce the friction they face when trying to access your services.

Here's how smart brands are planting their flag.

■ Partnering with LLMs

At the start of 2024, OpenAI started aggressively courting major publishers to become official content partners. Soon, other LLMs followed suit. As of now, OpenAI has deals with giants like *Condé Nast*, *The Atlantic*, *Time*, and *Vox Media*, while the Perplexity Publisher Program has locked in *Entrepreneur*, *The Texas Tribune*, and *Der Spiegel*.

If you are a major publisher, there is huge potential to align yourself directly with LLMs to guarantee your visibility. If you aren't? It's time to rethink your PR. You should be actively prioritizing your digital PR and link-building campaigns to secure mentions and features in the publications that *do* have these direct partnerships. If the LLM trusts *Time* magazine implicitly, and *Time* magazine trusts you, you have effectively hacked the pipeline.

■ Creating custom GPTs

Building a custom GPT isn't just a fun development exercise; it's a highly effective PR move. It allows you to carve out a curated, branded space right inside the ChatGPT ecosystem.

For instance, the *Diagrams: Flowcharts & Mindmaps* GPT has links built directly into its outputs that drive traffic straight back to the brand's website. Furthermore, these custom GPT URLs actually rank on traditional Google search, creating a fascinating cross-platform SEO loop. Don't underestimate this opportunity. Agencies that build custom GPTs to help their clients adopt these new technologies will quickly find themselves creating highly valuable, billable services.

■ **Embracing e-commerce and native experiences**

At Wix, we haven’t just watched this trend—we’ve actively built for it. We created custom GPTs and specialized apps directly with OpenAI. We also entered into e-commerce collaborations with OpenAI and Perplexity.

The goal here is simple: you need to look for opportunities to treat LLMs as both a referer *and* a conversion channel. As Wix CEO Avishai Abrahami summed it up when the Wix ChatGPT app was launched, explaining that these tools offer a “New entry point. Same great Wix experience.”

When you convey to your customers that you are partnering with these cutting-edge tools to make their lives easier, it builds profound brand trust while securing your place in the new ecosystem.

Prepare for the AI SERP: Check Your Favicons

When we talk about visual branding in search, marketers usually think about massive hero banners or perfectly curated social cards. But in the AI SERP, your most important visual asset is often the smallest.

A favicon is the tiny, 16 × 16-pixel icon that sits next to your site title. In traditional browsing, it helps visitors recognize your brand and stands out in a crowded row of browser tabs. In AI search, it is exponentially more important.

How can Base44 help me make an app?

 Answer  Sources · 20  Steps

 Base44

Base44: Build Apps with AI in Minutes

 Base44

Everything You Need to Build Apps - Base44...

 Work-Management.org

Base44 Review 2025: AI No-Code Builder Explained

 Base44

Base44 Enterprise | Build Apps at Scale

When your brand shows up as a citation in an AI response—whether it’s listed as a source at the top of a Perplexity answer or hyperlinked in a Gemini response—your favicon is sometimes the *only* visual visibility you get. It is shown across all platforms and attached to almost all sources and citations.

If your favicon is missing, broken, or fails to load properly, you won’t get a polite error message. You simply blend into the background, represented by a generic, default gray globe icon. In a dense user interface filled with competing text and references, a strong, recognizable favicon is what catches the user’s eye and drives the click-through rate (CTR).

This is a low-effort, high-reward technical optimization. You must ensure your favicon is:

1. **Present.** Actually uploaded to your site.
2. **Consistent.** Using the same recognized brand mark across all your digital properties.
3. **Loading properly everywhere.** Verified to show up correctly on different browsers and devices, meeting Google’s specific guidelines.

Do not let a missing 16×16 -pixel icon be the reason a user ignores your brand in an AI response. Dress your citations for the occasion.

Consider All Angles

AI search optimization is not a game of tricks or shortcuts. You can’t manipulate an LLM the way we used to stuff keywords into a footer or build spammy backlinks. The new rules of AI search require a much broader, more holistic approach.

Whether an AI is using lightning-fast RAG, reasoning through an MoE, or pulling from static training data, its ultimate

goal is to synthesize the truth. To become the definitive answer, you need to align every part of your digital presence.

Moving forward, your universal playbook for LLM optimization boils down to the four core pillars we just covered:

1. **Work smarter, not harder.** Phase your rollouts, starting with RAG, and maintain global consistency.
2. **Create AI-ready content.** Guide the bots, chunk your text, and make your visuals machine-readable.
3. **Strengthen your brand signals.** Own your entities, leverage your community's voice, and proactively correct AI hallucinations.
4. **Improve native platform presence.** Lean into platform partnerships and sweat the small visual details that drive clicks.

The ultimate takeaway? You aren't just optimizing a website anymore—you're managing a living ecosystem.

By structuring your domain cleanly, managing your global footprint, and treating your digital presence as one interconnected system, you stop chasing the algorithm. Instead, you give the AI the exact, undeniable blueprint it needs to confidently cite you as the authority.

10

Embrace Experimentation

In search, experimentation isn't optional—it's survival. This space moves at a blistering pace, and ideas fly around faster than any team can canonize into "best practice." SEO and search conventions that felt immovable just a few years ago are now being rewritten in real time.

A perfect example: crafting meta descriptions by hand was once a sacred ritual. Then Google began rewriting them with AI. Soon after, practitioners started generating them with AI at scale. Today, outside of a small set of high-value pages, many teams rely on automation for the task entirely. That is the nature of this industry: nothing is invulnerable, no matter how time-honored it may seem.

In a landscape where opinions from thought leaders abound, emerging tactics can sound promising, provocative, or downright chaotic. Some will be brilliant. Some will need refinement. And a few simply won't stick. But that doesn't undermine their

potential value—or the importance of exploring them. What does matter is how you explore them: safely, intentionally, and in spaces where experiments don’t jeopardize traffic, revenue, or resources you can’t afford to lose.

The truth is: best practices are being written every day. Those willing to test, learn, iterate, and embrace the “fail fast, win big” mindset often earn first-mover advantage. I like to remind teams of this simple truth: it’s called web development, not web establishment. Progress comes from trying things—not from protecting the status quo.

With that spirit in mind, this chapter looks at three emerging techniques that have sparked extensive debate across the search optimization community:

- `llms.txt`
- `llms-full.txt`
- Generating pages in Markdown

These practices come with real potential use cases as well as valid questions and reservations. But unless you’re willing to experiment, you won’t be able to assess their true value for your own site, your stack, or your strategy. So, let’s dig in, look at what the industry is discussing, explore where these tactics might help (or hinder), and equip you to run experiments that are safe, strategic, and genuinely informative.

What `llms.txt` Is

`llms.txt` is an emerging web protocol that calls for a plain-text, Markdown-formatted file placed on a website (e.g., `yourdomain.com/llms.txt`). The intention is to provide LLMs with a clean, structured, human-curated map of your most important content.

Think of it as:

- **Not a robots.txt.** It doesn't block crawlers from doing anything.
- **Not a standard sitemap.** Although it behaves like a highly curated one.
- **A crib sheet for AI models.** It can surface your most LLM-friendly pages in a highly compact format.
- **A tool to speak directly to agents.** This lightweight file type can be used to guide agents to more effectively navigate your website.

It was originally proposed in 2024 by Jeremy Howard as a way to help LLMs avoid HTML clutter, JavaScript bloat, and excess tokens—delivering only what's actually helpful for inference. This is crucial because Markdown is up to 95× more token-efficient than standard HTML. As mentioned, while a standard AI web crawler has to comb through all your raw HTML, CSS, and site navigation just to figure out what you do, an `llms.txt` file strips away the noise.

The primary reason tech giants like OpenAI, Google, and Microsoft are actively building with these files is simple: computational efficiency. Because these files contain only text, they require a fraction of the tokenization cost of a standard web page. In one test, processing an `llms.txt` file required 95 times fewer tokens than a relatively simple HTML homepage.

While it is not yet an official web standard, I'd like to acknowledge that it's only a few years old. It took 26 years to formalize `robots.txt`, so this really is early days. Adoption is currently being driven by humans. Web users and SEOs around the globe are fascinated by the idea. Google searches for the term "`llms.txt`"

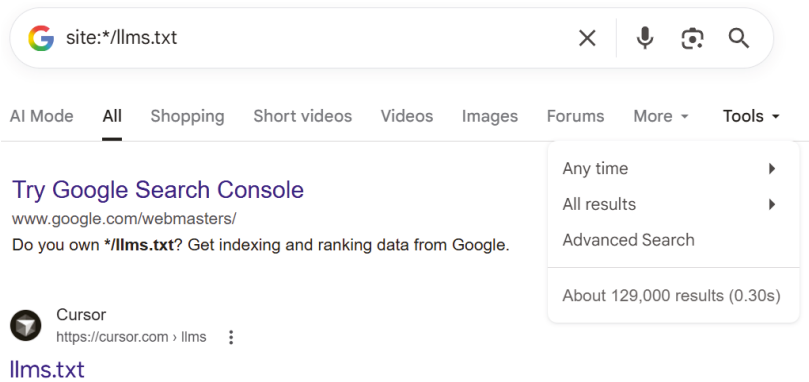
hit around 90,000 in August 2025 as Wix rolled out its automatic generator for e-commerce sites (Wix eCommerce). That contributed to a 94% increase compared to the previous three months. Plus, there are over 26,000 Reddit results on Google for the topic, suggesting genuine human interest in this tool.

This customer interest is something that drove my team to roll out `llms.txt` for our Wix and Wix Studio users. This scale of interest also suggests that if you have not implemented it, your competitors already may have.

Mythbusting the LLM File Hype

Whenever a new tactic drops, the SEO community tends to react like it is either the Holy Grail or completely useless—and `llms.txt` is no exception. Let’s break the biggest myth right now: simply creating this file is not a magic bullet that will instantly make you “rank #1” in ChatGPT. But dismissing the protocol as entirely useless for GEO is demonstrably false. The reality sits somewhere in the middle.

There is clear evidence these files can be crawled, indexed, and surfaced across major platforms. Research shows that Google currently indexes tens of thousands of `llms.txt` files globally.



March 2026 search of site:*/llms.txt shows 129,000 files on Google

After personally reviewing 1,400 files, I found that they actively surface in Google's AI Mode, ChatGPT, and Perplexity. Sometimes these files are pulled in via RAG because they rank well in classic search, and other times AI engines seem to bypass search entirely to access them as an independent grounding source.

While showing up in traditional search isn't the primary goal, it has been argued—and there is merit to this—that Google treats these files no differently from a standard `.txt`. The file type is well established and has been surfaced in classic Google Search for many years. Indeed, Google representatives have said as much. However, observationally, I've seen a small portion of these files ranking for semantic search results in a way that I've never seen `robots.txt` files do. As more of these files become live, the situation may evolve, just as it has done for other web phenomena.

Bear in mind, if you are seeking to test `llms.txt`, you should do so with the expectation that the ROI may be low, or it may be very long-term. Therefore, it's useful to implement it in a low-lift situation. For instance, at Wix, we have created automatic `llms.txt` files for our premium eCommerce users with custom domains. This means that users don't have to do anything to implement it; rather, they opt out. Other implementations can include automatic generation where you simply click a button to generate the file. In these cases, the risk in terms of resources is extremely low, and it allows you to gather primary data about the effectiveness of this technique.

If generating this file from scratch requires massive development resources, there are likely other techniques that will use your time more efficiently and yield faster results. That being said, the initial goal of `llms.txt` was never to drive AI search traffic. It was to make web content easier for LLMs to process, which has a few different applications that major companies are putting in place right now.

Six IRL Ways Marketers Can Use 11ms.txt Today

Moving from theory to practice, how do we actually put this protocol to work? I recently crawled and assessed over 1,400 publicly listed 11ms.txt files to see exactly what the architects of our AI future—companies like OpenAI, Google, Amazon, Microsoft, and Anthropic—are actually doing.

The data is wild: the average 11ms.txt file is just 9.8kb. That is 275 times smaller than the average web page weight of 2,600kb in 2024, proving exactly why AI bots love them. They are also hyper-connected, containing an average of 428 URLs per file to aggressively facilitate content discovery.

Based on this research, here are six core, data-backed ways you can use this protocol today.

1. Signal core context

Think of this as handing the AI your brand's resume before the interview even starts. The most common use for this file is to point LLMs exactly where to find your most relevant information. For example, Wix Premium eCommerce users automatically generate an 11ms.txt file that summarizes their core business offering, contact details, and product catalog. Files like this give the AI the exact context it needs without forcing the bot to scrape the visual HTML of the entire site.

2. Act as a markdown sitemap

Human users might enjoy browsing a beautifully spaced mega-menu, but AI bots just want the facts, fast. Tech giants are using 11ms.txt as a streamlined directory to point agents to plain-text Markdown pages. Google does this for Gemini, Chrome, and Firebase, linking to over 50 Markdown pages in a single file to keep crawlers efficient. Cloudflare takes the exact same approach, using

their file to link AI directly to raw Markdown versions of over 30 documentation topics.

3. Inform agent-to-agent architecture

We are moving rapidly towards an agent-to-agent (A2A) web, where AI tools talk directly to other AI tools to complete tasks. If you run a SaaS product, your `llms.txt` file acts as the ultimate instruction manual. Anthropic, AWS (specifically Powertools for Lambda), and Google's Agent Development Kit (ADK) actively advise developers to use these files to inform the knowledge base for AI agents. Anthropic even explicitly tells developers that LLM-friendly documentation should default to flat `llms.txt` files on official sites.

4. Reduce LLM hallucinations

Nothing is more frustrating than an AI completely hallucinating the capabilities of your product. You can use this file as a bouncer against bad data by providing direct instructions to the bots. For instance, Weather.com includes specific instructions on the URL structure for their forecast pages right in the opening paragraph of their file—and my tests show Perplexity actually surfaces it as a top-cited, and sometimes only, source for responses. Similarly, Zapier uses its file to explicitly instruct AI on how to properly make API calls, drastically reducing the chance of an LLM getting it wrong.

I find this use case particularly interesting because one of the criticisms of `llms.txt` is that it can repeat information already on the website, causing duplicate content. However, even if you don't get into agentic instructions, you can include information in this file that you would never share with human users. Weather.com, for instance, includes their URL formatting rules:

“All location-specific forecasts use the pattern `/weather/[forecast-type]/1/[location-code]`, where location-code is Weather.com’s unique identifier for cities and regions.”

This is incredibly useful to machines but would not make any sense to put on a human-facing website. Equally, it’s not something that would go into a `robots.txt` file. So, logically speaking, `llms.txt` is the best place for it.

5. Provide robust grounding context

GEO isn’t just about visibility; it’s about helping agents solve tasks right then and there. Some companies use this file to provide step-by-step grounding context for complex tasks. Vercel’s AI SDK documentation took a unique approach, creating an absolute beast of a document—over 200,000 words long—designed specifically to help agents troubleshoot migration errors. Microsoft does something similar in their GitHub repo for GenAIScript, explicitly telling the LLM to “add TODOs where you are unsure” during code generation.

6. Reduce token costs

Some of the most consistent uptake of `llms.txt` comes from the teams behind A2A tools. Building for an agentic future, teams at Google, AWS, Anthropic, Perplexity, Microsoft, and OpenAI have all implemented these files specifically for the purpose of improving A2A communication. Anthropic recommends using them in their *Writing for Agents* guidance, while OpenAI has them for their Agents SDK and their Agentic Commerce Protocol.

Cost of tokenization is critical for agentic users. For example, Gemini’s 2.5 Flash model costs \$0.10 per 1,000 tokens. If we look at the raw HTML of the relatively simple `llmstxt.org` homepage, the text file could be processed 286 times for the exact same price as one HTML homepage tokenization.

**11ms.txt versus homepage tokenization cost
(a review of llmstxt.org)**

File type	Tokens	Cost per tokenization (USD)	Cost per 1,000 tokens (USD)	Tokenizations per \$1
11ms.txt	158	\$0.0000158	\$0.10	63,291
Homepage HTML	15,085	\$0.0045255	\$0.10	221

The agents will bear the cost of accessing a website, and if the tokenization cost is too high, they may simply choose another site. This suggests that 11ms.txt could be useful for ensuring that websites are part of the agentic web of the future.

So, Should You Implement 11ms.txt?

This is a question I’ve been asked a lot, and my answer is a classic SEO response: “It depends.” If you can do so, knowing that this is a developing protocol, and you are testing the impact in a low-lift environment, then absolutely have a go. If it’s going to take extensive developer resources to implement, then you may want to focus on other things. But test, iterate, and listen to your human and AI users.

What About 11ms-full.txt?

If your standard 11ms.txt file acts as the high-level summary and directory for an AI agent, then 11ms-full.txt is the unabridged recipe book. This file is frequently paired with websites that are already implementing the standard 11ms.txt protocol. While the core file provides a neatly summarized directory pointing to other Markdown pages, 11ms-full.txt consolidates all of that rich, detailed information into one massive,

highly structured text document. It is designed specifically for modern LLMs that boast incredibly expansive context windows and can ingest hundreds of thousands of words in a single prompt.

Powering Agentic Development

This extended format is not just a theoretical concept; it's actively being used to power agentic development. For instance, the team behind Google's ADK uses this format. In their official resources, they explain that their ADK documentation is available as machine-readable files following this exact standard, noting that these files are automatically generated with every documentation update to remain perfectly up to date. Google teaches developers to use these files to provide robust grounding for custom AI agents, and they practice what they preach by offering their own `llms-full.txt` to help developers easily train models on their framework.

A Strategic Proof of Concept

Much like its smaller sibling, you need to approach `llms-full.txt` as an experimental proof of concept rather than a magic lever that is going to generate immediate, hockey-stick traffic results. These expansive files do the heavy lifting when they are built with a clear, intended audience and a highly specific intent. For example, if you are looking to instruct a third-party AI agent on exactly how to use your complex APIs, or you need to teach it how to seamlessly access specific software resources, you can unpack all of those instructions within this single file. It effectively hands the bot your complete, uninterrupted rulebook in one hyper-efficient sweep.

Backing It Up for Humans

However, I must stress that rolling out an `llms-full.txt` file does not replace robust technical documentation. This

machine-centric protocol must absolutely be backed up with documentation that allows human users to surface and read your content in normal, accessible ways. We are building for a hybrid future where humans and bots will navigate the web side-by-side, so your strategy needs to serve the machines without ever leaving your human customers in the dark.

So, Should You Implement `11ms-full.txt`?

You can implement `11ms.txt` without implementing `11ms-full.txt`. If you are looking at the full version, you should do so when you have an explicit use case for it (i.e., as a tool for agentic use or deep grounding).

The Markdown Efficiency Play

As we established with `11ms.txt`, Markdown is a hyper-efficient format for presenting web information. Because the files are incredibly small and strip out the visual bloat of HTML, they dramatically reduce processing complexity. This is exceptionally valuable for AI agents that are making rapid-fire, repetitive calls to a web page, as it saves massive amounts of context window space and computational overhead. In fact, serving raw Markdown instead of full HTML can result in a token reduction of up to 96%.

Spotting this trend, a growing number of vendors and networks—like Cloudflare with their recent “Markdown for Agents” rollout—are offering services that automatically convert your HTML pages into lightweight Markdown. For sites actively utilizing this, you will often see a literal toggle button on the page, allowing users to flip seamlessly between the human-facing HTML design and the bot-facing Markdown text. These same sites typically use their `11ms.txt` file as a master index, pointing agents directly to this repository of Markdown variants.

Addressing the Duplicate Content Elephant

Naturally, the loudest criticism of this technique hinges on SEO's oldest boogeyman: duplicate content. Having two seemingly identical versions of the same page floating around the web sounds like a canonical nightmare. While it is a highly legitimate reservation, I like to remind people that this dual-format approach is neither brand new nor reserved solely for experimental commercial enterprises.

Take the Internet Engineering Task Force (IETF) and their "Request for Comments" (RFC) series, which is where core web protocols are actually reviewed and formalized. For years, they have offered thousands of these highly technical pages in both rich HTML and plain .txt formats. When you search these repositories, Google manages the duplication effortlessly. This proves that search engines already possess the sophisticated filtering tools required to handle dual-format files without arbitrarily penalizing the host site.

Safe Execution and Monitoring ROI

If you are still losing sleep over the idea of your AI-focused Markdown files muddying up your classic search results, there is a simple, stress-free fix: you can confidently `noindex` the Markdown versions. This keeps them safely invisible to traditional search crawlers while remaining perfectly accessible to the direct, targeted agentic calls that actually need them.

Finally, if you are looking to test this by using a third-party vendor or service to generate and maintain these files, you must keep a close eye on the cost. Some platforms are rolling this feature out at no additional fee, which makes for a highly accessible, low-risk experiment. But if a vendor is charging you a premium to host parallel Markdown pages, you need to treat it like any other piece of martech. Monitor your server logs, track the bot traffic, and ensure you are seeing a tangible ROI in visibility or

task completions—just like you would evaluate the success of any other tool.

Building Your “Fail-Fast” Sandbox

Before you go and rewrite your entire site architecture to appease our new AI audiences, take a breath. I said experimentation is mandatory, but reckless experimentation is just bad business. You need a sandbox—a controlled, low-risk environment where you can throw things at the wall to see what sticks, without jeopardizing your primary revenue drivers or exhausting development resources.

Setting up a safe test kitchen for AI search protocols comes down to four practical steps:

1. Isolate the playground

Don't run your first `llms.txt`, Markdown, or other experiment on your highest-converting product page or your flagship homepage. Instead, pick a specific, informational subfolder. Your API documentation, your help center, or a dedicated customer service directory are perfect candidates. These pages are naturally rich in facts, making them prime targets for AI agents looking to answer user queries, but they aren't your main money-makers.

2. Keep the lift low

We are aiming for a “fail-fast” ethos, not a six-month, cross-departmental development sprint. Use the simplest tools available. If you are on an enabled platform with automated implementation—like Wix has for some users—activate the automatic `llms.txt` generation. If you are hand-coding it, start with a basic text file that points to just a handful of your best pages. If an experiment takes your engineering team a month to deploy, it's no longer a sandbox—it's a massive bet.

3. Monitor the bot action

You can't evaluate an experiment if you're flying blind. Once your files are live, you need to watch your server logs like a hawk. Look for hits from AI-specific user agents like OAI-SearchBot, Google-Extended, or PerplexityBot. Are they actually pinging your `11ms.txt` file or fetching your new Markdown pages? Additionally, keep a close eye on Google Search Console to see if Google's traditional crawlers are indexing these files.

4. Evaluate the actual ROI

Finally, track the real-world impact. If you built an `11ms-full.txt` file specifically to stop AI tools from hallucinating your pricing tiers, go into ChatGPT or Perplexity a few weeks later and ask about your pricing. Did the hallucinations stop? Are you appearing as a cited source more frequently? Treat this like any other piece of mar-tech. If you see a tangible improvement, scale the experiment. If it falls flat, you pivot—having lost nothing but a couple of hours of setup time.

Conclusion: Don't Wait for the Dust to Settle

The landscape of AI visibility and the agentic web is developing right in front of us. As a marketing leader, it is incredibly easy to look at experimental formats like plain-text protocols, `11ms.txt`, or Markdown pages and write them off as purely developer concerns. But at their core, exploring these formats is about communicating to key user groups; what's changed is that those groups now include both humans and AI tools.

Don't let the fear of the unknown, the lack of a 20-year industry precedent, or even a lack of industry consensus paralyze you. Find out for yourself.

Testing can no longer be a one-off event; setting up an environment for experimentation must become a regular, baked-in

part of your digital strategy. As this technology continues to evolve at breakneck speed, we are going to need to test everything. If you don't build the muscle memory to run these experiments yourself, and rely entirely on third-party data, you lose control of your own strategy.

The cost of running controlled, low-risk experiments is incredibly low, but the cost of ignoring them entirely is steep. Set up your sandbox. Build the habit of testing, monitor your own server logs, and evaluate your own ROI. Remember, we are working in a living, breathing web. The brands that win won't be the ones waiting for the dust to settle—they will be the ones out there actively kicking it up.

Get Involved with the Platforms

This might seem obvious to say, but AI platforms are driven by innovation. Not only does it bring new users to the platform, but it also attracts the media and investors. At this stage in the game, competition for AI users and market capture is fierce, so each new product feature can amount to millions of dollars in revenue and earnings for the platform providers.

How does a business “get involved”?

- Build content partnerships with AI search tools
- Become an authorized e-commerce seller on OpenAI
- Create platform-specific branded agents (like OpenAI's “Operator”)
- Partner with partners who are already AI-integrated

Marketing teams who are present within new features, strong collaborators, and adopt new elements for brand discovery early on are far more likely to see returns on their efforts than less innovative teams as the features are rolled out with high priority to platform users.

For each of these participation methods, the goal is to make your brand name visible to users on these platforms, to demonstrate that your brand is capable of making connections and isn't afraid to get involved with new ideas, and of course to help you satisfy your business goals.

But regardless of new features, you still need to make sure your brand is front and center. Many of the old rules still apply, so you should be looking for opportunities to:

- Link to your website
- Show your brand values and logo
- Recommend products
- Upsell
- Capture emails
- Encourage app downloads

Anyone who has been in the game long enough knows that with native integrations, you need to include as much information as possible on your brand, leave no field unfilled, in order to get maximum visibility. From Perplexity pages to agentic AI integrations, one digital marketing rule to live by is to use *all* the tools available to you to make sure your brand stands out.

Before You Begin

With the incredible scope of AI, there is potential for mission creep here and to simply chase after the shiniest new object, but before you start going Gollum, you will need to assess whether the channel is appropriate for your brand or not. This is simply because the technology is only of any use to you if it helps to showcase what brings the most value to your customers.

Equally, don't let technological prowess hinder you. Many of these new AI features can be made accessible by using AI to tell you how to use the AI.

11

Build Content Partnerships with AI Search Tools

AI tools need data and content, something legacy publishers have by the bucketload. This is why OpenAI, Gemini, and others are actively courting major publishers around the world.

Since 2023, OpenAI has made a distinct effort to acquire partnerships with some of the world's most renowned publishing groups for news and magazines. Partnering with the likes of Hearst, NewsCorp, Vox Media, even the World Association of News Publishers, the AI firm has collaborations that allow it “to enhance its products” and gain access to current and archival information to train its models.

But what is the marketing play?

Publishers with these relationships benefit from a potential stream of revenue and can also solidify brand authority within an AI platform and the wider web. These licensing deals can create steady income in a dynamic online reality.

So, if you are a major publication, then you may very well wish to get your publication included as a partner for OpenAI and/or other AI tools. This will mean that you can potentially get privileged in terms of citations, and this has been evidenced when you compare certain publishers to others.

Alternatively, if you are not a major publication but you are a commercial brand or other entity, then you may wish to prioritize where you seek to get PR and media coverage. That is to say, if you can take an interview, create content, or get a mention in a publication that has a partnership with OpenAI, you will be more likely to see that information pull through into the models than if you get coverage from another party.

Harness Agentic Solutions

In the same way that apps have transformed the way that we get stuff done on the internet, so too will agents in the very near future.

Unlike chatbots, agents can act independently of humans to carry out tasks autonomously. Brands that can facilitate agentic solutions for their customers will be more visible with AI because they will be built for AI to access at will.

For instance, if a smart refrigerator runs out of milk, then it could potentially use an agentic shopping assistant to purchase another carton. In this scenario, the human expecting to use the milk on their cereal wouldn't be involved in the purchasing directly at all.

Just as vendors create offerings specifically for Uber Eats or apps for the Apple App Store, your brand may need to create agentic solutions to capitalize on these new business opportunities.

You can capitalize on agentic solutions by

- Becoming an OpenAI e-commerce seller
- Utilizing agentic infrastructure
- Creating your own agents

Become an OpenAI E-commerce Seller

As of 2026, the strategy for selling on OpenAI has shifted. OpenAI is no longer trying to be a warehouse; they are a *discovery layer*.

By registering as an authorized e-commerce seller, your product catalog becomes discoverable to “Operator” agents. When a user tells ChatGPT, “I need a professional-grade camera for under \$1,000,” the AI doesn’t just search the web—it looks for verified merchants who can fulfill that need instantly.

How to do it (the simple way):

1. **Use AI-forward platforms.** If you are on a platform like Etsy, or Wix, eligible sellers can enable AI Discovery within your platform settings.
2. **App-based checkout.** OpenAI has moved towards app-to-app transactions. Instead of buying *inside* the chat, the AI deep-links the user directly into your brand’s app with the cart already filled.
3. **Factual over flowery.** AI agents prioritize “product truth.” Ensure your descriptions are literal (e.g., “100% Organic Cotton, Navy Blue, XL”) rather than just marketing fluff.

Utilize Agentic Infrastructure

Established tech teams like Google and AWS are creating A2A infrastructure and services to facilitate the agentic scaling and implementation. E-commerce payment systems like PayPal and Stripe are working with e-commerce retailers and OpenAI to create/trigger automated purchases based on AI parameters.

In the case of PayPal and Stripe, they will be seeing genuine business value coming from these AI platforms as transactions flow their way. Additionally, their collaboration will add more value to the brand.

For smaller businesses, you can increase visibility by taking advantage of platforms and services that are fully integrated with this agentic infrastructure. So yes, make sure you have payment options that include eligible platforms. But also think more widely.

In October 2025, the Agentic Commerce Protocol (ACP) was announced, ushering in an era of instant purchasing on ChatGPT and other platforms. The team at Wix joined this early cohort, taking steps to make instant purchases from Wix Stores in ChatGPT via Stripe, available to over 300 million Wix users. They followed up with agentic purchasing integration with Perplexity via PayPal, in the following month.

In this case, Wix users, both current and future, benefit from agentic architecture without having to build it themselves or apply for access individually. The lesson here being that when selecting tool providers, look for those who are AI forward and are actively building for an A2A future. This will mean that you can access more customers on more surfaces more consistently as this technology evolves.

Create Your Own Agents and Agentic Solutions

Creating your own agents and agentic solutions for your brand can help embed your brand as a native feature within AI tools themselves. This means, when a user wants to complete a task—either by direct prompt or through an AI agent—ChatGPT or similar platforms will select or recommend your brand’s agentic solution automatically. So, they basically do your marketing for you.

OpenTable is a prime example.

In March 2023, OpenTable launched a ChatGPT plugin as part of OpenAI’s early plugin rollout. This agentic plugin let ChatGPT give real-time restaurant recommendations from OpenTable’s network and included direct booking links,

enhancing discovery and making the brand visible within the AI tool because it was part of the tool.

Building on this, in July 2025, OpenTable launched their AI Concierge—a generative AI-powered dining assistant embedded directly within restaurant profiles on their own platform. Straight away, they connected to OpenAI’s Operator agent, which can carry out autonomous multistep tasks like booking reservations independently.

They’d moved from agentive to agentic.

Did that impact OpenTable’s AI visibility? Absolutely.

According to data from SimilarWeb, between July and September 2025, OpenTable increased their AI referral traffic by about 40%. During the same period, ToastTable, one of their closest competitors, saw their AI traffic decline by 20%.

Partner with Partners

You might be looking at this list of tactics thinking, “I don’t have a multinational SaaS company, how am I going to build direct partnership with OpenAI?” or “My team is still working through a long list of marketing initiatives, how are we going to find time to make an AI agent?” And this is fair.

We’ve all struggled with access and capacity at times, and when this happens we should look for a little help from our “friends”—or, more specifically, the vendors and services that our business already uses.

AI challengers to Google have fully embraced the adage of *if you want to go fast, go by yourself. If you want to go far, go together.* OpenAI, Perplexity, and Claude actively pursue partnerships with teams that help them to provide better services for their users. OpenAI doesn’t need to build a database of restaurants, because they can partner to use OpenTable. Perplexity doesn’t need to build its own medical database, it can partner with b.well, Connected Health, and fitness apps like Strava to provide a knowledge base for its health features.

This benefits you as a marketer.

It means that you can continue to thrive on platforms that you know and love, and if you choose your partners wisely, it can and *will* increase your presence on AI search.

Take the website builder Wix as an example. When OpenAI announced their instant shopping feature using the ACP, Wix was prepared. The ACP was developed with Stripe using a similar tech configuration that Wix used to set up MCP payments with them just a few months before. This meant that Wix's team was able to swiftly adapt and align with new protocol, building the agentic future of Wix users.

When Perplexity added a similar feature via PayPal, Wix, who had a long-standing relationship with PayPal, was able to capitalize on this to create a native integration for agentic shopping. Again, creating AI search opportunities for users of their products that are built into the existing product.

In these times of change, choose your tech partners to stack strategically. Look for teams that are:

- AI forward in ethos with regular investment in their own AI products
- Building partnerships, apps, and integrations with OpenAI and the like
- Actively innovating with MCPs and native agentic solutions

It might seem obvious to say, but teams that build with AI will also help you grow with AI. If the SaaS products you use are not locked into AI tools, you might want to log out.

All of this is to say that if you want to improve your AI visibility, you can't just sit on the sidelines, you can't just keep talking about it. You need to get active, get your hands dirty, and get involved.

12

Be Prepared for the Agentic Era of AI Search

The digital landscape is shifting beneath our feet again. We’ve moved from a world of blue links to the era of “answer engines,” and now, we’re crossing the most transformative frontier yet: the agentic web. In 2026, AI agents aren’t just answering questions or summarizing text; they’re actually *doing stuff*. They are browsing, booking, and buying on behalf of users.

For the savvy marketer, this isn’t a threat—it’s a massive industry shift to be mastered. If traditional SEO was about building a beautiful, welcoming storefront for human eyes, being “agent-ready” is about giving the digital concierge the exact tools they need to run errands for your customers. To be agent-ready is a core GEO strategy because it makes your brand explicitly eligible to surface—and transact—in the responses and workflows that take place entirely within AI environments.

Agents are users too, and it’s time we started optimizing for them.

The Rise of the Agentic Web

If you think agents are just a passing phase, the data says otherwise. We are seeing a massive paradigm shift driven by both executive buy-in and consumer demand.

The global AI agents market hit \$11.55 billion in 2026, and it's projected to explode to over \$236 billion by 2034—a staggering compound annual growth rate of nearly 46%. Gartner reports that 40% of executives will have adopted AI agents by 2026, while Statista notes that nearly a third of Americans (32%) actively want an AI personal assistant to automate everyday tasks.

Let's ground this in reality. This isn't about traditional search disappearing overnight; it's about a very practical shift in how people choose to interact with the web for complex tasks. Users are starting to take new routes: they are accessing dedicated agentic solutions right inside ChatGPT, deploying agentic browsers with Claude, and relying on Google to carry out multi-step shopping and booking workflows. These systems have real-world impact because they can actually navigate web pages, complete forms, access databases, and carry out cross-tab orchestration. They aren't just reading the web; they're actively using it.

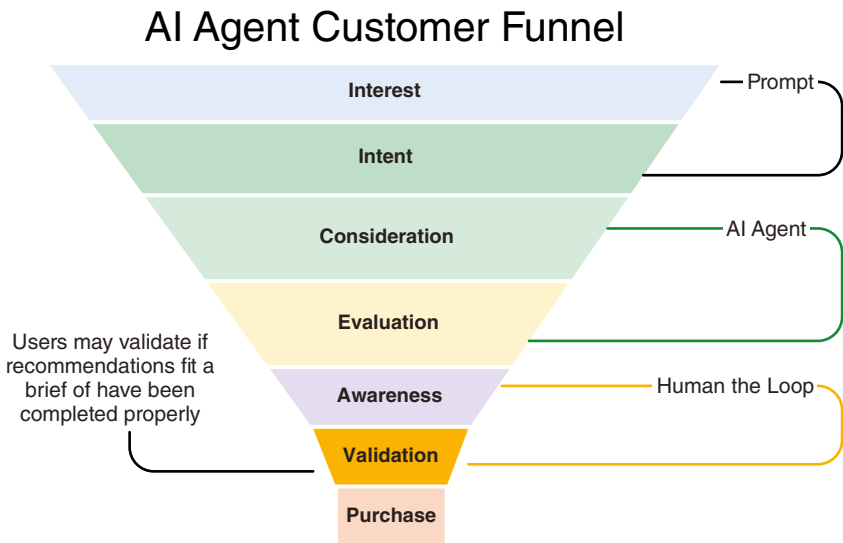
This means that focusing solely on showing up in text prompts for brand visibility could mean you miss out on actual conversions. Earning visibility in singular AI responses is incredibly valuable—it is the foundation of your strategy—but tapping into agentic opportunities gives you 360° gains. To win, you must adopt agent-ready tech solutions and practices to meet your customers in the next phase of the web, ensuring your brand is ready to act when the AI comes knocking.

Bracing for the Agent-Mediated Funnel

AI agents directly disrupt the classic customer funnel because they introduce a new gatekeeper. Historically, human users would enter the funnel at the *awareness* stage, evaluate their options, and progress to a purchase. Today, AI agents are making decisions

before those options are ever presented to your audience. Think of the AI as the Alfred to your customer's Batman—tireless, laser-focused, and incredibly protective of their time.

Consider the weary driver asking Android Auto to “navigate to the next rest stop.” The driver declares their *interest*, but it's the AI system that *evaluates* which stop is closest, considers what is open, and checks the reviews. The driver only becomes aware of the specific brand when the AI announces the destination. In this new reality, users may not directly engage with your brand until after the agent has already considered and evaluated you against the competition.



So, how can you prepare a website for AI agents?

Make Your Site a “Do-Thingsy” Destination

In an agentic ecosystem, a website that only *looks* good is a dead end. If your static website only provides information, the agent will simply bounce and find another site where it can actually satisfy a user's request. You have to shift from being a brochure to being a conversion engine.

Give agents the tools to book appointments, buy products, and make inquiries. This means moving beyond simple copy and adding integrated forms and booking tools across all relevant pages, not just your “Contact Us” page. You must also prioritize bot-friendly accessibility. Agentic browsers use ARIA tags (the exact same ones used for screen readers) to navigate and interpret your page. Following WAI-ARIA best practices improves human accessibility *and* agentic performance simultaneously.

Choose Your Tech Entourage Wisely

Your brand’s agentic future depends heavily on the company you keep. The agentic web is expanding rapidly, and you need to look at the partners that drive your business—your content management system (CMS), payment providers, and CRM—to assess their agentic readiness.

Tech providers moving towards our agentic future have security features, APIs, protocol compliance, and clear documentation showing how agents can use their tools. Platforms like Wix and Stripe are actively building agentic compliance directly into their systems, creating native integrations that allow users to easily tap into agentic shopping on platforms like Perplexity. If your martech stack is aligned with teams actively building for the agentic web, you’ll reap the benefits without having to code the infrastructure from scratch.

Optimize Your Feeds and APIs

Agents value data that is well-structured and offers real-time accuracy. If an agent can’t easily “talk” to your website or inventory, it simply won’t recommend your products.

Audit your shopping feeds to ensure your product data is complete, accurate, and compliant with the specific agentic channels you want to use, like those curated by OpenAI or Google Merchant Center.

Audit the APIs on your websites to see if they are compatible with relevant protocols. Agents will access APIs at scale so they can quickly achieve their goals. But if an agent can't "talk" to your website or inventory, it won't recommend your product or services.

Ask yourself, can AI agents find, understand, and safely use your APIs without confusion.

API goal	Ideal API feature	Why it helps agents
Findability	Lists your API spec in your documentation, or at a clear URL like <code>/openapi.json</code> ; has an <code>llms.txt</code> file pointing to it.	Agents discover which APIs you offer without hunting docs.
Clear instructions	Defines inputs/outputs with examples (success, errors); uses consistent error messages.	Agents know exactly how to call and handle responses.
Smart endpoints	Has actions like "create checkout" instead of tiny steps; allows safe retries.	Agents finish tasks in a few calls, not dozens.
Safety	Uses limited permissions, bot-friendly rate limits, and preview steps for big actions.	Stops mistakes from repeated or bold agent tries.
Protocol match	Checks protocol conformance for shopping (ACP) or tools (MCP).	Improves eligibility to connect directly to agent marketplaces like Google or OpenAI.

Serve Lightweight Markdown Content

Raw HTML is heavy, cluttered, and prone to “context rot” when processed by LLMs. During vectorization, LLMs strip out HTML to assess the context of a page. Providing your business data in markdown formats—like an `llms.txt` file or AI-friendly JSON files (`/.well-known/agent-card.json`)—gives web scrapers the exact context they need without the clutter.

Major players like Anthropic, AWS, and Google have already built `llms.txt` files into their A2A architecture, using them as a machine-readable index of documentation.

ADK Docs Index

The ADK documentation is available as machine-readable files following the [llms.txt standard](#). These files are generated with every documentation update and are always up to date.

File	Description	URL
<code>llms.txt</code>	Documentation index with links	google.github.io/adk-docs/llms.txt
<code>llms-full.txt</code>	Full documentation in a single file	google.github.io/adk-docs/llms-full.txt

By intentionally structuring your data this way, you ensure the AI has exactly what it needs to understand your brand context and execute tasks correctly, significantly reducing the chance of hallucinations.

Configure MCP Servers

Think of the Model Context Protocol (MCP) as a digital USB for the agentic web. Agents rely on interoperability—the ability to use multiple digital systems at once—and MCP allows LLM clients (like Claude or ChatGPT) to securely interact with your local data sources and remote services.

Configuring an MCP server forms the bedrock of your agentic stack. It provides deep grounding for agents by giving them direct access to live data and resources. Whether you build an MCP server from scratch using a Python SDK or use built-in solutions like the Wix Site MCP server, you are giving AI agents a clear, structured way to read and use the specific “tools” (like “check a price” or “book an appointment”) you have configured.

In my team, the Wix MCP is a tool that enables interoperability. By connecting to the MCP, users can generate websites on Wix from ChatGPT, connect to Stripe’s Agentic Commerce Suite, add content to blogs, or products to their e-commerce store from tools like Cursor or Perplexity.

As a brand, offering an MCP can create new channels for customer engagement and turn your content into a tool.

Adopt Protocols Directly

For enterprises operating at global scale, directly adopting agentic commerce protocols isn’t just a shortcut—it’s the new foundation for participating in AI-driven shopping ecosystems. Industry leaders such as Visa and Deloitte are already engaging with these standards, but doing so requires significantly more than simply following a GitHub checklist.

Modern agentic commerce is built on *multi-layers of agentic protocols, designed to give agents the guardrails to act independently*. You may have heard of protocols like:

- Model Context Protocol (MCP)
- Agent-to-agent (A2A) protocol
- Agentic Commerce Protocol (ACP)
- Universal Commerce Protocol (UCP)
- Agent Payments Protocol (AP2)
- NL Web

These frameworks *layer together* to support discovery, negotiation, checkout, payments, and fulfillment across AI-driven agents.

When an enterprise “adopts a protocol,” what they are really doing is implementing a set of *rigorous, interoperable requirements* covering:

- Real-time inventory exposure
- Standardized product discovery endpoints
- Machine-readable business rules
- Identity and account-linking mechanisms
- Provable, secure payment flows
- Order management and fulfillment surfaces

Adopting these protocols directly requires resource, but once a business meets the required specifications and proves protocol compliance, it becomes eligible for high-value agentic channels—such as *AI-native Instant Checkout* in ChatGPT (ACP/AP2) or *AI Mode* in Google Search (UCP). Instead of optimizing for AI search clicks, businesses begin optimizing to gain customers with agentic *interoperability*, ensuring that AI systems can discover, understand, and transact with them in real time.

For smaller businesses, utilize your vendor’s agentic compliance to ensure that you can make use of all the benefits of the agentic age.

Build Your Own Agents

You wouldn’t try to win a game without understanding the rules, and the best way to understand how agents work is to build one yourself. Marketers who actively test and deploy agents will be miles ahead of the competition because they gain firsthand experience with the technical dependencies of these systems.

You don't have to be a developer to start. Leverage the agent-creation features *within* your existing tools—like configuring custom agents in the backend of a Wix website or using portals in Salesforce. This gives you a sandbox to see exactly how agents respond to prompts, leverage multiple tools, and validate steps. Taking this final, hands-on step ensures you aren't just reacting to the agentic future, but actively shaping it.

The Personalization Playbook

Agentic AI doesn't just want the technically correct answer; it wants the highly personalized answer. Tools like Gemini and ChatGPT have memory and profile personalization built right into their interfaces.

If a user tells their AI travel agent they have a family of four, the agent will pair technical accessibility with affinity requirements. It will only look for websites that allow it to complete the booking *and* meet the family-friendly criteria. Brands that try to be everything to everyone will lose out to brands that clearly declare what they are best for. You must explicitly state your niche—whether through detailed FAQs, clear on-page copy, or targeted video content—because agentic AI makes affinity optimizations just as important as intent optimizations.

Your Agent-Ready Action Plan

The agentic web isn't a far-off vision; it's a strategic reality. To ensure you aren't just reacting to this shift, but actively shaping it, start with these steps:

- **Add conversion points to your website.** Shift from being a brochure to a conversion engine by ensuring your forms, booking tools, and ARIA tags are properly labeled and easily accessible for agentic browsers.

- **Choose your partners wisely.** Ask your CMS, payment providers, and CRM about their roadmap for agentic readiness so you can leverage their native integrations instead of building from scratch.
- **Optimize your feeds and APIs.** Ensure your product data is clearly structured and your APIs use “smart end-points” that allow agents to complete tasks efficiently in just a few calls.
- **Create and optimize related markdown content.** Draft an `llms.txt` or `agent-card.json` file to give web scrapers the exact, clutter-free context they need to understand your brand and avoid hallucinations.
- **Configure MCP servers and related tools.** Give LLM clients a structured, secure way to read your live data and use the specific digital tools you have connected to your business.
- **Adopt protocols directly.** For enterprises, directly adopt frameworks like UCP and ACP to unlock new channels like Instant Checkout; smaller businesses should actively lean on their vendor’s built-in compliance.
- **Build your own agents.** Use the native agent-creation features within your existing tools (like Wix or Salesforce) to get hands-on experience with how these systems respond and navigate.
- **Declare your affinity.** Explicitly state who your product is best for across your content so highly personalized AI agents can confidently match you with their users.

13

Measure with the Right Stick

As the old marketing adage goes: if it ain't measured, it doesn't matter. This is true for classic search, and it is absolutely true for AI search.

To make a long story short, we're going to need some better metrics. As marketers, we need to report to stakeholders, pitch for resources, and demonstrate results. This is impossible without KPIs. But this sudden shift in our digital world has dragged KPIs with it, meaning a new set of measurables is required.

For businesses, the essentials stay the same—revenue, sales, and the bottom line are still the core metrics anchoring everything else. But AI assistants have a unique user experience, meaning the path to those conversions looks completely different. Generative search visibility isn't just an SEO play; it's a confluence of PR, social media, technical web strategy, and UGC.

Translating the Metrics: Legacy SEO Versus AI Search

If you are presenting to a board that is used to traditional SEO reports, you need to translate your new AI metrics into a language they already understand. The table below acts as your cheat sheet. While these metrics might look like distant cousins at first glance, the mechanics of how you actually track and report on them are fundamentally different.

New AI search metric	Classic metric	Bridging the gap
Adoption market share	Total addressable market	Align your growth expectations with how much of your audience actually uses AI for discovery.
Brand mentions	Share of voice	Benchmark how often your brand surfaces in relevant, non-branded AI prompts.
Domain citations	Indexed pages	Identify your most valuable pages and ensure they are retrieved by LLMs.
Third-party citations	Backlinks	Audit external links to see if they come from domains included in LLM training data.
“Rank”	Keyword rankings	Measure the placement order of your brand name in an LLM response.

New AI		
search metric	Classic metric	Bridging the gap
Sentiment	PR sentiment	Track if your brand is discussed positively, negatively, or neutrally.
AI referrals	Organic traffic	Set up GA4 explorations to isolate traffic coming specifically from AI bots.
AI bot traffic	Search bot traffic	Monitor server logs to see how LLM crawlers behave compared to traditional Googlebots.
AI conversions	Conversions	Update intake forms to explicitly track leads that originated from AI platforms.

Let's explore why each of these KPIs matters for your AI search visibility.

Adoption Market Share Versus Total Serviceable Market

For LLM visibility, your audience's actual AI adoption rate dictates your true growth potential.

Think about the difference between someone searching for "programming" versus someone searching for "potatoes." Users searching for programming are 112% more likely to use ChatGPT as a discovery channel. But if you look at queries for potatoes? ChatGPT isn't even in the top seven discovery channels. Knowing your adoption market share tells you exactly how much and how fast you should invest in GEO. The rate at which your audience adopts AI will have a massive impact on your long-term reach, so it's absolutely worth paying attention to right now.

Within this metric, you must also track relative improvement. If your nearest competitor gets 1,000 referrals from ChatGPT a month and you are getting 1,500, that is a massive win. Even if 1,500 feels low compared to traditional organic traffic, proving that you own the lion's share of AI traffic demonstrates the immediate value of your efforts.

Assessing Prompt Presence

When you monitor your presence in LLMs, you are essentially assessing your prompt presence. In prompt tracking, there are four core metrics within a single output that you must understand:

- **Mentions.** When your name appears.
- **Citations.** When your website gets a link.
- **Rank.** Your mention placement order.
- **Sentiment.** The tone of the response.

Let's break down how to utilize each of these, along with the other crucial KPIs.

Brand Mentions: The Evolution of Share of Voice Share of voice is getting a major makeover. It's no longer just about how many non-branded keywords you rank for on page one; it's about how often your brand name actually appears in relevant LLM prompts compared to your competitors.

Why is there suddenly so much emphasis on mentions? Because within AI search, a mention acts like a highly qualified recommendation rather than just a blue link. People who use AI assistants regularly learn to trust them. A 2025 KPMG study showed that 65% of people have faith in the technical ability of AI systems to provide accurate outputs. When users treat these tools as coaches and companions, an AI's recommendation carries a similar weight to a human referral. This explains why

ChatGPT conversion rates are significantly higher than classic search engines.

Keep in mind, it is entirely possible to be mentioned without receiving a direct website citation, particularly when the AI is relying on its underlying training data or pulling in consensus reviews. For instance, if a user asks, “What is the best CRM software?” the LLM might highly recommend Monday.com in the text, but the actual citation link might point to a third-party review site like G2 or a forum like Quora. A mention is still a massive win for brand awareness, even if the direct click goes elsewhere.

To measure this effectively, you have to establish a baseline. If you run a local pottery studio, you need to know exactly what percentage of ChatGPT responses actually mention your business when someone asks for “best date night activities near me.” Whether you use dedicated AI search monitoring tools or a good old-fashioned manual tracking sheet, it is about setting a benchmark to monitor your growth in generative search, rather than just obsessing over SERP positions.

When you set up prompt trackers, they will typically show you the percentage and frequency of times you are mentioned across your tracked queries. If you log in and see that you are showing up 100% of the time, don’t pop the champagne just yet. It probably means you are only tracking branded prompts. If you tracked a prompt for vibe coding platform Base44 asking “Is Base44 a good platform?” it doesn’t count as a true GEO win for mentions (though it could help understand sentiment), because the query itself presupposes your brand.

To accurately assess whether your mentions are actually driving growth, you absolutely must filter out branded queries and establish your baseline using non-branded, discovery-based prompts.

Domain Citations Are the New “Plain Blue Links” In classic SEO, you prioritize making sure pages are crawled and indexed before they can rank. For LLMs, you need to understand which pieces of content are actually retrieved and used for responses. Domain citations are the new “plain blue links.”

If you are tracking prompts, filter out the pages from your domain to find out:

- Which pages show up in AI responses
- How frequently they appear
- Which parts of your site show up most regularly

Pages that show up frequently usually have highly citable attributes—they might be listicles or cover high-reach topics. From here, you can enhance the “rankability” of these pages or improve conversions with CRO and adding internal links.

You will often see that LLMs favor specific sections of your site. ChatGPT might love your blog, while Grok loves your support pages. Support, help, and documentation pages are prime candidates for LLM ingestion because they are text-heavy, packed with FAQs, and require very low design lift to update.

Third-Party Citations Are the New Backlinks Is there an overlap between website citations in LLMs and tracking backlinks in classic SEO? Absolutely. In both cases, you are measuring whether the wider web considers you a relevant, authoritative source of information. But in the generative era, not all links are created equal. Here’s how to audit your citations for maximum GEO impact:

■ **Audit your acquisition partners**

Have a look through your AI citations to see if you recognize any familiar faces. Are you pouring budget into

acquisition channels that yield absolutely zero visibility in LLMs? Or, conversely, are some of your long-standing affiliate partners quietly generating a massive amount of AI citations? Reviewing these links helps you understand exactly where to double down and where to pull back your budget and effort.

■ **Re-evaluate your social strategy**

Social media is no longer just an engagement play; it's a vital citation engine. If your goal is GEO, you need to review which platforms are actually fueling your AI mentions. For instance, if you notice you're getting significantly more citations from Quora than from Reddit, or that YouTube is feeding the models far better than Instagram, pivot your resources. These insights take the guesswork out of your content strategy, making your team's efforts more efficient and your results more effective.

■ **Target LLM-ready PR partners**

You need to tightly align your digital PR strategy with your backlink activity. Start by auditing your current profile to answer three core questions:

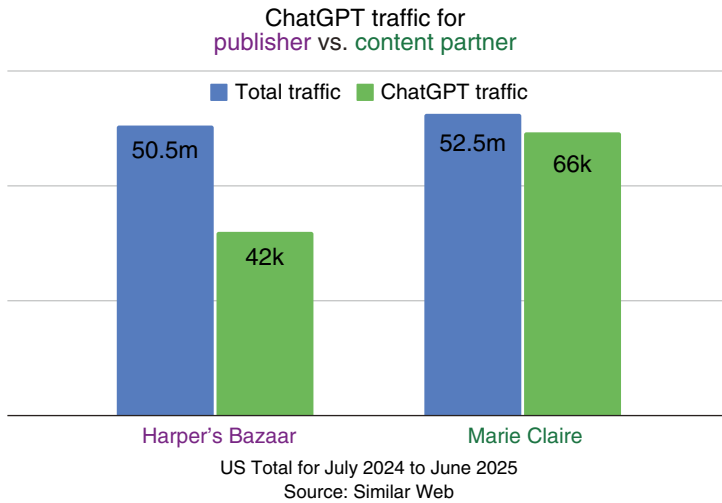
1. What percentage of your backlinks come from web sources known to be included in LLM training data?
2. How frequently do those backlink referrals actually appear in LLM responses?
3. Is the context of those mentions accurately reflected in the AI's output?

Major platforms have direct content and revenue-sharing partnerships with massive publishers. Perplexity, for instance, partners with ADWEEK, *Time*, *Der Spiegel*, and the *Los Angeles Times*. OpenAI has inked deals with some of the biggest media conglomerates in the world. Getting mentions in these specific publications feeds your brand directly into the AI's core memory.

Notable OpenAI media partnerships

Media company	Notable publications and websites
Axel Springer	<i>Bild</i>
Associated Press	AP News
Atlantic Media	<i>The Atlantic</i>
Vox Media	Vox
News Corp	<i>The Wall Street Journal</i> , MarketWatch, <i>The Times</i>
Dotdash Meredith	Travel + Leisure, <i>PEOPLE</i> Magazine, <i>Food & Wine</i> , <i>Health</i>
Financial Times	<i>Financial Times</i>
Le Monde	<i>Le Monde</i>
Prisa Media	<i>El País</i> , <i>Cinco Días</i> , <i>As</i> , <i>El Huffpost</i>
TIME	<i>TIME</i> Magazine
Condé Nast	<i>Vogue</i> , <i>The New Yorker</i> , <i>GQ</i> , <i>Wired</i> , <i>Vanity Fair</i>
Hearst	<i>Houston Chronicle</i> , <i>ELLE</i> , <i>San Francisco Chronicle</i> , <i>Runner's World</i>
Future	<i>Marie Claire</i> , <i>TechRadar</i> , <i>Cycling Weekly</i>
Axios	<i>Axios</i>
Schibsted Media Group	Various Nordic publications

These publishers are seeing a massive impact from these deals. For example, if we compare women's fashion magazines over a 12-month period, *Harper's Bazaar* and *Marie Claire* bring in a similar rate of traditional traffic. But the story changes entirely when we look at ChatGPT. Because *Marie Claire* (owned by Future) is an official content partner, they receive 44% more referral visits from ChatGPT than *Harper's Bazaar*.



The takeaway for marketing leaders is clear: if you are building brand authority through digital PR, prioritize coverage and links in these partnered publications to maximize your generative reach.

14

“Rank” in AI Responses

Ranking is a core tenet of classic SEO, but in GEO, ranking is a little different. Let’s break it down.

If a tracking tool tells you that you rank number three on a particular LLM response, what they are actually saying is that you are the third-mentioned brand in that specific output. Many tools will aggregate your position placements to give you an overall rank across a bank of queries.

But here is the challenge with that ranking system:

- The reasons why LLMs place brands in a specific order depend heavily on the query itself. For instance, I tested a query asking, “Can you list some significant country pop albums by women?” The list that came back was strictly chronological. Shania Twain was listed first, and Beyoncé’s *Cowboy Carter* was last. The response order was a reflection of chronology, not necessarily of quality or authority.
- Not every query triggers a mention (of anyone). In classic search, there is always something to rank. Search engines will always show a list of websites unless you somehow

trigger a Googlehack. In LLMs, this is not the case. You don't necessarily get a brand mention for every single prompt. If your prompt isn't the kind that triggers a brand mention, you won't be seen, nor will you be ranked—because no one is seen or ranked.

These limits also change depending on the platform. Tools like Perplexity, OpenAI, and Gemini, which have more commercial capabilities, are far more likely to mention brands than base conversational models. Specifically, an AI-forward Google that is directly connected to features like the shopping panel is highly likely to include brand mentions. Because of this, your mentions and your rank will undoubtedly vary by platform.

That said, being first is still ideal, simply for the psychology of it. Users are much more likely to read the top of the response than they are to scroll to the bottom of a lengthy reply. No matter how many times you explain to people that a generated list is not necessarily in a particular order, whoever is mentioned first is psychologically perceived as being of higher value.

Sentiment

How a brand is depicted online isn't often measured in traditional SEO, but it has a massive impact on the value of LLM mentions. Since LLMs use vector embeddings of words and phrases to identify patterns, brands with reputational challenges may struggle to see a fair representation of their products. In this situation, the goal of the optimizations may be to improve the quality of the response rather than the frequency.

Typical sentiment challenges include:

- Unfair product critiques
- The predominance of opinions from a single review
- Inaccurate information

When you see poor sentiment, take action. Choice words don’t materialize from nowhere. Review the citations triggering the poor sentiment. If the critique is fair, update your business practices and tell the reviewer. If a publication is spreading false information, offer them an update. If a single bad review is dominating the model, run a campaign to foster a more balanced review profile.

The RAPP Framework: Quality over Keywords

To easily analyze the value of your prompt presence, use the RAPP framework. Ensure your mentions are Regular, Accurate, Prominent, and Positive.

RAPP Framework for AI Citations		
Pillar	Core Question	Key Focus Areas
Regular	Does the brand name and website appear regularly for relevant LLM responses?	Frequency of brand and website mentions across AI outputs.
Accurate	Does accurate & up to date information about your brand show in LLM responses?	Correctness, currency, and reliability of the data retrieved.
Prominent	Does your brand or website show first on LLM lists alongside other brands?	Search visibility, ranking position, and competitive placement.
Positive	Does the mention show positive or fair sentiment towards your brand?	Sentiment analysis, tone of voice, and brand reputation in AI generation.

Bot Logs: The Ultimate GEO Goldmine

Your server bot logs can tell you an incredible amount about your LLM visibility. Just like you would analyze traditional user agents for SEO to see if Googlebot is encountering errors, your bot logs show you exactly how AI crawlers are behaving on your site behind the scenes.

Use your logs to get granular. Compare how the OpenAI Searchbot behaves versus the OpenAI Userbot, or look at the specific pages Perplexity accesses compared to traditional Googlebots. Monitoring this gives you a front-row seat to exactly what the AI cares about, allowing you to take immediate, data-backed action.

For instance, mining your bot logs empowers your team to make strategic GEO decisions, such as:

- **Prioritizing high-traffic real estate.** If you notice your homepage is visited by AI bots three times more often than your next most frequented page, you know exactly where to prioritize your content updates. Make sure the absolute most important details about your core business entities live on that page so the models digest them first.
- **Catching AI hallucinations.** If an LLM is inventing fake URLs for your site in its outputs, you will see those hallucinated links show up as 404 errors in your logs. Spotting these early lets you set up redirects and regain control of the user journey.
- **Pruning low-value retrieval.** You might see that LLMs are heavily crawling and retrieving outdated or low-value pages. Knowing this gives you the power to update your `robots.txt` file and block those crawlers, ensuring the AI isn't pulling the wrong data to train its answers.

Make no mistake—this log data is some of the absolute most valuable, unfiltered information your GEO team has at its disposal. It takes the guesswork out of technical optimization and shows you exactly how the machines view your brand.

Traffic

Traffic from LLMs is likely to be lower than traffic from classic search in the immediate future. However, what you should focus on is making sure that you are moving in the correct trajectory, that you are competitive within your niche, and that this high-intent traffic is landing on the right pages.

You can track traffic with the same site-analytics tools that you always have, but make sure you’re not only looking at organic. LLM traffic can show up in referrals, but some tools also have dedicated AI channels.

Sealing the Deal: Conversions and Leads from LLMs

Traffic is great, but stakeholders want revenue. You are already tracking goal completions and e-commerce transactions, but the challenge with generative search is a lack of established attribution tools.

The simplest, most effective hack for this? Update your “How did you hear about us?” forms. Seer Interactive is an AI search agency that has used this tactic with clients for years. Seer Interactive founder, Wil Reynolds, explains: “I was seeing traffic from ChatGPT and Perplexity growing, but I wanted to be able to track if that traffic was more likely to turn into qualified leads than search traffic” he says. “So, I had to add that in the drop-down to be able to track it through to closed deal.”

Looking Ahead Towards Paid AI Visibility

Right now, we are largely relying on “mirror metrics” from third-party tools, but the floodgates are about to open. OpenAI officially began testing ads within ChatGPT in early 2026, Microsoft has increased ad presence in Copilot, and Google is rolling out AI-forward ads. As these paid features scale, expect a flurry of new native metrics and dashboards directly from the platforms—because once they ask for your ad dollars, it behooves them to prove their worth.

As these features roll out, CMOs face a classic dilemma: How much do we spend on renting space versus owning our authority? Think of paid AI placements—like ChatGPT ads or sponsored spots in an AI-forward Google—like renting a luxury Airbnb for the weekend. It gives you immediate access to the best views and puts you right where the action is, but the minute you stop paying, you are locked out. It absolutely does not replace the need to build your own digital real estate.

To balance your efforts, use a two-pronged approach:

- **Paid AI for the “bottom of the funnel” sprint.** Use your ad dollars to target high-intent, transactional moments, like ChatGPT Shopping queries. Hold these campaigns to strict, traditional metrics like direct conversions and return on ad spend.
- **Organic GEO for the “base model” marathon.** You can’t buy your way into an LLM’s foundational training data. Organic GEO is about feeding the machine high-quality, verifiable information so that your brand becomes the default, trusted answer when someone asks a broader research question. Measure this using the RAPP framework.

Throughout my career, I’ve always kept a close eye on the paid space because platforms very often give advertisers far more robust data about how their content performs compared to organic users. Taking this into account, I highly recommend that any team committed to GEO signs up as an advertiser on these tools. Even if you only run campaigns on a minimum budget, it should give you unique, exclusive insights from their native dashboards that will undoubtedly inform your overall organic strategy.

Budgeting for the Blended Search Era

For the next 12–18 months, treat your paid AI budget as an experimental sandbox. As an AI-forward Google continues to push traditional text ads further down the page, your organic GEO strategy is your best insurance policy. As your suite of available data grows, don’t fall into the shiny object trap and get lost in the sauce. Keep your eyes on the prize, and always tie every new metric strictly to your core business objectives.

The AI Search Executive Checklist

If you are walking into a board meeting tomorrow, these are the three core AI search KPIs you need to present:

- 1. Adoption market share.** Are our customers actually using AI to find our specific services, and are we meeting them there?
- 2. RAPP mentions.** Is our brand appearing Regularly, Accurately, Prominently, and Positively in LLM outputs?
- 3. AI-influenced pipeline.** How many closed deals can we directly attribute back to Perplexity, ChatGPT, or Gemini referrals?

15

Track Your Performance: Making Prompt Tracking Work on Any Budget

Here's the thing about tracking your brand's visibility in AI: you don't need a massive budget or a fancy enterprise platform to do it well. You just need to be *smart* about what you're tracking—and more importantly, what you're *not* tracking.

Think of prompt tracking like setting up a home security system. You could install cameras on every wall, monitor every window, and track every shadow that crosses your lawn. Or you could focus on the doors people actually use, the valuables you actually care about, and the break-in patterns that matter in your neighborhood. Same principle here: efficient

tracking means qualifying your audience, solidifying your brand topics, understanding your competition, and ruthlessly refining your prompts.

Let's break down how to build a lean, actionable tracking system that doesn't waste your time—or your budget.

Qualify Your Audience First (Or You'll Track Everything and Learn Nothing)

Before you track a single prompt, you need to know *who* you're tracking for. Your AI-first customers aren't everyone—they're a specific slice of your audience who prefer ChatGPT over Google, who ask Perplexity for product recommendations, who browse with Gemini open in another tab.

Start with web analytics to confirm platform preferences. Use GA4 (or your analytics platform of choice) to figure out which AI platforms your users actually prefer and which content is getting AI-driven traffic. You might discover that 38% of your AI engagement happens in ChatGPT and 52% in Gemini—that's not trivia, that's your resource-allocation map. Don't waste time tracking Claude responses if your audience isn't using Claude.

Filter by region and identify topic trends. Your AI-first niches might be regional. Maybe your UK customers use AI search for product specs, while your US customers ask for comparisons. By filtering user data by geography and reviewing which topics trend where, you can focus your tracking efforts on the models *and* content that matter most to your highest-value AI-first customers.

The outcome? You're not tracking everything—you're tracking the *right* things. You'll focus your efforts, identify AI-first niches, confirm which models are most important, and understand which content your AI-first customers actually prefer.

Prepare Your Brand Topics (Know What You're Called Before You Track How Often)

You can't track your brand accurately if you don't know what counts as a mention. This is where a lot of marketers stumble—they assume “our brand” just means the company name, when in reality it includes product names, acronyms, regional variations, and half a dozen other identifiers.

Identify your branded terms using keyword research tools. Use any SEO or keyword research platform to analyze which branded terms are actually associated with your domain. For Toyota, that's not just “Toyota”—it's “Toyota Highlander,” “RAV4,” “Camry,” “Tacoma,” and “Corolla.” Those are the terms you need to track, because that's how real people (and real LLMs) refer to your products.

Specify regional keywords and confirm theme consistency. Your brand's key terms can vary by geography. Toyota.com in the United States surfaces car model names (Highlander, RAV4, Camry), while Toyota.com in the United Kingdom surfaces broader brand terms like “Toyota vehicle models” and “Toyota vehicle features.” Run keyword analysis for each major market to confirm what people—and AI systems—actually use when they talk about your brand. That's your baseline for tracking.

This step isn't glamorous, but it's foundational. If you don't know what counts as a brand mention, your visibility scores will be meaningless.

Identify and Track Your Real Competitors (Not Just the Ones You *Think* Matter)

In traditional SEO, you pick your competitors based on market share or product overlap. In AI search, your competitors are whoever the Knowledge Graph *thinks* you're similar to—and that's not always the same list.

Discover entity competitors from search engines. Your entity competitors are the brands that align with what your brand *means* in the broader web ecosystem. The easiest way to find them? Search for your brand on Google and look at the “People also searched for” section. For Toyota, that might be Honda, Nissan, and Mazda—these are the brands search engines (and by extension, most LLMs) consider related. These are the competitors you should be tracking, because they’re the ones you’re *actually* competing against in AI-generated responses.

Track entity competitors, not just market competitors. Your biggest market competitor might not show up in AI answers the same way you do. Meanwhile, a smaller competitor with better-structured content might be stealing your AI visibility. By tracking entity competitors, you get better insight into the *real* competitive landscape as it exists inside generative engines.

Refine Your Prompts for Valuable Output (What You Measure Defines What You Track)

This is where efficiency lives or dies: choosing the *right* prompts to track. Because here’s the hard truth—branded prompts will always show your brand, and competitor-branded prompts never will. If you track both, you’re skewing your data and wasting your budget.

Confirm your tracking approach’s parameters. Before you add prompts, understand what you’re actually measuring. Which models are you querying? How frequently will you collect data? What terms will you use to identify mentions? These parameters determine what kinds of prompts will give you valuable output. If you’re only checking responses once a week, don’t load up your list with time-sensitive prompts.

Prune branded prompts ruthlessly. Branded prompts—like “Do Toyotas have good resale value?”—will *always* trigger your brand in the response. That’s not insight, that’s confirmation bias.

Similarly, competitor-branded prompts like “What are the best Kia models for road trips?” are unlikely to ever mention your brand. Both types should be pruned.

Focus on non-branded comparisons and general queries. The most valuable prompts are non-branded comparisons (e.g., “How does the comfort of Hyundai cars compare to other cars for road trips?”) and non-branded, on-topic general queries (e.g., “What hybrid cars are better for long-distance driving?”). These are the battlegrounds where your brand’s true AI visibility is tested. If you’re showing up here, you’re winning. If you’re not, you know where to improve.

What to prune versus what to keep

Prompt type	Example	Track it?
Branded	“Do Toyotas have good resale value for families?”	✗ Prune (always shows your brand)
Competitor-branded	“What are the best Kia models for long road trips?”	✗ Prune (unlikely to show your brand)
Non-branded comparison	“How does the comfort of Hyundai cars compare to other cars for road trips?”	✓ Keep (tests your competitive visibility)
General, on-topic	“What hybrid cars are better for long-distance driving?”	✓ Keep (tests your category visibility)
General, off-topic	“How important is trunk space for road trips?”	⚠ Maybe (depends on how closely it relates to your products)

This pruning process is the difference between a bloated, expensive tracking setup and a lean, actionable one.

Test, Iterate, Grow (This Isn't Set-It-and-Forget-It)

Prompt tracking is a living, breathing function. The AI landscape shifts, models get updated, competitors launch new campaigns, and your own content evolves. Efficient tracking means constant testing, iteration, and refinement.

Group and test topics like keyword clusters. Start with high-priority topics, monitor performance, then iterate. Are you struggling with “Best” lists? Could you do better on review-style queries? Test prompts by type and see where the gaps are.

Align prompts with real customer queries. Your most valuable prompts are the ones your customers are actually asking. Pull queries from sales calls, CRM data, customer service logs, product reviews, Reddit threads, and social media. Organize them, clean them up (AI is great at this), and feed them back into your tracking. This ensures you're monitoring what *actually* matters to your audience, not just what you *think* matters.

Create and update content based on low-performing topics. If you're tracking non-branded prompts and consistently missing mentions in certain topic areas, that's your content roadmap. Improve your content and presence for those low-performing topics, then monitor growth over time.

The Payoff: What Efficient Tracking Gets You

When you follow these principles—qualify your audience, prepare your brand topics, identify entity competitors, and refine your prompts—you're not just tracking AI visibility. You're building a feedback loop that drives real outcomes:

- A more accurate visibility score (because you're not tracking junk prompts)

- More actionable data (because every prompt tells you something you can act on)
- Better competitor insight (because you're tracking entity competitors, not just market rivals)
- Greater AI visibility knowledge (because you understand *why* you show up where you do)

And here's the best part: you can do all of this on a budget. You don't need the most expensive tool or the biggest team. You just need to be disciplined about what you track, strategic about which prompts you monitor, and relentless about refining your approach as the landscape shifts.

Because in AI search, like in everything else, it's not about having the most data—it's about having the *right* data.

Your Prompt Tracking Workflow: From Seed Phrases to Actionable Data

Now that you know *what* to track, let's talk about *how* to build your tracking system from the ground up. This workflow takes you from raw ideas to refined prompts to measurable performance data—and it's designed to work whether you're using a dedicated tracking tool or just a well-organized Google Sheet.

Step 1: Source Potential Prompt Seeds

Your best prompts don't come from thin air—they come from the places your customers are already asking questions. Think of this as prompt archaeology: you're digging up the real language people use when they talk about your category.

Where to look for prompt seeds:

- **Long-tail keywords.** Pull these from your existing keyword research. The three to five-word phrases that feel too specific for traditional SEO? Perfect for AI prompts.

- **Social media content.** What questions do people ask in comments? What topics generate the most engagement? These are your signal fires.
- **Genuine customer FAQs.** Sales calls, customer service emails, live-chat transcripts. This is gold. It's literally what people ask.
- **Reddit threads.** Find the subreddits where your audience hangs out. Look for question posts with lots of upvotes and comments. That's where the pain points live.
- **"People Also Ask" boxes.** Google's PAA sections are a free insight machine. They show you the question chains people follow when researching your topic.

The key here is *authenticity*. You're not trying to guess what people might ask—you're documenting what they *do* ask. Keep a running list. Aim for 50–100 seed phrases before moving to the next step.

Step 2: Convert Seed Phrases into LLM Prompts with AI

Here's where it gets fun: you're going to use AI to generate AI prompts. It's like training for the Olympics by competing against yourself—effective and slightly meta.

The prompt template that works:

■ "Using the copy below, generate 15 questions that a [target customer] is likely to ask an LLM about [topic]."

Paste in your seed phrases, specify your target customer (e.g., "mid-market CMO," "first-time homebuyer," "SaaS product manager"), and let the AI expand your list. You'll get variations you hadn't considered, phrasing that feels more natural, and question angles you missed.

Pro tip: Run this prompt multiple times with different seed-phrase combinations. ChatGPT, Claude, and Gemini will each give you slightly different outputs—and that diversity is valuable. You want to capture the range of ways people might ask about your topic.

Step 3: Add Prompts to Your Tracker (or Test in Google Sheets First)

If you’re using a dedicated prompt-tracking tool, this is where you upload your refined list and configure your settings (models to test, competitor brands, tracking frequency).

But if you’re working on a tight budget or just getting started? Google Sheets is your friend. Set up a simple testing framework, as follows.

Prompt	ChatGPT result	Gemini result	Perplexity result	Brand mentioned?	Position in response
“What hybrid cars are better for long-distance driving?”	[paste response]	[paste response]	[paste response]	Yes/no	Top/middle/bottom

This manual approach is tedious, but it’s also incredibly instructive. You’ll start to notice patterns—which prompts trigger your brand, which competitors dominate certain question types, which models favor different content structures. That knowledge is worth the effort.

Once you’ve manually tested 20–30 prompts and understand the patterns, *then* consider investing in an automated tracking tool. You’ll know exactly what you need it to do.

Step 4: Carry Out GEO Optimization Tactics Around Your Topic

Now comes the content work. Your tracking data will reveal gaps—topics where your brand *should* show up but doesn’t. This is your GEO roadmap.

Focus your optimization efforts on:

- **High-value prompts where you’re absent.** If “best family SUVs for road trips” consistently shows competitors but not you, that’s a content opportunity.
- **Prompts where you’re mentioned late in the response.** If you show up in position #8 out of 10, you have visibility but not prominence. Strengthen your topical authority.
- **Non-branded comparisons.** These are the battlegrounds. Create comparison content, feature-focused guides, and use-case breakdowns that directly address these prompts.

Apply the GEO tactics we covered in earlier chapters: structured data, authoritative citations, clear entity definitions, FAQ content, and E–E–A–T signals. But do it *strategically*—one topic cluster at a time, aligned with your prompt-tracking priorities.

Step 5: Monitor Performance of Specific Prompts Related to Your GEO Activity

Here’s where the loop closes: you optimize content around specific topics, then track whether your visibility improves for the prompts related to those topics.

Set up your monitoring cadence:

- Weekly checks for high-priority prompts (the ones tied to active GEO campaigns)
- Monthly reviews for your full prompt list
- Quarterly audits to prune underperforming prompts and add new ones based on emerging customer queries

Track these key metrics:

- **Mention rate.** What percentage of tracked prompts mention your brand?
- **Average position.** When you're mentioned, how early in the response?
- **Competitive displacement.** Are you appearing *instead of* competitors, or just alongside them?
- **Topic-specific trends.** Are you gaining ground in "best" lists but losing in "reviews"?

The goal isn't just to see your numbers go up—it's to understand *why* they're moving and what levers you can pull to accelerate growth.

The Workflow in Action: A Real Example

Let's say you're a SaaS company selling project management software. Here's how this workflow might look:

Step 1. Source seeds

- Long-tail keyword: "project management for remote teams"
- Reddit thread: "What's the best alternative to [competitor]?"
- Customer FAQ: "Can your tool integrate with Slack?"
- People Also Ask: "What project management features do small teams need?"

Step 2. Convert to prompts

You use AI to generate a list of prompts: “What project management tools work best for fully remote teams?” “How do I choose project management software for a team of 10?” “What are the essential features in a project management tool for startups?”

Step 3. Test manually in Sheets

You run these through ChatGPT, Gemini, and Perplexity. You discover you’re mentioned in 2 out of 10 prompts, always in the bottom half of responses.

Step 4. Optimize

You create a comparison guide (“Top Project Management Tools for Remote Teams: An Honest Comparison”) with structured data, clear feature tables, and integration details. You also add an FAQ page addressing the specific questions from your prompt list.

Step 5. Monitor

Over the next month, you track the same 10 prompts. Your mention rate climbs to 6 out of 10. Your average position moves from bottom-third to middle-third. You’ve gained ground.

That’s the workflow in practice. It’s not magic—it’s systematic, it’s measurable, and it works.

16

Tracking LLM Traffic: Finding Your Brand in the AI Sandbox

The digital landscape has shifted under our feet. Users are increasingly turning to LLMs like ChatGPT, Gemini, and Perplexity—not just for writing emails, but as sophisticated engines to find information, products, and services. For CMOs and marketing leaders, this introduces a critical new traffic source. If LLMs are the new *Gossip Girl* of the internet, you need to know if your brand is the one they’re talking about. Tracking this traffic requires specialized monitoring to understand your brand’s visibility and growth potential. Here is how you can set up your tracking to see exactly who is arriving from the AI sandbox and what they are looking for.

Catching the AI Referrers

GA4 is your first line of defense in identifying users arriving directly from AI platforms. Because AI traffic often appears fragmented, it can feel like trying to keep up with a dozen different group chats. You must create custom filters to see the full picture.

- **Isolate growth potential.** Use traffic reports to estimate how much of your audience is migrating to AI platforms over traditional search.
- **Identify valuable channels.** Tracking referrers helps you understand which specific AI bots are your most valuable referral sources.
- **The master regex.** In the User Acquisition report, create a filter for “Session source/medium” using a “matches partial regex” match type. Use a comprehensive string to capture all major players: `openai|chatgpt|copilot|gemini|gpt|perplexity|bard|deepseek|claude|meta\.ai`.

Going Deep with Native Dashboards

Modern platforms like Wix Studio now offer native AI bot traffic dashboards that go deeper than standard referral headers by tracking server-side bot activity. This shows you exactly how AI engines are interacting with your content.

- **Distinguish “crawl” versus “query.”** These reports track how often bots crawl your site specifically in response to live user queries (like OpenAI’s OAI-SearchBot) versus general training crawls (like GPTBot).
- **Visits by page.** Identify exactly which of your pages are being cited or “read” by bots like OpenAI, Meta AI, and Amazon AI.

- **Zero-click insight.** Tools like the “User queries on AI” widget show you how often your brand is mentioned in AI interfaces, even if the user never clicks through.

Prompt Tracking Tools: Measuring Your Share of AI

While website analytics tell you who arrived, prompt tracking tells you who *almost* arrived, or how the AI is describing you. Think of it as auditing the AI’s perspective on your market—like setting up Google Alerts, but instead of tracking the entire web, you are essentially eavesdropping on the AI’s private group chat.

What Is Prompt Tracking?

Prompt tracking is the practice of using specialized SaaS tools to systematically run mass queries across multiple LLMs like ChatGPT, Gemini, Perplexity, and Claude. Because AI models rank content differently based on their unique underlying indexes and evolving re-ranking criteria, a true visibility score is achieved by running the same prompts across multiple models and calculating the metrics individually. For marketing leaders, this systematic prompting is the foundation for defining new AI search KPIs and proving to stakeholders that your GEO efforts are actively moving the needle.

LLM Tracking Tool Features

A new generation of tools—such as Limy, Peec AI, Profound, and the Semrush AI Visibility Toolkit—is emerging to help marketers bridge the gap between traditional SEO and agentic search. When assessing these tools for your tech stack, here are the core capabilities they typically include:

Feature category	How it works	Strategic value for brands
Visibility and share of AI	Tracks your brand mentions, citations, and positioning for both branded and non-branded prompts over time.	Proves the ROI of your GEO efforts by showing exactly how often you appear as the recommended solution.
Source and citation analysis	Identifies the specific external domains and websites the AI relies on when formulating answers about your industry.	Pinpoints the exact third-party publications and PR targets you need to secure to influence the AI’s “brain.”
Sentiment analysis	Evaluates whether the AI’s generated answers frame your brand in a positive, neutral, or negative tone.	Protects brand reputation in zero-click environments where the AI acts as an active decision-maker for the user.
Entity and topic mapping	Analyzes the long-tail keywords, related entities, and referenced questions the LLM associates with your domain.	Uncovers content gaps on your own site that are preventing you from being cited as an authoritative, top-tier source.

Third-Party Competitor Tools: Benchmarking Your Share of Voice

To win in the AI era, you need to know how you stack up against the competition. Many established SEO tools have started rolling out features that give you visibility into third-party AI traffic, which can help you set a baseline benchmark and establish your Share of Voice.

But there is a catch: accuracy. Different toolsets have entirely different configurations. Many of these third-party platforms rely on a fixed bank of predefined prompts; they simply run those set questions and track how often your brand appears. Because you don't always have eyes on their full, underlying collection of prompts—it is essentially a black box—it is difficult to assess exactly how accurate their total visibility score really is. It is a bit like trying to win a trivia night where you aren't allowed to see the question cards.

However, if you are just starting your AI search journey and already have access to these established SEO tools in your stack, they are an excellent, low-friction way to get going. You can use them to:

- **Review competitor traffic.** See how much total AI traffic your rivals are successfully pulling in.
- **Identify traffic sources.** Figure out which specific AI tools (e.g., ChatGPT versus Perplexity) are sending traffic your way.
- **Spot winning content.** Analyze exactly what types of content are resonating in AI search and earning those coveted citations.

The Dark Traffic Dilemma

Tracking this traffic is not without hurdles. The biggest villain in the AI tracking story is “dark traffic.” Many AI platforms do not pass clean referral data, causing visits to appear as mysterious

“direct” traffic in GA4. Furthermore, AI referrers change their domain structures frequently—like moving from `openai.com` to `chatgpt.com`—meaning your filters require constant maintenance.

The Phone-Book Funnel

But there is another, more human layer to this tracking challenge: the user journey itself. In reality, a massive chunk of your audience is spending their time chatting with an AI to weigh their options and make a final decision. Once the AI convinces them you are the right choice, they don’t always click a citation link. Instead, they open a new tab, type your brand name into Google, and click the first result.

They are essentially turning Google into the modern-day phone book.

If you have ever tried to attribute a spike in sales to a viral TikTok, a great podcast mention, or a stellar PR campaign, this probably sounds incredibly familiar. This multi-touch, hard-to-track funnel is something we already see consistently across social media and digital PR—AI is just the newest, highly influential member of the club. Because the actual decision happened off-site in an AI chat environment, it creates a massive attribution blind spot.

Returning to Zero-Party Data

This exact dilemma is why the humble “How did you find us?” dropdown on your checkout or lead generation forms is having an absolute renaissance. When standard attribution models fail to capture the invisible hand of the AI, simply asking the user is one of the most powerful tools in your arsenal.

If you can establish a consistent baseline for the types of queries and customers emerging this way—for instance, noticing a steady 5% of leads typing “ChatGPT told me to” or “Saw you on Perplexity” into your intake form—you can identify a

reliable trend to map against your direct and branded search growth. You cannot fix dark traffic completely, but by combining log file analysis, proactive prompt tracking, and good old-fashioned qualitative surveys, you can illuminate enough of the dark to make smart, data-driven decisions for your brand.

The Bottom Line: Triangulating Your AI Truth

Tracking AI traffic today can feel a bit like trying to solve a Rubik's cube in the dark. The platforms are constantly shifting, the rules are being written in real time, and a frustrating amount of the data is inherently obscured. But as marketing leaders, we know that perfection cannot be the enemy of progress. You do not need a flawless, 100% accurate attribution model to understand the broader narrative; you just need a reliable baseline.

By layering your GA4 regex filters, leaning into your native bot dashboards, deploying prompt tracking tools, and embracing zero-party data, you are essentially building a triangulation system for your brand. You are moving from flying blind to reading the room. Once you have a steady pulse on how these generative engines perceive you, you can stop guessing and start strategizing. Now that our tracking instruments are calibrated and we know who is talking about us in the AI sandbox, it is time to give them something better to say.

17

Really, Really Know Your Customers

If traditional SEO was about shouting your brand's name through a megaphone in a crowded town square, AI search is like sitting down with a hyper-attentive personal shopper. Today's AI engines—like ChatGPT and Google Gemini—don't just look at *what* a user is asking; they look at *who* is asking.

Users are actively building out “Personalization” profiles, giving these models direct instructions like, “I’m a vegan, health-conscious gym girly,” or “I’m a chocolate lover, a foodie, and a bit bougie.” When your ideal customers ask for recommendations, they get entirely different responses based on these exact preferences. For CMOs and marketing leaders, the question is simple: If an AI is playing matchmaker, does your digital footprint give it the right signals to match with your target audience?

Here is how you can optimize your brand for the era of hyper-personalized, agentic AI search.

The Anatomy of an AI Match: Moving Beyond Generic Content

To win in personalized AI search, you need to create detailed, data-rich content that acts as “hooks” for specific user identities. It’s no longer enough to be the “best athletic wear.” You need to be the “best cruelty-free athletic wear for tall women.”

Here are the four key on-site pillars of customer-centric AI optimization:

- **Entity affinity through value propositions.** Brands must explicitly state their stances on lifestyle and ethical choices. Whether it’s sustainability, veganism, or cruelty-free practices, spell it out across your website documentation. When a user tells their AI they only shop sustainably, the AI crawls for these specific ethical alignments to filter its recommendations.
- **User-specific metadata in reviews.** Generic social proof is out. Rich, specific data is in. Encourage and capture reviewer metadata like age, height, body type, and typical activities. This allows an AI to confidently tell a user, “This product is recommended for you because someone with your exact height and body type loved it.”
- **Declare specific audience programs.** Don’t hide your community affiliations. Explicitly list programs like student, military, or healthcare worker discounts in your website footer. This helps AI associate your brand with those specific entities.
- **First-person language in support hubs.** Users naturally prompt AI in the first person (e.g., “How do I track my order?”). Mirror this conversational tone in your support documentation. Framing your FAQs as first-person questions makes your content much more likely to be surfaced for personalized, long-tail queries.

The Offsite Echo: Fueling Agentic Personalization

Getting your website in order is just step one. If your site is your brand's home base, offsite platforms are the neighborhood associations where AI goes to run a background check.

When we talk about agentic personalization—the emerging phase where AI agents actively scout, filter, and make decisions on behalf of a user's highly specific persona—your offsite footprint becomes just as critical as your on-site content. LLMs don't just take your word for it; they look for corroborating evidence in the wild.

- **Facebook and Reddit as personalization goldmines.**

Traditionally, CMOs might view Facebook and Reddit purely through the lens of community management. In the era of AI search, these platforms are deeply valuable for agentic personalization because they are built on self-identified micro-communities. Subreddits and Facebook Groups are literal hubs of niche identities. When your brand is consistently mentioned in *r/veganfitness* or a localized “Sustainable Living” Facebook group, you feed the LLM the exact contextual entity associations it needs. The AI learns, “Ah, the people who hang out in this niche community love this brand.”

- **Echoing value propositions.** Your value propositions (like being cruelty-free or locally sourced) need to travel. When these propositions are discussed, debated, and validated in offsite forums, it strengthens the AI's confidence in matching your brand to a user who demands those values.

- **Cementing community positioning.** Agentic AI wants to recommend brands that feel like an authentic cultural fit for the user. By actively cultivating your community positioning on platforms like Reddit and Facebook, you create a web of validation. You aren't just a

sporting goods company; you are a verified pillar of the “gym girlie” community.

The Gymshark Playbook: High versus Low Personalization

When we look at the landscape of AI visibility, high customer centricity directly correlates with high AI traffic. Gymshark is a prime example of this, often outperforming larger competitors by aggressively focusing on user-specific details. Let’s look at how their high-personalization approach stacks up against a lower-personalization baseline, like Lululemon:

Feature	Low personalization (Lululemon example)	High personalization (Gymshark example)	AI search benefit
Review content	General fit information.	Includes height, body type, and activity level of the reviewer.	Allows AI to perform high-level matching based on a user’s physical profile.
Support hub	Formatted by broad topics; lacks exclusive search.	Dedicated hub with searchable, first-person Q&A.	Highly valuable for answering specific “long-tail” customer queries.
Audience identifiers	Discounts not clearly declared in the footer.	Specific footer links for students, NHS, and military.	Explicitly associates the brand with niche audience entities.

Feature	Low personalization (Lululemon example)	High personalization (Gymshark example)	AI search benefit
Niche content	Relies on high-level “story” blurbs without a dedicated blog.	Active blog addressing specific needs like “tall leggings.”	Creates top-of-funnel hooks for specific customer segments.

Gymshark’s investment in these granular, customer-centric details is exactly why they frequently emerge as the sole branded site linked in personalized Copilot responses.

The “Tale of Two Profiles” Test

If you want to see exactly how dramatically AI tailors its responses, try running this simple test:

1. Open two separate ChatGPT accounts.
2. Navigate to Settings > Personalization.
3. Fill out the “More about you” section with distinct personas. For my test, I used “Vegan, health conscious, gym girly” for Profile A and “Chocolate lover, bougie, foodie” for Profile B.
4. Ask the exact same query in both accounts.

When I asked for “London brunch recommendations,” the “gym girly” was served up the Black Cat café (a fully vegan spot), while the “bougie foodie” was pointed towards Sunday in Brooklyn, known for its decadent hazelnut maple praline pancakes. The exact same divergence happened for winter vacation destinations—the gym girly was recommended cozy, snowy

locales in Switzerland for skiing, while the foodie was directed to luxury spa retreats.

When you run this test for your own industry, pay attention to three things:

- **Variance.** How much do the responses differ? On this test there was some overlap, but on others, there is completely divergent output.
- **Translation.** Notice how ChatGPT translates a colloquial phrase like “gym girlie” into a broader operational concept like “health-friendly.”
- **Sourcing.** Where are they pulling this personalized data from? I’ve seen AI pull from Google Business Profiles, support docs, environmental statements, and offsite affinity declarations.

Your Action Plan: Own Your Narrative

How do you adapt your marketing strategy to this new reality? Start at the bottom of the funnel.

Make sure your brand affinities and communities are clearly and proudly stated on *your* website. Embed them in your support documentation, your FAQs, your About pages, and your footers. Then, make moves to get these values cited on third-party websites and niche community forums, because these considerations are make-or-break for modern consumers, and LLMs will build them into their models.

Do you offer free access for charities? A robust DEI policy? A sustainability pledge? Military discounts? Are your products cruelty-free, wheelchair accessible, locally sourced, organic, made in the United States, dishwasher safe, or backed by a life-long warranty?

Whatever your unique hooks are, make them crystal clear. Do not wait for a random Redditor to establish your brand values for you—lead the conversation so the AI finds exactly what it’s looking for. ChatGPT has already explicitly stated they use memory and personalized profiles, and they just announced a new feature to carry out personalized shopping queries. Back at Google I/O, we heard the exact same messaging from the Gemini team. The AI models are ready to recommend you—you just have to give them the data they need to make the match.

18

Be the Answer They Can't Ignore

The transition to AI search is not a distant prediction; it is happening right now, with platforms like ChatGPT and Gemini capturing billions of visits. In my 20 years of marketing, I have never seen a more voracious rate of change. Users are actively opting into conversational AI to seek practical guidance and decision-making support rather than just clicking blue links.

For marketing leaders, this is not a moment to panic about declining organic traffic, but an opportunity to confidently lead a fundamental shift in how information is managed and consumed. We are at the precipice of a new discipline: generative engine optimization (GEO).

Adapting the Playbook: From Predictability to Volatility

Many of today's marketing leaders came up when SEO was predictable: good content and backlinks equaled steady growth. With the rise of AI Overviews and fractured search experiences across TikTok, Reddit, and ChatGPT, that predictability is gone.

However, your core search skills are still your foundation. The shift simply requires adjusting your framework to prioritize how large language models (LLMs) ingest and recommend information. This is not about abandoning the old ways, but about moving towards a holistic, 360° approach. AI assistants do not just take your website's word for it; they look at what Reddit, reviews, and fan communities say to form a total brand response.

Use All Your Tools

You don't need a massive budget to start seeing returns on GEO. Attribution models are becoming increasingly inaccurate as users carry out deep searches in AI before ever visiting Google. To fight this attribution gap, you can capture immediate, high-intent data by updating your lead forms.

Adding "ChatGPT" or "Perplexity" to your "How did you find us?" question acts as an LLM lie detector that bypasses murky analytics and shows where the decision was actually made. Over time, these responses become one of the most reliable signals you have for how often AI assistants are truly driving your pipeline.

Get Into the Weeds for Rich Data

Diving into your server bot logs is one of the most powerful ways to measure your real AI visibility. By tracking user agents like OAI-SearchBot and ClaudeBot, you can see exactly which pages AI tools are actively crawling and which they are skipping.

This raw data gives you hard evidence to explain the shifting landscape to the C-suite and secure buy-in for technical GEO fixes. Instead of debating hypotheticals about AI, you can point directly to crawl patterns, gaps, and improvements over time.

Your Digital Footprint: All Killer, No Filler

AI models do not just evaluate your homepage; they scrape your entire digital ecosystem. That means your content must be optimized for tokenization and ingestion from the ground up, on your pages and beyond.

In the past, we chased junk traffic through bloated tag pages and filler content. In the AI age, that is a liability. You want your digital footprint to be all killer, no filler. By creating clean, semantic structures, you make it easier for bots to parse your information. When your data is structured and concise, models can more easily reason over it and feature your brand as a trusted solution.

The Path Forward: Embracing Change

Because the rules of the agentic web are still being written, your most effective strategy is a cycle of pragmatic experimentation. The brands that win will be the ones testing new standards, and protocols.

The shift to AI search is ultimately about removing friction. By consistently measuring your baseline, refining your digital footprint, and accepting that the last click rarely tells the whole story, your brand will remain highly visible. Start small, rely on your bot data, and focus on being the definitive answer that AI models can't ignore.

Resources

Jargon Buster

Here are 20 jargon-buster terms in plain list form.

1. Generative engine optimization (GEO)

A discipline focused on making your brand more visible and useful inside AI assistants and chat-based search, not just classic Google results.

2. Large language model (LLM)

An AI system trained on huge amounts of text that can answer questions, write content, and hold conversations in natural language.

3. AI search

Any search experience where answers are generated or summarized by AI, including tools like ChatGPT, Gemini, Perplexity, and Copilot.

4. AI Overview

Google's AI-generated summary that appears in some search results, separate from a full conversational chat experience.

5. Training data

The historical content (web pages, books, code, etc.) an LLM was trained on before deployment, which shapes what it already “knows.”

6. Knowledge cutoff

The date after which events and updates are not fully included in a model’s core training data.

7. RAG (retrieval-augmented generation)

When an LLM looks up fresh information (e.g., via web search) and then writes an answer that blends that with its existing training data.

8. Reasoning/mixture of experts (MoE)

A response mode where the model breaks a complex request into parts and consults specialist data sources or tools before answering.

9. Agentic web

An emerging version of the web where AI agents browse, click, compare, and act on behalf of users, rather than just showing them links.

10. llms.txt

A proposed machine-readable file (similar to robots.txt) that gives AI crawlers explicit instructions about how to use your content.

11. Ingestion

The process where AI systems crawl, read, and store your content so it can be used later when generating answers.

12. Token/tokenization

The way models break text into small units (tokens) so they can process, store, and bill for it efficiently.

13. Embeddings/vectors

Numeric representations of text that let models find content by meaning and similarity, not just by exact keywords.

14. Semantic triple

A simple statement in the format “thing—does/is—thing” (e.g., “Brand X provides payroll software”) that makes relationships clear to machines.

15. Entity

A distinct “thing” the model tracks, such as a brand, person, product, or place, which lives in its internal knowledge graph.

16. Entity understanding

The point at which an AI system clearly knows who you are, what you do, and whom you serve, and can recommend you accurately.

17. Chunking

Breaking long content into small, self-contained sections (headings, FAQs, lists) so AI tools can easily quote, reuse, and understand it.

18. AI referral traffic

Visits that ultimately happen because an AI assistant recommended your brand, even if they later show up as “direct” or “organic” in analytics.

19. “LLM lie detector”

Your updated “How did you find us?” field that explicitly lists AI tools and reveals when a chatbot actually drove the decision.

20. AI visibility KPIs

The set of metrics you track to prove GEO impact, such as AI-driven leads, LLM citations, brand mentions, and conversion rate from AI-referred users.

Get more resources at www.crystalcarter.ai.

Acknowledgments

I want to acknowledge the people who have encouraged, sustained, and inspired my work in this space. In my work life, Kiera Carter, Paula Ximena Mejia, and George Nguyen have been instrumental in shaping my ideas and moving them forward. Across the wider industry, conversations and support from Ross Simmonds, Aleyda Solis, Areej Abu Ali, Erin Simmons, Lidia Infante, Kristina Azarenko Sparks, Tori Gray, Chima Mmeje, Rejoice Ojiaku, Anu Adegbola, Miracle Inameti-Archibong, Lily Ray, Sam Torres, Rita Cidre, Nicole Ponce, Barry Schwartz, Mordy Oberstein, and Andi Jarvis have kept me grounded and motivated. During this period of transformation, the fearless exploration of Mike King, Wil Reynolds, Krishna Madhavan, R. V. Guha, and all the other AI search pioneers has been a constant source of inspiration. Finally, I am eternally grateful for my family and friends—Sam, Cass, Donna, Cordell, China, Kiesha, Yvonne, Mel, Mike, and Annie.

Index

Numerics

301 redirect, 109

404 error, 30, 109, 160

A

A2A. *See* agent-to-agent

Abrahami, Avishai, 112

ACP. *See* Agentic Commerce
Protocol

acquisition partners, 152–153

ADK. *See* Agent Development
Kit, Google

adoption of AI, 11–16

and age, 15

and gender, 14–15

and geography, 12–13

and language, 13–14

market share, 16–18, 149–150

and operating systems, 15–16

total serviceable market, 18–19

age, and AI adoption, 15

Agent Development Kit (ADK),
Google, 121, 124

agentic browsers, 16, 140

Agentic Commerce Protocol (ACP),
134, 143

agentic interoperability, 142–144

agentic solutions, 111, 131–132

agentic infrastructure, 133–134

becoming an OpenAI

e-commerce seller, 133

creating own, 134–135

and partnerships, 135–136

agent-to-agent (A2A), 121, 122,
133, 134, 142

AI agents

adoption of protocols, 143–144

auditing of shopping feeds and
APIs, 140–141

building, 134–135, 144–145

and choice of tech partners, 140

configuring MCP servers for,
142–143

customer funnel, 138–139

markdown content for, 142

personalization of, 145,
187–188

rise of agentic web, 137, 138
website for, 139

AI Concierge, OpenTable, 135

AI Mode, Google, 79, 119, 144

AI Overviews, Google, 1–2, 79

AI SDK, Vercel, 122

AI SERPs, 112–113

“all killer, no filler” approach,
101, 195

- analytics, for tracking LLM traffic,
 - 26, 161
 - GA4, 26–27
 - website, 28–29
- Android Auto, 139
- Anthropic, 121, 122, 142
- APIs, and AI agents, 141
- Apple, 15, 16
- ARIA tags, 140
- audience programs, 186
- AWS, 121, 133, 142
- B**
- Baidu, 6
- Bing, 6, 72, 74, 75, 78
- Bing Chat, 74
- bot logs, 29–31, 34, 76, 100, 101, 109, 128, 160–161, 194–195
- BPE. *See* byte pair encoding
- branded prompts, 151, 168, 169
- brand entities, 105–107
- brand mentions, 60–61, 150–151, 158, 175
- brand names, 130
 - ambiguity, semantic triples for solving, 44
 - tokenization of, 38–40
- brand topics, 167
- budget, paid AI, 163
- business offerings, semantic triples for defining, 45
- byte pair encoding (BPE), 38
- C**
- ChatGPT, 15, 72, 78, 99, 191
 - ChatGPT-3, 59, 62
 - content partner of, 154–155
 - custom GPTs, 111, 112
 - female user base of, 14–15
 - Instant Checkout, 143
 - metrics, 162
 - personalization, 145
 - plugin, of OpenTable, 134–135
 - search engine, optimization for, 73–75
 - use of `llms.txt` files by, 119
 - visits, growth in, 3, 4
- ChatGPT Atlas, 16
- chunking, 87–89, 101–103
- citations
 - and content partnership, 132
 - domain, 152
 - favicons in, 113
 - RAG, 54, 72, 78
 - RAPP framework for, 159, 162
 - reasoning-model responses, 87, 88, 90–92
 - reviewing, 159
 - third-party, 152–153
- Claude, 68, 135, 138
- click-through rate (CTR), 113
- Cloudflare, 120–121, 125
- community positioning, 187–188
- competitors
 - competitor-branded prompts, 169
 - identification and tracking of, 167–168
 - third-party competitor tools, 181
- complex questions, and reasoning models, 89
- content clusters. *See* topic clusters
- content strategies for LLM responses, 49–50
- content structure. *See* chunking
- context window of LLMs, 37–38
- conversational search, 1–3
- Copilot. *See* Microsoft Copilot
- cosine similarity, 42–43
- crawlers, LLM, 34–35
 - bot logs, 29–31, 34, 76, 100, 101, 109, 128, 160–161
 - optimization for, 99–101
- credentials, author/brand, 91
- CRM. *See* customer relationship management (CRM)
- CTR. *See* click-through rate

- current events, and reasoning models, 87, 88
- customer relationship management (CRM), 25
- customer(s), 185
 - adoption market share, 16–18, 149–150
 - and agentic personalization, 187–189
 - AI adoption rate of, 11–16
 - customer-centric AI optimization, 186
 - funnel, AI agent, 138–139
 - high vs low personalization, 188–189
 - identification of, 10–11, 166
 - queries, 170–172
 - semantic triples for specifying, 45
 - tailored responses to, 189–190
 - total serviceable market, 18–19
- custom GPTs, 111, 112
- D**
 - dark traffic, 181–182
 - dashboards, for tracking LLM traffic, 178–179
 - data ingestion, 33–36, 63–64, 107, 152
 - data pipeline
 - ingestion, 33–36
 - token IDs and vector embeddings, 41–48
 - tokenization, 36–41
 - Deep Research, Gemini, 56, 85, 87, 89–90
 - DeepSeek, 87, 88
 - Deloitte, 143
 - Diagrams: Flowcharts & Mindmaps* GPT, 111
 - digital footprint, 6–7, 35, 96, 107, 195
 - direct traffic, 28–29, 181–182
 - documentation page, 62
 - domain citations, 152
 - duplicate content, and Markdown pages, 126
 - E**
 - e-commerce, 112
 - becoming an OpenAI e-commerce seller, 133
 - payments, 133–134, 136
 - Wix eCommerce, 118, 120
 - entity competitors,
 - identification of, 168
 - entity management, 105–107
 - ethical choices, stances of
 - brands on, 186
 - experimentation, 115–116, 195
 - environment for, 127, 128
 - involvement with AI
 - platforms, 129–130
 - llms - full .txt, 123–125
 - llms .txt, 116–123
 - Markdown, 125–127
 - monitoring bot action, 128
 - ROI of, 128
 - targets, 127
 - time period, 127
 - F**
 - Facebook, 107, 187, 188
 - favicons, 112–113
 - feedback loops, 68–69, 97, 109
 - first-person language in
 - support hubs, 186
 - Fishkin, Rand, 17
 - G**
 - GA4. *See* Google Analytics 4
 - gating network, 84, 85
 - Gemini, 3, 5, 14, 99, 104
 - data ingestion in, 35
 - Deep Research, 56, 85, 87, 89–90
 - feedback to, 68–69
 - Gemini 2.5 Flash, 122

- Gemini (*continued*)
 - integration into Google ecosystem, 15–16
 - personalization, 145
 - search engine, optimization for, 73
 - GenAIScript, 122
 - gender, and AI adoption, 14–15
 - general queries, 169
 - Gen Z, AI adoption by, 15
 - geography
 - and AI adoption, 12–13
 - and keyword research, 167
 - GitHub, 122
 - global content, managing, 96, 97, 99
 - Google, 107, 115
 - Agent Development Kit, 121, 124
 - agentic infrastructure/ services of, 133
 - AI Mode, 79, 119, 144
 - AI Overviews, 1–2, 79
 - Android Auto, 139
 - multimodal assessment tools, 104
 - NotebookLM, 6
 - use of `llms.txt` files by, 118–121, 142
 - Google Analytics 4 (GA4), 26–27, 166, 178
 - Google Assistant, 15
 - Google-Extended, 35
 - Google Lens, 104
 - Google Merchant Center, 140
 - Google Search, 4–5, 23, 78, 109, 111, 119, 168, 172, 182
 - Google Search Console, 128
 - Google Sheets, 173
 - Google Vision AI, 104
 - GPT-4, 37
 - GPT-5.4, 38
 - Grok, 15, 74, 75, 108
 - Grokipedia, 75
 - Gymshark, 188–189
- H**
- hallucinations, AI, 108–110, 121–122, 128, 160
 - Harper's Bazaar*, 154–155
 - Howard, Jeremy, 117
 - “How did you find us?” question on high intent forms, 22–25, 182, 194
 - HTML, 122, 142
 - pages, conversion into Markdown, 125
 - tokenization on, 38
- I**
- IETF. *See* Internet Engineering Task Force
 - incorrect information, cleaning up, 66–67, 99
 - IndexNow, 72–73
 - information pipeline, 98–99
 - Instagram, 107
 - Internet Engineering Task Force (IETF), 126
 - interoperability, agentic, 142–144
- J**
- JSON-LD structured data, 78, 79
- K**
- keyword research, 167, 171
 - keyword search volume, 18–19
 - King, Mike, 43–44
 - knowledge cutoffs of LLM platforms, 61–62, 97
- L**
- language
 - and AI adoption, 13–14
 - first-person language in support hubs, 186
 - natural language processing, 105–106

- large language models (LLMs), 1–2, 22–23, 194. *See also* data ingestion; responses, LLM
 - bot logs, 29–31, 34, 76, 100, 101, 109, 128, 160–161
 - chatbots, growth of, 3–4
 - content partnership, 111, 131–136
 - context window of, 37–38
 - conversations and leads from, 161
 - crawlers, 29, 34–35, 99–101
 - experts in, 51–52, 55–57, 84–86
 - feedback to, 68–69, 97, 109
 - knowledge cutoff of, 61–62, 97
 - search engines, optimization for, 73–75
 - traffic, using analytics for tracking, 26–29, 161
 - user agents of, 35, 100
 - legacy domains, and LLM traffic tracking, 27
 - Levi's, 86, 89
 - lifestyle choices, stances of brands on, 186
 - LinkedIn, 108
 - listicles, 103
 - live topics, and reasoning models, 87, 88
 - LLMs. *See* large language models
 - `llms-full.txt`, 123–124
 - documentation backup for, 124–125
 - for powering agentic development, 124
 - as a strategic proof of concept, 124
 - `llms.txt`, 116–118, 125, 127, 142
 - for informing agent-to-agent architecture, 121
 - as markdown sitemap, 120–121
 - myths, breaking, 118–119
 - provision of grounding context by, 122
 - for reducing LLM hallucinations, 121–122
 - for signaling core context, 120
 - tokenization cost of, 117, 122–123
 - LLM traffic, tracking, 177
 - AI referrers, 178
 - analytics for, 26–29, 161
 - dark traffic dilemma, 181–182
 - native dashboards, 178–179
 - prompt tracking tools for, 179–180
 - third-party competitor tools for, 181
 - and user journey, 182
 - zero-party data, 182–183
 - logs, bot, 29–31, 34, 101
 - Looker Studio, 25, 27
 - Lululemon, 188–189
- M**
- Marie Claire*, 154–155
 - Markdown pages, 116, 125, 142
 - and duplicate content issue, 126
 - safe execution and ROI monitoring of, 126–127
 - markdown `.txt` file, tokenization on, 38
 - market share, AI adoption, 16–18, 149–150
 - “matches partial regex” match type, 178
 - MCP. *See* Model Context Protocol servers
 - media partnerships, of OpenAI, 154
 - Meta AI, 108
 - metadata, user-specific, 186
 - meta descriptions, 115
 - metrics, AI search, 147–149, 163
 - adoption market share, 16–18, 149–150
 - brand mentions, 150–151
 - domain citations, 152

- metrics, AI (*continued*)
 - from platforms, 162
 - ranking, 157–158
 - sentiment, 158–159
 - third-party citations, 152–153
- Microsoft, 74, 122
- Microsoft Copilot, 16, 73, 74, 162, 189
- Millennials, AI adoption by, 15
- mixture of experts (MoE), 51–52, 55–57, 83–86
- model card, 62
- Model Context Protocol (MCP)
 - servers, 142–143
- MoE. *See* mixture of experts
- multimodal data ingestion, 35
- multimodal optimization, 100, 103–104
- multi-perspective content, and reasoning models, 87, 91
- N**
 - named entities, 46–48, 64, 91, 106
 - native platform presence, 110–113
 - natural language processing, 105–106
 - NLWeb, 79
 - noindex tag, 126
 - non-branded comparisons, 169, 174
 - NotebookLM, 6
- O**
 - old information, cleaning up, 66–67, 99
 - OpenAI, 3, 107, 136, 140, 162
 - context window of LLMs, 37
 - and e-commerce payments, 133
 - e-commerce seller, becoming, 133
 - Operator, 135
 - partnerships of, 111, 131, 135, 153–154
 - tokenizer, 36, 39
 - use of `llms.txt` files by, 122
 - OpenTable, 134–135
 - operating systems, and AI adoption, 15–16
- P**
 - paid AI ad placements, 162, 163
 - PayPal, 133, 136
 - “People Also Ask” boxes, Google Search, 172
 - Perplexity, 121, 134
 - partnerships of, 135, 153
 - RAG responses of, 54, 78
 - reasoning-model responses of, 56
 - search engine, optimization for, 73
 - use of `llms.txt` files by, 119
 - Perplexity Publisher Program, 111
 - personalization
 - agentic, 145, 187–188
 - high vs low, 188–189
 - profiles, of customers, 185
 - phased approach to LLM
 - optimization, 97–99
 - platform preferences of
 - customers, 166
 - platform presence, 110–113
 - prompt tracking, 76, 150, 165–166, 179–180
 - brand mentions, 150–151
 - choice of prompts for, 168–170
 - creating and updating
 - content for, 170
 - domain citations, 152
 - efficient, results of, 170–171
 - identification and tracking of
 - competitors, 167–168
 - identification of audience, 166
 - preparation of brand topics, 167
 - testing, iteration, and
 - refinement, 170
 - third-party citations, 152–153
 - workflow, 171–176

PR partners, LLM-ready, 153
 publishers, partnership with, 111,
 131–132, 153–155

R

RAG. *See* retrieval-augmented
 generation

ranking, 157–158

RAPP (Regular, Accurate,
 Prominent, and Positive)
 framework, 159, 162

reasoning-model responses,
 51–52, 55–57

chunking, 87, 88

citations, 87, 88, 90–92

complex questions, 89

content rewarded by reasoning
 models, 86–88

current events, 87, 88

fanning out queries, 89–90

functioning of, 84–86

opinionated, multi-perspective
 content, 87, 91

optimization checklist, 92–94

social presence, 91–92

user-generated content, 87, 91–92

Reddit, 107, 172, 187, 188

referral traffic, 26–29, 76, 135, 178

regex filters, GA4, 26, 27, 183

Request for Comments (RFC)
 series, 126

responses, LLM

generation of, 50–52

RAG responses, 51–52, 54–55,
 71–81

reasoning-model responses,
 51–52, 55–57, 83–93

training-data responses,
 51–54, 59–70

retrieval-augmented generation
 (RAG), 34, 35, 39, 42, 51–52,
 54–55, 96–97, 119. *See also*
 reasoning-model responses

checking core queries, 75–76

chunking, 89

IndexNow, 72–73

low-value retrieval, 160

maintaining fresh content, 71–72

and MoE model, 85

optimization checklist, 79–81

optimization for LLM search
 engines, 73–75

prioritizing already cited pages,
 76–77

publishing in topic clusters and
 waves, 77–78

and UGC, 107

using structured data and rich-
 results friendly markup, 78–79

return on investment (ROI)

of experiments, 128

monitoring, and Markdown
 pages, 126–127

and testing `llms.txt` files, 119
 and token efficiency, 38

Reynolds, Wil, 161

RFC. *See* Request for
 Comments series

`robots.txt`, 31, 35, 100,
 101, 160

ROI. *See* return on investment

S

Salesforce, 145

search engines, 157, 168

handling of duplicate content
 by, 126

LLM, RAG optimization for,
 73–75

traditional, 100, 101, 109

Seer Interactive, 161

semantic triples, 43–46, 64

semantic web, 105–106

sentiment, 158–159

shopping feeds, auditing, 140

Siri, 15

social media
 citations in, 153
 content, as prompt seeds, 172
 presence, and reasoning-model responses, 91–92
 profiles in, 107–108
 South Korea, AI adoption in, 14
 space-prefix factor, in tokenization, 39–40
 Sparktoro, 17
 Stripe, 133, 135, 140
 structured data, 106, 195
 for AI agents, 140, 142
 and retrieval-augmented generation, 78–79
 support hubs, first-person language in, 186

T

target market. *See* customer(s)
 Tensorflow, 42
 third-party citations, 152–153
 third-party content, updating, 67
 TikTok, 108
 token IDs, 41
 tokenization, 36
 of brand names, 38–40
 and context window, 37–38
 cost, of `llms.txt`, 117, 122–123
 importance to marketers, 38–40
 role of case and context in, 37
 space-prefix factor in, 39–40
 tokens, 36
 efficiency, 38, 40–41, 117
 reduction, using Markdown for, 125
 topic clusters, 46, 77–78, 90
 topic trends, 166, 175
 total serviceable market, 18–19

training-data responses, 51–54, 59–60, 162
 brand mentions, 60–61
 building foundations for data ingestion, 63–64
 cleaning up incorrect/old information, 66–67
 feedback to LLMs, 68–69
 foundational questions exercise, 65–66
 new businesses waiting to be recognized in training data, 64
 optimization checklist, 70
 preprocessing pipelines of training datasets, 78
 training data updates, 61–64, 97

U

user-generated content (UGC), 87, 91–92, 107–108
 user-specific metadata, 186

V

value propositions, 104, 186–188
 vectorization, 41–42, 142
 cosine similarity, 42–43
 named entities, 46–48
 semantic triples, 43–46
 topic clusters, 46
 vector search systems, 42
 Vercel, 122
 Visa, 143
 visual content, 100, 103–104, 107

W

WAI-ARIA best practices, 140
 Weather.com, 121–122
 website analytics
 for confirming platform preferences, 166
 for tracking LLM traffic, 28–29, 76

- Wikipedia, 105–106
 - Wix, 135, 140, 145
 - Wix AI Visibility Overview tool, 28
 - Wix ChatGPT app, 112
 - Wix eCommerce, 118, 120
 - Wix MCP, 143
 - Wix Stores, 134
 - Wix Studio, 63, 118, 178
 - women, AI adoption by, 14–15
 - workflow, prompt tracking
 - adding prompts to tracker, 173–174
 - conversion of seed phrases into LLM prompts, 172–173
 - example, 175–176
 - monitoring prompt performance, 174–175
 - optimization tactics, 174
 - sourcing potential prompt seeds, 171–172
 - writing, for token efficiency, 40–41
- X**
- X, 75, 108
- Y**
- YouTube, 102, 103
- Z**
- Zapier, 121
 - zero-party data, 182–183

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access
Wiley's ebook EULA.