

 PLAYBOOK

AI Search **Visibility** Playbook

How to get mentioned and cited when buyers ask ChatGPT, Google AI, Perplexity, Copilot and Claude.

Best accountant for a small online store?



① **yourbusiness.com**

② directory

ChatGPT



AI Overviews



Perplexity



Gemini



Copilot



52 pages

8 page templates (.docx)

Prompt tracker (.xlsx)

8 before-and-after rewrites

Contents

	How to use this playbook	3
01	How AI assistants decide what to cite	5
02	Let the right crawlers in	8
03	Get indexed where the answers come from	13
04	Pages that get cited: seven formats	17
05	Before and after: eight rewrites	21
06	Make your business unmistakable	26
07	Earn the mentions assistants trust	30
08	Products in AI shopping answers	34
09	Measure AI visibility	37
10	llms.txt, honestly	42
11	Mistakes that cost visibility	44
12	The 30/60/90-day plan	46
	If you want help applying this	48
	Glossary	49
	Sources	50

01

How AI assistants decide what to cite

Every assistant in this book works the same way at its core: it turns a question into searches, retrieves pages from an index, and writes an answer from what it reads. If you understand that loop, most of the tactics you'll hear about sort themselves into "useful" and "noise".

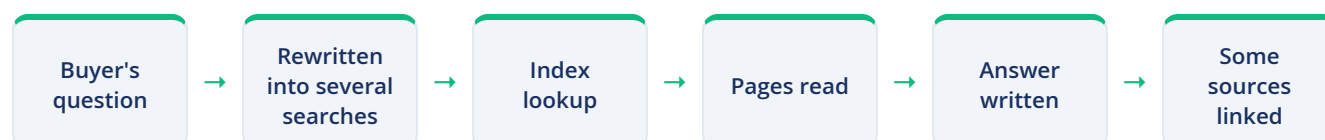
Retrieve first, then write

Language models have two ways of "knowing" about your business.

- **Training memory.** What the model absorbed from web pages when it was trained. It is months old, you can't edit it, and it's shaped by what *other* people wrote about you.
- **Live retrieval.** When an assistant searches the web while it answers. This is current and it comes with links. It's the part you can influence this month.

Google calls live retrieval *grounding* or retrieval-augmented generation (RAG). Its systems "rely on our core Search ranking systems to retrieve relevant, up-to-date web pages from our Search index", then generate a response with links to the pages that support it.

How a cited answer gets built



The second step matters more than most people think. Google calls it **query fan-out**: the model issues "a set of concurrent, related queries" to fetch more results. ChatGPT does the same; its help center says it "typically rewrites your query into one or more targeted queries" that it sends to its search providers.

So a buyer asking "which bookkeeping service is best for a small online store?" might trigger searches like "bookkeeping for ecommerce businesses", "online bookkeeping pricing" and "bookkeeper vs accounting software". Your page doesn't need to match the buyer's exact words. It needs to be the best result for one or more of the searches the assistant actually runs.

WHAT THE RESEARCH SAYS

Google describes fan-out with an example: for "how to fix a lawn that's full of weeds", fan-out queries might include "best herbicides for lawns", "remove weeds without chemicals" and "how to prevent weeds in lawn". — [Google Search Central, 2026](#)

Where each assistant gets its pages

The assistants don't share one index. Here is what each one searches, which crawler you must allow for it to read your pages, and where you can see the results. A "training crawler" collects content for future models; blocking it does not, by itself, remove you from live answers (Chapter 2).

Assistant	What it searches	Crawler to allow	Separate training control	Where you see results
Google AI Overviews and AI Mode	Google Search index	Googlebot	Google-Extended (no effect on Search)	Search Console: Generative AI performance report
Gemini app	Google Search, used for grounding	Googlebot	Google-Extended also covers grounding in Gemini Apps	GA4: AI Assistant channel
ChatGPT search	OpenAI's own index plus partner providers, including Microsoft	OAI-SearchBot	GPTBot	GA4: referrals from chatgpt.com
Microsoft Copilot	Bing index	Bingbot	—	Bing Webmaster Tools: AI Performance
Perplexity	Perplexity's own index	PerplexityBot	Perplexity says PerplexityBot isn't used for foundation models	GA4 referrals
Claude	Anthropic's search index plus a third-party provider (Brave Search was added as a subprocessor in March 2025)	Claude-SearchBot	ClaudeBot	GA4 referrals

Three conclusions follow from this table.

- 1. Google and Bing indexes reach most of these answers.** Google's own AI features run on its index. Copilot runs on Bing's, and OpenAI lists Microsoft among ChatGPT's search providers. A site that is weak in Bing is weak in more places than Bing.
- 2. Each AI company has its own search crawler.** If a firewall or an old robots.txt rule blocks OAI-SearchBot, PerplexityBot or Claude-SearchBot, you can't be cited by that assistant, however good your pages are.

3. **The training question is separate.** You can stay in live answers and still opt out of training. Chapter 2 sets out the trade-off.

How big is this, honestly?

900M

weekly ChatGPT users, February 2026

TechCrunch, 2026

8% vs 15%

of Google visits ended in a result click with vs without an AI summary

Pew Research, 2025

0.17%

of the average site's traffic came from AI chatbots (3,000 sites)

Ahrefs, 2025

Read those three numbers together. The audience asking assistants is very large. The clicks they send are still small, and AI summaries reduce clicks on classic results. Ahrefs notes its figure is likely an undercount, because some AI visits arrive as "direct" traffic. Pew also found people clicked a link *inside* the AI summary on just 1% of visits.

So the main prize is not referral traffic. It's being **named in the answer** when a buyer is forming a shortlist. Bing says it plainly in its guidelines: "A decline in clicks does not always indicate a loss of visibility." Google, for its part, reports that clicks from results pages with AI Overviews "are higher quality", meaning people spend more time on the site.

KEY IDEA

You can't buy or engineer your way into an answer. You can make sure you're retrievable, quotable and talked about, then measure whether it's working. That's the whole method.

WATCH OUT

Be careful with anyone promising a "#1 spot in ChatGPT". Google states that "no third-party tool has access to our internal ranking or AI systems", and Bing's guidelines say GEO "does not guarantee grounding or citations in AI experiences". Answers also change from one run to the next, so a single screenshot proves very little.

DO THIS

- ☐ Write down the five questions a buyer asks before choosing a business like yours. Use their words, not yours. You'll turn these into your prompt panel in Chapter 9.
- ☐ Ask three assistants each question today and save the answers. That's your informal "before" picture.

04

Pages that get cited: seven formats

Assistants quote passages, not whole pages. A passage gets quoted when it answers one question plainly, names who and what it's about, and carries a fact the assistant can't get elsewhere. Seven page formats do this reliably.

What the evidence says, and what it doesn't

The most cited research on this is a 2024 paper from Princeton and collaborators that coined the term "generative engine optimization". The researchers rewrote source pages in different ways and measured how much of each page made it into AI answers.

WHAT THE RESEARCH SAYS

The three best methods were adding citations to credible sources, adding quotations and adding statistics. Each "achieved a relative improvement of 30-40%" on the paper's main visibility measure. Keyword stuffing offered "little to no improvement". Pages ranked lower in search gained the most: citing sources lifted visibility 115.1% for pages ranked fifth, while the top-ranked page's visibility fell 30.3%. — [Aggarwal et al., KDD 2024](#)

Read this with care. It was a lab benchmark built on the AI engines of 2023–24, with one live test on Perplexity. It tells you what kind of writing models prefer to quote. It doesn't tell you any specific rewrite will move a specific answer.

The platforms' own guidance points the same way, with one disagreement worth knowing about:

- **Google** says you don't need to "chunk" content into tiny pieces or write "in a specific way just for generative AI search". Its main advice is to create **non-commodity** content: "unique expert or experienced takes that go beyond common knowledge."
- **Bing** is more prescriptive. It asks you to "place essential information near the top", make "facts and definitions explicit", keep each URL focused on one topic, and write content that's "easy to understand without external context".

These don't conflict once you see them side by side. Neither engine wants pages written for machines. Both reward pages where a human can find the answer fast and verify it. That's the standard this chapter uses.

The quotable passage test

Run any important section of a page through these five questions. Each "no" is a fix.

The quotable passage test

- ☐ Does the first sentence or two answer the question directly, without a preamble?
- ☐ Does it name the subject ("Acme Bookkeeping charges...") rather than "we" or "it"?
- ☐ Does it include a specific: a number, a range, a condition, a time frame or a named standard?
- ☐ Would it still make sense if it were the only paragraph someone read?
- ☐ Does it carry something only you know: your data, your process, your experience, your prices?

Seven formats that earn citations

Format	Questions it answers	Must include	Avoid
Answer block	"How long...", "Can I...", "What happens if..."	A 1–2 sentence direct answer, then conditions and detail	Opening with background or a sales pitch
Definition page	"What is...", "What does X mean"	"X is a [category] that [does what], used by [who]" plus one worked example	A dictionary line with nothing of your own
Comparison page	"X vs Y", "Is X better than Y"	Criteria table and a "choose X if... choose Y if..." verdict	Pretending your option wins on every row
Original data	"How much...", "What's typical...", "How often..."	Your own numbers, sample size, date and method	Unsourced round numbers
Pricing explainer	"How much does X cost"	Real ranges, what moves the price, two worked examples	"Contact us for a quote" as the only answer
FAQ set	The long tail of buyer questions	Real questions from sales calls and support, each with a direct answer	Questions nobody asks, invented to stuff keywords
Step list	"How to...", "What's the process"	Numbered steps, time per step, what you need	Steps that depend on images to make sense

1. Answer block

The workhorse format. Put the answer in the first one or two sentences under a question-shaped heading, then add the conditions, the exceptions and the detail. As a house rule, we aim for a first answer of about 40–60 words: long enough to be complete, short enough to quote whole. That range is our practice, not a platform rule.



AdvanSEM

— ADVANCE YOUR MARKETING —

Advance your marketing.

advansem.com · info@advansem.com