



FREE GUIDE • AI SEARCH • 2026 EDITION

AI Search Visibility.

Most of what is sold as GEO is SEO with a new invoice. This guide shows what actually decides whether ChatGPT, Gemini and Google's AI Overviews name your firm, what the published evidence says, and the 24 checks you can run this week.



WHAT IS INSIDE

- | | | | |
|----|-------------------------------|----|----------------------------|
| 01 | How an AI answer gets built | 04 | The part that is different |
| 02 | Five claims that are not true | 05 | The 24 checks |
| 03 | What the evidence shows | 06 | What to fix first |

PUBLISHED BY

Nimbus Media
nimbusmedia.io

EDITION

September 2026
24 checks, 16 pages

TIME TO COMPLETE

About 70 minutes
No paid tools

How to use this.

The boxes are live. Tick them in any PDF reader, save the file, and your progress stays put.

Read the four parts in order

Access, then rank, then mentions, then measurement. A page a crawler cannot fetch will not be cited at any rank, so the order is not arbitrary. Each part ends with checks you can run in a browser, and each part is written to make sense on its own if somebody forwards you a single page.

Tick the boxes as you go

They are live form fields, not pictures of boxes. Open this in Preview, Acrobat or most browsers, tick them, and save. Your answers stay in the file when you send it to whoever runs your site, which saves you writing the same list out again in an email.

Answer for the whole site, not the homepage

A check counts only if it holds for every page that matters to you. Half done is not done, and half done is how most sites got into this. If one service page fails a check, leave the box empty and note which page it was.

Take the failures to whoever owns the site

Most of what fails in Part One is a single line in a template or one setting in a firewall. The person who can fix it in an hour is usually not you, and the fastest thing you can do is hand them a short list with the exact setting named.

Do not buy anything until you finish

Every check here can be answered with a browser and Search Console. If you still want a tool afterwards you will at least know what you are asking it to do, which is the only way to tell a useful platform from an expensive dashboard.

Run Part Four again in ninety days

The first three parts are fixes. The last one is a habit. Twenty minutes a quarter turns a set of impressions into a trend line, and a trend line is the only thing that will tell you whether any of this worked for your firm.

Print it if that helps

Every page carries the URL in the footer, because these get forwarded and printed a page at a time. If you print it, the boxes still work as boxes. If you fill it in on screen and then print, your ticks come with it.

BEFORE YOU START

Every figure in this guide comes from a source published by Google, OpenAI, Anthropic or Perplexity, or from a named study we cite on page 10. Where two studies disagree, we show both numbers.

How an AI answer gets built.

Before a model writes a word about your firm, something has to go and find your page. That step is a search problem. It is where AI visibility is won and lost, and it is why this guide spends its first pages on crawlers rather than on writing.



38%

CITED AND RANKING TOP 10

76%

SAME FIGURE, JULY 2025

161%

LIFT FROM FAN-OUT RANK

- Retrieval: something searches for pages that might answer the question.
- Reading: the model is handed the pages it found, and only those.
- Writing: it answers from what it was handed, and names some sources.

THE SENTENCE THIS GUIDE RESTS ON

Lily Ray put it plainly in March 2026. Content that is not indexed and ranking cannot enter the model's context window in the first place.

Every tactic in this document is downstream of that one fact. So is every tactic anyone else will sell you, whether or not they say so.

It also means the reverse is true. Damage your organic rankings chasing an AI shortcut and you lose the AI visibility as well, because you have removed yourself from the step that comes first.

Retrieval comes first.

Ask ChatGPT which restructuring advisor to call in Houston. It does not know. It goes and looks, the same way you would, then writes up what it found.

The looking is retrieval, and every major AI answer product does it. That is why your search performance and your AI performance are not two performances. A page outside the index is outside the running. A page ranking tenth is in the running behind nine others.

Google describes AI Mode as breaking your question into subtopics and running many searches at once on your behalf. It calls this query fan-out. One question from a client becomes a spread of searches, so you are no longer competing for a single ranking. Ahrefs measured the effect across 76.7 million AI Overviews and found pages ranking for those spun-off questions were 161% more likely to be cited, at a correlation of 0.77.

WHICH CRAWLER DECIDES WHICH SURFACE

AI Overviews, AI Mode	Googlebot	You leave Search too
ChatGPT search	OAI-SearchBot	No ChatGPT answers
Perplexity	PerplexityBot	Not surfaced or linked
Claude search	Claude-SearchBot	Not used for results
Gemini training	Google-Extended	No effect on Search

Read that last row twice. Google-Extended is the control for training and grounding Gemini. Google states that it does not affect your inclusion in Google Search and is not a ranking signal, and Google's AI features documentation treats it as a control for its other systems rather than for AI Overviews. Firms block it, believe they have opted out of AI, and are then surprised to find themselves quoted in an AI Overview the following week.

WHY THIS GUIDE IS IN THIS ORDER

Access, then rank, then mentions. A page a crawler cannot fetch will not be cited at any rank.

A page that ranks nowhere is rarely fetched in the first place. Mentions on other people's sites come last, because they raise your odds of being named. They do not create access.

Checks 1 to 8: access.

Every one is a yes or no. Every one can be answered in a browser in the next twenty minutes.

Crawler access

Four separate switches. Most firms have never set any of them on purpose.

- ☐ 1. Your robots.txt does not disallow OAI-SearchBot
Open `yoursite.com/robots.txt` and read every line. OpenAI's documentation says sites opted out of OAI-SearchBot will not be shown in ChatGPT search answers.
- ☐ 2. Your robots.txt does not disallow PerplexityBot
Perplexity says PerplexityBot exists to surface and link sites in its results, and is not the crawler used for training models.
- ☐ 3. Your robots.txt does not disallow Claude-SearchBot
Claude-SearchBot is the search crawler. ClaudeBot is the training one. They are separate lines and separate decisions.
- ☐ 4. Your firewall does not block AI crawlers by IP
Ask whoever runs your Cloudflare or WAF whether AI bot blocking is on. It often is by default, and nobody chose it.

Preview controls

The three settings almost nobody thinks to check.

- ☐ 5. No template outputs a nosnippet robots meta tag
View source on a service page and search for nosnippet. Google says it also stops the content being used as a direct input for AI Overviews and AI Mode.
- ☐ 6. No template sets max-snippet to a fixed character count
Search the source for max-snippet. Google says the value also limits how much of your content those features may use.
- ☐ 7. No data-nosnippet attribute wraps your main answer
Search the source for data-nosnippet. It gets left on pricing blocks and disclaimers and then quietly spreads.
- ☐ 8. Your main service pages are indexed
Run URL Inspection in Search Console on each one, or search `site:yourdomain.com` and compare the count to your page count.

THE CHECK ALMOST NOBODY RUNS

max-snippet started life as a click-through tactic. Somebody set it years ago to stop Google showing too much of the page in a result.

Google now says the same value also limits how much of your content AI Overviews and AI Mode may use as a direct input. A setting made for one reason in 2019 is quietly capping you in 2026, and nothing in your analytics will report it.

Cyrus Shepard's May 2026 review of 54 studies scored preview controls 9.2 out of 10, fourth of 23 factors. It is the highest-ranked thing on his list that most firms have never looked at.

Five claims that are not true.

Each of these is being sold right now, some of it at eight thousand dollars a month. Each one is contradicted by the documentation of the company that owns the surface, or by a study large enough to argue with.



- You need a GEO budget separate from your SEO budget.
- You need an llms.txt file.
- Schema markup is the lever that gets you cited.
- Blocking Google-Extended keeps you out of AI.
- Publish enough content and the citations follow.

Two of these come from reading the wrong control. Two come from products that need to sound new to justify a price. One is simply old advice wearing a new name. None of them survives ten minutes with the documentation.

HOW TO TEST ANY CLAIM LIKE THESE

Ask who owns the surface, then read what they published about it. Google, OpenAI, Anthropic and Perplexity all document their own crawlers and their own requirements in public. None of it is a secret, and none of it costs anything.

The first three.

All three are answered by Google's own documentation, which is free and takes ten minutes to read.

1. "You need a separate GEO budget."

Google's documentation on AI features states that there are no additional requirements to appear in AI Overviews or AI Mode, and no other special optimizations necessary. Gary Illyes of Google said the same thing out loud at a Search Central event in July 2025: to get content into AI Overviews, use normal SEO practices. When the company that owns the surface says the work is the same work, a second invoice for it needs a very good explanation. Ask for one, in writing.

2. "You need an llms.txt file."

Illyes said at the same event that Google does not support llms.txt and is not planning to. Google's AI features documentation goes further and says you do not need to create new machine readable files, AI text files or markup at all. In May 2026 Cyrus Shepard reviewed 54 experiments, patents and case studies across ChatGPT, Gemini and Perplexity and scored llms.txt 2.0 out of 10, the lowest score on his list of 23 factors. He reported finding no credible evidence it changes anything. It takes ten minutes to make, which is the only reason the advice survives.

3. "Schema markup is the lever."

Google's documentation says there is no special schema.org structured data you need to add for these features, and that structured data should match what a visitor can see on the page. Shepard scored structured data 5.6, nineteenth of 23. Schema still earns rich results and still helps a machine work out what your firm is. That makes it housekeeping. It is being sold as a lever, and the gap between those two words is where the money goes.

WHAT GOOGLE'S DOCUMENTATION ACTUALLY ASKS FOR

- Let crawlers reach the page, in robots.txt and in your hosting.
- Make content findable through internal links.
- Give the page a strong experience for the person reading it.
- Keep important content in text, not locked inside images or scripts.
- Match any structured data to what is visible on the page.
- Keep Google Business Profile and Merchant Center details current.

WHAT THE THREE CLAIMS HAVE IN COMMON

Each one sells you a file or a tag. Files and tags are easy to sell because they are easy to deliver, easy to invoice and easy to show on a report at the end of the month.

Ranking a page is none of those things. That is why the market keeps inventing new files, and why Google keeps publishing documentation saying it does not need them.

And the two expensive ones.

One of these makes firms invisible by accident. The other makes them publish their way into a penalty.

4. "Blocking Google-Extended keeps you out of AI."

Google-Extended governs whether your content is used to train future Gemini models and to ground them at prompt time. Google states that it does not affect a site's inclusion in Google Search and is not a ranking signal. Google's own AI features documentation describes it as a control for its other systems, and points at nosnippet and max-snippet for limiting AI Overviews and AI Mode. So firms block Google-Extended, believe they have opted out of AI, then appear in an AI Overview anyway and conclude the controls are broken. The controls work fine. The wrong one was used. If you genuinely want less of your content in those features, the honest levers are the snippet controls, and they cost you snippets in ordinary search too.

5. "Publish more and the citations follow."

Ahrefs studied 75,000 brands in December 2025 and found almost no relationship between how much content a site has and how often AI systems name it. A separate Ahrefs analysis, this one of 76.7 million AI Overviews, put the correlation between word count and being cited at about 0.04, which is another way of writing zero. Meanwhile Google's scaled content abuse policy, published in 2024, exists precisely for sites that publish at volume. Lily Ray tracked several sites featured in AI content tool case studies through 2025 and 2026 and found the same shape each time: a rise, then a fall. In one case the articles celebrated in the case study were later removed from the site entirely.

FOUR QUESTIONS THAT KILL THE SIXTH CLAIM TOO

- Who owns this surface, and what does their documentation say?
- Would this tactic also help the person reading the page?
- What was this case study's domain ranking for the year before?
- Did this advice exist in 2019 under a different name?

WHY THE CLAIMS SELL ANYWAY

Most AI visibility case studies are run on sites that already ranked. The site was found because it was findable.

Lily Ray's phrase for this is that correlation is being sold as causation. Before you believe a case study, look at what that domain was doing in organic search the year before it started.

That single check disqualifies most of what you will be shown this year, and it takes about four minutes in any rank tracking tool you already pay for.

What the evidence shows.

Four studies published between December 2025 and May 2026, covering millions of AI answers and tens of thousands of brands. They disagree about the size of the effect. They agree about the direction.



9.5

TOP FACTOR: URL ACCESS

0.66

MENTIONS TO VISIBILITY

2.0

LLMS.TXT, LOWEST OF 23

- Ahrefs, 75,000 brands, December 2025. Which signals track with mentions.
- Ahrefs, 863,000 SERPs, March 2026. Where AI Overview citations come from.
- Cyrus Shepard, 54 studies and patents, May 2026. Twenty-three factors.

WHAT NONE OF THESE STUDIES FOUND

Not one of them found a technique that works on AI search and does nothing for organic search.

The factors at the top of every list are the factors that were at the top of the SEO list already. That is not a disappointing result. That is the result.

The factors, ranked.

Cyrus Shepard scored 23 factors across ChatGPT, Gemini and Perplexity from 54 experiments, patents and case studies, published May 2026. Top six and bottom three.

FACTOR	SCORE	RANK
Can the URL be fetched	9.5	1st
Search rank	9.4	2nd
Rank for fan-out queries	9.3	3rd
Preview controls	9.2	4th
Query and answer match	9.2	4th
Answer near the top of the page	8.8	7th
Structured data	5.6	19th
Domain authority	5.0	21st
An llms.txt file	2.0	Lowest

Ahrefs ran a different test on 75,000 brands in December 2025, measuring which signals track with being named at all. Correlation is not cause, and the study says so. But the shape is hard to ignore. Being talked about tracks with being named far more closely than being linked to does, and it is not close.

SIGNAL	CHATGPT	AI MODE	AI OVERVIEWS
Branded web mentions	0.66	0.71	0.66
Branded anchor text	0.51	0.63	0.53
Branded search volume	0.35	0.47	0.39
Domain Rating	0.27	0.29	0.33
Backlinks	0.19	0.20	0.18

Sources

Google Search Central: AI features and your website, the robots meta tag reference, and the Google crawlers reference. Gary Illyes at Search Central Deep Dive, July 2025, reported by Search Engine Land. OpenAI, Anthropic and Perplexity crawler documentation. Cyrus Shepard, Zyppy, May 2026. Ahrefs brand visibility study, December 2025, and AI Overview citation study, March 2026. Lily Ray, January and March 2026. Spearman correlations rounded to two places. The backlink figures are read from the study's chart rather than its text.

Checks 9 to 19: rank, mentions.

Access got you eligible. These decide whether you are the one who gets chosen.

Rank and structure

- ☐ 9. Every service page ranks in Google's top 10 for its main term
Search Console, Performance, filter by page, read average position.
- ☐ 10. Each service page answers one question, not five
Say the page's question out loud. If it needs an 'and', it is two pages.
- ☐ 11. The answer sits in the first 100 words
Read your opening paragraph alone. Does it answer, or introduce?
- ☐ 12. Each answer paragraph makes sense read on its own
Cut every 'as mentioned above'. A model quotes a passage, not a page.
- ☐ 13. Your headings are questions a client would actually ask
Compare your H2 headings to the questions in your last five calls.
- ☐ 14. You rank for the follow-up questions, not only the head term
Ask the head question in AI Mode, note the sub-questions it answers, then search each one.

Being nameable

- ☐ 15. Your firm name is spelled identically everywhere online
Search the name in quotes. Watch for LLC versus Ltd, ampersand versus 'and', and old office names.
- ☐ 16. At least ten sites you do not own mention you by name
Search your firm name in quotes and exclude your own domain.
- ☐ 17. Your About page says what you do, where, and for whom, in plain sentences
Read it as though you knew nothing about the firm.
- ☐ 18. Your Google Business Profile details are current
Google's AI features documentation names it as something to keep up to date.
- ☐ 19. You appear on a list, award page or directory you did not pay for
Paid listings still count as mentions. Earned ones count for more.

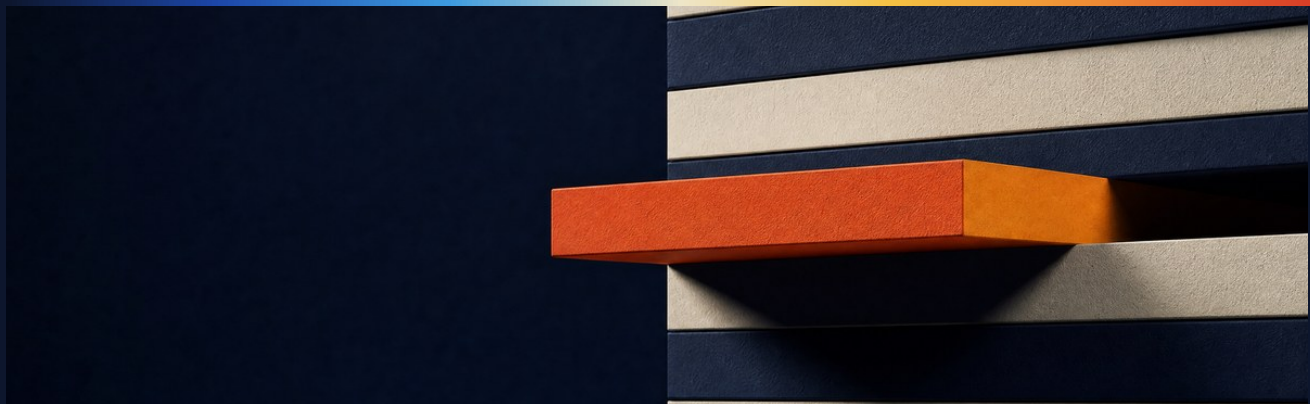
THE ONE NUMBER WORTH ARGUING ABOUT

In July 2025, about 76% of AI Overview citations came from pages ranking in Google's top 10. By March 2026, across 863,000 SERPs and four million cited URLs, Ahrefs put the same figure at 38%.

Both numbers come from the same team. The overlap is real and it is shrinking, which is the honest version of this story. Of the citations that ranked nowhere in the top 100, 18.2% were YouTube URLs. That is where check 19 and the video line on page 15 come from.

The part that is different.

Saying it is all just SEO is only nine tenths true. Four things genuinely work differently, and naming them honestly is the only way to tell them apart from the twenty things that do not.



- Crawler access is now four separate decisions, not one.
- Passages get quoted. Pages get ranked. That is a writing change.
- Being mentioned matters more than being linked to.
- There is no rank to report, so there is nothing to rank track.

That is the whole list. If somebody is selling you a fifth item, ask them which of these four it is a version of, and what happens to it when the platform notices.

WHAT THAT ADDS UP TO

Four changes. None of them needs a new budget line, a new platform or a new agency. Three of them make the site better for human readers as well, which is usually the sign that a tactic will still be working in two years.

Four things that really differ.

This is the honest delta. Everything else on the GEO menu is either SEO renamed, or a tactic with a shelf life.

1. Access is four decisions now.

Googlebot used to be the whole question. OAI-SearchBot, PerplexityBot and Claude-SearchBot are separate switches, each with its own line in robots.txt and its own rules. You can allow a search crawler and block a training crawler, and for many firms that is exactly the right answer. Perplexity's documentation notes that Perplexity-User generally ignores robots.txt, because a person asked for that page directly.

2. Passages travel. Pages do not.

A model lifts a paragraph, not a document. If your best paragraph only makes sense after the two above it, it does not travel. Shepard scored self-contained passages 8.0 and answer near the top 8.8, both in his top dozen. This is the single writing change worth making, and it happens to make the page better for the client reading it on a phone in a car park.

3. Mentions beat links.

In the Ahrefs study, branded web mentions correlated 0.66 with ChatGPT visibility. Backlinks came in near 0.19. A link is a mention with a URL attached, and the URL turns out to be the less important half. Awards, podcasts, local press and named client stories all count. So does YouTube, which was the most cited domain in AI Overviews in the March 2026 study and has grown 34% in six months.

4. There is no rank to report.

You cannot be position three in an answer that names one firm. Responses vary by person, by session and by day, and they are increasingly personalised in ways an outside tool cannot reproduce. Any product selling you a stable AI rank is selling you a sample and calling it a position. Track whether you are named and who is named instead, month after month. That is the honest metric, it costs nothing, and it is the one your competitors are not writing down.

WHERE THE AGGRESSIVE END OF THIS LEADS

Charles Floate has run AI search campaigns for over a year and teaches levers including source stacking, citation hijacking and prompt injection through hidden text. Some of it works, for a while.

In February 2026 Microsoft documented over 50 examples of hidden prompt instructions from 31 companies across 14 industries and classified the practice as prompt injection. Google classifies mass comparison pages as scaled content abuse. Both were working tactics right up until they were named.

The question to ask about any of it is not whether it works today. It is what happens to your firm's name when it stops.

Checks 20 to 24: measurement.

You do not need a GEO platform to know where you stand. You need five questions and twenty minutes a month.

Your monthly twenty minutes

Five checks, run monthly. Written down, they become a trend line instead of a feeling.

- ☐ 20. You have asked ChatGPT, Gemini and Perplexity your five most valuable client questions this month
Use the words a client would use, not the words your industry uses. Log the date, the answer and who got named.
- ☐ 21. You have written down which competitor gets named when you are not
That name is your gap analysis. Go and read the page that got cited.
- ☐ 22. You can name the page that was cited when you were cited
Click the citation. It is often not the page you would have guessed, and that tells you where to put the next hour.
- ☐ 23. Your analytics separates referrals from chatgpt.com, perplexity.ai and gemini.google.com
Add them as a channel group in GA4. Most setups bury them inside Direct, where they are invisible.
- ☐ 24. You know AI referral traffic as a share of your total
Get the number before you form an opinion about the number.

THIS MONTH'S LOG

THE QUESTION YOU ASKED	WHO GOT NAMED	PAGE CITED	DATE
Example: who handles Chapter 11 in Houston	A rival firm	Their blog	03 Sep

WHAT THAT NUMBER WILL PROBABLY SAY

In January 2026 Lily Ray reported Google holding over 90.6% of global search market share, with AI platforms accounting for under 1% of global referral traffic. In her own poll, 70% of respondents said ChatGPT sent them less than 2% of their site traffic.

That is not an argument for ignoring AI search. It is an argument for not moving your budget out of the thing that still sends the other 98% of it.

What to fix first.

In order. The first two cost nothing and take an afternoon between them. Do not start at number four because it is the interesting one.

ORDER	DO THIS	WHO	TIME
1	Read robots.txt and the firewall rules	Developer	1 hour
2	Strip nosnippet and max-snippet from templates	Developer	1 hour
3	Rewrite the opening of five service pages	You	Half a day
4	Lift the pages already at positions 4 to 15	SEO	Ongoing
5	Get mentioned: awards, lists, press, podcasts	Marketing	Ongoing
6	Put your best explanation on video	Marketing	Ongoing

One and two are the only items here that can be silently costing you everything. A firewall rule blocking OAI-SearchBot removes you from ChatGPT search answers completely, and nothing in your analytics will report that it happened. You will simply never appear, and you will never know why. Check both before you spend a dollar on anything else in this document.

Three is the only writing task on the list, and it is half a day. Four, five and six are the rest of your year. Notice that four, five and six are also just SEO, content and public relations, which is the point the whole guide has been making.

THE UNCOMFORTABLE PART

Nothing on that list is new. Every item was good SEO in 2019 and is good SEO in 2026.

That is the finding, not a disappointment. The firms getting named in AI answers are not doing a different job. They are doing the same job, and they started earlier.

Worth a conversation?

You can run all 24 checks yourself, and you should. Nothing here was held back to make you call us. If you want a second pair of eyes on what you find, that is what the call is for.

What the call covers

Thirty minutes. You tell us which checks failed. We tell you which are actually costing you traffic, which can wait a quarter, and roughly what each is worth fixing.

Who tends to book it

Firms between 20 and 200 people whose site was built once, three years ago, and has been quietly drifting since. Usually somebody has just asked them what they are doing about ChatGPT.

How we actually work

We build in Webflow and we run the search and paid work around it. If your problem is a firewall rule, we will tell you to go and fix the firewall rule.

What it is not

Not a pitch deck, and not a GEO retainer. If the answer is that your site is fine and you simply need to be mentioned in more places, we will say so and you keep your money.

What you should bring

Your domain, the checks that failed, and the five client questions from check 20. Twenty minutes of preparation makes the call worth twice as much to you.

If the answer is no

Plenty of firms read this, fix items one and two, and never speak to us. That is a good outcome. The guide was written to be worth the download on its own.

BOOK A FIT CALL

Thirty minutes at cal.com/nimbus-media/30min, or just reply to the email this arrived in. We are in Houston, and we work with professional services firms across Texas and beyond.

team@nimbusmedia.io

nimbusmedia.io



September 2026 edition