

FIELD GUIDE · SEPTEMBER 2026

Answer engine optimisation: how to get your business quoted by AI

Fewer than one in three Google searches now sends a click to a website. This guide sets out what the evidence actually says about becoming citable inside ChatGPT, Google's AI Overviews, Gemini, Claude, Perplexity and Copilot, and which popular tactics the data does not support.

PREPARED BY

Whitehat
Agency

EDITION

September
2026

WRITTEN FOR

Australian business
owners

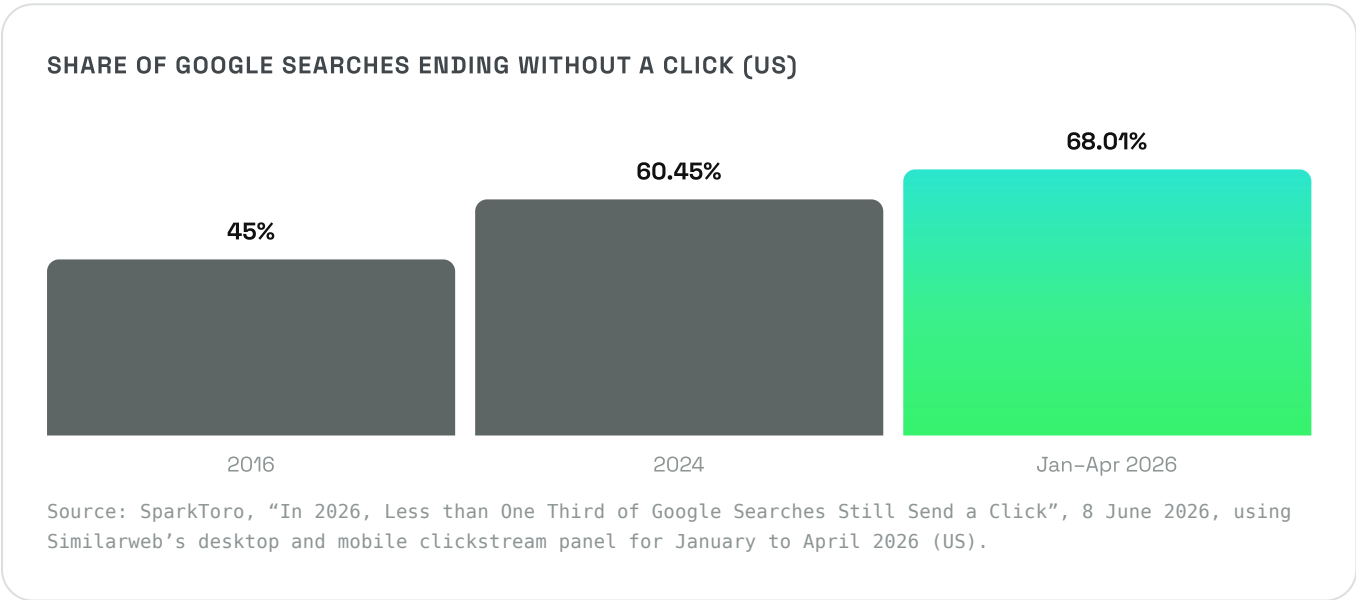
EVIDENCE

Sources listed on pages 23 and
24

Search stopped sending clicks, and the answer moved above the results

In the first four months of 2026, 68.01% of Google searches in the United States ended without a click to the open web. In 2024 that figure was 60.45%, and ten years ago it was around 45%. Fewer than one in three searches now sends a visitor to a website at all.

The cause is visible on the page. Google’s AI Overviews now appear on roughly half of tracked queries, up from about 30% a year earlier. Pew Research followed 68,000 real searches and found people clicked a result on 8% of those carrying an AI Overview, against 15% of those without one.



58%

of Australians aged 14+ used an AI tool in an average four weeks, March quarter 2026 (13.6 million people)

2.5b

monthly users of Google’s AI Overviews, by Google’s own count

1b

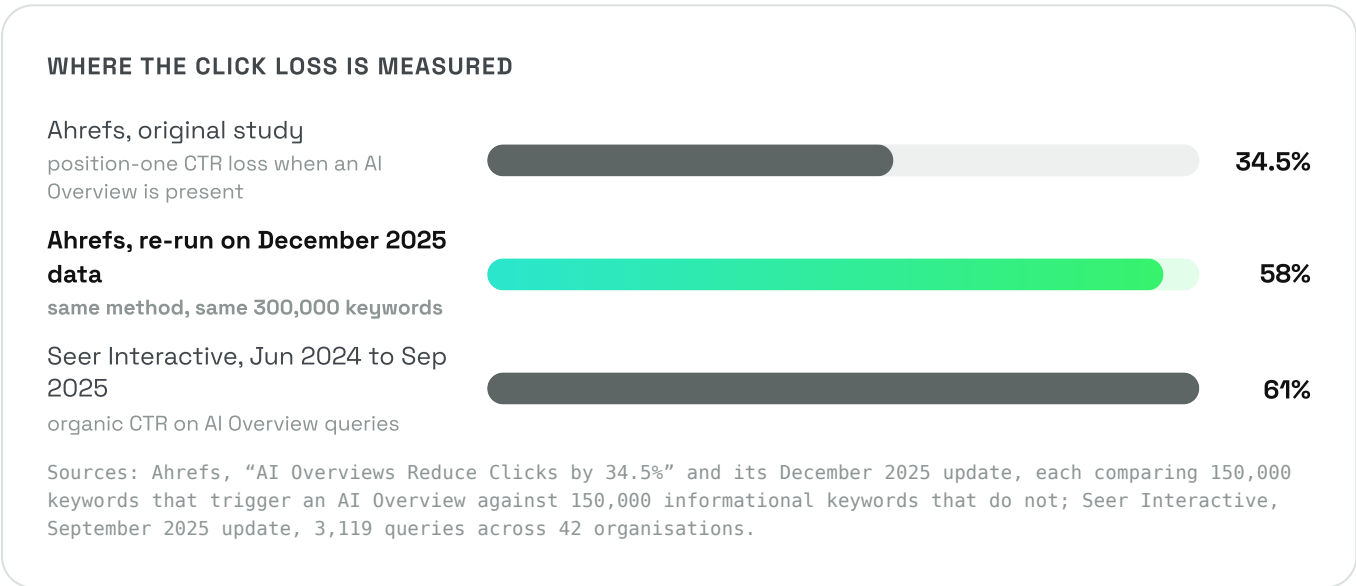
monthly users of Google’s AI Mode, passed in 2026

Sources: Roy Morgan Single Source, January to March 2026, n=14,646; Google, “New opportunities, control and insights for website owners”, 2026; Pew Research Center, 22 July 2025; BrightEdge AI Overview coverage tracking, 2026.

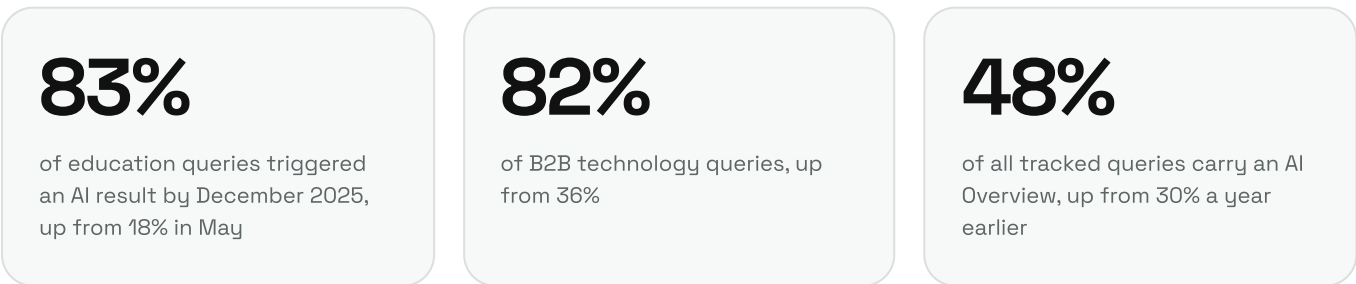
The traffic is small today, and each visitor is worth several ordinary ones

Two things are true at once, and most coverage of this topic reports only one of them. AI referrals are still a low single-digit share of total traffic for most websites. The visitors who do arrive convert far better than average.

Semrush puts the average AI search visitor at 4.4 times the worth of a traditional organic search visitor. The reason is intent. Somebody arriving from an AI answer has already had the basic question resolved and is following a citation to verify, go deeper, or buy.



We report the range rather than the headline because the studies disagree, and any agency quoting a single number to you has picked the one that suits the argument. The honest summary: the click loss on affected queries is large, it is getting larger, and it varies by industry.

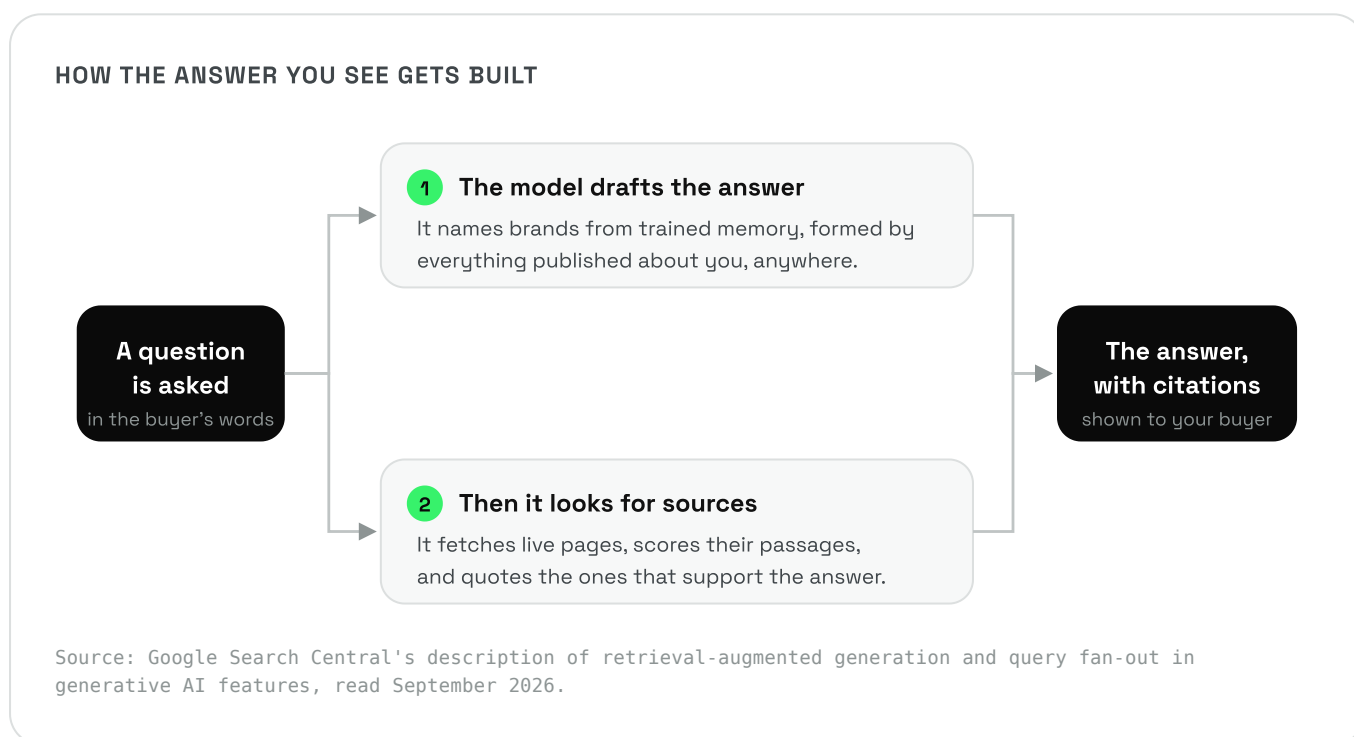


Source: BrightEdge AI Overview coverage tracking, 2025 to 2026. Healthcare, education, B2B technology and insurance lead; check your own vertical before setting a strategy.

An AI answer is assembled in two separate steps, and they reward different work

Published descriptions of how these systems work set out the sequence this way: the model drafts its answer first, choosing which brands to name from what it already knows, and only then runs a retrieval step to find sources that support those choices. The citations are the bibliography, not the brainstorm.

That single finding explains why so much AEO advice contradicts itself. Being named and being cited are two different outcomes, earned by two different bodies of work.

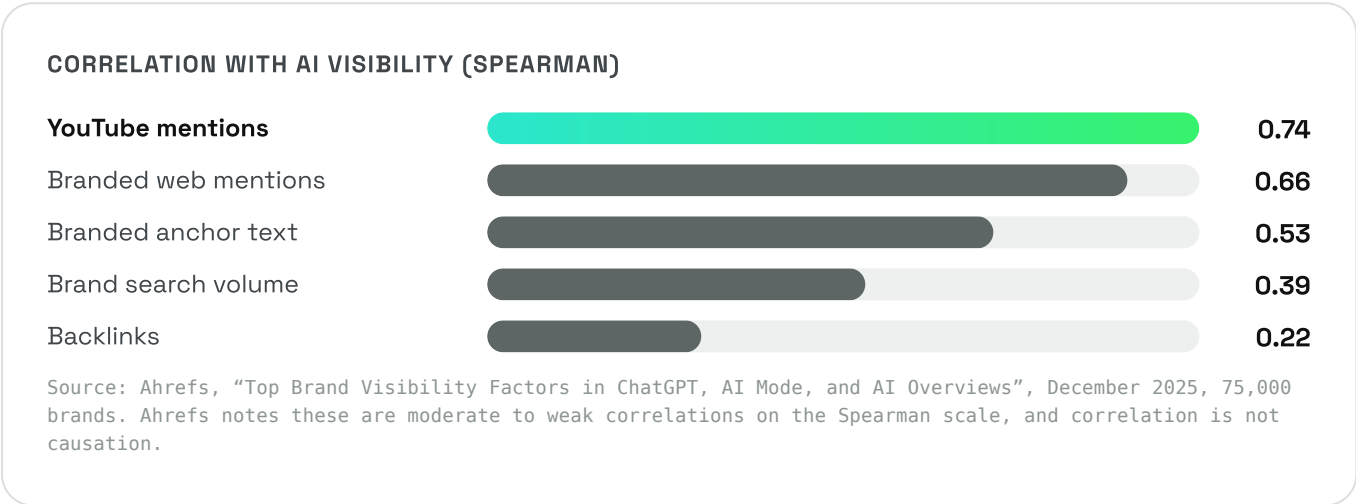


What this means for the work

Step one is won away from your website, in the places that write about you. Step two is won on your website, by being fetchable and quotable. A business that does only one of the two gets named without a link, or gets crawled without ever being mentioned.

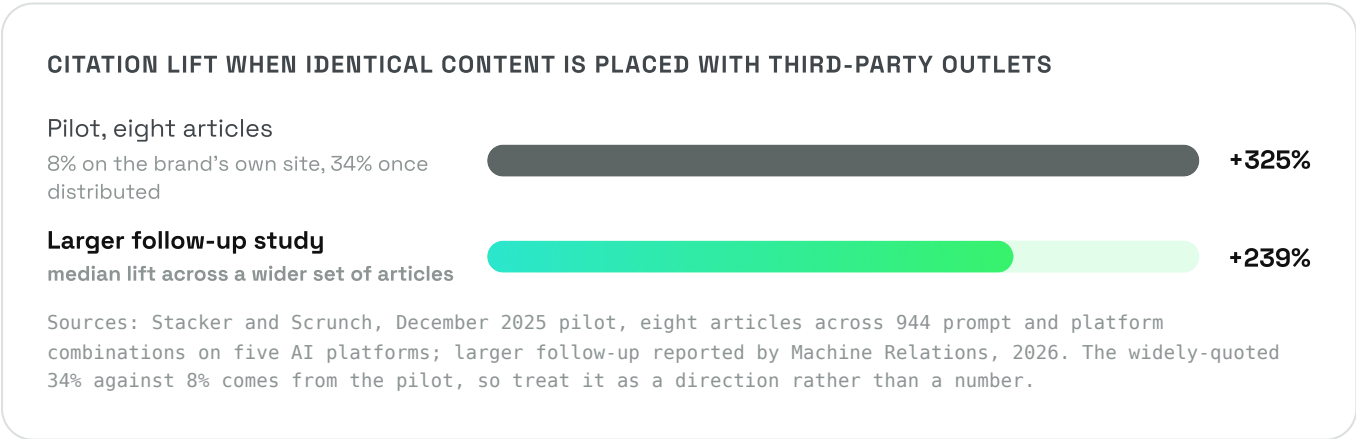
Across 75,000 brands, the strongest signals sit outside your website

Ahrefs measured which signals track most closely with a brand appearing in AI answers, across 75,000 brands. The three strongest are all off-site, and branded web mentions correlate roughly three times more strongly than backlinks.



Where the same content is published changes how often it is quoted

Two studies have tested this directly and disagree on the size of the effect. Both are shown, because the smaller number is the one with more evidence behind it.



Backlinks still matter for ordinary search rankings, and nothing here says stop building them. What the data says is that a link and a mention are no longer the same currency. A model appears to weigh being talked about above being linked to, and most marketing budgets are still allocated the other way around.

The only controlled test of page-level tactics found five that moved citations

Researchers at Princeton built a benchmark of roughly 10,000 queries across nine datasets and tested content changes against generative engines. It remains the most rigorous published work on what a page owner can change. Visibility rose by 22% to 41% depending on the tactic and the query set.

WHAT WAS CHANGED ON THE PAGE	EFFECT ON CITATION VISIBILITY
Citing sources	Among the three strongest, in the 30% to 40% band. Produced a 115.1% lift for a site ranked fifth in Google, and a 30.3% decrease for pages already ranked first.
Adding direct quotations	Among the three strongest, in the 30% to 40% band.
Adding statistics	Among the three strongest, in the 30% to 40% band.
Improving fluency and readability	Positive, smaller effect.
Writing in an authoritative voice	Positive, smaller effect.
Keyword stuffing	No benefit. Slightly negative on Perplexity.

Source: Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan and Deshpande, "GEO: Generative Engine Optimization", ACM SIGKDD, 2024. Read the paper before quoting the 40% figure: it is an upper bound, not an average.

Read that table the right way

The tactics that worked are the ones that make a passage independently checkable: a number, a quote, a named source. A model has no way to verify a claim written as an adjective, so it skips it and quotes the page that gave it something concrete. This is the same instinct a careful journalist has, expressed as software.

Note the second column on citing sources. The lift went to the page ranked fifth, not to the page ranked first. For a business that is not already the biggest name in its category, the evidence here is unusually encouraging.

Write so a single paragraph can be lifted out and still make sense

A retrieval system quotes a passage, not a page. It scores the parts of your page and lifts the part that wins, so a paragraph depending on the three above it is hard to quote however good it is. Google is explicit that no page length is required and no chunking is needed, so this is about writing clearly rather than reformatting.

Seven rules that follow from that

RULE	WHY IT WORKS
Answer in the first 40 to 60 words under every heading	The opening of a section is what gets scored and lifted. A conclusion at the bottom of 1,200 words is rarely what gets quoted.
Phrase each heading as the question a customer asks	People type whole questions into an assistant rather than the two or three keywords they give a search box, so the heading that matches is a question.
Make every passage self-contained	Repeat the subject rather than saying "it". A chunk with an unresolved pronoun is a chunk that cannot be quoted.
Use a number instead of an adjective	"Three boutique wineries" is checkable. "A selection of excellent wineries" is not, so it gets skipped.
Put comparisons in a table	A table is a clean, discrete block. Prose comparisons get mangled or dropped.
Put a real named author on the page	A named, credentialed human is the clearest signal of genuine first-hand experience behind the writing.
Publish something only you could publish	Your own data, photographs and process. Google singles out non-commodity content as the thing that matters most over time.

Source: Google Search Central, "Optimizing your website for generative AI features on Google Search", 15 May 2026. The rules above are ours, drawn from that guidance and from how retrieval systems select passages.

What Google itself says

Google's own guidance is blunter than most agency advice: write for people rather than machines. No special schema, no separate file and no AI-specific rewrite is required. What it asks for is content only you could have written.

If a crawler cannot read it, none of the writing matters

Vercel and MERJ analysed more than 500 million GPTBot fetches and found no evidence of JavaScript being executed. AI crawlers download the raw HTML and read that. Googlebot renders JavaScript. The AI crawlers do not.

So product descriptions, pricing tables, FAQ answers and comparison content that only appear after JavaScript runs are invisible to ChatGPT, Claude and Perplexity, while the same page still ranks normally in Google. A business can be doing well in search and be absent from AI answers for this one reason.



Source: Vercel and MERJ, “The rise of the AI crawler”. ChatGPT and Claude’s crawlers do download JavaScript files, roughly 11.5% and 23.8% of their fetches, and never run them.

The checklist

CHECK	WHAT GOOD LOOKS LIKE
Rendering	Server-side rendered or static HTML. View the page source, not the inspector, and confirm the words are in it.
Semantic HTML	Real headings, lists and tables. Google recommends it, and it is what makes a passage extractable.
Bing indexing	Bing’s index is the retrieval layer behind ChatGPT search and Microsoft Copilot. A page missing from Bing is invisible to both.
IndexNow	Push new and updated URLs to Bing on publish. Google has tested IndexNow and does not support it.
Duplicate content	Reduce it. Google lists this among its generative AI recommendations.

Sources: Google Search Central generative AI guidance, May 2026; Bing Webmaster Tools and IndexNow documentation, 2026.

Decide deliberately which AI crawlers you let in

Training crawlers and retrieval crawlers are different bots doing different jobs, and most websites block them by accident rather than by decision. Blocking a training crawler keeps your content out of a model. Blocking a retrieval crawler removes you from the answers that model gives today.

USER AGENT	OPERATED BY	WHAT IT DOES	CALL
GPTBot	OpenAI	Collects content for model training	Yours
OAI-SearchBot	OpenAI	Builds the index ChatGPT search cites from	Allow
ChatGPT-User	OpenAI	Fetches a page live when a question needs it	Allow
ClaudeBot	Anthropic	Collects content for model training	Yours
Claude-SearchBot	Anthropic	Search indexing for citation	Allow
Claude-User	Anthropic	Live fetch on a user's request	Allow
PerplexityBot	Perplexity	Search indexing for citation	Allow
Google-Extended	Google	Gemini training. Does not affect Search or AI Overviews	Yours
CCBot	Common Crawl	Open dataset used in training corpora	Yours
Bingbot	Microsoft	Powers Bing, Copilot and ChatGPT search retrieval	Allow

Sources: vendor crawler documentation, read September 2026; Cloudflare, "New options to manage AI traffic", 1 July 2026. Platform behaviour changes. Re-check before relying on this table.

Two traps worth knowing about

One line per bot. Blocking ClaudeBot leaves Claude-SearchBot and Claude-User untouched. Every user agent needs its own rule.

Your host may decide this for you. From 15 September 2026, domains new to Cloudflare are set up to block training bots and live fetch bots on any page carrying ads, with search crawlers still allowed. If your site sits behind Cloudflare, open the setting and read which of the three presets is active rather than assuming nobody has touched it.

Google now has a switch for AI answers, and it is on by default

Google’s rule for appearing in AI Overviews and AI Mode has two parts. The page must be indexed and allowed to show a snippet, and the site must be included in the Search generative AI control. That control went live to every website in the world on 31 August 2026, and it sits in Search Console under Settings.

CONTROL	WHAT IT DOES	WHAT IT COSTS
Search generative AI control	Decides whether your site can appear in AI Overviews, AI Mode and generative AI features in Discover	Nothing in ordinary search. Google states it is not used as a ranking or inclusion signal elsewhere in Search
<code>nosnippet</code>	Stops any text being shown from the page, in ordinary results and AI answers alike	Your ordinary search snippet
<code>data-nosnippet</code>	Excludes one marked section of a page	Only that section
Google-Extended	Opts out of Gemini model training	Nothing in Search or AI Overviews
<code>noindex</code>	Removes the page from Google altogether	All search traffic

Source: Google Search Console Help, “Search generative AI control”, rolled out to all websites worldwide on 31 August 2026; Google Search Central, “Optimizing your website for generative AI features on Google Search”.

The check worth doing this week

Every property starts on **include**, so most businesses are already eligible and have nothing to change. The reason to look anyway is inheritance: a property follows its parent’s setting by default, so a site can be switched off by a decision taken on a parent domain nobody thought of as yours.

Where to look

Search Console, then Settings, then Search generative AI. Confirm it reads “Include my site’s links and content in Search generative AI features” rather than an inherited setting you did not choose.

Structured data is infrastructure, not a growth lever

Schema markup is the code that tells a machine what a page is about: this is a business, that is a review, this number is a price. It is worth having, and it is the most oversold tactic in this field.

1,885

pages tracked by Ahrefs as they added structured data, between August 2025 and March 2026

No lift

in citations on any platform, measured across ChatGPT, AI Mode and AI Overviews

Source: Ahrefs, “We Tracked 1,885 Pages Adding Schema. AI Citations Barely Moved.” A widely-repeated claim that most cited pages carry schema has no traceable source, so it is left out.

A high share of cited pages does carry schema, and that is correlation rather than cause. Sites that add schema are also doing the technical, content and authority work that earns the citation.

What we still recommend, and why

TYPE	REASON TO IMPLEMENT
Organization and LocalBusiness	Anchors your entity. Name, address, phone and identifiers stated once, matching every other profile you own.
Product, Service, Event	Still powers rich results in ordinary search, which is where most of your clicks still come from.
Article with author	Attaches a named, credentialed human to the writing.
FAQPage	Google deprecated FAQ rich results on 7 May 2026 and still uses the data to understand pages. Keep it where a question and answer format fits.

Sources: Google Search Central announcements, May 2026; schema.org type registry.

One caveat matters more than the schema debate. When a model fetches your page live it reads what a person would see. Testing found ChatGPT, Claude, Perplexity, Gemini and AI Mode all ignored information that existed only in hidden markup, so facts you want quoted belong in the visible words.

Most citations come from somewhere other than your own domain

Ahrefs tracks every domain the major assistants cite and republishes the ranking monthly. The pattern is consistent: the sources these systems lean on hardest are platforms, not company websites. In Google AI Mode the top two domains alone take more than a third of the citations among its leading sources.



What this changes about the work

The platforms carrying these citations are ones you can be present on without owning them. A YouTube channel, honest participation in the communities your customers use, a Wikipedia entry that survives editorial scrutiny, and accurate retail listings are all inside the citation pool. Your own website is one input among many, and for most businesses it is not the input carrying the most weight.

The engines disagree, and that is the useful part

Reddit leads two of those three lists and appears nowhere in Copilot’s top fifty. Wikipedia is the only domain in all three. A business visible in one assistant can be invisible in another, which is why a single AI visibility score from one tool tells you very little on its own.

ENGINE	WHAT ITS CITATION MIX SUGGESTS IT FAVOURS
Google AI Mode and AI Overviews	Its own ranking and quality systems, applied through retrieval and query fan-out, with a heavy tilt towards user-generated platforms.
Perplexity	Discussion and video, then encyclopedic reference. Its top two sources are the same two as AI Mode, at even greater concentration.
Microsoft Copilot	Retail and transactional sources. Its list reads like a shopping index, which matters a great deal if you sell products.

Sources: Ahrefs Brand Radar monthly citation studies, September 2026; Google Search Central's description of retrieval-augmented generation and query fan-out, 2026.

Which one matters to you

IF YOUR BUSINESS IS	WATCH, AND WORK ON
Selling products	Copilot, where retail sources dominate. Accurate product listings, a clean Merchant Center feed and presence on the marketplaces already cited do more than blog posts.
Selling professional services	ChatGPT and Perplexity, where reference, editorial and discussion sources lead. Being written about and being discussed is the lever, as page 5 sets out.
Serving a local area	Google’s AI surfaces, which lean on its own local systems. Your Business Profile and your reviews are the material they read, covered on page 16.

Our reading of the citation data above, together with Google's own guidance naming Merchant Center and Business Profile as inputs to its generative AI features.

Five plays that build the half of this you cannot control directly

PLAY	WHAT IT LOOKS LIKE IN PRACTICE
Earned media	Genuine coverage in publications your industry reads. This is the play with the largest measured effect: the same content placed with third-party outlets was cited several times more often than when self-published (page 5).
Video	YouTube mentions carry the single strongest correlation Ahrefs measured, at 0.74. A short explanation of what you do, published properly and described in text, is inexpensive and durable.
Being discussed honestly in communities	Answer questions in the places your customers already ask them, as yourself. Do not manufacture threads. Planting synthetic discussion is detectable, reversible, and a reputational risk that outlives any short-term gain.
Entity consistency	Your name, address, phone number, description and service list should read identically on your website, Google Business Profile, directories and every social profile. Contradictions make a model less confident about naming you.
Third-party review platforms	Independent sources describing your business in words a model can read and attribute to somebody other than you.

Sources: Ahrefs, “Top Brand Visibility Factors in ChatGPT, AI Mode, and AI Overviews”, December 2025, 75,000 brands; Stacker and Scrunch distribution pilot, December 2025, and the larger follow-up reported by Machine Relations, 2026.

Why this order

The list is ranked by measured effect, not by how comfortable each one is. Most businesses start at the bottom because review platforms and profile tidying are easy, then never reach the top two. The evidence points the other way. One piece of genuine industry coverage is worth more than a quarter of directory listings.

The uncomfortable version, said plainly

If you have budget for one thing this quarter, spend it on being written about somewhere other than your own website. Everything else in this guide compounds that, and none of it substitutes for it.

The one lever Google hands you, and almost nobody uses it

Google lets a person nominate the sites they want to see more of. Once somebody picks you, your content is more likely to appear for them in Top Stories, and in AI Mode and AI Overviews it can carry a "preferred" badge. The feature is live in every language and locale where those features run, so an Australian business can use it today.

What it is, said plainly

This is not a ranking factor. Nothing about it changes what a stranger sees. It lifts you for people who have already chosen you, which makes it an audience loyalty lever rather than an SEO one. Treat any agency selling it as a way to outrank competitors with suspicion.

How to use it

STEP	WHAT TO DO
Check you are eligible	Only domain-level and subdomain-level sites can appear in the tool. A site at example.com or blog.example.com qualifies. A subdirectory such as example.com/blog does not.
Add the button	Google publishes a two-line JavaScript snippet that renders a localised, Google-styled button. A custom design option exists if your brand needs one.
Or link to it directly	Point a link at Google's preferences tool with your own domain in the query, in the form <code>google.com/preferences/source?q=yourdomain.com.au</code>
Ask the people who already like you	Your newsletter list, your repeat customers, the people who open your emails. This lever rewards an existing audience rather than building a new one.

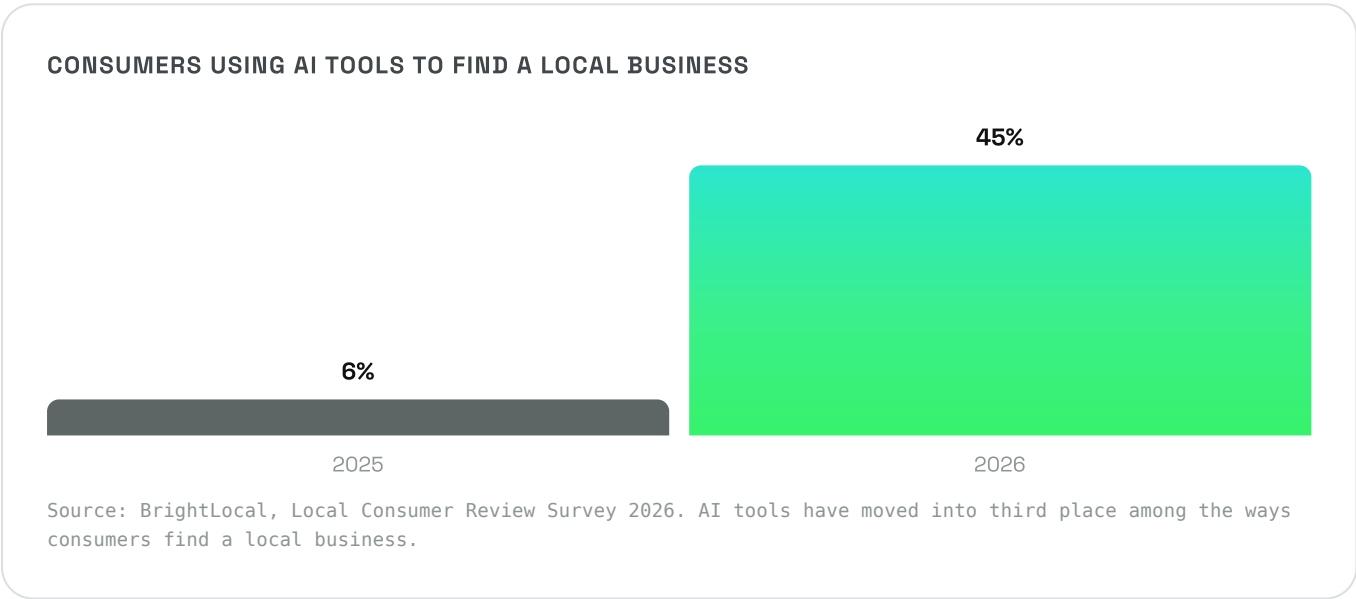
Source: Google Search Central, "Help your readers find your site through preferred sources in Google Search", read in full September 2026.

Worth being realistic about the size of this. It moves what your existing audience sees, not what the market sees, and the number of people who will go and set it is small. It is on this list because it costs an hour, it is first-party, and it is one of the few switches in AI search that a business owner can flip themselves.

Your reviews are now source material for a recommendation

The change here has been fast. BrightLocal’s 2026 Local Consumer Review Survey found 45% of consumers now use AI tools to find local businesses. A year earlier that figure was 6%.

AI systems read a local business as one data stream: the Google Business Profile, the website, the reviews and the social profiles, together. What the reviews actually say is now source text, not just a star count, and consumers themselves discount older ones.



68%

of consumers will only use a business rated 4 stars or higher

88%

of consumers fact-check the reviews an AI tool quotes at them

90 days

how recent consumers expect a review to be before they count it

Source: BrightLocal, Local Consumer Review Survey 2026. Only 18% of people who get an AI recommendation feel ready to contact that business without checking further.

The practical version is unglamorous and it works. Ask every satisfied customer for a review, every month, without a gap. Reply to all of them in words that describe what you actually did, because those replies are also text a model reads. Keep your opening hours, service list and service area accurate everywhere they appear.

Freshness moves ChatGPT, and does almost nothing for Google

Ahrefs compared the age of pages cited by AI assistants against the age of pages ranking in ordinary Google results, across 17 million citations. AI assistants cite pages that are 25.7% fresher on average. That is the headline, and on its own it would send you publishing every week.

The useful finding is underneath it. The preference is not shared evenly, and the engine your buyers use decides whether a publishing habit is worth paying for at all.

AVERAGE AGE OF A CITED PAGE, IN DAYS — LOWER MEANS FRESHER

Cited by AI assistants

ChatGPT, Perplexity, Copilot and Gemini averaged



1064

Ranking in ordinary Google results



1432

Source: Ahrefs, "AI Assistants Prefer to Cite Fresher Content", 17 million citations analysed.

Ahrefs found ChatGPT shows the strongest preference for fresh content of any assistant, while **Google's AI Overviews and ordinary search results are the most likely to cite older pages**. So a publishing cadence is a real lever in ChatGPT and close to irrelevant in AI Overviews. Anyone selling you a content schedule as the route into AI Overviews is selling against the evidence.

What updating actually means

Changing a date and republishing fools nobody and costs trust. A real update adds this year's numbers, removes what is no longer true, answers the questions customers have started asking since you wrote it, and says plainly what changed. Recency only helps on questions with a time dimension. An evergreen explanation does not decay, and rewriting it on a schedule is wasted effort.

A workable rhythm for most businesses

Review your ten highest-value pages every quarter. Update the three that have moved. Leave the rest alone until a price, a process or a fact on them changes.

Six tactics being sold right now that the evidence does not support

This field is young and the advice market has run ahead of the research. Each row below is something a client has asked us about, with what is actually known about it.

THE CLAIM	WHAT THE EVIDENCE SAYS
"Add an llms.txt file so AI can find you"	Google's Search team does not use or endorse it, and John Mueller has said it cannot help a model tell one website from another. Ahrefs looked at 137,000 domains and found 97% of published llms.txt files received zero requests in May 2026.
"Rewrite your content for AI"	Google's guidance includes a mythbusting section dismissing AI-specific rewriting, chunking files and special machine-readable formats as unnecessary. It asks you to write for people.
"Repeat the keyword so the model picks it up"	Tested in the Princeton work. No benefit, and slightly negative on Perplexity. These systems read for meaning, not density.
"Submit your site to the AI engines"	There is no submission process for any of them. Anyone selling one is selling nothing.
"Buy mentions, reviews or Reddit threads"	Fabricated third-party signal is detectable and reversible, and platforms remove it. It also breaches the Australian Consumer Law provisions on misleading conduct and fake testimonials.
"We guarantee you a place in AI answers"	No ranking factors are published for any generative engine, there is no submission route, and the output is generated fresh per query. A guarantee is the clearest signal you are being sold something else.

Sources: Google Search Central guidance and Google statements on llms.txt, 2025 to 2026; Ahrefs llms.txt analysis, May 2026; Princeton GEO paper, 2024.

We hold ourselves to the same line. What Whitehat Agency sells is a measurement of where you stand today, work that demonstrably makes your pages more citable and your brand more mentioned, and the same measurement again in ninety days. Every one of those is something we can show you on a report.

Nearly everything here can be measured, and one thing cannot

Measurement here is partial, and any agency saying otherwise is guessing. Five sources report real data. One number is invisible by design.

WHAT YOU WANT TO KNOW	WHERE IT COMES FROM	HONEST LIMITATION
Visits from ChatGPT, Gemini, Copilot, DeepSeek and Grok	Google Analytics 4. Its AI Assistant channel, added 13 May 2026, sets the medium <code>ai-assistant</code> from the referrer.	Not retroactive. Earlier sessions stay in Referral, so a comparison crossing that date looks distorted.
Clicks from Google AI Overviews and AI Mode	Nowhere. Google states its AI Assistant channel excludes both.	The referrer arrives as ordinary organic search, so these clicks cannot be separated out.
Appearances in AI Overviews and AI Mode	Search Console’s Generative AI performance report, live worldwide since 31 August 2026. Impressions, pages, countries, devices and dates.	Reports visibility rather than what people did next, so it complements analytics.
Citations in Microsoft Copilot	Bing Webmaster Tools AI Performance report.	Covers the Bing and Copilot side only, which is also the ChatGPT search retrieval layer.
Whether the AI crawlers are visiting at all	Your own web server access logs, filtered to the user agents on page 9. First-party, free, and rarely looked at.	Shows the visit rather than what came of it. Some hosting plans make logs awkward to reach.
Whether AI names you when asked about your category	Brand monitoring. We use Ahrefs Brand Radar: the answers a brand appears in, the pages cited, and share of voice against named competitors.	Sampled rather than exhaustive, and no external tool sees inside Google. Read it as a trend line.

Sources: Google Analytics 4 channel definitions, read September 2026; Google Search Console Help, generative AI performance reports, rolled out worldwide 31 August 2026; Bing Webmaster Tools documentation, 2026.

The measurement that decides the budget

Set a baseline before anything starts. Ask the engines the twenty questions your buyers ask, record who gets named today, then record it again in ninety days. That comparison beats any single traffic number.

Ninety days, in the order that matters

Every item below is drawn from the evidence in this guide. The sequence is deliberate: the technical floor first, because nothing above it counts until a crawler can read the page. The first five are foundations, and none of them needs a budget approval.

1 Confirm your pages are readable without JavaScript

View the page source on your five most valuable pages and check the words are in the HTML. If they are not, this is the whole project until it is fixed.

2 Get into Bing, and turn on IndexNow

Verify in Bing Webmaster Tools, submit the sitemap, push new URLs on publish. This is the retrieval layer behind ChatGPT search and Microsoft Copilot.

3 Decide your crawler policy on purpose

Allow the retrieval bots. Make a considered choice on the training bots. Write a separate line for each user agent, because a rule for one is ignored by the others. While you are there, confirm Search Console's generative AI control reads "include", and check what your host's own crawler settings are doing (pages 9 and 10).

4 Baseline where you stand

Record who the engines name today for your twenty most commercially important questions, before any work starts. Without this you cannot tell later whether anything worked.

5 Rewrite the openings of your top ten pages

The question as the heading, a direct 40 to 60 word answer underneath, a real number in it, and no pronoun that depends on the paragraph above.

None of the first five needs a budget

Steps one to four are checks. Step five is an afternoon of editing by whoever writes your website copy. If your pages fail step one, stop and fix that before spending a dollar on anything else in this guide, because nothing above it counts until a crawler can read the words.

The second five build the half you cannot control directly

Steps six to ten move from your website to your reputation. They take longer, they cost more, and on the measured evidence they carry more weight. Start them in parallel with the first five rather than after them.

6 Put a named author and real evidence on your best content

A byline with genuine credentials, your own data, your own photographs, and sources cited inline. Citing sources was among the three strongest tactics in the Princeton test (page 6).

7 Fix your entity across every profile

Identical name, address, phone, description and service list on the website, Google Business Profile, directories and socials. Contradictions make a model less confident about naming you.

8 Start a review habit and keep it running

Ask every satisfied customer, every month, without a gap. Reply in words that describe the work you did, because those replies are also text a model reads.

9 Earn coverage somewhere other than your own website

One genuine piece of industry media or one properly published video beats another page on your blog, by the measured margin on page 5.

10 Measure again, and change one thing

Re-run the baseline at ninety days. Keep what moved, drop what did not, and resist adding five new tactics at once. One change at a time is the only way to learn what your market rewards.

How long before any of this shows

Faster than search rankings, and slower than anyone selling it will tell you. Retrieval engines re-read pages on their own cycle, and a refreshed page can move from invisible to cited within a single index cycle, which is days rather than months. The off-site half moves on a different clock. A mention published this month is a signal a model may only carry into its answers after its next training or index refresh.

So expect the technical and on-page work to show first, and judge the reputation work at ninety days rather than at thirty. That is the reason step ten exists, and the reason we baseline before anything starts.

NEXT STEP

Find out where you stand before you spend anything

Your AI Search Score measures where your business stands in AI answers today. It is free, it runs on your real website, and we walk you through what it found.

WHAT YOU GET

WHAT IS IN IT

The twenty questions

The questions your buyers actually ask an AI before they choose a supplier in your category, written for your business rather than a template.

Who gets named today

Which businesses the engines recommend on those twenty questions right now, and where you sit among them.

Your score against the ten steps

Which of the ten steps on pages 20 and 21 your website already passes, and which one is costing you the most.

One recommendation

The single change we would make first, with the reason and the expected effect. Not a list of thirty.

1

Enter your website on the AI Search Score page

A short form. It takes about a minute, and the audit starts straight away.

2

We run the twenty questions across the major engines

ChatGPT, Google AI Overviews and AI Mode, Gemini, Perplexity and Copilot, against your real site.

3

We walk you through the result on a call

Thirty minutes with a strategist. The full report is yours afterwards, to keep and act on.

Call 1800 465 300 or email hello@whitehatagency.com.au

Whitehat Agency has run on a white-hat-only rule since 2013, with zero penalties and over \$650M in attributed revenue generated for clients. We are a Google Premier Partner and a Meta Business Partner.

Every source behind the numbers in this guide

We publish our sources for the same reason the Princeton research found it works: a claim a reader can check is worth more than one taken on faith. Where studies disagreed we showed the disagreement rather than picking the flattering number.

SOURCE	WHAT WE USED IT FOR	PAGES
Google Search Central	"Optimizing your website for generative AI features on Google Search", 15 May 2026, read in full; retrieval-augmented generation and query fan-out; snippet controls and Google-Extended; FAQ rich results deprecation, 7 May 2026	04, 07, 08, 10, 11, 12, 18
Google Search Console Help	"Search generative AI control", rolled out to all websites worldwide 31 August 2026; generative AI performance reports	10, 19
Google Search Central	"Help your readers find your site through preferred sources in Google Search", read in full	15
Google	"New opportunities, control and insights for website owners", updated 31 August 2026: 2.5 billion monthly AI Overviews users, 1 billion AI Mode users	02
Google Analytics 4	Channel definitions: the AI Assistant channel, added 13 May 2026, and its stated exclusion of AI Overviews and AI Mode	19
Ahrefs	AI Overviews CTR study and its December 2025 re-run; brand visibility correlations across 75,000 brands; schema test across 1,885 pages; llms.txt analysis of 137,000 domains; "AI Assistants Prefer to Cite Fresher Content", 17 million citations	03, 05, 11, 17, 18
Ahrefs Brand Radar	Monthly most-cited-domain studies for Google AI Mode, Perplexity and Copilot, September 2026, broad United States queries	12
SparkToro and Similarweb	Zero-click clickstream analysis, 8 June 2026, January to April 2026 United States panel	02
Cloudflare	"New options to manage AI traffic", 1 July 2026, and the crawler defaults taking effect 15 September 2026	09
Vercel and MERJ	"The rise of the AI crawler": JavaScript execution and download behaviour across AI crawlers	08

Sources, continued

SOURCE	WHAT WE USED IT FOR	PAGES
Aggarwal et al., ACM SIGKDD 2024	"GEO: Generative Engine Optimization", the only controlled test of page-level tactics	06
BrightEdge	AI Overview prevalence and industry coverage tracking, 2025 to 2026	02, 03
Pew Research Center	Click behaviour with and without an AI Overview, 68,000 searches, 22 July 2025	02
Roy Morgan	Single Source, Australian AI tool usage, January to March 2026, n=14,646	02
Semrush	Relative value of an AI search visitor against an organic search visitor	03
Seer Interactive	Organic CTR on AI Overview queries, June 2024 to September 2025, 3,119 queries across 42 organisations	03
Stacker and Scrunch	December 2025 content distribution pilot, eight articles across 944 prompt and platform combinations	05, 14
Machine Relations	Larger follow-up analysis of citation lift from third-party distribution, 2026	05, 14
BrightLocal	Local Consumer Review Survey, 2026	16
Bing Webmaster Tools	AI Performance report and IndexNow documentation, 2026	08, 19
Whitehat Agency	Our own published figures: over \$650M in attributed client revenue, zero penalties since 2013, and partner status	22

Compiled by Whitehat Agency, September 2026. Platform behaviour changes quickly. Re-check any figure before it informs a spending decision, and ask us for the current edition rather than relying on an old copy.