

GENERATIVE ENGINE OPTIMIZATION · EDITION 5.0 · JUNE 2026

The GEO White Paper v5.0

Why AI Visibility Is Becoming a B2B
Publishing Operations Discipline

Shane H. Tepper

Co-founder, Resonate Labs · Author of *Cited*

ChatGPT · Perplexity · Claude · Gemini

— WHAT'S NEW IN V5

The fifth edition, corrected to the source.

This is the fifth edition of our state-of-the-field report on AI-native brand visibility. You don't need the earlier editions to read it; everything load-bearing is defined here. Three things are new this edition.

01: A CORRECTED BASE

Several widely circulated GEO statistics, including ones earlier editions of this report repeated, turned out to be misattributed to the wrong study or platform. This version traces every figure back to its primary source and corrects the record. Section 14 shows the specific fixes and the standard behind them.

02: A NEW ORGANIZING THESIS

GEO has been sold as a content discipline. The data now says it's a publishing operations discipline. Two prerequisites decide whether a brand is eligible to be cited at all: a site the models can read, and enough off-page presence that they trust and surface it. Once a brand clears that bar, the variable it most controls, and most underuses, is how fast it can publish. Citation runs on a clock measured in days while the top of search takes years, so speed becomes the lever. Section 3 sets up the argument; Section 10 pays it off.

03: FIRST-PARTY EVIDENCE

Earlier editions rested entirely on third-party research. This one adds a named client engagement, Insynctive, that shows the core arguments playing out inside one brand over a single cycle (Section 11), and along the way makes a dent in the field's biggest open question (Section 12).

CONTENTS

What's inside.

FRONT MATTER

Introduction	04
Five key findings & how to use this document	05

THE REPORT

01	What has held: the high-confidence baseline	07
02	The model is already on the buying committee	09
03	Two clocks	11
04	What predicts citation	13
05	Read versus credited: the ghost-citation funnel	15
06	Platform divergence	18
07	Dark traffic and the branded-link shift	20
08	Defensive GEO	22
09	Measurement architecture	25
10	Cadence and operations	27
11	First-party evidence: the Insynctive engagement	29
12	Open questions the field hasn't answered	33
13	Trajectories	35
14	Methodology and corrections note	36

REFERENCE

Works Cited	38
About	40

— INTRODUCTION

The shortlist forms inside the answer.

For two decades, B2B vendor discovery ran through a search box. A buyer typed a query, scanned a page of links, and clicked the ones that looked credible. The brands that ranked got considered. The rest didn't exist.

That layer is being replaced. Buyers now ask ChatGPT, Perplexity, Gemini, or Claude to name the vendors worth evaluating, and the model answers in a paragraph instead of a page of links. [G2's survey](#) of more than 1,000 software buyers found half of them now starting vendor research in an AI assistant rather than a search engine. The shortlist forms inside the answer, before a buyer visits a single site and long before sales hears about the deal.

This is the problem generative engine optimization (GEO) addresses. Not search rankings. Not website traffic. Whether your brand is visible, credible, and accurately represented in the AI-generated answers where your buyers now do their research.

Marketing leaders can feel the shift. In [Modus and Semrush's 2025 B2B CMO survey](#), 85% of marketing leaders called GEO a critical priority, and nearly half admitted they had no reliable way to measure it. The urgency is widely felt. The method is not yet settled.

Earlier editions of this report made the case that the shortlist now forms inside AI. This one takes the next step. It shows how fast that surface moves, how little of what models read ever reaches a buyer as a recognizable brand name, and what a serious program does about both. This is a state-of-the-field report, not a how-to guide. It's built to give marketing leaders the evidence, the framework, and an honest read on what's known and unknown before they commit budget to a category that didn't exist two years ago.

Five key findings.

- 1 The pre-sales research surface has moved into AI, and the click-loss mechanism is now causal.** Nearly all B2B buyers use AI tools during the purchase process, and most shortlist seats are filled before sales engages. Three claims stack here, and the newest one is the firmest: as of early 2026, the click-loss mechanism underneath it has moved from correlation to causal evidence (Section 2).
- 2 AI visibility runs on a different clock than search.** Models cite a good new page in about a week. The top of Google's results takes years. The speed gap falls out of how the two systems work, which makes it durable (Section 3).
- 3 Reading is not crediting.** Models retrieve and evaluate far more content than they ever attribute. The majority of brand appearances are uncredited, and only a small fraction of pages a model retrieves turn into a name a buyer hears (Section 5).
- 4 One "AI visibility" number averages signals that point in opposite directions.** The engines have built opposite citation architectures. A strategy tuned to one produces the wrong visibility on another (Section 6).
- 5 Once the foundation is in place, the constraint on citation speed is usually internal.** Crawlability and off-page presence decide whether a brand is eligible to be cited at all. After that, the engines can cite a good page in about a week, while most B2B content processes take a month or more to ship one. That gap is the part a team controls, which is why GEO is now a publishing operations problem (Section 10), shown end to end in a single client engagement (Section 11).

SYNTHESIS Taken together, these point to one conclusion: AI visibility is close to the cheapest it will ever be to establish, because the advantage compounds for whoever moves first and the cost climbs as the rest of a category operationalizes around the same citation clock. Most of the trajectories in Section 13 point the same way.

How to use this document.

Section 1 sets the stable baseline, the findings that have survived across versions and methodologies. Section 2 sets the stakes and Section 3 sets the spine. Sections 4 through 9 are the working evidence base: what predicts citation, where the engines diverge, how to defend your own narrative, and what to measure. Sections 10 and 11 cover execution and proof. Section 12 is honest about what nobody knows yet, Section 13 looks forward, and Section 14 documents sourcing and the corrections resolved in this version.

| A note on evidence

GEO is young and vendor data is everywhere, so this report separates evidence types and tries to signal which is which as it goes. Section 1 collects what's high-confidence. Section 12 collects what isn't.

■ PEER-REVIEWED

Independent, peer-reviewed research. Carries the most weight.

□ PRE-REGISTERED EXPERIMENT

A registered, randomized field experiment (may be a working paper). The strongest *causal* design here, even before peer review.

■ VENDOR STUDY

Useful, but flagged as the vendor's own measurement.

■ FIRST-PARTY CASE

Case evidence from our own engagements.

■ INFORMED INFERENCE

The direction is clear, but the proof isn't in yet.

These tags reappear beside the anchor stat on each section, so the basis of every headline figure is visible at a glance.

To discuss how these findings apply to your brand, visit resonatelabs.co.

RESONATE LABS

§ 01 THE HIGH-CONFIDENCE BASELINE

BRAND MENTIONS

0.664

correlation with AI Overview visibility
(about 3× the 0.218 for backlinks)

AHREFS · 75,000 BRANDS

VENDOR STUDY

What has held

In a field this young, what's stable matters more than what's trending. Anyone can spot a new statistic. The harder and more useful thing is identifying which findings have survived across multiple methodologies, multiple research teams, and successive editions of this report. Those are the load-bearing walls. You can build a program on them without worrying they'll move next quarter. The rest of this document goes deep on each. This section is the short version of what you can trust.

The shortlist still forms before sales, and the zero-click shift behind it is now causal

Every version of this paper has argued that the vendor shortlist forms before sales engages. The supporting data has only strengthened. [6sense's study](#) of nearly 4,000 B2B purchase decisions still anchors it: 95% of winning vendors were on the buyer's Day One shortlist. What changed in 2026 is that the zero-click shift underneath it moved from correlation to cause (Section 2).

Mentions move citation more than links

Across research teams and sample sizes, the same hierarchy keeps appearing: how often the web talks about your brand predicts AI citation better than how many links point at you. The cleanest measurement, [Ahrefs across 75,000 brands](#), puts brand web mentions at a 0.664 correlation with AI Overview visibility versus 0.218 for backlinks, and a December 2025 follow-up extended the same pattern to ChatGPT and Google AI Mode. [Muck Rack's analysis](#) of more than 25 million AI-cited links lands in the same place from a different direction, with roughly 84% of citations coming from earned media. The skills transfer from SEO. The scoreboard doesn't (Section 4).

| Platform divergence is structural, not noise

The engines cite different sources, in different formats, for different intents. The roughly 11% citation overlap between ChatGPT and Perplexity was a single-source claim in earlier versions. Multiple independent 2026 datasets now land in the same range. This is a structural feature of the landscape, not a temporary quirk of immature models (Section 6).

| Citations are volatile; the verdict underneath is stable

Track AI citations day to day and you'll watch them churn. Track the model's actual opinion of your category and it barely moves. Ahrefs found cited URLs changing 46% of the time between consecutive captures while the meaning of the responses held at 0.95 cosine similarity. Any program that mistakes surface churn for real change is measuring noise (Section 9).

| Freshness shifts citation at the margin

AI retrieval skews toward recent content. Ahrefs' 17 million-citation analysis found AI-cited pages running 25.7% fresher than organic results on average, and Seer's bot-crawl logs point the same way from a different method, with 65% of AI bot hits landing on content under a year old. The absolute average is still 2.9 years, so this is a margin, not a mandate to republish everything weekly. The margin is where competitive queries get decided (Sections 4 and 10).

These five have held. Everything that follows either deepens them or reports what's genuinely new since the last edition.

The model is already on the buying committee

95%

of winning vendors were on the buyer's Day One shortlist

6SENSE · ~4,000 B2B DECISIONS

● VENDOR STUDY

The surface you can't see

Discovery and evaluation now happen on a surface your brand can't see and didn't design. Buyers open an AI assistant, describe their problem, and get a shortlist back before they've visited a single vendor site. The model assembles that list from whatever it can retrieve, trust, and read as relevant. The brand has no window into the conversation and no seat at the table.

The clearest read on where this starts comes from [Similarweb's market-research panel](#): 35% of consumers using AI at the product-discovery stage versus 13.6% for traditional search, with AI holding a 2:1 or greater edge through evaluation and the gap closing to near parity only at the final "where to buy" step. This is consumer panel data, so treat the precise split as directional for B2B rather than a B2B measurement. The shape matches what B2B buyers report: 6sense found [94% of buyers using LLMs during the buying process](#), and [TrustRadius](#) puts 56% of tech buyers relying on AI chatbots as a top source for vendor discovery. These aren't early-adopter numbers. This is the center of the market.

| The evidence went causal

For a long time this argument rested on correlation. Buyers use AI, and zero-click behavior is rising, so the two looked related. That changed in early 2026.

A pre-registered randomized field experiment by [Agarwal and Sen](#) PRE-REGISTERED EXPERIMENT (Indian School of Business and Carnegie Mellon, 1,065 U.S. participants, January to February 2026) found that when Google AI Overviews appeared, they reduced outbound organic clicks by 38% and raised the probability of a zero-click search by about a third, from 54% to 72% in their sample. This is the strongest *causal* evidence type in the paper: a registered experiment with a control condition, not a vendor's own dashboard read. (It is a working paper, posted to SSRN in April 2026, not yet peer-reviewed, but the design carries more causal weight than any observational dataset here.) The authors flag the AI Mode arm as exploratory because of attrition and extension uninstalls, so read that piece as suggestive. The headline causal result is solid, and it's the first time the zero-click shift has been measured as cause rather than coincidence. It's worth being precise about what the experiment proves: it shows AI summaries cut clicks, not that AI hands out shortlist seats. That's the mechanism feeding the shortlist, and the shortlist evidence is a separate argument, which is next. The causal read lines up with the observational one: [Pew's browsing-data study](#) of about 900 U.S. adults found users clicking a traditional result 8% of the time when an AI summary appeared, versus 15% when it didn't. One experiment, one large observational dataset, the same direction.

| The shortlist was always the gate

The structural point sits underneath the adoption numbers. [6sense studied](#) nearly 4,000 B2B purchase decisions and found 95% of winning vendors on the buyer's Day One shortlist, the consideration set formed before any sales engagement, with 3.6 of roughly 5 (5.1) seats filled on Day One. Those numbers predate the AI shift, and that's the point. The shortlist was always the gate. AI didn't replace it. AI matters most for the remaining one or two seats, the ones still in play while buyers expand or validate the list, and in competitive deals that margin is where outcomes get decided.

Picture a category with four or five credible vendors, say cloud-based performance management tools. The AI's recommendations aren't random. They concentrate. The same four or five names surface repeatedly across buyers and sessions. If you're in that set, every appearance compounds into the next. If you're not, you're arguing your way onto a list that closed before the call.

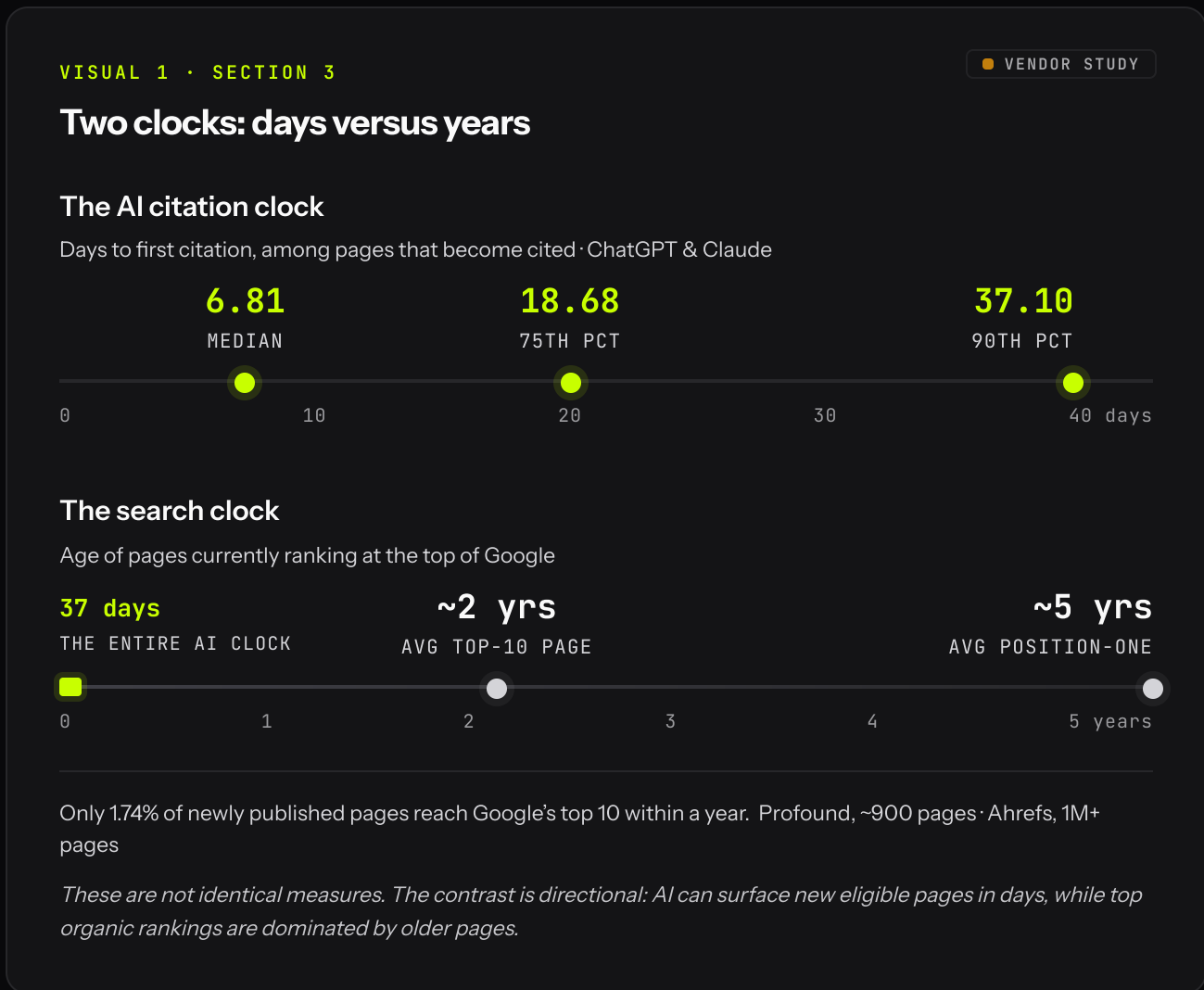
| The committee member nobody invited has a vote. It cast it before your name came up.

Two clocks

Search and AI retrieval run on different timelines, and the gap between them is the most useful thing a B2B content leader can understand right now.

Why the gap exists

Search rewards accumulated trust. Backlinks, engagement history, domain age, the signals that compound over years. AI retrieval rewards freshness and access. A system that prefers recent, accessible, specific content will cite a new page quickly almost by design. A system that rewards accumulated authority will not. The speed difference falls out of how the two systems work, which is why it holds across model generations rather than disappearing with the next release.



| The clock for getting noticed

In May 2026, [Profound](#) published the median time from publishing a page to first citation by ChatGPT or Claude: 6.81 days, across roughly 900 pages and billions of response logs. The full distribution is more useful than the median. Three-quarters of cited pages within 18.68 days, ninety percent within 37.10. Past day 37, patience stops explaining much; pages still uncited that late are usually held up by something time won't fix, and the remaining problem is typically technical.

Two boundary conditions hold the claim in place. It measures ChatGPT and Claude only, with Perplexity, Gemini, and AI Overviews left out. And it measures first citation, the moment a model first pulls a page, which is a lower bar than being cited consistently or recommended. The clock also applies only to pages that become citable at all: it describes how fast citation tends to happen, not whether a given page earns one. So this is the clock for getting noticed, not the clock for winning.

| The clock you can't beat

[Ahrefs' 2025 analysis](#) of more than a million pages is the cleanest counterweight. The average page in Google's top 10 is over two years old. The average page at position one is around five years old, up from two in 2017. When Ahrefs tracked newly published pages directly, only 1.74% reached the top 10 within a year. The part of the SERP that drives B2B pipeline is the top two or three, and that real estate is older and harder still, held by entrenched incumbents and aggregators.

Page age is a proxy for time-to-rank, not a stopwatch, and a high-authority domain can rank fast on low-competition terms. The contrast lives at a specific altitude: the top of the SERP on the queries your buyers actually run takes years to reach, when it's reachable at all.

Put the two clocks side by side and the strategic read is plain. You can influence AI answers on high-intent buyer queries faster, and against incumbency you could never crack in traditional search, than the SERP top three will ever open up. That advantage sets up everything in Sections 10 and 11.

~84%

of AI citations come from earned media

MUCK RACK · 25M+ CITED LINKS

VENDOR STUDY

What predicts citation

Models cite what they can retrieve, trust, and read as fresh and specific. Strip away the tactics and that's the whole mechanism. Each word in it points at a different lever, and the off-page lever is the one most teams underweight.

Mentions over links

The strongest single signal in the published research is brand mentions, not links. [Ahrefs' analysis of 75,000 brands](#) found brand web mentions correlating with AI Overview visibility at 0.664, roughly three times the 0.218 correlation for backlinks, with the top three correlating factors all off-site. That's an SEO data company's finding, and a PR analytics company reaches the same conclusion by a different route: [Muck Rack's analysis](#) of more than 25 million AI-cited links found roughly 84% coming from earned media and a fraction of a percent (0.3%) from paid placement. Two instruments, one result. These are correlation studies, not causal proof, but the convergence across an SEO dataset and a PR dataset is what makes the direction hard to dismiss. A program that pours effort into link acquisition and ignores how often the brand gets talked about across the web is optimizing the weaker variable. The off-page work that moves AI citation looks more like earned media and category presence than like a backlink campaign.

The one controlled, independent study in this area comes at the question from the content side. The peer-reviewed [GEO paper](#) ● PEER-REVIEWED from Princeton and collaborators (KDD 2024), which coined the term, tested content-side changes across a benchmark of thousands of queries and found that adding statistics, citations, and quotations raised a source's visibility in generative-engine answers by up to 40%, with the effect varying by domain. Read it as a benchmark result about which content properties engines reward, not a field citation rate your brand should expect to hit. It complements the off-page picture rather than competing with it: off-page presence governs whether you're eligible to be cited, and on-page structure governs how readily a model extracts and credits you once you are. It's also the rare non-vendor, causal-leaning evidence in a field still short on it.

■ Freshness as a margin, not a mandate

Freshness belongs here too, as a mechanism rather than a headline number. [Ahrefs' study of 16.975 million cited URLs across ChatGPT, Perplexity, Gemini, Copilot, AI Overviews, and organic results](#) found AI-cited content averaging 1,064 days old versus 1,432 for organic, so AI citations run 25.7% fresher. The absolute average is still 2.9 years. This doesn't translate to a mandate to publish constantly. Recency shifts citation probability at the margin, and that margin is where competitive queries are won. Section 10 turns that into a cadence.

■ The takeaway is uncomfortable for teams built around technical SEO. The strongest predictor of your citation rate is how widely and specifically the rest of the web talks about you.

The ghost-citation funnel

Two filters

Two filters sit between your content and a buyer's recognition. Retrieval doesn't mean citation. Citation doesn't mean recognition. A model can retrieve and evaluate your page, decline to cite it, and even when it cites you, attribute the claim to a source link the buyer never reads as your brand. Most GEO programs measure the top of that funnel and call it visibility.

The scale of consumption

Reading is cheap. [Akamai reports](#) billions of daily AI bot requests across its network, up nearly 300% year over year. [Ahrefs' analysis of 1.4 million ChatGPT prompts](#) found the model retrieving many pages and citing only about half of them, with 67.8% of all non-cited URLs coming from Reddit. The models retrieve and evaluate vastly more than they credit, and the gap between consumption and credit is where most brand visibility quietly disappears.

The ghost rate

The anchor number for the recognition gap comes from [Kevin Indig's Growth Memo analysis](#) (April 2026), drawn from the Semrush AI Visibility Toolkit: across 3,981 domains spanning 115 prompts, 14 countries, and four AI engines, the data shows 74.9% of appearances as citations, 38.3% as text mentions, and only 13.2% as both. Indig puts the ghost-citation rate at 61.7%, brands cited as a source link without being named in the answer the buyer reads. A citation the buyer never connects to your name does far less work than one they do.

VISUAL 2 · SECTION 5 · THE CENTERPIECE

● VENDOR STUDY

The ghost-citation funnel

A THE PATH, STAGE BY STAGE



The web consumes far more than it credits, and credit is not the same as recognition. These are conceptual stages; only the breakdown below is measured.

B HOW ONE APPEARANCE BREAKS DOWN



● **Citation only**

cited as a source link, never named: the ghost-citation rate

● **Mention only**

named in the answer text, not cited as a link

● **Both**

cited and named: the signal that behaves like a search result

Three mutually exclusive states, derived from the study's citation (74.9%), mention (38.3%), and both (13.2%) figures; they sum to 100%. KEVIN INDIG / GROWTH MEMO · SEMRUSH AI VISIBILITY TOOLKIT · 3,981 DOMAINS · 115 PROMPTS · 14 COUNTRIES · 4 ENGINES

Mention plus citation compounds

This is where mentions and citations reinforce each other. Seer found brands named in the response text earning roughly 5x the citation rate of brands that aren't. AirOps, in its own analysis, found brands earning both a citation and a mention 40% more likely to resurface across consecutive runs than brands cited only as a link. AirOps is a vendor self-study, and its own both-signal rate sits on a different dataset than Indig's 13.2%, so the two shouldn't be merged into one figure. The direction is consistent across both: the citation-mention, not the bare citation, is the signal that behaves like a search result to a buyer.

SOURCING NOTE One sourcing note belongs in the open. The platforms most likely to be cited and not credited are the large user-generated and reference domains, and the major AI companies have struck data-licensing agreements with several of them. The agreements are real and material. The specific dollar figures that circulate for them are not reliably sourced, so this paper describes the arrangements rather than quoting a number (Section 14).

A model can crawl and retrieve everything you publish and credit almost none of it. Plan for the funnel, not the front door.

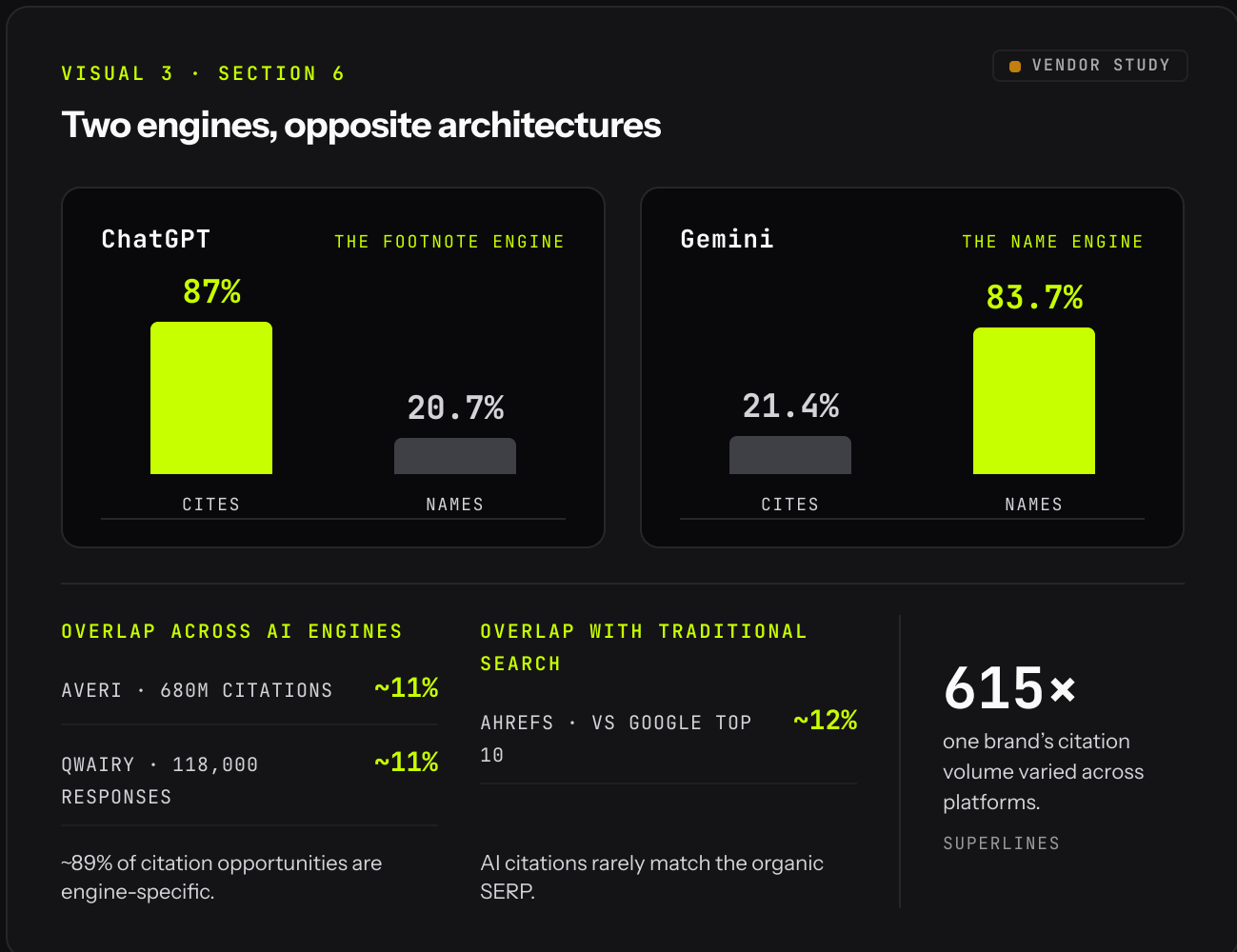
§ 06 THE ENGINES DISAGREE

Platform divergence

The engines have built opposite citation architectures. A single “AI visibility” metric averages signals that point in different directions, and a strategy tuned to one engine can produce the wrong visibility on another.

| The citation-mention inversion

The cleanest illustration is the inversion between the two largest surfaces. In Indig’s Growth Memo data, ChatGPT cites a source 87% of the time but names the brand in text only 20.7% of the time. Gemini runs the opposite way, naming brands 83.7% of the time while generating a citation link only 21.4% of the time. Same brand, same query, two engines: one reads you as a footnote, the other as a name. Optimizing your link profile for the first does little on the second, and earning narrative mentions for the second does little on the first.



Low overlap is structural

The structural fact underneath this is low overlap. Earlier editions reported the roughly 11% domain overlap between ChatGPT and Perplexity from a single 100,000-prompt study. Multiple 2026 datasets now point to the same low-overlap range: [Averi](#) (680 million AI citations, ~11%) and [Qwairy](#) (118,000 responses, ~11%), both vendor measurements, while [Ahrefs' query-level analysis](#) independently finds about 12% overlap with Google's top 10. Roughly 89% of citation opportunities are platform-specific, and that figure is now corroborated rather than single-sourced. The same [Superlines analysis](#) that tracks this found a single brand's citation volume differing by up to 615x across platforms.

Content type splits by intent

Content type splits the same way, by intent. [Wix's Studio AI Search Lab](#) (March 2026) found articles drawing 45.48% of informational-intent citations and listicles drawing 40.86% of commercial-intent citations. The intent label matters: articles are a smaller share of citations overall, so the 45.48% reads as “for informational queries,” not “in general.” On length, [Growth Memo](#) (March 2026) found pages over 20,000 characters averaging 10.18 ChatGPT citations versus 2.39 for pages under 500. Read alongside Indig's follow-up, which found shorter, focused pages winning, the length finding is really about breadth of coverage in a single URL (“what is it,” “who uses it,” “how to choose,” “pricing”) rather than raw word count. Both Wix and Growth Memo are vendor studies, so treat them as well-instrumented signals from interested parties.

One number averaged across platforms hides the strategy. The work is per-engine, and the engines disagree on purpose.

0.4% → 6.2%inline brand links in ChatGPT answers, in
a single day

QWAIRY · 140,000+ ANSWERS

VENDOR STUDY

Dark traffic and the branded-link shift

When an AI assistant sends a visitor to your site, that visit often arrives with no referrer attached. It shows up in your analytics as direct traffic, as organic, or as nothing at all. That invisible referral stream is what the field calls dark traffic, and it's the reason most brands can't see how much business AI already sends them. In 2026, ChatGPT started partially un-hiding its own share of it, the most significant mechanism update of the past year.

ChatGPT moves toward the front door

The interface is moving toward brand discovery. Where source links once sat buried in footnotes, prominent clickable brand links now appear inline in answers. The effect is to send users to the front door of a brand's site rather than to a deep page, and that narrows the attribution gap on ChatGPT specifically while it stays wide everywhere else.

The magnitude, measured several ways

The cleanest single read comes from Profound's own measurement. Across the brand sites Profound monitors for live referral traffic, daily ChatGPT referrals rose 60 to 65%, beginning May 7, 2026, as the share of ChatGPT responses containing a URL jumped from about 4.5% to 20 to 24% (roughly 5x) on the same day. That's the primary referral measure for the section, and it should carry its caveats rather than be stacked on the other magnitudes.

The other figures corroborate the direction without agreeing on the size, which is normal for a clickstream event measured several ways. Similarweb's panel found total ChatGPT referrals up 157.7% week over week, homepage referrals up 354.7%, and homepage share of referrals settling around 60% and holding. Qwairy's analysis of more than 140,000 answers measured the same change from the answer side: inline brand links jumped from about 0.4% to 6.2% of answers in a single day. Three independent methods, a referral basket, a clickstream panel, and an answer-text analysis, all land on the same May 7 event, which is the credibility case for a shift OpenAI never announced. Qwairy adds the control that matters most: only ChatGPT moved, while Perplexity, Gemini, and Copilot stayed flat across the same window, which rules out a web-wide artifact.

VISUAL · SECTION 7

Three methods, one event: May 7, 2026

ANSWER BEHAVIOR

Inline brand links in answers

0.4% → 6.2%

QWAIRY · 140,000+ ANSWERS

REFERRAL BEHAVIOR

PRIMARY

Daily ChatGPT referral volume

+60-65%

PROFOUND · LIVE REFERRAL BASKET

DESTINATION MIX

Homepage referrals

+354.7%

SIMILARWEB · CLICKSTREAM PANEL

A referral basket, an answer-text analysis, and a clickstream panel, all dating the shift to the same day. Only ChatGPT moved; Perplexity, Gemini, and Copilot stayed flat. PROFOUND IS THE PRIMARY REFERRAL MEASURE; THE OTHERS CORROBORATE DIRECTION, NOT MAGNITUDE.

One shift, three denominators

The shift got reported as a “60%” jump in several places. But the figure depends entirely on what’s being measured against what, and three credible measurements land in three different places.

Total daily referral volume	+60-65%	PROFOUND
Homepage <i>share</i> of referrals, from a 26-32% base	~60%	SIMILARWEB
Homepage share, industry-wide	3.5% → 24%	PROFOUND

Same direction, three denominators. Two more caveats. The measured brands skew toward consumer and SaaS, so treat the precise magnitude as directional for any single B2B category, even though early reads suggest B2B software saw gains on the stronger end. And OpenAI never named this as a product update, so describe it as observed behavior, not an announcement.

For the first time, one major AI surface is telling brands where its traffic goes. The window into that data is open on ChatGPT and closed elsewhere.

40%

vs 16% open-ended

brand-direct prompts trigger a site-scoped query

PROFOUND · 3,380 PROMPTS

● VENDOR STUDY

Defensive GEO

Most of this document is about offensive GEO: getting your brand into AI responses and earning citations. This section is about the other side. What happens when someone else shapes what the model says about you, and what it takes to control your own narrative on a surface you don't own.

| When buyers name you, the model reads your site

When a buyer names your brand in a prompt, the model's behavior changes in a way you can use. A brand-direct query turns your own site into the model's primary retrieval target. The measurement comes from Profound, surfaced by [Nick Lafferty](#), Profound's Head of Marketing: across 3,380 prompts on ChatGPT 5.4 in March 2026, brand-direct prompts triggered at least one site-scoped query 40% of the time versus 16% for open-ended prompts, a 2.5x difference. Within the first two searches the model runs, the gap holds, around 71% versus 42%. When someone asks the model about you specifically, the model goes looking at your site first. Treat the exact figures as directional, since this is a single platform, a single model version, and a single month.

| The information gap doesn't stay empty

That reframes a familiar problem. The gap on your own domain doesn't stay empty. The mechanism underneath it is that models are drawn to specificity as a proxy for credibility. A source with concrete numbers, even wrong ones, gets weighted over a page that says "contact sales." If your pricing page speaks in generalities and a comparison blog names a figure, the model uses the blog's figure, accurate or not. The bar for defensive content runs higher than having an answer somewhere on your site. The answer has to be more specific than anything else the model can find.

A 2025 account [shared by Chris Long](#) shows how clean this gets. A founder shopping for a PEO (a professional employer organization, the kind of firm that runs payroll, benefits, and HR compliance on another company's behalf) described asking Google's AI Mode whether Gusto, the payroll tool he already used, could handle PEO services. The answer the model gave was pulled from a Rippling FAQ page: Gusto doesn't offer PEO. The model put a competitor into the conversation at the exact moment he was evaluating options. Rippling didn't lie. They published a factually accurate FAQ answering a question their competitor's site didn't address. The model handed the moment to whoever had bothered to write the page.

| The trust exploit and the platform split

The more dangerous version is the partial debunk. A third-party article first corrects a piece of obvious misinformation about a brand, which earns the model's trust, then introduces new claims that are misleading or fabricated. The model has already classified the source as credible, so it accepts the new claims with less scrutiny. A controlled brand-manipulation experiment by [Ahrefs \(2025\)](#) showed exactly this: a Medium "investigation" that debunked a fake brand's obvious rumors first, earning the models' trust, then slipped in fresh fabrications that several engines repeated as the corrected record. The same experiment ran a parallel source, a fabricated Reddit AMA that models treated as credible firsthand testimony. It's the informational version of a social-engineering attack: establish trust, then spend it. Platforms aren't equally exposed, so a brand can be well-protected on one engine and misrepresented on another using the same available sources. Defensive GEO, like offensive GEO, has to be platform-aware.

| The Reddit paradox

Reddit is the cleanest illustration of why citation tracking alone misses narrative risk. In ChatGPT retrieval analysis, Reddit-style forums are heavily retrieved but rarely surface as formal citations, on the order of a couple percent for B2B-style queries. Reddit rarely shows up as a citation for B2B. Where it carries weight is in the narratives that form around a brand. Two channels do that work: Reddit's content feeds model training through data-licensing agreements with major AI companies, so its sentiment is baked into the baseline understanding of a brand independent of query-time citation, and models treat Reddit posts as firsthand practitioner testimony. The operational read for B2B: monitor category subreddits for narrative risk, but don't fund a Reddit content strategy at the expense of vendor-site, review-platform, and earned-media work. For B2B the citation value is low and the narrative risk is high.

| What defensive GEO requires

NARRATIVE GAP IDENTIFICATION

Map every question a buyer naturally asks about you (pricing, implementation time, customer count, limitations, competitive differences) to a specific, crawlable, official answer, because "we don't publish that" is worse than silence.

NARRATIVE MONITORING

Track what the models actually say about you across platforms, not just whether you're cited.

RAPID RESPONSE

When monitoring surfaces an error, publish something more specific, more authoritative, and more crawlable than the source the model is currently using, and let the freshness bias from Section 4 do the displacing within a refresh cycle.

| The cost of prevention is a fraction of the cost of remediation.

0.95

cosine similarity between consecutive AI responses

AHREFS · 43,000 KEYWORDS

VENDOR STUDY

Measurement architecture

Most GEO monitoring measures the middle of the funnel and reports it as the end. Presence is not winning. A brand can appear somewhere in a response and still not register with a buyer the way a search result does, and a single snapshot of that presence is close to meaningless because the surface churns constantly while the underlying verdict barely moves.

The surface churns; the verdict doesn't

The anchor finding makes the distinction concrete. [Ahrefs' study of 43,000 AI Overview keywords](#) found that between consecutive captures, the response text changes about 70% of the time, cited URLs change 46% of the time, and entities change 54%, while cosine similarity between consecutive responses sits at 0.95.

70%	of the time the response text changes between consecutive captures
54%	of the time the entities change
46%	of the time the cited URLs change
0.95	cosine similarity: the meaning of the answer holds steady

Cosine similarity measures how close two texts are in meaning, where 1.0 is identical. The words shuffle and the URLs turn over, but the model's overall read of the category holds steady.

One precision worth keeping: cosine similarity captures the topical shape of the answer, not your exact position within it, so a brand's rank or inclusion can still move run to run even as the response stays semantically the same. A team tracking week-to-week citation changes is mostly tracking noise. The churn itself is well-corroborated: [Profound's tracking](#) puts citation drift at 40 to 60% of cited domains changing month over month across the major engines, a fourth independent read on the same instability.

One snapshot is variance, not position

Two more studies guard against single-snapshot thinking. [SparkToro and Gumshoe](#) ran 2,961 prompts and found the probability of getting an identical recommendation list twice from the same prompt was under 1%, yet top brands in narrow categories still appeared in 70 to 90% of runs. The deck reshuffles; the same cards keep coming up. [AirOps](#), in its own analysis of more than 45,000 citations, found only 30% of brands holding visibility from one run to the very next, and just 1 in 5 holding from the first run through the fifth. AirOps is a vendor self-study, so flag it as such, but the structural point is independent of the source: any metric built on one capture is reporting variance, not position.

What to actually track

So the right measurement question runs past "did we appear." It asks how often, across repeated runs, you earn a citation-mention on the queries that actually shape a shortlist. In practice that means a stable query set tied to real buyer intent, repeated sampling rather than single snapshots, scoring for the citation-mention rather than the bare appearance, narrative accuracy alongside presence (per Section 8), and share of voice tracked against named competitors over time. The instrument has to match the volatility of what it measures.

You can't manage a position you only sampled once. Measure the verdict, not the surface.

§ 10 VELOCITY AS THE VARIABLE

25.7%

AI citations run fresher than organic: the margin this section turns into cadence

AHREFS · 17M CITATIONS

VENDOR STUDY

Cadence and operations

This is where the two clocks pay off. If the engines cite a good page in about a week, then for a brand that’s already readable and has some presence, the binding constraint on citation speed lives inside the company. The engines can be ready in days. In our experience, many B2B content operations still need a month or more to move a page through drafting, review, legal, design, and deployment. That internal gap is what slows a brand down, and it’s the variable a team can actually control.

The 48-hour refresh was always wrong

A piece of advice spread last year: refresh high-priority pages every 48 hours. The rigorous data doesn’t support it. [Ahrefs’ study of the top 1,000 ChatGPT-cited pages](#) found a strong recency bias, but [Seer’s log-file analysis](#) of more than 5,000 URLs gives the cleanest planning numbers: 65% of AI bot hits target content published in the past year, 89% within three years, 94% within five. A 48-hour cycle runs far faster than anything the citation pattern rewards, and it tends to produce cosmetic edits the systems can detect anyway. [Google’s John Mueller has said repeatedly](#) that artificially freshening a publication date without substantive change is “just noise” and won’t move rankings, and the directional read for AI citation is the same.

A defensible refresh framework

The anchor for this section is a three-tier cadence.

MONTHLY

HIGH-PRIORITY PAGES

Comparison guides, pricing and packaging, integration documentation: the buyer-stage assets that drive pipeline.

QUARTERLY

CATEGORY EDUCATION

Thought leadership, where content is meant to age into authority rather than respond to news.

TRIGGER

EVERYTHING ELSE

Product launches, competitor moves, compliance changes, customer wins worth naming.

A platform note for sophisticated teams: if Perplexity matters to your category, monthly is closer to a floor than a ceiling because it weights recency hardest; if AI Overviews dominate your category, quarterly is often sufficient because it has the weakest freshness bias of the major surfaces. Treat these tiers as an informed starting point to calibrate against your own re-audit data, not a validated protocol.

| The staffing the wrong cadence produced

The resourcing implication is the part the 48-hour advice got most wrong. A serious program needs one writer on a recurring refresh schedule, not three writers on a daily treadmill. The wrong cadence didn't just waste effort. It produced the wrong staffing model. The same architecture that makes a fast refresh possible (server-rendered pages, pre-structured briefs, a deployment process that ships and verifies in a day) is what lets a team feed the clock at all, because presence behaves like a subscription: stop paying and it lapses.

| GEO is a publishing operations discipline. Architecture, briefs, deployment cadence, and re-audit rhythm move citations more than any individual page does.

§ 11 FIRST-PARTY EVIDENCE

The Insynctive engagement

The aggregate curves in this paper describe thousands of pages across many brands. Watching the same pattern inside a single company is more useful, and it surfaces the mechanism the averages hide: crawler traffic is the leading indicator, citation visibility is the lagging one, and both move in the first cycle when the foundation is unblocked.

| Invisible, and not for the obvious reason

Insynctive is a B2B HRIS platform, the software companies use to manage payroll, benefits, and employee records. At baseline, the brand surfaced on 2 of 150 buyer-research queries, 1.3% visibility across ChatGPT, Claude, Gemini, and Perplexity. Effectively invisible. The site was technically open to crawlers, which made the real problem easy to miss. The platform was built on a client-side-rendered framework, so AI crawlers were requesting pages and receiving empty HTML shells. The engines' crawlers, GPTBot, ClaudeBot, and PerplexityBot, were getting pages whose content only appeared after JavaScript ran, and these crawlers don't run it. Vercel and MERJ's analysis of more than 500 million GPTBot fetches found zero JavaScript execution: the crawler downloads a JavaScript file about 11.5% of the time but never runs it, and the same holds for ClaudeBot and PerplexityBot. Past the 90th-percentile threshold with no citation, more waiting wouldn't have helped, because the problem was technical from the start.

One engine doesn't fit that mechanism. Google renders JavaScript when it crawls, and Gemini grounds on Google's index, so Gemini could read the pages the other three couldn't. Gemini was invisible for a different reason: the site had very few pages indexed, and what was indexed had gone stale, so there was little current content to surface. The rebuild closed both gaps at once. Server-rendering opened the pages to the three blind crawlers, and the 38 fresh, internally-linked pages gave Google current content to index, which is what Gemini needed. Two reasons for invisibility, one rebuild that cleared both.

| The fix, and the leading indicator

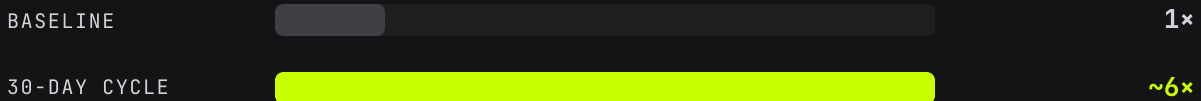
The fix was an architectural workaround, not a re-platform: a server-rendered version of the same page content, delivered consistently to both human visitors and crawlers through an edge worker (a lightweight server layer that sits in front of the site). The content was identical for both audiences. The only change was making it readable without running JavaScript. Then a 48-hour content sprint against a prioritized action plan, then a structured internal-linking rebuild. Across the 30-day cycle, the team shipped 38 new optimized pages. ChatGPT bot traffic moved first, with GPTBot pulling roughly 6x its baseline volume, a +504% increase. That’s the leading indicator: the crawler shows up before the citation does.

VISUAL 4 · SECTION 11

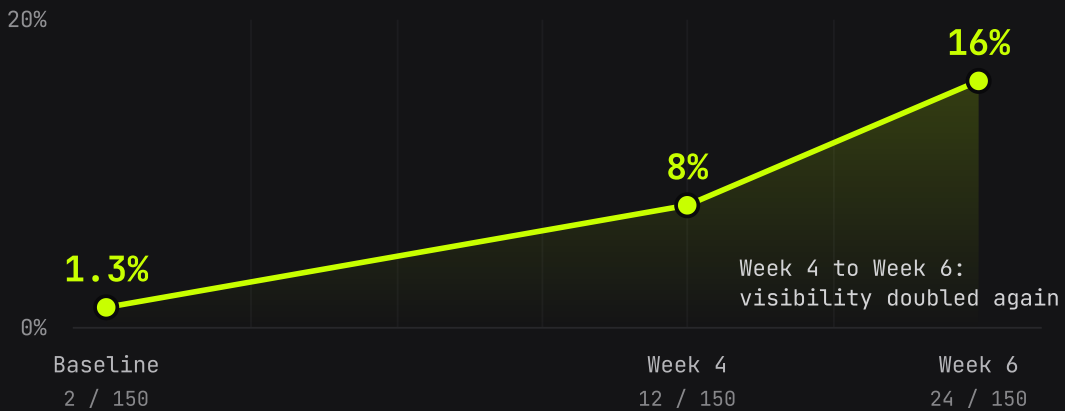
FIRST-PARTY CASE

The Insynctive trajectory

LEADING INDICATOR: GPTBOT CRAWL VOLUME



LAGGING INDICATOR: VISIBILITY ON THE 150-QUERY PANEL



<p>6.2%→25%</p> <p>requirements-building queries, the largest stage jump</p>	<p>38</p> <p>optimized pages shipped in the 30-day cycle</p>	<p>2 of 150</p> <p>queries surfacing the brand at baseline</p>
---	---	---

X-axis is time-proportional; the baseline is near zero by design. Measured on a fixed 150-query panel across ChatGPT, Claude, Gemini, and Perplexity. RESONATE LABS · INSYNCTIVE ENGAGEMENT

| The lagging indicator follows

Then the lagging indicator followed. At the four-week re-audit, on the identical 150-query set with the identical surfacing definition, visibility moved from 1.3% to 8%, six times the baseline presence, a 6x lift from a deliberately low starting point, across ChatGPT, Claude, Gemini, and Perplexity on tracked queries. The largest gains landed on the queries closest to purchase: requirements-building visibility rose from 6.2% to 25%, while shortlisting, problem identification, and validation each moved from a standing start, and capability queries like compliance management went from 0 to 28.6%. The part most case studies can't produce: every newly surfacing query traced back to a specific page the team had shipped, and the model was citing that page. Audit identifies the query, brief gets written, page goes live, query surfaces at re-audit citing the exact page. A closed loop.

| The clock, made visible

A second snapshot at six weeks showed visibility moving again, from 8% to 16%. It doubled in two weeks after taking four weeks to get from 1.3 to 8. A curve that accelerates between week four and week six is the citation latency resolving: pages shipped in the first sprint crossing the citation threshold weeks after they went live, consistent with what Profound's distribution would predict, a median around a week and a long tail past day 37. The public distribution describes about 900 pages in aggregate. This engagement describes one brand over six weeks. They're the same phenomenon at two scales, and the fit is close enough that the public distribution stops being abstract and becomes a planning tool.

TWO HONEST CAVEATS First, read this as foundation-first evidence, not proof that publishing speed is the universal bottleneck. Insynctive's binding problem was crawlability: the site was unreadable to three of the four engines' crawlers, and for those engines no amount of publishing speed would have mattered until that was fixed. The velocity dynamic only showed up once the foundation was in place.

Second, this is one engagement, and the magnitude is a function of where it started. A brand that's already readable and somewhat visible has far less room to move, so the 6x owes as much to the near-zero baseline as to the method. The sequence is what generalizes, foundation first, then tempo. The multiple does not.

| Case-study methodology

The detail that separates evidence from anecdote, stated plainly so a skeptical reader can judge it.

QUERY SET

150 buyer-research queries, locked before the sprint and re-run unchanged at every audit, so each cycle measures the same ruler. Built around six evaluator-stage buyer personas: HR and benefits administrators, HR directors, CFOs, broker principals, TPA operations leads, and benefits account managers.

PLATFORMS MEASURED

ChatGPT, Claude, Gemini, and Perplexity, held constant across cycles.

VISIBILITY DEFINITION

A query counts as surfacing the brand when the model names or cites it in the answer. Each newly visible query is then traced to the page shipped to serve it, with three outcomes recorded: the model cited the named page, the brand was visible but the citation couldn't be mapped to a URL, or no citation was captured.

INTERVENTION

Client-side rendering replaced with a server-rendered version of the same content via an edge worker; a 48-hour content sprint; an internal-linking rebuild; 38 pages shipped across the 30-day cycle.

SIGNIFICANCE

Per-query visible/not-visible changes tested with McNemar's exact test on the paired panel.

COMPETITIVE SET

The brand was tracked by mention share against a defined competitive set, with new entrants surfaced each cycle rather than assumed away.

ATTRIBUTION STANCE

The standard is "consistent with," not "caused by." Content shipped and visibility moved on the same dimension and timeline, and the evidence chain runs deployed, then cited, then visible, but model drift across the window can't be fully ruled out, so this is attribution, not strict causation.

LIMITS

One engagement, no holdout group, and model versions weren't logged cycle to cycle, so a model update inside the window can't be cleanly separated from a content effect.

§ 12 WHERE THE EVIDENCE RUNS OUT

Open questions the field hasn't answered

The sections above present what's known, with evidence and confidence levels. This section presents what nobody knows yet. None of these are edge cases. They're structural gaps in the field's understanding, and any organization claiming a complete GEO strategy is working around them, not through them. Knowing where the evidence runs out is as important as knowing what it shows.

01 **OPEN** The longitudinal intervention gap is narrowing, not closed

The field's single biggest gap has been the absence of controlled before-and-after evidence in a live competitive landscape: did specific GEO interventions actually cause visibility gains for a real brand against real competitors, or do we only have cross-sectional snapshots of what tends to get cited. The peer-reviewed work that exists, chiefly the Princeton GEO study, is controlled but runs on a synthetic query-and-source benchmark rather than a live brand in a live market.

Section 11 makes a real dent in this. The Insynctive engagement is a documented before-and-after with page-level attribution: a specific intervention, a measured baseline, a re-audit on an identical query set, and every new appearance traced to a shipped page. It's the closest thing to causal first-party evidence this paper has carried. It's also a single engagement with no holdout group. One attributable case is a long way from a published longitudinal trial. The gap is smaller than it was a few months ago. It isn't closed.

02 **OPEN** Nobody has mapped the freshness ceiling

Recency shifts citation probability, but only up to a point. Below some threshold, staleness gets penalized; above some other threshold, additional freshness stops adding value. The 48-hour advice failed precisely because it overshot a ceiling nobody has measured. We know the ceiling exists, from the gap between the strong recency bias in the ChatGPT Top 1,000 data and the 2.9-year average age in the 17 million-citation study. We don't know where it sits, and it almost certainly varies by platform. Until someone maps it, refresh cadence is an informed guess calibrated to avoid the obvious error.

03 **OPEN** Whether the branded-link shift generalizes to B2B

The branded-link change in Section 7 is the most important mechanism update of the past year, and its cleanest evidence comes from a sample that skews consumer and mixed. Whether ChatGPT sends the same front-door traffic to a mid-market B2B SaaS vendor as it does to a large consumer brand is still an inference. If it generalizes, the attribution gap on ChatGPT narrows for B2B too. If it doesn't, B2B keeps operating in the dark-traffic environment while consumer brands get a clearer signal.

04 **OPEN** How training data and real-time retrieval interact

Every major platform blends pre-trained knowledge with real-time retrieval, but the weighting between them, and how it varies by model, query type, and topic, is poorly documented. The practical stakes are large. If training-data dominance is durable, incumbents with years of authoritative content hold a structural advantage that real-time optimization can't easily overcome. If real-time retrieval dominates, a new entrant with excellent current content can move fast, and the case for an ongoing GEO program is stronger. Both are plausible. Neither is isolated in a published study.

05 **OPEN** Whether agentic discovery produces citations at all

As buying shifts toward agent-mediated research, the narrow GEO question is whether an agent acting for a buyer produces citations at all, or interacts directly with APIs, product feeds, and structured data instead. If agents bypass the citation layer, the measurement framework in this paper (citation share, mention tracking, platform visibility) shifts toward machine-readability scoring and data-accessibility auditing. The current paradigm assumes a human reading an AI response with visible sources. Agentic commerce may make the audience another machine.

06 **OPEN** International and non-English dynamics are essentially unstudied

Nearly every finding here is English-language and US-centric. Whether citation mechanics, platform preferences, volatility patterns, and source hierarchies hold in other languages and markets is unknown, and there are reasons to expect they don't: differing regulation, uneven platform availability, and language-specific training data. Any global program is extrapolating from US English data to markets where the dynamics may differ at the root.

These gaps don't invalidate the evidence in the preceding sections. They bound it. Act on the well-supported findings with confidence, and treat the rest as bets placed in the dark, sized accordingly.

§ 13 THE NEXT 6 TO 12 MONTHS

Trajectories

Where this goes over the next 6 to 12 months. These are informed projections, not measurements.

● INFORMED INFERENCE

→ **ChatGPT keeps moving toward a brand-discovery interface.**

The branded-link shift in Section 7 reads less like a one-time change and more like a direction. If the largest assistant keeps surfacing clickable brand links, the attribution gap on that one surface narrows further while it stays wide on the others, and brand-name recognition inside the answer becomes even more valuable than the bare citation.

→ **The velocity arms race compounds.**

Once a few teams in a category realize the engines cite in a week and the only constraint is their own publishing speed, the advantage of moving fast erodes as more teams move fast. The cheap window in this paper's conclusion is partly a window on competitor inertia. It closes as that inertia does.

→ **The citation-mention becomes the consensus metric.**

The funnel in Section 5 and the measurement argument in Section 9 point the same way. As tooling matures, “did we appear” gives way to “how often do we earn a citation-mention across repeated runs,” and the field converges on the end-of-funnel signal that actually behaves like a search result.

→ **A paid layer forms on top of organic AI visibility.**

Ads have started appearing in AI Overview and AI Mode results, still early and uneven. This looks like the SEO-to-SEM split playing out on a compressed timeline, and organic GEO and paid GEO will likely become separate disciplines with separate budgets before long.

The field is roughly where SEO was in the mid-2000s: clearly important, directionally understood, without the longitudinal evidence base that would make optimization a science rather than an informed craft. The organizations that invest in measuring their own interventions rigorously will define best practices for everyone else.

Methodology and corrections note

This paper holds to a few sourcing standards. Every statistic traces to a named primary source. Vendor self-studies are flagged as such rather than treated as neutral on their own. Figures circulating in derivative form are traced back to the primary before use. The discipline isn't new to this version; our published writing already owned this stat-cleanup work in public, including the admission that even our own materials had to be corrected before they could be repeated. This section continues that work rather than introducing an apology. Three statistics that earlier editions of this report repeated turned out to be misattributed to the wrong study or platform. Each is corrected below.

CORRECTION 01

The freshness misattribution

An earlier edition carried the claim that “Ahrefs analyzed 1.9 million AI Overview citations and found 76.4% of cited pages had been updated within 30 days, and 89.7% within the current year.” The numbers were close to right; the attribution was wrong on three counts. The figures come from [Ahrefs’ analysis of the top 1,000 ChatGPT-cited pages](#) (September 2025), not a 1.9 million-citation study, and they describe ChatGPT, not AI Overviews. Of the 564 URLs with detectable dates, 76.4% were updated within 30 days and 89.7% within calendar year 2025. Two caveats travel with it: Wikipedia is 53% of the dataset (excluding it, about 82% of pages were still updated in 2025), and “last modified” timestamps sometimes reflect a data-request date rather than a real content change.

CORRECTION 02

The SERP-overlap conflation

The 76% figure that got entangled with freshness was actually SERP overlap: the share of AI Overview citations coming from pages ranking in Google’s top 10 ([Ahrefs, July 2025](#), the original 1.9 million-citation study). That figure has since [dropped to about 38%](#) ([Ahrefs, March 2026](#)). AI surfaces are increasingly drawing from outside the traditional SERP, which is a different and more interesting finding than freshness.

CORRECTION 03**The conversion-multiple stack**

Earlier editions stacked several conversion multiples (4.4x, 5x, 9x, 23x). The 23x figure carried a Microsoft Clarity misattribution and shouldn't be repeated. The cleanest usable conversion point is from [Visibility Labs](#), reported by [Search Engine Land](#): ChatGPT referral traffic converting at 1.81% versus 1.39% for non-branded organic, a 31% lift, across 9.46 million organic and 135,000 ChatGPT sessions. Note that this is an ecommerce dataset and the absolute conversion volume is acknowledged as small, so read it as directional for B2B. One figure, properly sourced, beats a stack of mismatched ones.

On data licensing

The major AI companies have agreements with several large user-generated and reference platforms. The arrangements are real and shape which sources get cited. The specific dollar figures that circulate for them are not reliably sourced, so this paper describes the agreements qualitatively rather than quoting a number.

Figures left out

A few widely repeated figures were left out of this version because they couldn't be traced to a primary source in the verification pass, including an Amsive 13-week freshness rule, a ConvertMate 3.2x claim, and a Position Digital Perplexity freshness figure. Freshness is already carried by the verified 17 million-citation study and Seer's bot-cadence data, so these were dropped rather than chased.

Ref. EVERY FIGURE TRACES TO A PRIMARY SOURCE

Works Cited

BUYER BEHAVIOR AND THE SHORTLIST

- 6sense, [“B2B Buyer Experience Report 2025.”](#)
- Modus and Semrush, [“B2B CMO Pulse 2025: The State of Generative Engine Optimization”](#) (November 2025).
- G2, [“How AI Chat Is Rewriting B2B Software Buying”](#) (2025).
- Similarweb, [“The AI Consumer Journey”](#) (2026).
- TrustRadius, [“Bridging the Trust Gap: B2B Tech Buying in the Age of AI”](#) (2025).
- Pew Research Center, [“Google Users Are Less Likely to Click on Links When an AI Summary Appears in the Results”](#) (July 2025).
- Saharsh Agarwal and Ananya Sen, [“Google AI Overviews and Publisher Traffic: Evidence from a Field Experiment,”](#) SSRN working paper (April 3, 2026).

THE TWO CLOCKS

- Josh Blyskal (Profound), [“Time to first citation”](#) analysis (~900 pages, March–May 2026), 2026.
- Ahrefs, [“How Long Does It Take to Rank in Google, and How Old Are Top-Ranking Pages?”](#) (2025).
- Profound, [“AI Search Volatility”](#) (2026).

WHAT PREDICTS CITATION

- Pranjal Aggarwal, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, Karthik Narasimhan, and Ameet Deshpande, [“GEO: Generative Engine Optimization,”](#) KDD 2024 (peer-reviewed; arXiv preprint 2311.09735, 2023).
- Ahrefs, [“An Analysis of AI Overview Brand Visibility Factors \(75,000 Brands Studied\)”](#) (2025); December 2025 follow-up extending the pattern to ChatGPT and Google AI Mode.
- Muck Rack, [“What Is AI Reading? \(May 2026 edition\)”](#) (25 million+ AI-cited links).
- Ryan Law and Xibeijia Guan, [“AI Assistants Prefer to Cite ‘Fresher’ Content \(17 Million Citations Analyzed\),”](#) Ahrefs (July 28, 2025).

READ VERSUS CREDITED

- Akamai, [“State of the Internet \(SOTI\): Digital Fraud and Abuse Report 2025”](#) (November 2025).
- Ahrefs, [“Why ChatGPT Cites Pages”](#) (1.4 million-prompt analysis, April 15, 2026).
- Kevin Indig, [“The Ghost Citation Problem,”](#) Growth Memo (April 2026).
- Seer Interactive, [“LLM Ghost Citations: Why Your Content Is Working and Your Brand Isn’t”](#) (2026).
- AirOps, [“How Citations and Mentions Impact Visibility in AI Search”](#) (45,000+ citations, 2025).

PLATFORM DIVERGENCE

- Growth Unhinged, [“Get Recommended by ChatGPT”](#) (11% overlap, original source).
- Averi, [“AI Citation Tracking Across ChatGPT, Perplexity, and Claude”](#) (2026).
- Ahrefs, [“Only 12% of AI-Cited URLs Rank in Google’s Top 10”](#) (2026).
- Superlines, [“AI Search Statistics”](#) (2026).
- Wix Studio AI Search Lab, [“Content Types Most Cited by LLMs”](#) (March 2026).
- Kevin Indig, [“The Science of How AI Picks Its Sources,”](#) Growth Memo (March 2026).

DARK TRAFFIC AND THE BRANDED-LINK SHIFT

- Profound, [“Is Zero-Click Marketing Dead? The Branded Link Update”](#) (2026).
- Similarweb, [“ChatGPT Referral Traffic Near Triples Overnight”](#) (2026).
- Qwairy, [“The ChatGPT Linking Shift, May 2026: Inline Brand Links”](#) (140,000+ answers).

DEFENSIVE GEO

- Nick Lafferty (Profound), [“Why the Same Page Gets Cited Differently: The Citation Asymmetry Problem”](#) (2026).
- Chris Long, [“SEO practitioner post on a Rippling FAQ surfacing in Google AI Mode for a Gusto / PEO query,”](#) LinkedIn (2025).
- Mateusz Makosiewicz, [“I Ran an AI Misinformation Experiment. Every Marketer Should See the Results,”](#) Ahrefs (December 10, 2025).

MEASUREMENT AND CADENCE

- Ahrefs, [“AI Overviews Change Every 2 Days”](#) (43,000-keyword study).
- SparkToro and Gumshoe, [“AIs Are Highly Inconsistent When Recommending Brands”](#) (January 2026).
- Ahrefs, [“67% of ChatGPT’s Top 1,000 Citations Are Off-Limits to Marketers”](#) (Linehan, October 28, 2025).
- Seer Interactive, [“Study: AI Brand Visibility and Content Recency”](#) (October 2025).
- Search Engine Journal, [“Google’s John Mueller: Changing Dates on Pages Won’t Improve Rankings”](#) (Southern).

CORRECTIONS (SECTION 14)

- Ahrefs, [“76% of AI Overview Citations Pull From Top 10 Pages”](#) (July 2025).
- Ahrefs, [“Update: 38% of AI Overview Citations Pull From the Top 10”](#) (March 2026).
- Visibility Labs, via Search Engine Land, [“ChatGPT Ecommerce Traffic Converts 31% Higher Than Non-Branded Organic Search”](#) (2026).

FIRST-PARTY

- Resonate Labs, [“Insynctive engagement”](#) (2026).
- Vercel and MERJ, [“The Rise of the AI Crawler”](#) (500 million+ GPTBot fetches; JavaScript execution analysis).

About WHO PUBLISHED THIS

About Resonate Labs

This white paper is published by [Resonate Labs](#), a GEO consultancy for B2B SaaS. It starts from the premise that runs through this paper: buyers now build their vendor shortlists inside ChatGPT, Claude, Gemini, and Perplexity, often before sales is ever involved, and most companies have no way to see whether those answers include them. Resonate Labs makes that visibility measurable, running the questions buyers actually ask against the engines answering them, then showing teams what to publish to win the conversations they're missing today. The Insynctive engagement in Section 11 is a Resonate Labs client engagement.

Shane H. Tepper is co-founder of Resonate Labs and author of [Cited: How B2B Brands Win in the Age of AI-Generated Answers](#), the first practitioner-grade book on GEO. His background spans content and marketing leadership at SoFi, Udacity, and IDVerse (acquired by LexisNexis Risk Solutions in 2025), with earlier work in film and advertising. Co-founder **Kolin Simon** leads the firm's technical build; he's a lawyer, engineer, and operator.

GET IN TOUCH

WEB	resonatelabs.co
EMAIL	hello@resonatelabs.co
LINKEDIN	linkedin.com/company/resonate-labs

Version 5.0 © 2026 Resonate Labs. All rights reserved. This document may be freely distributed in its entirety. Excerpts may be quoted with attribution to Resonate Labs.

THE GEO SNAPSHOT · FREE

See where you stand in **AI** search.

The GEO Snapshot is a fast, free positioning read built from your public content footprint and competitive set. It's a first look at where your brand stands in AI search and the narrative gaps where buyers can't find you. A full GEO audit then validates it across **150 buyer-intent queries** run on the platforms your buyers actually use.

| The Snapshot is yours whether or not you go further.

Get your free GEO Snapshot →