CITED.

Why some companies keep appearing in AI answers — while yours doesn't.

A field report on how AI systems decide which companies to recommend, and the mechanics that separate the cited from the invisible.

THE SHIFT

The service firms winning long-cycle mandates — contractors, professional practices, regional providers — are not the ones ranking on Google. They are the ones being named by AI assistants when buyers ask.

Field analysis across the service sector reveals firms with decades of history and substantial regional scale absent from AI answers for their own markets. Competitors are recommended in their place — some smaller, some newer, some without notable market share. **The pattern is not about size or age. It is about a specific characteristic that the incumbents' websites do not have, and their smaller competitors sometimes accidentally do.**

The mechanics behind this shift are not the mechanics of SEO. Rank does not predict recommendation. Content volume does not drive citation. What matters is something entirely different — and most firms in the sector have not started measuring it, let alone building it.

This report is a field guide to that characteristic, the market pattern producing it, and the window of opportunity it has opened for firms willing to act on it.

FINDING 01

Google rank and AI visibility have decoupled.

88% of Google AI Mode citations do not come from the organic top ten. 73% of first-page rankers receive zero AI mentions in their category.

FINDING 02

Third-party signals matter more than what companies say about themselves.

Brand mentions across the web correlate with AI visibility at 0.66 to 0.74 — roughly three times stronger than backlinks or domain authority.

FINDING 03

Visibility is a surface-area game.

Firms that dominate AI recommendations have expanded their presence deliberately — from a well-structured website outward to directories, listings, and third-party mentions. Most firms have started building on only a fraction of the surface available to them.

SECTION 01

Google Rank No Longer Predicts Discovery

68% of Google searches in 2026 end without a click. When an AI Overview appears on the results page, that figure climbs to 83%.

In Germany, Sistrix quantified the local effect in February 2026: 265 million clicks per month are absorbed by AI Overviews before they reach a website. Position 1 in Google, which used to earn a 27% click-through rate, now earns 11% when an AI Overview is present.

The click loss is not the story. The story is that a new visibility layer has formed above the traditional results page. The buyer no longer picks which link to click. The AI has already picked.

For long-cycle service businesses, this shift lands harder than in most sectors. The buyer journey — a homeowner selecting a contractor, a small business owner choosing an accountant or lawyer, an operations lead picking a long-term maintenance partner — has always begun with research. What changed is where the research happens. It now happens inside ChatGPT, Perplexity, Google's AI Mode, and Gemini. Increasingly, it also happens inside TikTok, YouTube, and Reddit search. **The website is no longer the destination. The recommendation is.**

And the companies AI recommends are not the ones ranking on page one of Google. Moz's 2026 analysis found 88% of Google AI Mode citations do not come from the organic top 10. Independent studies place the share of page-one rankers with zero AI mentions at 73%.

> The two systems have decoupled. A firm can dominate Google's first page and be entirely absent from every AI answer in its category.

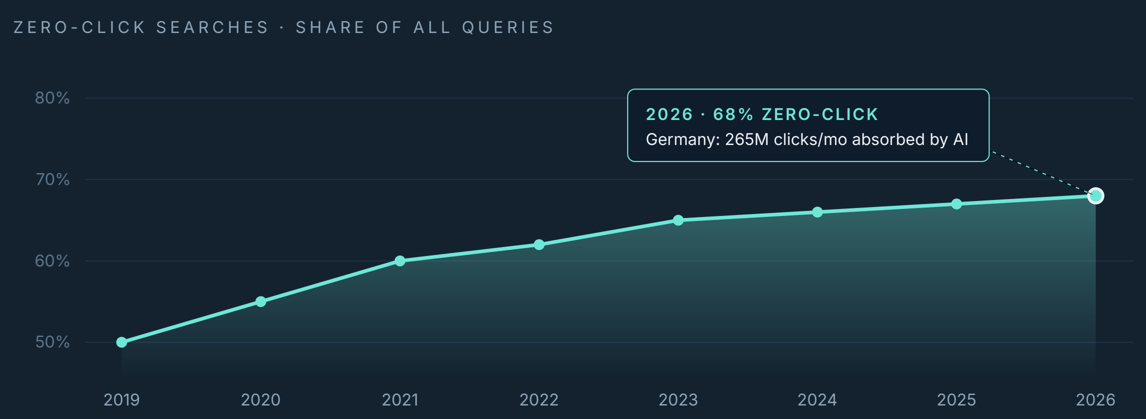

FIG. 1 · THE SHARE OF SEARCHES ENDING WITHOUT A CLICK, 2019–2026. SOURCES: SPARKTORO, SIMILARWEB, SISTRIX (DACH).

SECTION 02

How AI Chooses What to Cite

A modern AI assistant does not retrieve "the best page" and quote it. It performs two mechanically distinct operations, and both reward different properties than SEO does.

Stage One · Query Fan-Out

The AI decomposes the buyer's question into five to ten sub-queries, most of which the buyer never types. Analysis of 15,000 prompts (AirOps, 2026) found 89.6% of prompts triggered two or more fan-out sub-queries, generating 43,000 total queries. 95% of those sub-queries have zero traditional search volume. 32.9% of pages ultimately cited by the AI appear only through fan-out — they would never have been retrieved by the original question.

The consequence: the retrieval pool for AI answers is not the Google SERP. 83.4% of pages retrieved by AI do not rank in Google organic at all.

Stage Two · Selection

From the retrieval pool, the AI picks which pages to cite. 85% of retrieved pages are never chosen. Selection is driven by semantic fit to the sub-query — cited pages score 0.656 similarity versus 0.484 for non-cited pages — and by extraction cost, meaning how easily a clean, self-contained answer can be lifted from the page.

Selection largely ignores traditional authority. Sites with domain authority 40 to 80 match or beat DA 90+ citation rates post-retrieval. In one 82,000-citation dataset, approximately 74% of citations went to sites with domain authority below 80.

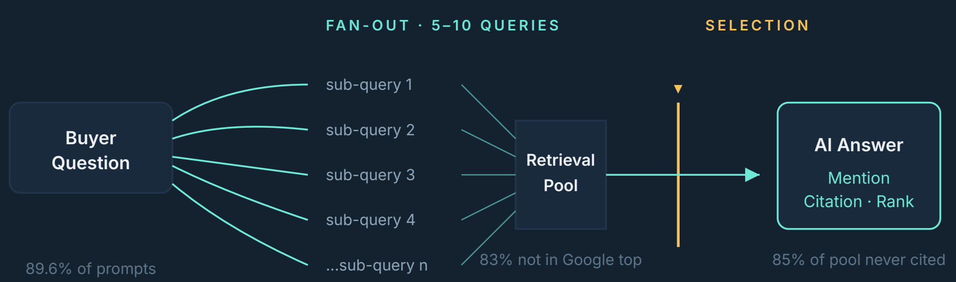

FIG. 2 · THE TWO-STAGE AI CITATION PIPELINE. RETRIEVAL EXPANDS THE POOL; SELECTION CONTRACTS IT.

Three Outcomes, Not One

What the buyer sees is not a single result. It is three things at once, each driven by different signals.

MENTION	CITATION	RANK
The company is named in the answer. Drives recommendation and shortlisting. No click required.	The company's URL is provided as a source. Drives traffic and research-stage engagement.	Being named first in a list of three is worth substantially more than being named third. Buyers see the first name and stop reading.

A company can be mentioned without being cited — the buyer hears the name but the source link points elsewhere. It can be cited without being mentioned — used as a source but not recommended. Two companies both mentioned are not equal; order does most of the work.

Any serious analysis of AI visibility must distinguish these three. Most vendor scoring tools collapse them into one metric, producing a misleading picture of dominance.

SECTION 03

Surface Area — What AI Actually Rewards

Both stages of the pipeline reward the same underlying quality: the breadth and depth of a company's semantic presence across the web.

Not how many pages exist. Not how much content is published. How much of the query-space the company plausibly answers, and how consistently.

Call this **surface area**.

Fan-out queries are probes across a semantic space. Whether a company gets retrieved depends on whether its surface covers the sub-query the AI happens to generate — pricing, comparison, specific service variant, location, alternative-to query, small-scale versus large-scale variant. Broader surface, more probes hit. Narrow surface, most probes miss.

The Five Layers of Surface Area

Surface area lives on five layers. Each answers a different type of fan-out query. Firms that dominate AI recommendations occupy all five. Firms invisible to AI typically occupy only one or two.

▲ HIGH-INTENT QUERIES

05 · COMPARATIVE
Content positioning the company · listicles, "X vs Y", "alternatives to Y" — on your own site or externally.

04 · DISCUSSION
Where humans talk · forums, Reddit, LinkedIn threads, YouTube reviews. Source of multi-perspective texture.

03 · CORROBORATING
Independent sources · associations, directories, review platforms, Google Business Profile, LinkedIn.

02 · STRUCTURED
Machine-readable labels · schema markup, sameAs, verified profiles that link the entity across surfaces.

01 · OWNED
Your website · pages, services, credentials, team, FAQ. The foundation on which every other layer builds.

▼ FOUNDATIONAL

FIG. 3 · THE FIVE LAYERS OF AI SURFACE AREA. EACH ANSWERS A DIFFERENT CLASS OF FAN-OUT QUERY.

Each layer requires different investment and answers a different type of query. A firm dominant in one layer but absent from the others has a lopsided surface. It wins the queries that hit its dominant layer and loses everything else.

Coverage, Not Volume

Publishing more pages that restate what existing pages already cover does not expand surface. It dilutes it. This is what the widely-cited "content volume" correlation of 0.194 actually measures — undifferentiated publishing.

What does expand surface is **coverage**: pages that plausibly answer sub-queries the site does not already answer. A single deep page answering one high-intent sub-query cleanly is worth more than fifty thin pages restating a variation of the same one. And multiple dedicated pages — one for each distinct sub-query — will outperform a well-written consolidated page on comparison and pricing queries, because semantic similarity is scored per page in retrieval, not per section.

Depth on a topic and breadth across topics are both legitimate expansions of surface. Undifferentiated volume is not.

Why Statistics Lift Citations

The Princeton / Georgia Tech / IIT Delhi GEO study, tested against a 10,000-query benchmark, found that adding concrete statistics to content lifted visibility by up to 41%. Quoting credible authoritative sources added further gains.

Statistics are atomic units of extractable surface: small, precise, self-contained, hard to fabricate, easy to attribute. Prose describing a company's expertise is diffuse and difficult to lift. A statistic — *"served 4,500 clients across three cities since 2011"* — is a discrete piece the AI can extract confidently. Statistics also implicitly claim authority: only a company that has counted knows the number.

The same principle explains why keyword stuffing produces near-zero effect. It inflates surface volume without adding extractable substance.

THE FORMULA

AI visibility = semantic clarity + entity clarity + answer-ready content + topical coverage + external corroboration

Semantic clarity — each page signals unambiguously what it is about. **Entity clarity** — who the company is, what it does, where it operates, consistent across surfaces. **Answer-ready content** — questions answered directly, front-loaded, with concrete substance. **Topical coverage** — the site plausibly answers the range of sub-queries buyers ask. **External corroboration** — independent sources confirm the picture.

SECTION 04

What Doesn't Move the Needle

The vendor market has produced a set of universal prescriptions for AI visibility. Most do not survive controlled measurement.

WHAT IS BEING SOLD	WHAT THE EVIDENCE SHOWS
× Publishing more of the same content	Correlation between site page count and AI visibility: **0.194** (Ahrefs, 75,000-brand study). Restating existing coverage does not scale citation probability.
× Building more backlinks	Correlation **0.218** — roughly one-third the strength of unlinked brand mentions. The link graph is not the citation graph.
× Generic schema deployment	Controlled Ahrefs study of 1,885 pages: **no significant citation lift**. Exception: schema with populated concrete attributes (real address, ratings, specs) lifts citation rates from 41.6% to 61.7%. Empty schema is decorative.
× Press wire releases	**0.04%** of AI citations across a 3,600-prompt, ten-industry study. Not adopted at meaningful scale.
× llms.txt files	300,000-domain study returned **null-to-negative** results. Not adopted by major engines at scale.
× Ranking on Google page one	**73%** of page-one rankers have zero AI mentions in their category. Rank is a precondition, not a driver.
× A long operating history	Firms with fifty years of operation and substantial market presence are **frequently absent** from AI answers. The models cannot read what the ecosystem does not surface.

SECTION 05

What Actually Drives Citations

Five levers. Each maps to one term in the formula and to one or more surface layers.

1. SEMANTIC CLARITY · surface layer: owned

Every page on the site should signal unambiguously what it is about. Title, H1, URL slug, opening paragraph, and content beneath should agree with each other. A page titled *"Our Services"* that covers plumbing, HVAC, electrical, and general contracting has no clear semantic identity — the AI cannot classify it against a specific sub-query. A page titled *"Emergency Plumbing in Denver"* focused on that single topic has high semantic clarity and is retrievable for the sub-queries that match it. Descriptive URL slugs correlate 89.78% citation rate versus 81.11% for opaque ones.

2. ENTITY CLARITY · surface layers: owned + structured

Can the AI answer, unambiguously: who is this company, what does it do, where does it operate, and what evidence confirms all three? Named team members. Verifiable credentials. Precise service descriptions. Dedicated location pages for each market served. Consistent name, address, and phone data across every external surface. This is the mechanism behind a pattern that surprises SEO practitioners: small, young companies with thin content sometimes outperform larger incumbents. They win not by publishing more but by being unambiguously one thing.

3. ANSWER-READY CONTENT · surface layer: owned structure

Answer-first paragraphs — the direct response to the buyer's question in the first paragraph. Concrete statistics. Quoted authoritative sources. Literal buyer questions used as headings. 44.2% of AI citations come from the first 30% of a page's text. If the answer is not near the top, the extraction fails.

4. TOPICAL COVERAGE · surface layers: owned + comparative

The site plausibly answers the range of sub-queries buyers ask about the core service. Dedicated pages for pricing, comparisons with named competitors, service-plus-location combinations, alternative-provider queries, small-scale versus large-scale variants. Depth on individual topics matters; breadth across topics matters more.

5. EXTERNAL CORROBORATION · surface layers: corroborating + discussion

Independent sources naming and describing the company. Industry associations, chamber listings, sector-specific directories, review platform profiles, Google Business Profile, LinkedIn, community mentions. This is the single strongest measurable correlate of AI visibility. Models learn what to say about a company from what others say about it. Your own pages are one voice. The ecosystem is the consensus.

The Signal Strength Ranking

Across 75,000 brands measured by Ahrefs and cross-validated in independent studies from Muck Rack, Bain, and Semrush, the correlations are consistent: third-party signals dominate. Backlinks and domain authority — the pillars of SEO for two decades — correlate roughly one-third as strongly.

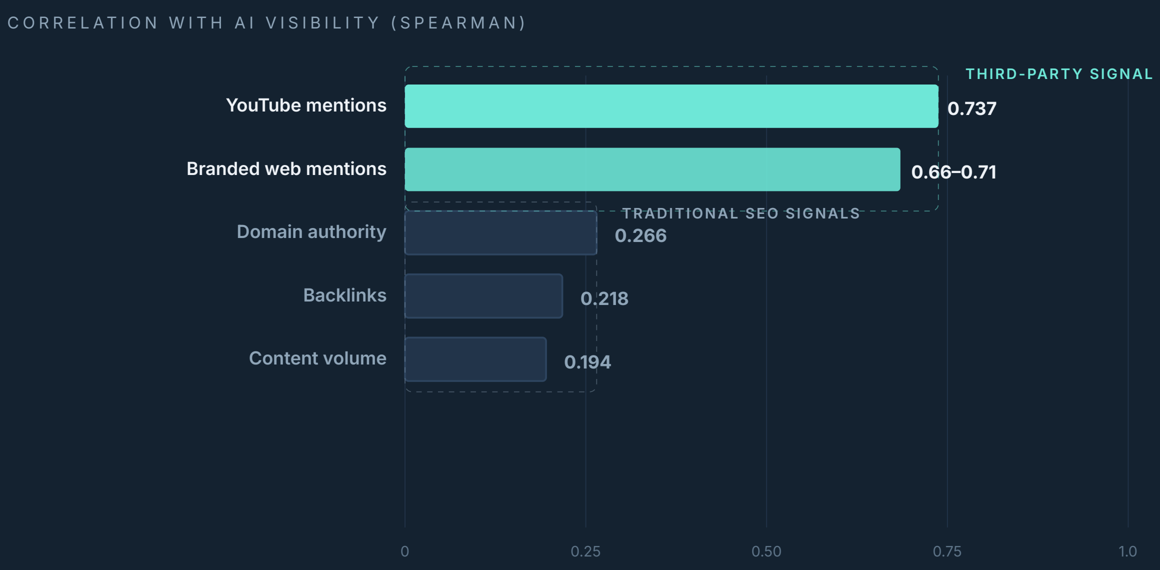

FIG. 4 · SPEARMAN CORRELATION OF EACH SIGNAL WITH AI VISIBILITY. SOURCE: AHREFS, 75,000-BRAND STUDY, 2026.

Retrieval eligibility — pages must be indexed, renderable, and accessible without JavaScript walls or paywalls — sits underneath all five as a precondition. A blocked page cannot be cited regardless of how well it satisfies the levers above.

SECTION 06

Three Firms, Three Outcomes

Across service industries examined in the field, firms cluster into three recurring patterns. Each shows a different way of ending up somewhere on the visibility spectrum — from invisible to hidden winner.

The Invisible Veteran

DECADES OLD · UNSEEN

A firm operating for decades. Real scale — a substantial client base, regional recognition among peers. Absent from every AI query tested for its city. History and scale are not signals AI can read directly. When the site does not assert identity, services, and locations — and independent sources do not corroborate them — the model has no material.

Semantic clarity	weak
Entity clarity	weak
Answer-ready	weak
Topical coverage	weak
External corroboration	weak

The Google Winner

RANKS · BUT ABSENT

A firm with strong search visibility. Ranks for the terms buyers historically typed. Fully optimised for Google. AI visibility close to zero. The decoupling in one company. The owned surface is well-developed for SEO but not structured for AI extraction. Structured, corroborating, and discussion layers are absent. Google traffic is arriving. The buyer's decision is being made elsewhere.

Semantic clarity	mixed
Entity clarity	weak
Answer-ready	weak
Topical coverage	mixed
External corroboration	weak

The Hidden Champion

CITED · QUIETLY

A firm mentioned as a top competitor across a significant portion of its peers' AI audits. Recommended for the flagship query in its market. Under audit: no LocalBusiness schema, no sameAs, no FAQ, no answer-first content. And still winning. Two things done right: clean semantic focus on the owned surface, and thorough external presence. The position is not permanent. Any firm that closes the missing layers can take it.

Semantic clarity	strong
Entity clarity	mixed
Answer-ready	weak
Topical coverage	weak
External corroboration	strong

SECTION 07

The Opportunity of a Lifetime

The layers of surface area compound — and the fundamentals that trigger the compounding are within reach of firms of any size.

A firm with a semantically clear website, well-structured schema, and thorough coverage of the sub-queries buyers ask becomes discoverable by AI systems from the moment those pages are published. As presence expands beyond the site — a maintained Google Business Profile, industry association listings, sector directories, review platform profiles, occasional third-party mentions — the classification becomes stronger. Stronger classification means retrieval across more sub-queries. More retrieval means more recommendations. More recommendations mean more buyer exposure, which produces more references over time.

Not every layer accumulates at the same speed, and this matters for what a firm can realistically do. **Owned-surface work** — schema completeness, page-level clarity, answer-ready structure, topical coverage — can be installed deliberately in weeks. **Google Business Profile, directory registrations, and social presence** take similar effort and are within reach of any firm willing to do the work. **Review accumulation and organic third-party mentions** build over months and years and are the layer where larger firms have a natural head start. But volume of reviews is not a prerequisite for appearing in AI answers. Firms with modest external presence but sharp fundamentals on their own site regularly outperform larger competitors whose sites are muddled.

Sustained recommendation frequency itself becomes a form of authority. To an industry expert, a company that shows up in every AI answer is not necessarily the best. The expert can name three technically superior competitors. But to the buyer — who is almost never an industry expert — a company appearing in every AI answer, on every directory, on every review platform *is* the top-tier provider. Perceived authority is the only kind of authority that converts.

> Here is the asymmetry that matters. In every market examined, incumbents with decades of history and scale are still investing in the previous era's playbook. The AI visibility positions are largely unclaimed. This is a temporary condition.

For any firm willing to move now — regardless of size — this is the widest strategic window that will exist in the category for the next decade. The fundamentals are within reach. The competition has not yet started. Firms that install the foundation first become the reference AI systems anchor to for years, while later entrants face the harder task of displacing an established name.

Every quarter of delay costs disproportionately more than the last.

FIG. 5 · SURFACE EXPANSION COMPOUNDS. EACH CYCLE MAKES THE NEXT EASIER.

SECTION 08

How to Get Cited

Six moves. Each expands one term in the formula. Ordered so that each move builds on the ones above it — and so that the moves a firm can execute on its own site come first.

0 — Fix the accessibility floor — PRECONDITION

Critical service pages indexed, renderable, free of staging leaks or duplicated content. A floor, not a lever. Every move below is worthless while this is unmet.

1 — Sharpen semantic clarity, page by page — SEMANTIC CLARITY

Every important page signals unambiguously what it is about. Title, URL slug, opening paragraph, and content agree. One page, one topic.

2 — Answer the four entity questions — ENTITY CLARITY

Who the company is. What it does. Where it operates. What evidence confirms these facts. Consistent representation across every surface.

3 — Structure content for extraction — ANSWER-READY

Front-loaded answers. Concrete statistics. Buyer questions used as headings. Populated schema with real data — not empty schema fields.

4 — Cover the fan-out — TOPICAL COVERAGE

Dedicated pages for pricing, comparisons with named competitors, service-plus-location combinations, alternative-provider queries. Turns recommendation wins into citation depth across intent categories.

5 — Extend external corroboration — EXTERNAL CORROBORATION

Industry association memberships. Directory profiles. Review platforms. Google Business Profile complete and verified. LinkedIn maintained. Lives outside the website and takes longer than the moves above, but it is the multiplier that compounds everything else.

Freshness sits underneath all six. AI-cited content runs approximately 25.7% fresher than organic top-10 results. Aging pages lose citation eligibility over time.

SECTION 09

Buyers of long-cycle services form their shortlist before the first phone call.

That shortlist is built inside an AI answer. The companies on it are the companies that will win a decade of contracts.

The mechanics governing the shortlist are not the mechanics of SEO. Rank is a precondition. Content volume is noise. Backlinks are secondary. What matters is whether the AI can classify the company unambiguously, corroborate it through the ecosystem, and extract a clean answer from its pages.

For most service firms today, none of the three are fully in place.

That is the opportunity.

Continue the argument at smartypant.xyz →

A note on this report — findings draw on field analysis conducted across service-industry businesses, supported by cited third-party research from Ahrefs (75,000-brand study, 2026), Moz (AI Mode citation analysis, 2026), Bain (B2B citation patterns, 2026), Muck Rack (media citation origins, 2026), SparkToro / Similarweb (zero-click search, 2026), Sistrix (DACH AI Overview impact, February 2026), AirOps (15,000-prompt fan-out study, 2026), Zyppy (54-study meta-analysis, May 2026), and the Aggarwal et al. GEO benchmark from Princeton / Georgia Tech / IIT Delhi.