

#SEO Article 1

#Platform Claude (Sonet 5 High, 20th August ~4.57pm, MP, India)

Claude finished the response

Searched the web

Searched the web

Here's how I'd split this for a team that already does solid technical and content SEO — organised by your four buckets, with the specific topics you named folded in where they actually belong rather than treated as a separate checklist.

Continue doing — this is still the foundation

Clean sitemaps, correct canonicals, fast rendering, no orphaned pages, no crawl traps. AI answer engines still largely retrieve from the same indexed web (often via Bing or Google's index, or their own crawl), so if Googlebot can't parse it cleanly, nothing downstream works either.

Nothing in the AI-search shift changes the fact that thin, derivative, inaccurate content underperforms. If anything it's punished harder, because LLMs re-rank retrieved passages for factual specificity and will skip a vague page for a more concrete competitor.

Still correlates with both organic rank and AI citation — it's just no longer the strongest correlate (more on that below). Don't defund it; reweight it.

Query fan-out (below) adds a layer on top of this; it doesn't replace it. You still need to know what people search for before you can anticipate what an LLM will fan that query into.

These remain ranking factors for the organic results AI systems often draw from, and there's no evidence AI crawlers are more tolerant of slow or broken pages — several AI crawler guides note the opposite, that real-time answer crawlers (OAI-SearchBot, PerplexityBot) have low tolerance for redirect chains and slow pages.

Change or add — AI discovery genuinely requires this

OpenAI and Anthropic now run separate bots for training versus live answering: GPTBot (training) vs OAI-SearchBot (answers ChatGPT search), ClaudeBot (training) vs Claude-SearchBot. Blocking the search-specific bot means you won't appear in ChatGPT search answers, though navigational links may still surface. The 90-day action: audit robots.txt against the current bot list (it drifts — recheck quarterly), decide separately on training vs. answer bots, and then check your CDN/WAF, because a meaningful share of sites accidentally block major AI crawlers at the CDN layer while their robots.txt says allow. One data point on the stakes: a Q1 2026 audit of B2B sites found 41% still blocking at least one major AI bot, usually a leftover from 2023–2024 blanket-block policies, with each blocked bot estimated to cost 18–34% of potential citations on that engine. [Anagram + 2](#)

This is the one piece of the AI-search shift with genuinely convergent evidence across independent sources: AI systems chunk a page and score passages independently, so a page needs multiple self-contained, citable sections rather than one narrative arc. A March 2026 study from the University of Tokyo and University of Tsukuba found that document structure — chunking, heading design, answer placement — affects citation probability independently of what the content actually says. Practically: answer-first opening sentences per section, descriptive (not clever) H2/H3s, one idea per paragraph, comparison tables for anything comparative. Treat this as a content-ops standard, not a one-off project. [Machine Relations](#)

This is the strongest "genuinely new" finding across the data: branded web mentions correlate with AI Overview visibility at roughly 0.66, versus 0.22 for backlinks, and a separate Ahrefs analysis of 75,000 brands found YouTube mentions were the single strongest signal at 0.737. Concretely: a Wikidata entry with verified attributes, a consistent canonical "About" entity home, consistent naming/sameAs across LinkedIn, Crunchbase, review

platforms, and pursuing genuine Wikipedia notability if you qualify. This is entity SEO becoming load-bearing rather than a nice-to-have. [Digital Applied TeamBetterAISearch](#)

Brands with third-party trust signals are reportedly cited in 75% of AI answers versus 1% without them, and earned media reportedly accounts for the large majority of AI citations — a surface pure on-page optimisation doesn't touch. Different engines skew to different third-party surfaces: ChatGPT leans heavily on Wikipedia and Reddit, Google AI Overviews on Reddit and YouTube, and Perplexity on Reddit above all else. This means digital PR and community engagement (Reddit threads, G2/Trustpilot reviews, YouTube presence) move from "nice for brand" to "input to the citation pipeline." Build it into the content calendar, not just comms. [Shadow + 2](#)

A large share of AI answers never produce a click, and the traffic that does convert is often misattributed as direct or unlabeled referral. The practical minimum: a repeated prompt set run across ChatGPT, Perplexity, Gemini/AI Overviews and Copilot on a weekly or biweekly cadence (a single snapshot is close to meaningless given run-to-run variance), tracking citation frequency, share of voice, and sentiment, not just presence.

Google and OpenAI's retrieval systems decompose a query into related sub-queries before answering — breaking a question down into multiple search queries around related subtopics that the user didn't explicitly ask for. The practical shift is from "one page targets one keyword" to "one topic cluster needs to win multiple sub-queries" — pricing, comparisons, edge cases, objections, not just the top-of-funnel definition. Map your money queries in AI Mode/ChatGPT, see which sub-queries your content currently wins, and fill the gaps. [Digiday](#)

Test — promising, not conclusive

The strongest controlled study here found close to zero lift from adding schema — an Ahrefs study tracking 1,885 pages found no statistically significant increase in ChatGPT or AI Mode citations after adding schema, and AI Overviews citations actually dropped 4.6%. But a separate cross-platform academic study found nuance worth testing: pages with Product or Review schema carrying concrete attribute fields — pricing, ratings, specifications — were cited at substantially higher rates than generic schema types, and the effect was strongest for lower-authority domains. So: don't roll out blanket JSON-LD expecting citation lift, but do A/B test attribute-rich Product/Review schema on category pages where you're competing against higher-authority sites. Keep it for rich results and knowledge-graph disambiguation regardless — those benefits are separately established. [FractionalseoSSRN](#)

The underlying mechanism (passages get scored and cited independently) is well-supported; the specific numbers circulating (134–167 words per chunk, 40–60 word answer blocks, etc.) are repeated almost verbatim across vendor content without a traceable, replicated source. Worth testing on your own content with your own citation tracking rather than adopting the numbers as fact.

A Princeton GEO study presented at KDD 2024 found that adding statistics to content improved AI citation visibility by up to 41% across 10,000 queries — a real, citable study, but a single one, on one tactic among nine tested. Worth testing on a page cohort before treating it as doctrine. [ALSEO Agency](#)

Generate one automatically if your CMS supports it (Yoast, Webflow) and move on — don't fund custom engineering for it.

Do not treat as proven, despite the claims

This is the clearest case of hype outrunning evidence. In September 2025, John Mueller confirmed Google does not use lms.txt for search or AI Overviews, and Gary Illyes reiterated this later that year; Google's 2026 generative-AI documentation lists it among unnecessary tactics. One study using an XGBoost model to predict citation frequency found that removing the lms.txt variable actually improved the model's prediction accuracy — the file added noise, not signal, and among the fifty most AI-cited domains in one analysis, only one had the file at all. Don't let a team invest real engineering time here. [Layer3Labs + 3](#)

You'll see stats like "schema boosts AI citations 30–36%" circulating in vendor content; the best controlled evidence available directly contradicts this for generic schema types. The correlation between AI-cited pages and schema presence is real but reflects a confound — sites that add structured data also tend to do the technical SEO, content, and link work that actually earns the citation — not causation. [MADX Digital](#)

Treat any number more specific than "keep sections self-contained and lead with the answer" as marketing copy dressed as research until you've replicated it on your own content.

This gets repeated as settled fact but the underlying figures conflict sharply across sources — anywhere from roughly 68% of AI Overview citations sitting outside the organic top 10 to

roughly 75–92% of citations coming from top-10/top-12 pages, depending whose study you read. The honest read is that ranking remains a strong but no longer sufficient signal — don't deprioritize core ranking work on the strength of a stat this contested.

Every GEO/AI-visibility platform samples a finite prompt set and none have disclosed methodology that's independently verified; the platforms with more statistical rigor run millions of prompts monthly specifically because small samples don't reveal reliable patterns, which most tools don't do. Use these tools for directional trend tracking against your own baseline, not as a scoreboard against competitors or across vendors. [Evertune](#)

If I had to compress this into one sentence: the 90-day shift isn't a new SEO discipline, it's a reweighting — from "backlinks + schema + rank position" toward "entity clarity + earned third-party mentions + passage-level content structure + crawler access hygiene," while most of the tactics being sold hardest right now (llms.txt, schema-as-citation-hack) are the ones with the weakest evidence behind them.

Searched the web

Searched the web

Fair challenge — that's exactly the audit worth doing before a team spends budget on any of it. Here's each claim from the "genuinely changes" list, broken into the three tiers, with what's actually behind it.

Tier 1 — Verified by platform documentation or strong experimental evidence

This is the most solidly sourced claim on the list.

Google's own search documentation names query fan-out as a core technique for retrieving content to generate AI answers, and a Google engineer described it publicly as a technique that triggers multiple searches at once and synthesizes them into one response. It was also stated on the record at Google I/O by Google's Head of Search. This isn't inferred from behavior — Google put a name on it and explained the mechanism themselves. [SemrushGoogle](#)

OpenAI's own documentation states that sites blocking OAI-SearchBot will not appear in ChatGPT search answers, though navigational links may still surface, and both OpenAI and Anthropic publish separate, named user-agents for training crawlers versus real-time answer crawlers. This is first-party, verifiable in your own server logs, and not a matter of interpretation. [Anagram](#)

Google's own fan-out documentation describes decomposing a query, retrieving from multiple sources in parallel, then generating a synthesized answer — that's RAG, described by the platform itself, not reverse-engineered by SEOs.

A controlled study running identical prompts repeatedly found cited-source sets overlapped only 34–42% between consecutive days under identical conditions, with brand-mention sets somewhat more stable at 45–59% but still highly variable. I'll flag one honest caveat: this is a preprint (arXiv), meaning it hasn't been through formal peer review — treat it as strong evidence, not gospel. [arxiv](#)

, not inference — each platform discloses whether it browses live or answers from a static training snapshot. Nothing contested here.

This isn't really an empirical claim needing a study — it follows from how those tools are built (they count sessions and clicks; a mention with no visit produces neither). It's closer to a logical necessity than a research finding.

Tier 2 — Observational correlation or reasonable inference, not causally proven

Multiple independent large-sample studies point the same direction: Ahrefs measured a 58% drop in click-through for the top organic result when an AI Overview is present, up from 34.5% a year earlier, and Pew Research's behavioral study of nearly 69,000 real searches found an 8% click rate with an AI Overview present versus 15% without. That's about as good as observational web-behavior evidence gets — independent methodologies, large samples, consistent direction. But notice the magnitude disagreement (34%, 47%, 58%, ~60% across different studies) — that spread is itself a sign these are correlational measurements on different samples and time windows, not a single controlled experiment. The causal story (an on-page answer removes the reason to click) is plausible and even matches Google's stated product intent, but no platform has published first-party data confirming the causal size of the effect. Treat direction as solid, magnitude as uncertain. [SEO-KreativArfadia](#)

Off-site brand mentions correlating with AI citation more strongly than backlinks — genuine data, but a live confound

The oft-cited figure (mentions correlating around 0.66, backlinks around 0.22) comes from real correlational analyses, not fabrication. But the number itself isn't stable — different sites citing what's ostensibly the same research report it as 0.664, 0.709, and other variants, which is a sign of a stat being passed hand-to-hand without anyone checking the original methodology. More importantly, there's an obvious confound the industry summaries never address: brands that are already more AI citations, for the same underlying reason (they're bigger, more covered, more trusted) — not necessarily because the mentions themselves cause the citations. This is a textbook case where correlation is being narrated as "mentions matter more than links now," when the honest read is "these two things travel together and we haven't isolated why."

Tier 3 — Commonly promoted by the SEO/GEO industry without strong evidence

(e.g., 134–167-word passages, 40–60-word answer blocks). The underlying mechanism — that AI systems retrieve at passage level — is well-supported (see Tier 1). But the precise numbers attached to it are repeated near-verbatim across dozens of GEO blog posts with no consistent traceable primary source, no disclosed sample, and no replication. This has the signature of content marketing copying content marketing, not research.

(e.g., a specific tool's finding that only 35% of domains repeat across runs, or that a brand needs to hit a certain mention percentage). These are usually single-vendor internal analyses run on proprietary, undisclosed prompt sets, published to promote that vendor's tracking tool. The general phenomenon they're illustrating (instability, need for repeated sampling) is real and backed by the Tier-1 arXiv study — but the specific numbers from any one vendor blog shouldn't be treated as an industry constant.

Claims that purpose-built AI-visibility platforms are meaningfully more accurate than bolt-on features in existing SEO

This shows up constantly in comparison content, but it's coming from the vendors selling the dedicated tools, comparing themselves favorably against competitors, with no independent, methodology-disclosed benchmark behind the claim. Reasonable to expect some difference in prompt-sampling scale; unreasonable to treat any specific accuracy claim as verified.

platform-documented mechanics (fan-out, bot separation, RAG architecture) are the most trustworthy layer of this whole topic, because they don't depend on anyone's interpretation. Large-scale behavioral measurements (CTR decline, answer instability) are the next tier down — real and converging, but observational, with magnitudes that shift by study. And almost anything with a suspiciously precise number attached to it (a correlation coefficient to three decimal places, an exact optimal word count, a specific percentage lift) is worth tracing back to its original source before you build a strategy on it — in this space, that source is very often another vendor's blog post citing yet another vendor's blog post.



Searched the web

Searched the web

Ranking well in Google answers a different question than any of these five stages ask. A page can be #1 organically and still fail at every one of them, for different and mostly unrelated reasons. Here's the pipeline broken apart, with what's actually confirmed versus what's reasonable inference.

Retrieval — does the page even enter the candidate pool?

This is where "but we rank #1 on Google" most often turns out to be irrelevant, because most AI answer engines aren't querying Google's index at all.

ChatGPT Search retrieves primarily through Bing's index plus OpenAI's own OAI-SearchBot crawler — if a page isn't in Bing's index, ChatGPT cannot cite it, and if it blocks OAI-SearchBot in robots.txt, ChatGPT won't surface it even when the page ranks well on Bing. A site can be flawlessly optimized for Google and functionally invisible to ChatGPT for reasons that have nothing to do with content quality — Bing simply hasn't indexed the page, or a robots.txt rule written years ago for a different purpose is silently blocking the one bot that matters. (Independent researchers have also found evidence ChatGPT

draws on Google's index in some configurations, something OpenAI has not officially confirmed — worth flagging as contested, not settled.) [SearchIntelalblueprints](#).

JavaScript rendering is a second, separate gate. A joint analysis by Vercel and MERJ, based on over 500 million GPTBot fetches, found zero evidence of JavaScript execution by GPTBot, ClaudeBot, or PerplexityBot — they fetch raw HTML and don't wait for client-side rendering. Vercel's own crawl analysis confirmed the same pattern across the AI crawlers they measured. If a site's main content, pricing, or FAQ text is injected client-side (common in React/Vue single-page apps), Googlebot renders it fine and ranks the page — but an AI crawler receives an empty shell. This is a purely technical failure mode

[Blick](#)

[AlpacaVercel](#)

Crawler access is bot-specific, not domain-wide. As covered earlier, blocking a training bot doesn't block the search bot and vice versa — a robots.txt written to keep GPTBot out of training data can accidentally also exclude OAI-SearchBot from live answers if the rules aren't scoped precisely.

the model silently generates for a given prompt. OpenAI has never published a complete architecture diagram for this pipeline, so which sub-queries get generated, and how many candidate documents get pulled per sub-query, is inferred from practitioner testing, not confirmed. A page can rank well for the head term and still never be retrieved for the three or four sub-queries that actually determined the answer. [Lemniscate Growth](#)

Source selection — of the pages retrieved, which get pulled into context?

This is a reranking step, separate from retrieval, and it's where "ranks well" most clearly stops mattering.

Practitioner-documented behavior describes a reranking stage where a smaller model scores retrieved candidates before any get used in generation — a step with no equivalent in classic organic ranking. What that reranker actually weights isn't published by any platform, so specific claims about its criteria are inference from observed patterns, not confirmed fact. [Lemniscate Growth](#)

Different engines seem to systematically prefer different types of sources over brand-owned pages regardless of the brand page's Google rank — one analysis found ChatGPT disproportionately citing Wikipedia and Reddit, Google AI Overviews leaning on Reddit and YouTube, and Perplexity favoring Reddit above all else. This — trust heuristics, training data composition, an

[Discovered](#)

[Labs](#)

Passage-level extractability (self-contained, answer-first sections) likely matters more here than at the retrieval stage, since this is specifically the step where a reranker judges relevance of a specific chunk of your page. Plausible given the RAG architecture Google itself describes, but no platform has confirmed reranker weightings publicly.

Citation — of the sources actually used to generate the answer, which get named?

This is a distinct step from selection, and it's genuinely under-documented.

ChatGPT Search's standard browsing mode typically returns a handful of numbered, clickable citations per response rather than every source consulted — meaning the visible citation list is a curated subset, not a full accounting of what informed the answer. Products cap how many citations they display. [Yoast](#)

Whether a source that was retrieved and used to inform the synthesized text simply doesn't make the final citation list — because of deduplication (multiple sources say the same thing, one gets credited), because of a display cap, or for some other selection logic — is not something any provider has documented in detail. This is the single least transparent stage in the whole pipeline. Anyone claiming to know the exact algorithm here is speculating from observed outputs, not citing a disclosed mechanism.

Brand mention — separate from citation entirely

in an answer with no link and no attributed source at all.

Models operate in (at least) two distinct modes — answering from parametric training knowledge, or grounding in live retrieval. A routing step decides whether a question needs live data or can be answered from the model's weights alone, and when the search step doesn't trigger, the answer comes from training data with no sources shown. Presence in training data reflects years of accumulated web corpus, while citations depend on the state of a live search index right now — these are two separate memory systems with separate failure modes. A brand can be well-optimized on today's live web and still be

poorly represented in an older training snapshot, or vice versa. [Lemniscate GrowthVistrix Labs](#)

How strongly a brand shows up in a model's parametric memory likely sources during the training period (news, Reddit, Wikipedia, forums) rather than by its own site — consistent with the observed correlation between third-party mentions and AI visibility discussed earlier. But no provider discloses how any specific brand is weighted in training, so this is inference from a plausible mechanism, not a documented fact.

Entity ambiguity (a generic brand name, no disambiguated Wikidata/knowledge-graph record) plausibly makes it harder for a model to reliably associate scattered mentions with one consistent entity. This is standard entity-SEO reasoning and mechanically plausible, but not something any platform has confirmed as a specific mechanism.

Recommendation — separate again from mention

Being named is not the same as being endorsed. A model can accurately describe a brand while recommending a competitor, or decline to recommend anyone confidently.

Recommendation -type answers (comparisons, "best of" queries) appear to draw disproportionately on sources that already do comparative evaluation — third-party reviews, "best X" roundups, Reddit threads with community consensus — rather than a brand's own product page, which structurally can't make credible comparative claims about competitors. This tracks with the broader pattern of AI systems favoring earned/community content, but the causal reason (safety-tuning against unsubstantiated claims, training data patterns, or something else) is not disclosed by any provider — genuinely unknown, not just under-documented.

Google's own fan-out description states that sub-queries are processed based on search history and the user's account, when consent is given — meaning recommendation -stage output can be personalized per user in ways that have no equivalent in classic organic rank tracking. A brand recommended to one user may not be recommended to another running the identical prompt, independent of anything on the brand's site. [Julien-gourdon](#)

retrieval and crawler-access failures are the best-documented and most fixable (index architecture, JS rendering, bot-specific robots.txt — all verifiable in your own logs). Source selection and citation are where the trail goes cold — real steps with real consequences, but undisclosed mechanics that the industry is inferring from output patterns, not reading from documentation. Brand mention and recommendation involve a second, mostly separate memory system (training-time parametric knowledge) that current-state SEO has very limited ability to directly influence at all — which is worth saying plainly rather than implying there's a lever for it that simply hasn't been found yet.

[Claude is AI and can make mistakes. Please double-check cited sources.](#)