ESSAY

In ChatGPT, the brand name moves the ranking before the buyer asks.

By Mike Millett with Digilu

October 27, 2026 · 8 minute read

ChatGPT writes the brands it already knows into its own searches and reorders identical products by name alone, so the head start in AI answers is set before the buyer types.

A buyer who asks ChatGPT to compare three products assumes the comparison starts with the question and rests on the pages the engine reads. Mahe Chen and Nick Koudas measured where it actually starts. Their preprint, posted to arXiv on October 10, 2026, covers 158,099 answers and 468,037 citations from ChatGPT, Gemini and Perplexity. The largest factor they found is not anything on a page. It is the search the engine writes for itself before it reads, and in ChatGPT that search already carries the names the model knows.

Brand names alone reorder ChatGPT's rankings

The name-only ranking shift is the Chen and Koudas finding that matters most to any brand an AI engine might compare. In controlled experiments, the authors report that "brand names reorder its rankings of products with identical facts," and the engine is ChatGPT. The design holds the product information fixed and changes only the name. Whatever moved the order, it was not the page, the specifications or the copy.

The claim here is existence, not size: the abstract reports that the reordering happens, not how far it moves. The test covers ChatGPT; Gemini and Perplexity appear in the citation analysis only.

Existence is enough to break one working assumption: that an AI ranking is a neutral summary of the pages the engine read. In ChatGPT, part of the ranking summarizes what the model already associated with each name.

A known brand enters ChatGPT's retrieval twice

Retrieval is the largest factor in the Chen and Koudas data. Whether a page appeared in the engine's own searches explained 32% of the variation in citations; page properties explained 2.8%. The two figures are not a full accounting of why pages get cited. They put two candidate causes on one scale, and on that scale the search outweighs the page more than tenfold.

A skeptic's first objection is that the 32% is partly mechanical. The paper reports that ChatGPT cites only pages its own search retrieved, and a page that was never found cannot be cited. A gate every citation must pass will explain a large share of citations whether or not anyone can push on it. On its own, the 32% shows that retrieval is a precondition, not that it is a lever.

The objection moves the contest rather than ending it. If being found is the gate, what decides the outcome is who appears in the searches the engine writes. Those searches belong to the engine: the authors report that the sub-queries the model rewrites on every answer "rarely coincide with Google's or Bing's results." Google documents the same habit in AI Overviews and AI Mode, which issue multiple related searches of their own, a technique Google calls query fan-out. Ranking well in a conventional index does not settle who is inside them.

The paper also finds that ChatGPT writes brands it already knows into those searches. A known brand therefore enters retrieval twice: once as a page that might match a query, and once as a word inside the query. An unknown brand enters only the first way, competing on searches written without it. That is where the head start sits, and it is why 2.8% is the right comparison. Page properties measure what a team can change on the day, after the gate. The name in the query acts on the gate itself, before any page is read.

The paper does not split the 32% into the share steered by names and the share a well-matched page can win. Generic searches remain contestable on the day. Named searches belong to brands the model already knows.

Brand trust in ChatGPT is stored before retrieval begins

Brand trust in ChatGPT, in the Chen and Koudas data, is knowledge the model holds before its search runs. Two findings require it. The name-in-query path: ChatGPT writes known brands into its searches, so the name comes from the model before any result exists. Unsupported mentions: ChatGPT names some brands with no supporting page at all, so the mention did not come from retrieval at the moment of the answer. The name-only ranking shift adds a third case, where the facts were held fixed and the order moved anyway.

Knowledge held before retrieval cannot be edited at retrieval. A page published today changes what this week's searches can find. It does not change what the model knew when it wrote the search, and that knowledge updates only when the model's own knowledge does. The schedule follows from where the familiarity acts: before the first page is fetched, so no page fetched today can reach it.

The Marketing Helix treats trust and relevance as separate forces, and the engine tests them at different times: relevance at retrieval, against today's pages, and trust at the model's last update, against the record that existed then. A brand can be aligned on one force and out of alignment on the other. Brand gravity is the trust half: the attraction a brand holds by condition rather than by spend, already in place when the question arrives. The funnel model ends at conversion and has no field for the input that matters here, last year's customer writing this year the account a later model reads when someone else asks.

The work on trust starts today and compounds. Every review, comparison and trade account published this quarter is material a later model may read. A smaller brand does not wait to start. It waits to collect.

The case against brand gravity: checklist GEO works where it claims to

The strongest case against brand gravity is that checklist GEO never claimed to change what a model already knows. It claims to change retrieval and selection, and the Chen and Koudas paper supports parts of both. Retrieval is the largest measured factor. Among pages on a single website, content that matches the questions people ask is cited more. Between websites, familiar on-page features such as FAQ blocks, structured data and page speed do make a difference, and a team can build every one of them. Add the fairness point, that "earned over time" sounds like "wait" to a newer brand, and the thesis starts to look like an incumbent's argument.

Most of that should be conceded. Generic retrieval is contestable. Site quality is a real lever. The answer lies in where each lever stops. The question-matching effect was measured among a site's own pages; the paper does not show it bringing a site into an answer the site was otherwise absent from. Site-level quality is a property of a whole domain, built across months of pages, which places it on the update cycle of reputation, not of a single edit. And because ChatGPT "prefers websites it knows," the between-site effect and the familiarity effect overlap in ways the paper's summary does not separate.

None of those levers reaches an order that moved while the facts stayed fixed. On fairness, the authors report that their findings replicate in a market of small B2B vendors, so the mechanics are not reserved for household names. A newer brand can build on them from today. What it cannot

do is collect on them the same day.

Citation share needs three checks that can come out wrong

Citation share is the proportion of answers to a fixed set of prompts that cite or name a brand. Page properties explained 2.8% of citation variation in the Chen and Koudas data, so a report that credits a rise in citation share to last month's schema update is probably naming the wrong cause. Three checks, run by the team in its own category, separate the causes.

The name test. Write a comparison prompt that states the same facts for the brand's product and two competitors. Run it with real names, then with the names replaced by neutral labels, then with the names swapped. If the brand's position holds across all three, the facts are doing the work. If the position follows the name, brand gravity is, and the size of the move is the category's own measure of an effect the paper's abstract does not quantify.

The unsupported mention count. For each answer that names the brand, check whether any cited page supports the claim. A rising count of unsupported mentions indicates that the model's familiarity is growing apart from retrieval. A flat count while citations rise places the gain in retrieval, where page work could be the cause.

The reading list test. Run the comparison prompts a buyer in the category would run, log every third-party domain the engine cites, and rank the domains by how often they recur. That ranking is the engine's own reading list for the category. Record the brand's presence on the top domains every quarter and compare it with citation share at a fixed lag, for example presence in the second quarter against citation share in the fourth. Stored familiarity is gathered before a model answers, so a review published in October cannot explain an October answer through that path. If fourth quarter citation share rises while second quarter presence was flat, reputation did not produce the rise, and page work in retrieval deserves the credit. If second quarter presence rose and fourth quarter answers name the brand without a supporting page, the gain is brand gravity, and the top of the reading list is where to build next.

> A known brand therefore enters retrieval twice: once as a page that might match a query, and once as a word inside the query.

A buyer arrives at an answer engine already in motion, and the brand arrives with whatever standing the model gave its name before the question. In ChatGPT, that standing reorders products whose facts are identical and writes known names into the searches that decide which

pages are seen. Day-of work decides which page speaks for a brand and competes for the searches no name has claimed. The search with a brand's name in it was drafted from the record the model last read, and customers, reviewers and trade press wrote that record quarters before the buyer typed.

Further reading

1. Mahe Chen and Nick Koudas, "What Drives AI Citations? Content, Retrieval and Model Knowledge in Generative Engine Optimization (GEO)," arXiv preprint: the primary study of 158,099 answers and 468,037 citations on which this essay rests arxiv.org/abs/2610.15033
2. Google Search Central, AI features and your website: Google's own documentation on query fan-out and on the absence of special optimization requirements for AI Overviews and AI Mode developers.google.com/search/docs/appearance/ai-features
3. The Marketing Helix: the model and its three forces of trust, relevance and timing marketinghelix.com

marketing helix

marketinghelix.com

A customer behavior model describing how a customer in motion gives a message attention, and how trust, relevance, and timing then shape whether it draws closer, drifts away, or returns.