



CiteHawk

For smart non-engineers

— THE MECHANISM, DEMYSTIFIED

How AI actually chooses what to recommend.

Most GEO advice is written by people who have never looked inside the machine they are optimizing for. This paper explains the mechanism in plain language: how an AI model composes an answer, how retrieval actually treats your pages, why the answer and the citations come from partly different processes, and what all of it means for how you write a page.

An AI model has two memories

When a buyer asks "what's the best CRM for a small agency," the AI model draws on two very different stores of knowledge, and the distinction explains most of what confuses people about this field.

The trained memory

What the model absorbed about your category during training: which brands appeared across its corpus, in what contexts, described how. It is vast, compressed and slow-moving. It updates only when the provider ships a new model, which is why long-established brands often over-perform in AI answers relative to their current market position, and why a reputation, once encoded, has momentum in both directions.

The live lookup

Most AI models now search the web mid-answer for anything current or specific: they issue searches, fetch a handful of pages, and read them before responding. This store moves in days, not quarters. It is the lever you can actually pull this month, and it behaves by rules closer to search than to language modeling, which is why classic SEO strength still matters here: pages that rank get retrieved, and pages that get retrieved get read.

"Training decides who the model already believes in. Retrieval decides who gets a hearing today. You are optimizing two systems with two clocks."

Retrieval does not read pages. It reads pieces.

Here is the mental shift that changes how you write: retrieval systems generally break content into **chunks**, score each chunk for relevance to the question, and hand the winning chunks to the model. The model often never sees your page. It sees fragments, competing against fragments from every other source, ranked by how directly each fragment addresses the question.

The practical implications are concrete:

Every important section must survive amputation. A chunk that says "as mentioned above, this makes it ideal" carries nothing. A chunk that says what your product is, for whom, at what price, wins on its own.

The heading is part of the chunk. A question-shaped heading followed by a direct two-sentence answer is the most retrievable structure there is.

One idea per section. Chunks that mix three topics rank mediocre for all three. Focused sections rank high for one.

Tables and lists survive extraction better than narrative. Structure is not decoration here; it is transport packaging.

None of this is a trick. It is writing so that any fragment of the page, read alone by a machine in a hurry, still says something true and complete about you.

Why the answer and the citations disagree

The most common confusion in this field: a brand gets cited but not recommended, or recommended but not cited. That is not a bug in your measurement. The **recommendation** is composed by the model from everything it has (trained memory plus every retrieved fragment), while the **citation list** reflects which retrieved sources the system chose to attribute. Two partly different processes, two partly different winners.

This is why review sites and directories dominate citation lists while the brands they discuss dominate the recommendations, and why chasing citations of your own domain can be the wrong goal entirely. The durable play is to be well represented *inside* the sources that get cited: the directories, the comparisons, the community threads. When those fragments enter the context, they speak for you.

It is also why we removed citation counting from our own ranking formula after our July 2026 audit: citations measure "is this a source," not "is this recommended," and conflating the two rewarded the wrong pages. The full reasoning is public in our methodology.

What this means for your next quarter

- 1 • Fix your facts at their sources.** The live lookup reads directories and profiles before it reads your homepage. Stale facts there become stale answers everywhere.
- 2 • Write answer-shaped pages.** Question headings, direct first sentences, self-contained sections, structured comparisons.
- 3 • Enter the citation graph.** Presence inside the sources answers cite beats links to your own domain for recommendation share.
- 4 • Measure per AI model, weekly.** Each AI model weighs the two memories differently, so the same work lands differently on each scoreboard. Watch all 8.

See how the machine currently reads you.

CiteHawk's free report asks all 8 AI models your category's real buying questions, stores every answer, and shows exactly which sources those answers trust. The mechanism in this paper, measured on your brand.

Run your free report → citehawk.com/signup

