

OCTOBER 2026

Decoded: ChatGPT's Ranking Formula Reciprocal Rank Fusion and .Best Domains

How AI search may merge several result lists into one before it cites a source, why consistency across many searches beats a single first place, and where a .Best name fits a ranking formula built on the word "best".

$1/(60+r)$

the RRF score of one result list
at rank r

Cormack et al., SIGIR 2009

+14.8%

all that ranking #1 adds over
#10 in one list, at $k = 60$

This paper, computed

24.3%

of ChatGPT advice fan-outs
add "best"

Peec AI, ~5M fan-outs

15%

of pages ChatGPT retrieves end
up cited

AirOps, 548,534 pages

CONTENTS

What this paper covers

01 Executive summary

The argument in one page, and the evidence behind each claim

02 One question, many searches

Query fan-out: what OpenAI and Google say, and what 2026 data shows

03 Reciprocal Rank Fusion, explained

The 2009 formula, the constant $k = 60$, and how search engines ship it

04 The arithmetic of a citation

Why breadth beats a single first place

05 Decoded: what is claimed about ChatGPT

The DevTools finding, its parameters, and what is not confirmed

06 From retrieval to citation

Fusion, reranking, and why most retrieved pages are never cited

07 "Best" across the fan-out

The words ChatGPT adds, and how they shift

08 Where a .Best name meets RRF

Five mechanisms, and what the name does not do

09 The RRF content playbook

Coverage, page architecture, freshness, entity

10 Measure your RRF footprint

A method you can run with a spreadsheet

11 Use cases

hotels.best, software.best, restaurants.best, worked through

12 Limits and open questions

What would prove this paper wrong

13 A 30-day plan, conclusion and sources

Four weeks to a measured baseline

ABOUT THIS PAPER

How to read the evidence

This paper examines one specific claim about AI search: that ChatGPT merges the results of its hidden searches with a classic formula called Reciprocal Rank Fusion (RRF), and that this formula decides which pages are in the running to be cited. It explains the formula in detail, sets out how strong the evidence for the claim is, and shows where a .Best domain fits.

The subject needs care. RRF itself is thoroughly documented: it was published in 2009 and ships today in Elasticsearch, OpenSearch and Azure AI Search. Its use inside ChatGPT is not documented by OpenAI. It rests on one independent researcher’s reading of ChatGPT’s web client in July 2025, repeated and extended by others since. We present that claim at full strength and with its limits, because a strategy built on it should know both.

The title deliberately echoes the article that started the discussion, Metehan Yesilyurt’s “Is RRF the Secret to Dominating AI Citations? I Decoded ChatGPT’s Ranking Formula” (July 2025). This paper builds on his finding and on the 2026 research that followed; it does not claim the discovery.

THE FIVE EVIDENCE TAGS

Every claim carries one tag, so you always know how firm the ground is.

MEASURED · .BEST	Measured directly on .Best domains in a controlled test.
MEASURED · GENERAL	Measured by an independent third party on AI search in general, not on .Best.
REPORTED	Observed by independent practitioners in ChatGPT’s client or traffic. Not confirmed by OpenAI.
MECHANISM	Our reasoning from the formula and the measured behaviour. Plausible; not yet measured.
OPEN QUESTION	Not measured by anyone yet. Section 12 proposes how.

Disclosure: this paper is published by the registry operator of .Best. Third-party figures are reported as published by their sources, listed at the end; several were read through secondary coverage because the original pages were not reachable when this paper was prepared, and are marked as such in the sources. Readers should consult the originals before quoting a figure. AI search changes quickly; every figure is dated.

01 EXECUTIVE SUMMARY

ChatGPT does not rank one list. It may fuse several.

When ChatGPT searches the web, it rewrites the buyer's question into one or more searches of its own. Each search returns a ranked list. Something has to turn those lists into the handful of sources the answer cites. The leading public hypothesis is Reciprocal Rank Fusion, a formula that rewards pages for appearing in many lists more than for topping any one of them.

$$\Sigma 1/(k+r)$$

RRF adds one small score per list a page appears in

Cormack et al., 2009

3.05×

score of #20 in four lists vs #1 in one list, $k = 60$

Computed, section 04

24.3%

of ChatGPT advice fan-outs add "best"

Peec AI, April 2026

~2

searches per prompt on average, rising with harder questions

Peec AI, AirOps, Ahrefs

Six findings

- 01 ChatGPT turns one question into several searches.** OpenAI says ChatGPT "typically rewrites" a query into "one or more targeted queries". Independent 2026 measurements put the average around two per prompt, and far more for complex or reasoning-mode questions. MEASURED · GENERAL
- 02 Several lists need a fusion rule, and RRF is the standard one.** It scores each page by summing $1/(k + \text{rank})$ across the lists it appears in, with $k = 60$ the default in Elasticsearch, OpenSearch and Azure AI Search. MEASURED · GENERAL
- 03 RRF in ChatGPT is reported, not confirmed.** RRF settings were observed in ChatGPT's web client in July 2025; one practitioner reproduced the observation; OpenAI has never described its ranking in these terms, and $k = 60$ was assumed, not seen. REPORTED
- 04 Under RRF, breadth beats peak.** At $k = 60$, a first place is worth only 14.8% more than a tenth place in the same list, while each additional list a page appears in adds a full new term. Ranking #20 in four searches scores three times a single #1. MECHANISM
- 05 The searches speak the language of "best".** "Best" is the word ChatGPT adds most often to its fan-out searches, in 24.3% of searches for advice questions, alongside "top", "comparison" and "reviews". MEASURED · GENERAL
- 06 A .Best name repeats that word in every list.** On a .Best site, the domain, the brand mention and many titles carry "best", so every "best" sub-query finds the same entity again. RRF rewards exactly that repetition. The name cannot make a weak page relevant; it helps a relevant one be the same, recognisable answer everywhere. MECHANISM

WHAT WE ARE NOT CLAIMING

We are not claiming that OpenAI has confirmed RRF, that $k = 60$ is ChatGPT's setting, or that ChatGPT gives any weight to a top-level domain as such. No published study measures how often ChatGPT cites .Best sites compared with others. The argument of this paper is narrower and testable: if fusion works as reported, the pages that win are the pages that appear across many related searches, and a name that states the searches' own key word is a structural fit for that. Section 12 sets out how to test it.

The evidence at a glance

FINDING	SOURCE	STATUS
ChatGPT rewrites a search into “one or more targeted queries”	OpenAI Help Center	MEASURED · GENERAL
AI Overviews and AI Mode may use “query fan-out” across subtopics	Google Search Central	MEASURED · GENERAL
RRF beats individual systems and Condorcet fusion; k = 60 near-optimal	Cormack, Clarke & Büttcher, SIGIR 2009	MEASURED · GENERAL
RRF ships with rank constant 60 by default	Elasticsearch, OpenSearch, Azure AI Search docs	MEASURED · GENERAL
RRF settings observed in ChatGPT’s web client	Yesilyurt, July 2025; one replication	REPORTED
ChatGPT adds “best” to 24.3% of advice fan-outs	Peec AI, ~5M fan-outs, April 2026	MEASURED · GENERAL
About 2 fan-out searches per prompt on average	Peec AI 2026; AirOps 2026; Ahrefs 2026	MEASURED · GENERAL
Only ~15% of retrieved pages are cited	AirOps, March 2026	MEASURED · GENERAL
Google + OpenAI’s API recover ~28% of cited URLs, ~50% of hostnames	arXiv 2610.15382, October 2026	MEASURED · GENERAL
Natural-language URL slugs cited more often (89.8% vs 81.1%)	Ahrefs, 1.4M prompts	MEASURED · GENERAL
Identical ads earned 20% more clicks and paid 8.1% less on .Best	GlobeRunner, small sample	MEASURED · .BEST
A .Best name matches “best” sub-queries in every list	This paper, section 08	MECHANISM
.Best sites are cited by ChatGPT more often than equivalent sites	Not yet measured	OPEN QUESTION

What to do

NAME

Pair your category with .best: hotels.best, crm.best, movers-lyon.best. The name should state what the site is the best answer to.

COVER

Map the sub-queries a buyer’s question fans out into, and build one strong, current page for each, so you appear in many lists rather than one.

MEASURE

Rank-track the sub-queries, compute your own fused score with the formula in section 10, and judge progress against your baseline.

02

QUERY FAN-OUT

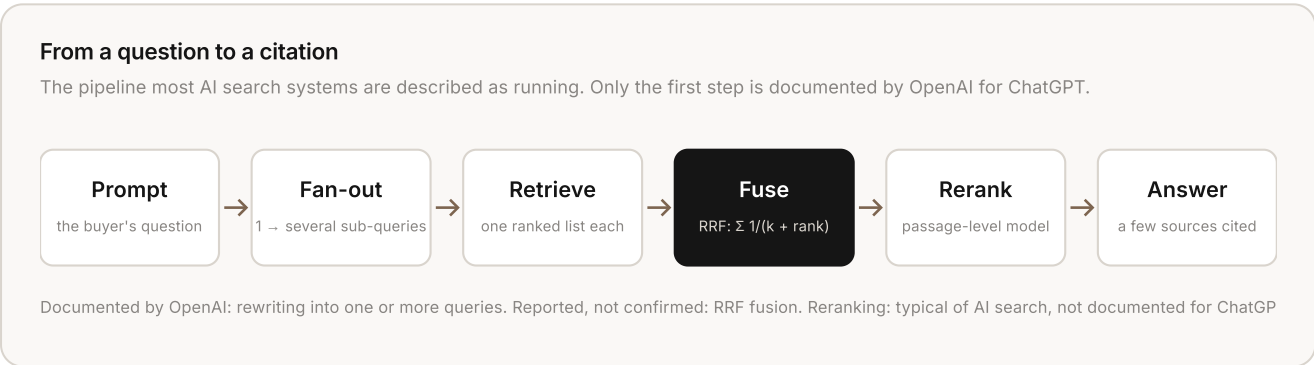
One question, many searches

Classic search answered the words you typed. AI search answers what it thinks you meant, and to find out it runs searches of its own. That step, called query fan-out, is the reason a fusion formula is needed at all.

What the platforms say

OpenAI’s help documentation describes the mechanism plainly: when ChatGPT search uses a search provider, it “typically rewrites your query into one or more targeted queries” and sends those to the providers, and it may use saved memories when rewriting. OpenAI also states that ChatGPT search draws on third-party search providers as well as content from partners. It does not describe how the results of several queries are combined. MEASURED · GENERAL

Google is explicit about its own version. Its Search Central documentation says AI Overviews and AI Mode “may use a ‘query fan-out’ technique — issuing multiple related searches across subtopics and data sources — to develop a response.” MEASURED · GENERAL



What 2026 measurements show

Several independent teams have logged ChatGPT’s fan-out queries at scale. Their figures differ because their bases differ (all prompts or only prompts that triggered a search; free or paid tiers; standard or reasoning models), so they should be read as a range, not one number.

STUDY	WHAT WAS MEASURED	RESULT
Peec AI, April 2026	~5M fan-outs from ChatGPT, Perplexity and Grok	2.1 fan-outs per prompt for ChatGPT (Perplexity 1.4, Grok 6.8)
Peec AI, Oct 2025 – Jan 2026	20M+ fan-outs	2.3–2.8 per prompt; fan-outs doubled in length, from ~6 to ~12 words
AirOps, March 2026	15,000 prompts → 43,233 queries	~2.9 queries per prompt
Ahrefs, October 2026	15M fan-outs from free, logged-out responses	14.4% of responses searched; ~1.6 fan-outs per searching response
Practitioner traffic studies, 2026	Reasoning-model sessions	Reported at 15 to 40 sub-queries for the thinking model; small samples

WHY FAN-OUT CHANGES THE GAME

With one query, there is one ranking, and the top result wins. With several, a page can be first for one sub-query and absent from the rest, or solidly placed in all of them. Which of the two profiles wins depends entirely on how the lists are merged. That is the question RRF answers.

The fan-outs are not the prompt. A 2026 preprint that replayed 1,600 runs per model found no instant-model run that searched with the user’s exact wording. Fan-outs add words the user never typed, such as a year, “best”, “reviews” or “official” (section 07). MEASURED · GENERAL

How many lists can a page collect?

The number of fan-outs searches sets a ceiling on any page’s fused score: with n searches, the most a page can earn is n first places, n/61 at k = 60. The ceiling explains why fan-out depth matters so much under fusion, and why the working threshold reported for ChatGPT (section 05) cannot be reached from a single search. MECHANISM

FAN-OUT SEARCHES FOR THE QUESTION	MAXIMUM FUSED SCORE, $K = \frac{}{60}$	CAN THE REPORTED 0.020 THRESHOLD BE REACHED?
1 (the prompt only)	0.0164	No: one first place is not enough
2 (Peec AI’s April 2026 average is 2.1)	0.0328	Yes, by appearing in both lists within the top 40
3	0.0492	Yes, by appearing in all three within the top 90
5	0.0820	Yes, with room to spare: four mid-page placements suffice
8 (reasoning-mode questions)	0.1311	Yes; breadth across the cluster dominates

Deeper fan-outs widen the gap between a page that covers the cluster and a page that wins one query. As questions get harder and models search more, breadth becomes worth more. MECHANISM

03 THE FORMULA

Reciprocal Rank Fusion, explained

Reciprocal Rank Fusion was introduced in 2009 by Gordon Cormack, Charles Clarke and Stefan Büttcher of the University of Waterloo, in a two-page SIGIR paper titled "Reciprocal Rank Fusion outperforms Condorcet and Individual Rank Learning Methods". It solves a simple problem: how to combine several ranked lists into one, when the systems that produced them score documents on incomparable scales.

$$\text{RRFscore}(d) = \sum_{r \in R} 1 / (k + r(d))$$

R: the set of result lists · r(d): the position of document d in list r (1 = first) · k: a constant, 60 by default · a list that does not contain d adds nothing

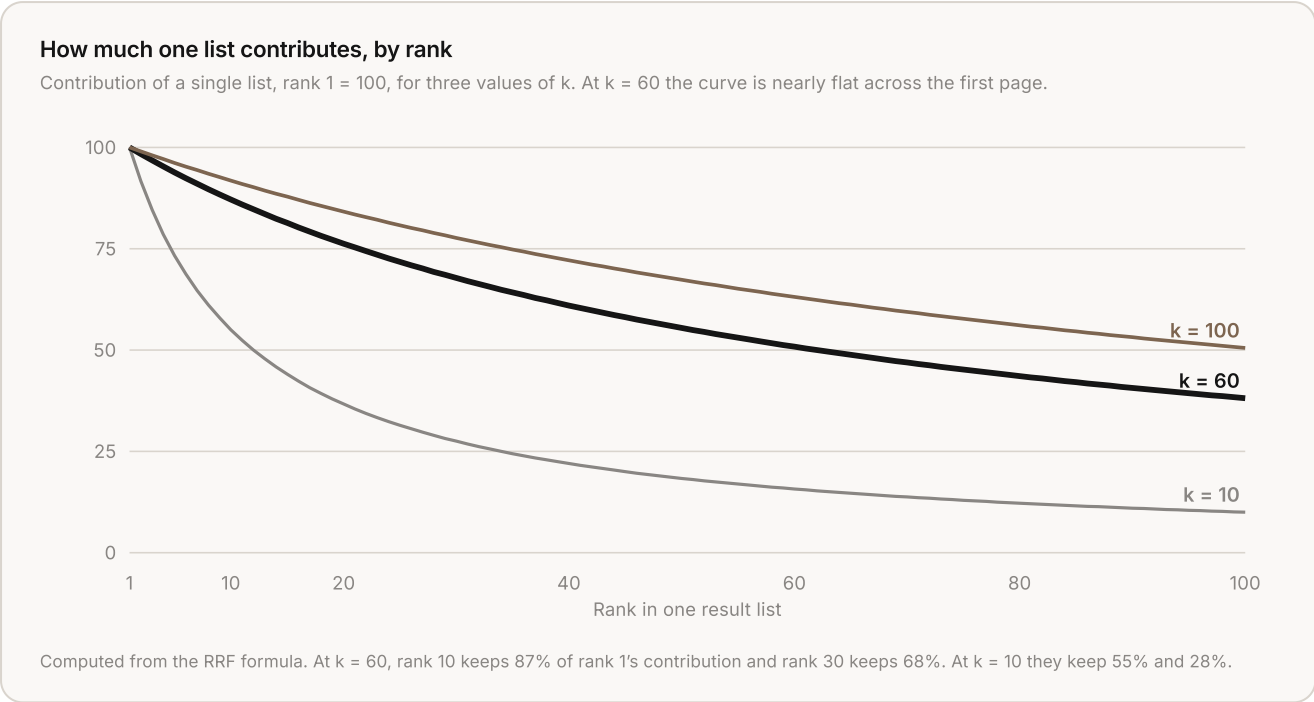
Three properties that matter

- 01 **It uses ranks, not scores.** A keyword engine and a vector engine score relevance in units that cannot be compared. RRF ignores the scores and keeps only the order, so any retrieval system can be fused with any other without normalisation. Elastic's documentation sums it up: RRF "requires no tuning, and the different relevance indicators do not have to be related to each other".
- 02 **The constant k flattens the top.** Without k, first place (1/1) would be worth ten times tenth place (1/10). With k = 60, first place is worth 1/61 and tenth place 1/70: a difference of 14.8%. The authors describe k as mitigating "the impact of high rankings by outlier systems": one list's enthusiasm cannot outvote the others.
- 03 **Every list adds a term.** A document that appears in five lists collects five reciprocal scores; a document that appears in one list collects one. Agreement across lists, not dominance in a single list, is what accumulates.

Why 60

The value was not derived from theory. The authors fixed k = 60 during pilot experiments, fusing the results of 30 configurations of their search engine over four TREC collections, and reported that 60 was near-optimal but that the choice was not critical. It was then held constant through validation, where RRF beat every individual system and the Condorcet fusion method. Sixteen years later, 60 remains the default almost everywhere.

MEASURED · GENERAL



How search engines ship it

RRF is not an academic curiosity. It is the default fusion method of the three most widely used hybrid search platforms, which is one reason it is a plausible choice for any AI search pipeline that merges several queries or retrieval methods.

PLATFORM	IMPLEMENTATION	RANK CONSTANT	WEIGHTS
Elasticsearch	rrf retriever; result window per list set by rank_window_size	rank_constant, ≥ 1, default 60	Equal by default; per-retriever weight (default 1.0) since 9.2
OpenSearch	Score ranker processor, since 2.19	rank_constant, 1–10,000, default 60	Optional weights that must sum to 1.0
Azure AI Search	Applied automatically to hybrid and multi-vector queries	k = 60, described as performing best	Per vector query

Weighted RRF generalises the formula: each list’s term is multiplied by a weight w , so that $\text{score}(d) = \sum w_r / (k + r(d))$. Elastic’s documentation gives exactly this form. A system that trusts one kind of retrieval more than another can say so without abandoning rank-based fusion.

```
# Reciprocal Rank Fusion, as documented by Elastic (pseudocode)
score = 0.0
for q in queries:
    if d in result(q):
        score += 1.0 / ( k + rank( result(q), d ) )
return score
```

What RRF is not

RRF decides the order of candidates; it does not decide relevance on its own. Pages first have to be retrieved by at least one list, and in modern pipelines the fused list is often re-scored by a reranking model before anything is shown. A 2026 industry paper from Dell Technologies found that fusion’s gains in recall can be “largely neutralized after re-ranking and truncation”. Fusion opens the door; it does not guarantee the seat. Section 06 returns to this.

04

THE ARITHMETIC

The arithmetic of a citation

The consequences of the formula are counter-intuitive enough to be worth working through. All figures in this section are computed directly from RRF with $k = 60$ and equal weights. They describe the formula, not any measured ChatGPT behaviour.

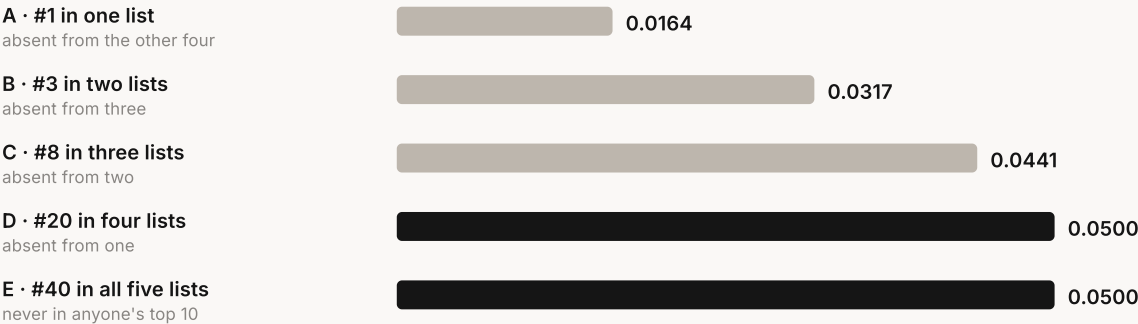
MECHANISM

Five pages, five sub-queries

Imagine a buyer’s question that ChatGPT fans out into five searches. Five pages compete. Each has a different profile across the five result lists.

Fused score of five ranking profiles

RRF, $k = 60$, five sub-queries, equal weights. A page that is absent from a list scores nothing for it.



A: $1/61 = 0.0164$ · B: $2/63 = 0.0317$ · C: $3/68 = 0.0441$ · D: $4/80 = 0.0500$ · E: $5/100 = 0.0500$. The page that never reaches anyone's top 10 (E) ties with D and triples the page that wins one list outright (A).

Profile A is the classic SEO trophy: first place for one query. Under RRF it finishes last. Profile E would never be noticed in a rank tracker that only watches the top 10, yet it collects five small scores that add up to three times A's single large one.

The rule of thumb

COMPARISON, $k = 60$	RATIO	WHAT IT MEANS
#1 vs #10 in the same list	1.15×	Climbing nine places in one list adds 15%
#1 vs #30 in the same list	1.48×	Even page three keeps two-thirds of the value
#61 in two lists vs #1 in one	1.01×	Two appearances anywhere in the top 61 beat one first place
#20 in four lists vs #1 in one	3.05×	Breadth multiplies; position barely adds
#5 in ten lists vs #1 in one	9.38×	The example used in the July 2025 finding: 0.154 vs 0.0164

THE STRATEGIC CONSEQUENCE

Under RRF, the unit of competition is not a keyword. It is a cluster of related searches. A site that is solidly present across a cluster outranks a site that wins one query and is missing from the rest. Every additional list you appear in adds a full new term to your score; every place you climb within a list adds only a fraction of one.

How sensitive is this to k?

The flatness comes from k. If a system used a small constant, first places would matter more. The 2026 SemEval shared task offers an example: one strong team tuned k down to 10 for its hybrid retrieval. The table shows how the #1-versus-#10 premium moves with k.

RANK CONSTANT	#1 VS #10, SAME LIST	#20 IN FOUR LISTS VS #1 IN ONE
k = 10	1.82×	1.47×
k = 60 (default)	1.15×	3.05×
k = 100	1.09×	3.37×

Even with k = 10, breadth still wins: four mid-page appearances beat one first place. The direction of the advice does not depend on the exact constant, only its strength does. That matters, because ChatGPT’s constant has not been observed.

05

THE FINDING

Decoded: what is claimed about ChatGPT

Everything specific about RRF inside ChatGPT traces back to one observation, made in July 2025 and discussed widely since. This section sets out exactly what was observed, by whom, and what remains inference.

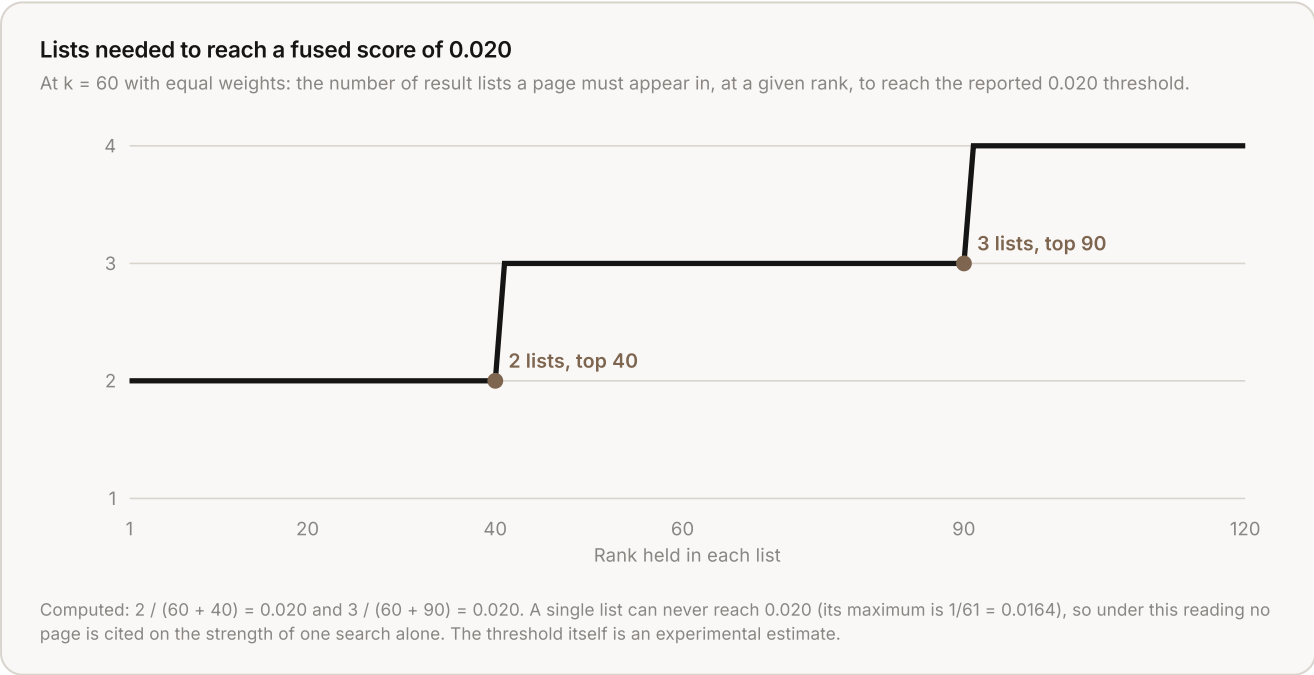
The July 2025 observation

On 22 July 2025, SEO researcher Metehan Yesilyurt published a thread and an article reporting that, while inspecting ChatGPT’s web client in Chrome’s developer tools, he had found settings indicating that ChatGPT uses Reciprocal Rank Fusion to combine search results. Secondary write-ups of his finding list configuration values such as `rrf_alpha: 1` and `rrf_input_threshold: 0`. REPORTED

Two points in his own account matter for anyone building on it. First, the constant *k* was not visible; he used 60 because no specific setting appeared in the code, and 60 is the standard default. Second, he wrote that he was not claiming to have found the exact formula, which could change. REPORTED

The follow-up: a top-*n* playbook

In September 2025 the same author published “The RRF Top-*n* Playbook”, which he labelled highly experimental. Using a browser extension that exposes ChatGPT’s search path, he reported that ChatGPT retrieved between 38 and 65 results, and proposed a working threshold: a fused score of about 0.020. The arithmetic of that threshold translates into a practical rule, shown below. REPORTED



How strong is the evidence?

CLAIM	EVIDENCE	STATUS
ChatGPT rewrites a prompt into one or more searches	OpenAI documentation; large independent logs	MEASURED · GENERAL
The client carries RRF settings	One researcher’s DevTools inspection; one independent practitioner found the same fields	REPORTED

CLAIM	EVIDENCE	STATUS
Those settings decide citations	Not demonstrated; the replicating practitioner said he could not determine it	OPEN QUESTION
k = 60	Assumed from the default; not observed	OPEN QUESTION
Threshold ≈ 0.020; 38–65 results retrieved	Experimental estimates from one author	REPORTED
OpenAI uses its own index alongside external providers	Vendor observation of server fields, mid-2026	REPORTED

OpenAI’s own public statement on ranking is brief: its publisher FAQ says ChatGPT ranks results using “multiple factors intended to help users find relevant, reliable information”, and that placement is not guaranteed. Nothing in it confirms or rules out RRF. A practitioner who studied ChatGPT’s network traffic in 2026 reported seeing no ranking scores in the data sent to the browser, which is consistent with fusion happening server-side, out of view. MEASURED · GENERAL

OUR READING

The RRF claim is credible but unconfirmed. It is credible because fan-out demonstrably produces several lists, because RRF is the industry-standard way to merge them, and because the client-side settings point that way. It is unconfirmed because the evidence is one observation, the key constant is assumed, and OpenAI can change its pipeline at any time.

The useful question is therefore not “is it exactly RRF?” but “what kind of page wins under any rank-fusion rule?”. The answer is robust across fusion methods: pages that appear across many related searches. That is the ground this paper builds on.

Why the client would carry fusion settings

A plausible explanation, offered as inference, is that ChatGPT’s search tool is configured per request: the client or the orchestration layer passes parameters for how many results to fetch and how to merge them, and the merging itself runs on OpenAI’s servers. That would explain why parameter names are visible while scores are not. MECHANISM

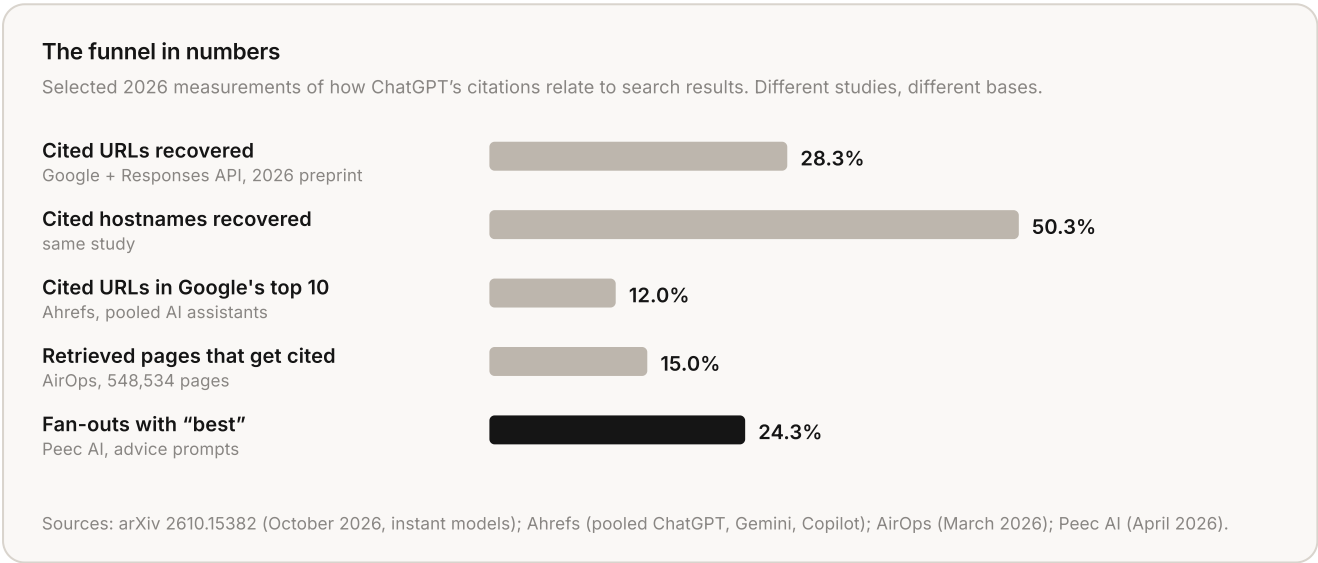
06 THE PIPELINE

From retrieval to citation

Fusion is one stage in a funnel. Pages are retrieved, fused, often reranked, and only then do a few become citations. The 2026 data shows how steep that funnel is.

Most retrieved pages are never cited

AirOps analysed 15,000 prompts that ChatGPT expanded into 43,233 queries and 548,534 retrieved pages. Only 15% of the retrieved pages were cited. A companion analysis reported that a page’s position in ChatGPT’s own result list was the strongest predictor of citation, falling from 58% at the first position to 14% at the tenth. MEASURED · GENERAL



Your Google rank is not your ChatGPT rank

Several studies have compared ChatGPT’s citations with classic search results. The overlap is partial and unstable:

- A preprint posted in October 2026 replayed 80 questions, 20 times each, on two ChatGPT instant models. Google results plus OpenAI’s own search API recovered only 28.3% and 26.2% of the cited URLs, but about half of the cited hostnames. MEASURED · GENERAL
- Ahrefs found that about 12% of URLs cited by ChatGPT, Gemini and Copilot also rank in Google’s top 10 for the original prompt. MEASURED · GENERAL
- AirOps found that pages ranked first on Google for a fan-out query were cited 43.2% of the time, about 3.5 times more than pages beyond the top 20, and that 32.9% of cited pages appeared only in the results of a fan-out query, not of the original prompt. MEASURED · GENERAL

Read together, these figures fit a fusion model well. Citations correlate with ranking (position helps), but the relevant rankings are those of the fan-out queries, not of the prompt, and the same site can be cited through a different page than the one that ranks. Hostname overlap being twice URL overlap is exactly what you would expect if sites, more than individual pages, accumulate presence across lists. MECHANISM

Rerankers: the second gate

In most modern AI search systems, the fused list is passed to a reranking model that reads passages and re-scores them for the question. Peec AI's 2026 explainer puts the order of operations bluntly: if a page was never retrieved, rewriting a paragraph for a reranker cannot fix it. Fusion and reranking reward different things. Fusion rewards presence across lists; the reranker rewards a passage that answers the question cleanly. A page needs both. MEASURED · GENERAL

GATE 1 · FUSION

Be in many lists. Rewards breadth of presence across the sub-queries of a question. Driven by coverage, consistent naming and topical depth.

GATE 2 · RERANKING

Be the clearest passage. Rewards a direct, quotable answer near the top of the page, with the entity and the claim stated together.

What the funnel means for strategy

- 01 **Visibility is earned per cluster.** A buyer's question produces a small family of searches. The pages that appear in most of that family form the candidate pool. Strategy should be planned at that level.
- 02 **Sites, not just pages, accumulate.** When different pages of one site appear in different lists, a system that groups by site or brand sees the same source again and again. A consistent, recognisable name helps that grouping be obvious. MECHANISM
- 03 **Position still matters, less than expected.** Moving from tenth to first in a list is worth having, but it is worth less than appearing in one more list.
- 04 **The last mile is the passage.** Once in the pool, the page that states the answer plainly is the page that gets quoted.

A NOTE ON VOLATILITY

Two ChatGPT runs of the same question often cite different pages. A 2026 preprint comparing engines found exact-URL overlap below 1% across engines for the same prompt. Single checks are anecdotes; only repeated measurement over many runs says anything. Section 10 builds that into the method.

07

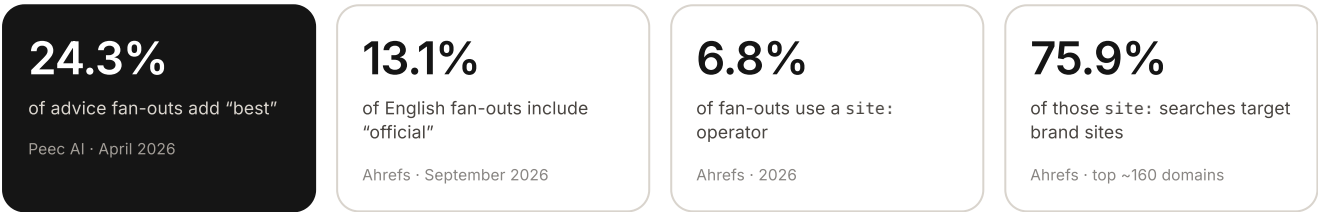
THE VOCABULARY

“Best” across the fan-out

Under fusion, what matters is which lists a page appears in. Which lists exist depends on which words ChatGPT adds when it rewrites a question. The most frequent of those words is “best”.

The words ChatGPT adds

Peec AI analysed about 5 million fan-out queries collected between 1 and 21 April 2026. The word ChatGPT most often added to its hidden searches was “best”: it appeared in 24.3% of fan-outs generated for advice-style prompts, where the user had not typed it. Other frequent additions were “top”, “comparison”, “reviews”, “tools”, “software” and “features”. MEASURED · GENERAL



A vocabulary that moves

The vocabulary is not fixed. Peec AI data discussed by Lily Ray indicates that after a mid-2026 model update, “best”, “top”, “vs” and “reviews” became less frequent while site: searches and “official” surged, and the listicle share of citations fell from 15.77% to 7.80%. Ahrefs found “official” and the current year rising through 2026. MEASURED · GENERAL

TWO SIGNALS, ONE DIRECTION

The 2026 shift is sometimes read as “best is fading”. A closer reading is that ChatGPT is looking for two things at once: **the best options** for a need, and **the official source** for each option. The first produces “best” searches; the second produces brand and site: searches. A site whose own name is the category plus “best” is addressed by both kinds of search at once. That is the subject of the next section.

What a “best” fan-out looks like

THE USER ASKS	CHATGPT MAY SEARCH (PATTERNS REPORTED BY PEEC AI; RUNS VARY)
Which Samsung Galaxy should I get?	best Samsung Galaxy phones comparison 2026
Where can I find candidates outside LinkedIn?	best tech recruiting platforms alternatives to LinkedIn
Who are the leading SEO solution providers?	best SEO agencies
A romantic hotel in Paris?	best romantic hotels Paris · Paris boutique hotel reviews

Why this matters under fusion

If one in four advice sub-queries contains “best”, then for a typical cluster of three to five searches, there is a good chance that at least one list is a “best” list, and often two. Under RRF, every list a page appears in adds a term. Pages that naturally match the “best” phrasing gain a list; pages that do not, lose one. MECHANISM

08

THE .BEST ANGLE

Where a .Best name meets RRF

A domain name cannot change a ranking formula. It can change which lists a site appears in, how consistently it is recognised across them, and whether people click when it is cited. Under fusion, those are the levers that count.

Five mechanisms

- 01

The word in every “best” list. A page on hotels.best titled “Best boutique hotels in Paris” states “best” in its address, its title and its brand. Keyword retrieval for a “best” fan-out finds an exact lexical match; semantic retrieval finds a page whose identity is the query’s intent. Each “best” sub-query is an additional list the site can appear in.

MECHANISM
- 02

One entity across all lists. When different pages of a site rank for different sub-queries, a name that states the category makes them read as one source: hotels.best for “best hotels”, hotels.best for “hotel reviews Paris”, hotels.best for the brand search. Hostname-level overlap in 2026 data (about half of cited hostnames recovered, against about a quarter of URLs) suggests sites, not just pages, are what persists across searches.

MECHANISM
- 03

The brand search is a category search. When ChatGPT adds “official” or runs a site: search to verify an option, the brand name it searches for is, on a .Best site, also the category and the word “best”. The verification list and the recommendation list point to the same place.

MECHANISM
- 04

Readable addresses get cited. Ahrefs’ study of 1.4 million prompts found natural-language URL slugs cited at an 89.78% rate against 81.11% for other slugs. A .Best domain is the most natural-language part of a URL: crm.best reads as words.

MEASURED · GENERAL
- 05

The click after the citation. Being cited is not the end: people still choose which source to open. In the only controlled test of .Best, identical Google ads earned 20% more clicks and paid 8.1% less per click on .Best than on .com, on a small sample. Clicks feed back into how classic search engines judge a page, and classic rankings feed the fan-out lists.

MEASURED · .BEST

WITHOUT A DESCRIPTIVE NAME

acme-hospitality.com ranks well for “Paris hotel reviews” and nowhere else. One list, one term: 1/(60+r).

WITH A .BEST NAME

hotels.best appears for “best romantic hotels Paris”, “Paris hotel reviews” and “hotels.best official”: three lists, three terms, one recognisable source.

Illustration of the mechanism, not a measurement. The second profile still requires pages that deserve to rank for each search.

A worked comparison

Take a question that fans out into four searches, and two sites with equally good content. The only difference is that one site’s name and titles carry the word “best”, which earns it two additional lists: the “best” and “top” phrasings of the question. The ranks below are illustrative.

SUB-QUERY	GENERIC NAME	.BEST NAME
best romantic hotels Paris	—	#12
boutique hotels Paris couples	#8	#9

SUB-QUERY	GENERIC NAME	.BEST NAME
Paris hotel reviews 2026	#15	#15
top hotels Paris comparison	—	#22
Fused score, k = 60	0.0280	0.0539

The generic site scores $1/68 + 1/75 = 0.0280$. The .Best site scores $1/72 + 1/69 + 1/75 + 1/82 = 0.0539$: about 1.9 times higher, without ranking first anywhere, because it is present in two more lists. MECHANISM

What the name does not do

CLAIM	REALITY
ChatGPT boosts .Best domains	No evidence that any top-level domain is weighted as such.
A .Best name puts you in every list	Only pages that are relevant for a sub-query are retrieved for it. The name helps a relevant page match the phrasing; it does not create relevance.
Fusion guarantees a citation	Fused candidates are typically reranked; the passage still has to answer the question.
The advantage is measured	The click advantage is measured on a small sample; the fusion advantage is a mechanism, not yet measured. Section 12 proposes the test.

THE .BEST ADVANTAGE, IN RRF TERMS

RRF adds one term for every list a source appears in. A .Best name is built to be found by the most common phrasing ChatGPT adds, to be recognised as the same source in every list, and to be chosen when it is shown. It does not replace good pages. It makes good pages count more times.

09 THE PLAYBOOK

The RRF content playbook

If fusion rewards presence across a cluster of related searches, content strategy changes shape: from one page per keyword to one well-covered cluster per buyer question.

1 · Map the fan-out, not the keyword

For each buyer question you want to win, list the searches ChatGPT is likely to run. Use the patterns the 2026 data shows: the category plus "best" or "top"; a "comparison" or "vs" form; a "reviews" form; a form with the current year; and the brand or "official" verification search for each option. Five to eight searches per question is a practical size.

2 · Give every search a page that deserves to rank

SUB-QUERY PATTERN	PAGE THAT ANSWERS IT	EXAMPLE ON HOTELS.BEST
best / top + category	A ranked list with stated criteria	Best boutique hotels in Paris, 2026
comparison / vs	A side-by-side table	Le Marais vs Saint-Germain: where to stay
reviews	Original, dated assessments	Hotel reviews, updated monthly
year / "2026"	A visibly current page	What changed in Paris hotels in 2026
official / brand	A clear entity page	About hotels.best: who we are, how we select

3 · Win the middle of many lists

Do not pour every resource into one first place. Under fusion, moving a page from #25 to #12 in three lists does more than moving one page from #4 to #1. Spread effort across the cluster.

4 · Make the answer quotable

Fusion gets a page into the pool; the reranker picks the passage. Put a one-paragraph answer at the top of each page, with the entity, the claim and the figure together, and keep comparisons in real tables.

5 · Keep it fresh and dated

The current year appears in a large share of 2026 fan-outs. Pages that show a visible update date and current information match those searches; pages that do not are filtered out of them.

6 · Be one entity everywhere

Use the same name, the same description and the same facts on your site, your profiles and the pages that mention you. On a .Best site, the name already states the category, so every mention reinforces the cluster.

7 · Let the crawlers in

OpenAI states that sites that opt out of its search crawler, OAI-SearchBot, will not be shown in ChatGPT search answers. Check robots.txt, render content in HTML, and avoid hiding answers behind scripts.

```
# robots.txt: allow AI search crawlers on a .Best site
User-agent: OAI-SearchBot
Allow: /

User-agent: Googlebot
Allow: /

Sitemap: https://hotels.best/sitemap.xml
```

Common mistakes under fusion

- **Chasing one trophy keyword.** A first place in one list is worth 1/61 and no more.
- **Splitting the entity.** Different names on different pages and profiles make one source look like several weak ones.
- **Writing for the prompt instead of the fan-out.** The searches add words the user never typed; pages that only match the prompt miss most lists.
- **Undated content.** Fan-outs carrying the year skip pages that look stale.
- **Burying the answer.** In the pool, the clearest passage wins the citation.

THE .BEST VERSION OF THE PLAYBOOK

Choose a name that states the category and the word the searches already use: hotels.best, crm.best, ebike.best. Build the cluster of pages behind it. Every page then carries “best” in its address, every mention of the brand repeats the category, and every list the site appears in shows the same, recognisable source.

10 MEASUREMENT

Measure your RRF footprint

Nobody outside OpenAI can see ChatGPT's fused scores. You can, however, compute a proxy from data you control, and track whether your cluster strategy is working. The method below needs a rank tracker and a spreadsheet.

Step by step

- 01 **Pick ten buyer questions** that matter to your business, phrased the way people ask ChatGPT.
- 02 **Write the likely fan-out** for each: five to eight searches, using the patterns in section 09. Where you can, capture real fan-outs with a fan-out logging tool or a browser extension that shows ChatGPT's searches, and replace your guesses.
- 03 **Track rankings** for every search in a classic search engine, top 100.
- 04 **Compute your fused score** per question with $k = 60$, and repeat with $k = 10$ and $k = 100$ to check the conclusion does not depend on the constant.
- 05 **Compare with competitors** on the same searches.
- 06 **Check real citations:** ask each question in ChatGPT several times a week, log which sources are cited, and see whether fused scores and citations move together.

```
# Fused score per question, k = 60 (Python)
def rrf(ranks, k=60, weights=None):
    """ranks: list of positions (1 = first), None if absent."""
    weights = weights or [1.0] * len(ranks)
    return sum(w / (k + r) for r, w in zip(ranks, weights) if r)

rrf([None, 8, 15, None]) # generic site → 0.02804
rrf([12, 9, 15, 22])   # hotels.best → 0.05391
```

Five numbers to track

METRIC	DEFINITION	WHY
Coverage	Share of fan-out searches where you rank in the top 100	The number of RRF terms you collect
Fused score	$\sum 1/(60 + \text{rank})$ per question	Your proxy for the candidate pool
Threshold hits	Questions where the fused score is ≥ 0.020	The reported working threshold
Citation rate	Share of ChatGPT runs citing your site	The outcome
Cited-site clicks	Sessions from <code>utm_source=chatgpt.com</code>	OpenAI adds this tag to referral links

HOW TO READ IT

The fused score is a model of the candidate pool, not of the answer. If it rises and citations do not, the bottleneck is the second gate: make passages clearer. If citations rise with it, the cluster strategy is working. Judge only trends over several weeks; single runs are noise.

11

USE CASES

Three .Best sites, worked through

The cases below apply the method to three category names. Fan-outs follow the patterns reported in 2026 data; ranks and scores are illustrative targets, not measurements.

hotels.best · “Where should I stay in Paris for a romantic weekend?”

LIKELY FAN-OUT	PAGE	TARGET RANK	TERM
best romantic hotels Paris	Best romantic hotels in Paris, 2026	10	0.0143
boutique hotels Paris couples reviews	Boutique hotel reviews	15	0.0133
Paris hotels with Eiffel Tower view	Hotels with a view, compared	20	0.0125
hotels.best official	About hotels.best	1	0.0164
Fused score			0.0565

software.best · “Which CRM should a ten-person startup use?”

LIKELY FAN-OUT	PAGE	TARGET RANK	TERM
best CRM for startups 2026	Best CRM software for startups	12	0.0139
HubSpot vs Pipedrive vs Attio	CRM comparison table	18	0.0128
CRM pricing small team	CRM pricing, updated monthly	25	0.0118
top CRM alternatives to Salesforce	Best alternatives to Salesforce	14	0.0135
Fused score			0.0520

restaurants.best · “Where should I eat in Paris tonight?”

LIKELY FAN-OUT	PAGE	TARGET RANK	TERM
best restaurants Paris	10 best restaurants in Paris	20	0.0125
top restaurants Paris open tonight	Open tonight, by district	16	0.0132
best restaurants Paris reviews	Restaurant reviews, dated	22	0.0122
Fused score			0.0379

None of the three sites needs a first place. Each clears the reported 0.020 threshold by appearing in several lists in the middle of the first two pages, with a name that matches the “best” phrasing of most of them. The customer journey that follows is the one GEO is built on: question, fan-out, retrieval, fusion, citation, visit, booking.

MECHANISM

12

LIMITS

Limits and open questions

A white paper about an unconfirmed mechanism should say clearly what would prove it wrong. These are the limits we see, and the study that would settle the .Best question.

What could be wrong

- 01 **ChatGPT may not use RRF.** The evidence is one client-side observation and one hedged replication. OpenAI may use another fusion method, or none. The breadth-over-peak logic holds for most rank-fusion methods, but not for a system that, say, takes only each sub-query’s top results. OPEN QUESTION
- 02 **The constant is unknown.** A small k would restore much of the value of first places. Section 04 shows breadth still wins at k = 10, by a smaller margin.
- 03 **Rerankers may dominate.** If a strong reranker re-scores a wide pool, fusion matters mainly for entry into the pool. Industry research in 2026 found fusion gains largely neutralised after reranking in some settings.
- 04 **The pipeline changes.** 2026 data shows ChatGPT building its own index, changing fan-out counts and shifting vocabulary within months. Any formula decoded today can change tomorrow.
- 05 **Fan-out data is vendor data.** Most 2026 measurements come from companies that sell AI visibility tools. Their samples are large but not independent, and their bases differ.
- 06 **No study isolates the domain name.** We found no 2026 study that tests whether a descriptive or keyword domain changes AI citation rates. The .Best fusion advantage is a mechanism until measured.

The .Best fusion study

HYPOTHESIS	METHOD	OUTCOME MEASURED
H1 · .Best pages appear in more fan-out lists than matched sites	Log real fan-outs for 200 category questions; rank-track every fan-out for matched .Best and non-.Best sites with comparable authority	Coverage and fused score
H2 · Higher fused scores predict citation	Ask each question 20 times per week for 8 weeks; log citations	Correlation of fused score with citation rate
H3 · .Best citations are clicked more	Track utm_source=chatgpt.com sessions per citation	Click-through per citation

The registry intends to publish such a study and invites .Best registrants to contribute anonymised data. Until then, the honest position is the one this paper takes: the formula is public, its use in ChatGPT is reported, and the .Best advantage under it is a reasoned hypothesis with one measured click effect behind it.

13 ACTION PLAN

A 30-day plan for a .Best site

WEEK 1 · MAP

Choose ten buyer questions. Write or capture the fan-out for each. List the pages you have and the searches they miss.

WEEK 2 · BUILD

Create or fix one page per missing sub-query: best list, comparison, reviews, current-year page, entity page. Put the answer first.

WEEK 3 · ALIGN

Make the name, description and facts identical everywhere. Open robots.txt to OAI-SearchBot. Date every page.

WEEK 4 · MEASURE

Rank-track every fan-out, compute fused scores at $k = 10, 60$ and 100 , log ChatGPT citations over repeated runs, and set the baseline.

Conclusion: count more times

AI search does not hand out one first place. It runs several searches, merges what they return, and cites the few sources that hold up across them. If the merging works as reported, by Reciprocal Rank Fusion, the arithmetic is stark: a first place is worth barely more than a tenth, and every additional list a source appears in adds a whole new share of the score.

That changes what a name is for. A .Best domain cannot make a page relevant, and nothing suggests ChatGPT weighs the extension. What it can do is state, in the address itself, the word ChatGPT adds to its searches more than any other, and present the same recognisable source in every list where the site's pages appear. Under a formula that rewards being found again and again, that is a structural advantage worth testing, and worth building on.

WHEN AI SEARCHES FOR THE BEST

Be in every list it builds. Map the fan-out, cover it with pages that deserve to rank, keep one name that says what you are the best answer to, and measure it. Discover .Best domains at **domains.best**.

REFERENCES

Sources

Read directly: Elastic and OpenSearch documentation (source files). Read through search coverage and secondary reports, because the original pages were not reachable from our network when this paper was prepared: all other sources. Figures are reported as published; check the originals before quoting.

Cormack, G. V., Clarke, C. L. A. & Büttcher, S. Reciprocal Rank Fusion outperforms Condorcet and Individual Rank Learning Methods. SIGIR 2009, pp. 758–759. doi:10.1145/1571941.1572114

Elastic. Reciprocal rank fusion; RRF retriever. Elasticsearch reference documentation (rank_constant default 60; per-retriever weight since 9.2).

OpenSearch. Score ranker processor (introduced 2.19; rank_constant default 60).

Microsoft. Relevance scoring in hybrid search using Reciprocal Rank Fusion (RRF). Azure AI Search documentation.

OpenAI. Introducing ChatGPT search, 31 October 2024. Searching the web with ChatGPT (Help Center). Publishers and developers FAQ (Help Center).

Google Search Central. AI features and your website (query fan-out in AI Overviews and AI Mode).

Yesilyurt, M. Is RRF the Secret to Dominating AI Citations? I Decoded ChatGPT’s Ranking Formula. metehan.ai, 22 July 2025; and X thread of the same date.

Yesilyurt, M. The RRF Top-n Playbook: How to Get Cited by ChatGPT’s Web Mode (k≈60). metehan.ai, September 2025.

Peec AI (Rudzki, T.). Patterns we see in ChatGPT query fanouts, May 2026; Words ChatGPT adds to fanouts (docs.peec.ai/research).

Peec AI. ChatGPT fan-outs have doubled in length in 4 months (20M+ fan-outs); Rerankers for GEO/AEO: how AI search chooses passages and sources; ChatGPT built its own search index, September 2026.

Ray, L. ChatGPT is changing which page types it cites in AI responses (new Peec AI data). X article, 2026.

Ahrefs (Linehan, L.). How ChatGPT does query fan-out (15 million queries studied), October 2026; Only 12% of AI cited URLs rank in Google’s top 10 for the original prompt.

Ahrefs (Linehan, L. & Guan, X.). Why ChatGPT cites one page over another (study of 1.4M prompts).

AirOps. The influence of retrieval, fan-out, and Google SERPs on ChatGPT citations, March 2026.

Understanding ChatGPT Citations: Evidence from Production Web Search Across Model Generations. arXiv 2610.15382, October 2026 (preprint).

Scoring With the Engine: Retrieval Exposure, Cross-Engine Divergence, and the Limits of Engine-Agnostic GEO Scores. arXiv 2609.22655, 2026 (preprint).

Medrano, L., Verma, A. & Chhabra, M. Scaling Retrieval Augmented Generation with RAG Fusion: Lessons from an Industry Deployment. arXiv 2603.02153, March 2026 (preprint).

Meng, S. & Ilvovsky, D. Sifei at SemEval-2026 Task 8: Hybrid Retrieval and Query Rewriting for Multi-Turn RAG. arXiv 2606.28352, June 2026.

Mohanadasan, S. How ChatGPT actually picks sources (I read the network traffic, not the outputs). 2026.

Discovered Labs. How ChatGPT uses Reciprocal Rank Fusion for AI citations.

GlobeRunner. .BEST vs .COM case study: controlled Google Ads test, four Dallas sectors, small sample.

About .Best

Type	Generic top-level domain (gTLD), delegated in the Internet root zone
Registry operator	Best TLD Pty Ltd
Registry website	domains.best
Author	Cyril Frémont, Founder & CEO, internet.Best

.Best

AI searches for the best. Be in every list it builds.

Find a .Best name for your category, build the cluster of pages behind it, and measure your fused score to see where you stand today.