

A third of Perplexity's citations don't contain the number they're cited for

Daniel Rossi

Founding Director - Haus Research
United Kingdom
research@hausresearch.com

Abstract—Perplexity's search models mark each sentence with the URL it came from. We put 310 factual questions about 210 technology companies to both models, collected the 2,511 citation markers they placed, fetched every cited page, and checked whether it contained the figures in the sentence attached to it. Of the 1,826 markers on a sentence stating a figure, 34.7% pointed at a page that would not open to an ordinary reader, or opened and contained none of that sentence's numbers. Scored per claim instead, 14.4% of 872 claims fail. Dead links account for little of this: only 1.3% of cited URLs were dead. The large categories are gated pages (16.1% of cited URLs) and readable pages that do not carry the figure. A quarter of cited URLs have never been archived. We report the deterministic check alongside a model-judged groundedness rate of 40.0% end to end, and treat the gap as the uncertainty a model judge introduces.

■ **WE ASKED** Perplexity's two search models 310 factual questions about 210 technology companies, collected every source they cited, fetched all of them, and checked whether the page said the thing it was cited for. Of the 1,826 citations attached to a sentence stating a figure — the ones checkable without a second opinion — **34.7%** pointed at a page that would not open to an ordinary reader, or opened and contained none of the numbers in the sentence they were attached to. The models placed 2,511 citation markers in all.

The unit above is the citation, not the claim. Two thirds of the 872 claims carrying a figure have more than one marker on them, and we score each marker separately. Score instead per claim, counting a claim as passing when any one of the pages it points at carries one of its figures, and **14.4%** fail. We lead with the citation because a marker is an individual claim of provenance: this sentence came from that URL.

The failure is not mainly dead links. Only 1.3% of cited URLs were dead. The two large categories are pages a reader cannot get into, and pages a reader can

get into that do not say it.

What we did

Ten question templates, each a fact somebody would actually look up: founding, latest funding round, headcount, entry price, headquarters, revenue, disclosed breaches, current CEO, acquisitions, paid-tier uptime SLA. Every company got one; 100 of them got a second on a different template. 310 questions, put at temperature 0 to `perplexity/sonar`, `perplexity/sonar-pro` and, as a control, GPT-4.1 with a web plugin.

Both Perplexity models mark their claims inline as `[n]`, and `n` indexes the citation array they return. That is the part that makes an audit possible: it is not a bibliography at the bottom of the answer, it is a specific assertion that this sentence came from that URL. We split each answer into sentences and produced one claim–citation pair per marker. Neither model ever emitted a marker pointing past the end of its own citation list.

Then we fetched every unique cited URL — 2,915 of them for `sonar` alone — and classified each as dead, gated, empty, unreachable or live. Anything that failed

Date of publication 02 09 2026; date of current version 02 09 2026

got two more chances: a longer timeout, then a retry through a rotating proxy so that no page was recorded as blocked merely because one datacentre address was unwelcome. That third pass rescued 192 URLs. The classification can only ever move in a page's favour.

The headline check needs no model at all. From each claim we pulled its specifics — money amounts, percentages, magnitudes, years, any run of three or more digits — and asked whether the cited page's visible text contains at least one of them, normalising so that \$185 million, \$185M and 185000000 all match. One figure is enough to pass. A bare year is enough to pass. The 34.7% is therefore a floor: every failing pair is one where the page contains not a single number from the sentence that cited it.

The citations that do not open

Across perplexity/sonar's 2,915 unique cited URLs:

Class	Share
Live and readable	78.7%
Behind a login, paywall, 403 or bot wall	16.1%
Client-rendered shell we could not read	2.5%
Dead (404, 410, DNS failure, soft 404)	1.3%
Still unreachable after three passes	1.4%

One citation in six is gated. That is not a fault of the source — PitchBook, ZoomInfo, Crunchbase and Reuters are entitled to charge — but it is a fault of the citation. A footnote a reader cannot open is a claim of provenance with no way to test it, which is the condition a citation exists to prevent.

Aggregated to the answer, 84.2% of sonar's 310 answers cited at least one URL an ordinary reader could not open, and 10.6% cited at least one that was outright dead.

The dead ones are worth naming, because about half of them are the same kind of page — 20 of sonar's 38 — and their URLs give them away. `komo.ai/directory/<company>-offices`. `temperstack.com/plans/<company>`. `devhelm.io/sla/<company>`. `apollo.io/where-is/<company>`. `portersfiveforce.com/blogs/brief-history/<company>`, and the identical path on `matrixbcg.com` and

`canvasbusinessmodel.com`. These are pages minted per company per question type, published at scale to catch exactly the query we asked, and taken down as cheaply as they went up.

Three we re-fetched on the day of writing. Asked where Elastic is headquartered, sonar cited `komo.ai/directory/elastic-offices` [1]: 404. Asked for Reddit's head office, both models cited `apollo.io/where-is/reddit` [2]: 410 Gone. Asked for Discord's cheapest paid plan, sonar-pro cited `temperstack.com/plans/discord` [3]: 404.

The citations that open and do not say it

Of the pairs whose page did open and was readable, 16.1% contained none of the claim's own figures.

The cleanest example is a price. Asked for the entry price of Vercel's cheapest paid plan, sonar answered that "the free Hobby plan is \$0/month, so the first paid tier starts at \$20/month", and cited the Vercel plans documentation [4]. We fetched that page at write time. It returns HTTP 200, it names the plans, and the strings \$20, \$20/month and 20/month do not appear anywhere in it. The number is probably right. The citation is not evidence for it.

The second pattern is more revealing, because it repeats across companies. Asked for headquarters, both models produce a street address and attribute it to the company's Wikipedia article:

Claim	Cited page	Address on that page?
Docker at 3790 El Camino Real #1052, Palo Alto, CA 94306	Wikipedia: Docker, Inc. [5]	No
Rippling at 430 California Street, San Francisco, CA 94104	Wikipedia: Rippling (company) [6]	No
Substack at 111 Sutter Street, San Francisco, CA 94104	Wikipedia: Substack [7]	No
SentinelOne at 444 Castro Street, Mountain View, CA 94041	Wikipedia: SentinelOne [8]	No

All four articles were fetched at write time and none contains the street number, the street name or the postal code attributed to it. Several of the answers say so themselves, in phrasing like "multiple sources list" or "several business directories list", and then attach a marker to Wikipedia anyway. The claim and the

citation were produced by the same process, and that process is not retrieval.

A softer version of the same thing: `sonar` said GitLab’s CEO is Bill Staples and that he took the role on 5 December 2024, citing GitLab’s own executive team page [9]. That page names Bill Staples. It does not carry the date. Half the sentence is sourced.

Where it fails worst

Pooling both Perplexity models, by question type, share of pairs whose cited page contained one of the claim’s figures:

Question	Pairs	Passed
Who is the current CEO	235	44.3%
Headquarters address	202	53.0%
Entry price	171	62.6%
Security incidents	205	66.8%
Uptime SLA	145	69.0%
Acquisitions	278	69.1%
Latest funding round	103	70.9%
Revenue or ARR	192	72.4%
Founding	128	75.0%
Headcount	167	82.0%

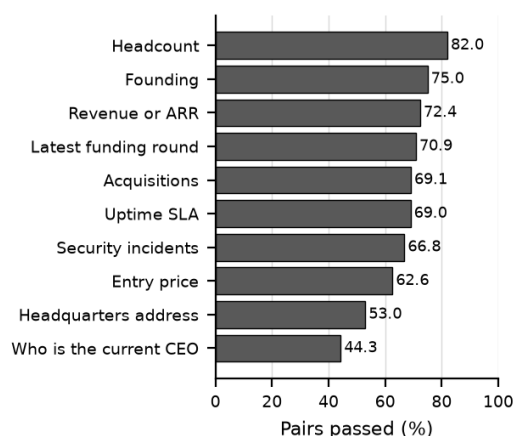


Figure 1. Pooling both Perplexity models, by question type: share of claim–citation pairs whose cited page contained at least one of the claim’s figures.

The ordering is not random. It tracks how well a fact is written down in one canonical place. Headcount and founding year sit in structured fields on pages built to hold them. A CEO’s start date and an office’s street number are the kind of thing everyone repeats

and nobody publishes, so the model reproduces the consensus and then points at a page that never carried it.

The premium model is not better, and not worse

Our pilot suggested that `sonar-pro` grounded its claims less well than `sonar`. At full scale that gap disappears. `sonar` passes on 65.9% of numeric pairs (95% interval 62.8–68.9), `sonar-pro` on 64.7% (61.5–67.7). The intervals overlap comfortably, and the two models cite at nearly identical rates: 9.8 and 9.7 sources per answer. On this measurement they are the same product. The pilot result was a small sample telling us what we wanted to hear, and it did not replicate.

GPT-4.1 with a web plugin behaves differently in one respect worth noting: it cites 2.0 sources per answer rather than 9.8, and 36.4% of them are the company’s own domain against Perplexity’s 23.4%. It emits no inline markers, so no claim-level check is possible on it, which is itself the finding — an answer whose citations are a list at the end cannot be audited sentence by sentence.

Who gets cited

Across `sonar`’s 3,031 citations, spread over 989 distinct hosts, 23.4% point at the company’s own domain and 23.1% at a B2B directory, revenue estimator or lead list — Tracxn, PitchBook, Clay, GetLatka, Zoom-Info, CB Insights, Growjo, Crunchbase and their many imitators. The largest single host is `linkedin.com` at 5.6%, then `en.wikipedia.org` at 4.3%, then `tracxn.com` at 3.0%.

Those directory pages are also the least durable material in the set: 66.0% of them opened, against 78.7% of citations overall. They are generated from databases, published at scale to rank, gated or retired without notice, and they are where a quarter of the sourcing for questions about real companies now goes.

The model-judged number, for comparison

We also ran a conventional groundedness judgment as a secondary metric: a separate model shown each claim and its cited page, asked whether the page supports it. On 400 randomly sampled readable pairs per model it called 50.8% of `sonar`’s claims supported, 24.5%

partial and 24.8% unsupported. Chained with the share of pairs whose page opens at all, that is an end-to-end rate of 40.0%.

The two metrics are not measuring quite the same thing, and the distance between them is the point. On a readable page, the deterministic check clears 84.6% of sonar's pairs, because one matching figure is enough to pass it; the judge, which is asked whether the page supports the whole assertion, clears 50.8%. End to end that is 65.9% against 40.0%, a spread of twenty-six points.

We lead with the deterministic figure anyway. A model grading another model's work is the weakest joint in this design, and the spread is the honest size of the uncertainty it introduces. The generous number is the one a reader can reproduce with `grep`, and it is already bad enough to be the finding.

Nobody is keeping copies

We looked up a random 1,500 of sonar's cited URLs in the Wayback Machine's CDX index. Of the 1,432 that resolved, **25.1%** (95% interval 23.0–27.5) have never been captured at all — not once, at any point in the life of the page.

For the directory and lead-list pages the figure is 39.3%. Two in five of them exist only as long as the company hosting them keeps them up, and we already know how that ends: the dead URLs in this study are almost entirely from that same population.

So the citation layer under AI answers is being assembled largely out of pages built to be found rather than kept, and a quarter of it has no copy anywhere. When one of these URLs goes, the claim that leaned on it does not go with it. The sentence stays, the marker stays, and the only thing that has disappeared is the ability to check.

What this does not show

It is one snapshot, taken on 2 September 2026. It is not a decay rate, and we have not re-run the same URLs later.

It covers English-language questions about technology companies. Nothing here establishes what happens in other languages or sectors.

A gated page is not a wrong page. A subscriber to PitchBook may well find the claim fully supported

there; we are measuring what a reader without an account can verify, which is a narrower and more useful question, but a different one.

We fetch HTML without running JavaScript, which is why the `empty` class exists and is reported separately rather than folded into the dead count.

The answers came from `perplexity/sonar` and `perplexity/sonar-pro` through OpenRouter, not from the Perplexity consumer product, which retrieves and cites under its own settings.

We fetch as an identified bot from datacentre addresses, so the gated share is larger than the share a person with a browser and a home connection would meet.

The 210 companies were picked by hand as well-known technology firms rather than drawn from a defined universe, so the mix of subjects is ours and not a population.

And a claim can be true with a bad citation, which most of these are. We are not measuring whether Perplexity is right. We are measuring whether the thing it offers as proof functions as proof. On a third of the citations we checked, it does not.

Data availability

The full dataset is published under CC BY 4.0:

- questions: <https://hausresearch.com/data/perplexity-citation-audit/questions.csv>
- every cited URL and how it resolved: <https://hausresearch.com/data/perplexity-citation-audit/citations.csv>
- every claim–citation pair with both verdicts: https://hausresearch.com/data/perplexity-citation-audit/claim_citations.csv
- the Wayback lookups: <https://hausresearch.com/data/perplexity-citation-audit/wayback.csv>

The collection and analysis scripts (<https://hausresearch.com/data/perplexity-citation-audit/README.md>), the full method (<https://hausresearch.com/data/perplexity-citation-audit/METHOD.md>), the computed figures (<https://hausresearch.com/data/perplexity-citation-audit/numbers.json>) and the write-time evidence (<https://hausresearch.com/data/perplexity-citation-audit/evidence.json>) for every named example are in the same directory.

■ REFERENCES

1. Komo, *Elastic offices*.
<https://komo.ai/directory/elastic-offices>
2. Apollo, *Where is Reddit*.
<https://www.apollo.io/where-is/reddit>
3. Temperstack, *Discord plans*.
<https://www.temperstack.com/plans/discord/>
4. Vercel, *Plans*. <https://vercel.com/docs/plans>
5. Wikipedia, *Docker, Inc*.
https://en.wikipedia.org/wiki/Docker,_Inc.
6. Wikipedia, *Rippling (company)*.
[https://en.wikipedia.org/wiki/Rippling_\(company\)](https://en.wikipedia.org/wiki/Rippling_(company))
7. Wikipedia, *Substack*.
<https://en.wikipedia.org/wiki/Substack>
8. Wikipedia, *SentinelOne*.
<https://en.wikipedia.org/wiki/SentinelOne>
9. GitLab, *E-Group*.
<https://about.gitlab.com/company/team/e-group/>