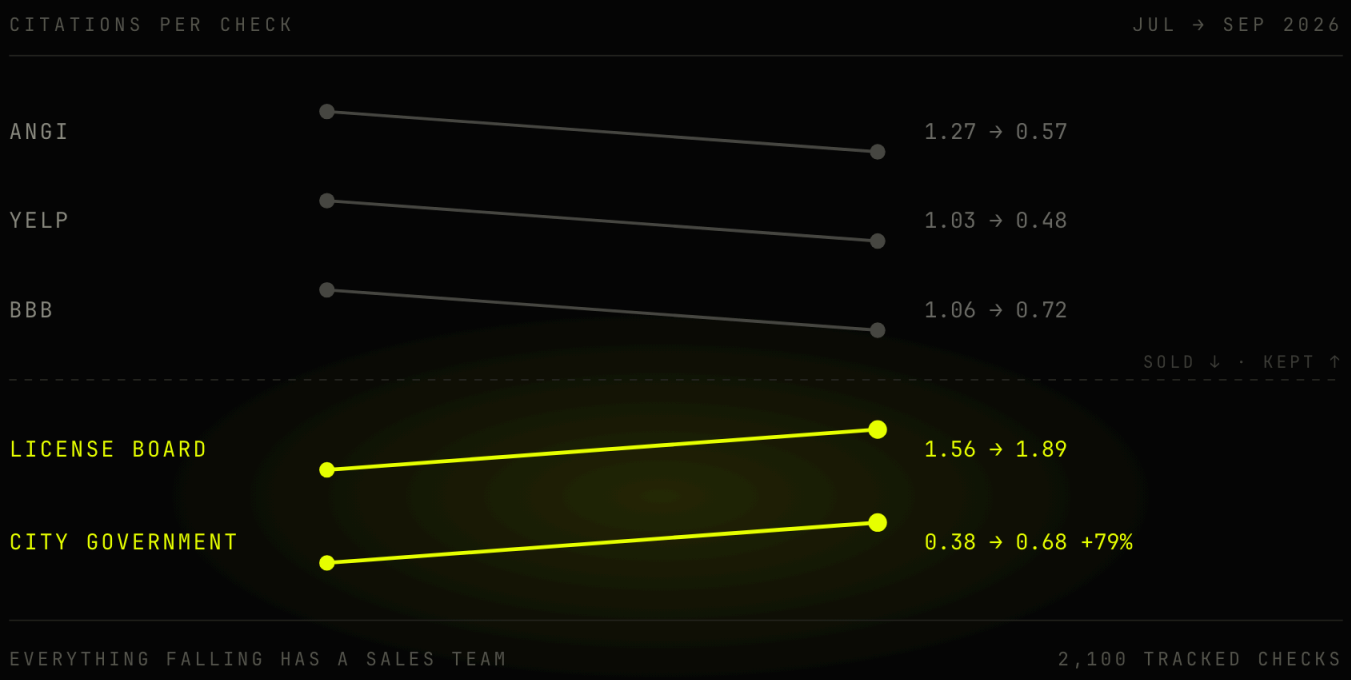
CITATION ANALYSIS × AI VISIBILITY

18 MIN READ

THE DIRECTORY *COLLAPSE.*

What 2,100 tracked AI checks in one Dallas-Fort Worth market reveal about the collapse of directory citations — and the institutional records quietly replacing them.



32.78°N · THE TRUST STACK, RE-WEIGHTED

97.33°W · DFW

AUTHOR

The 6 Signal Research Team

AI Visibility · Citation Analysis

PUBLISHED

October 2026

READ

6signal.co/research/the-directory-collapse →

WHITE PAPER · 18 MIN READ · OCTOBER 1, 2026

The Directory Collapse: 2,100 Tracked AI Checks Show Yelp, Angi, and BBB Losing Their Grip

Between July and September 2026 we ran 2,100 automated AI visibility checks against one Dallas–Fort Worth plumbing market and stored every source trail. Normalized per check, citations to Yelp fell 53%, Angi 55%, and BBB 32% — while the state license board rose 21% and the city government page rose 79%. The trust stack is re-weighting, and the direction is away from anything a contractor can buy.

BY MATT VINCENT WALKER · 6 SIGNAL

In September 2025, the advice to a contractor who wanted to be found by AI was straightforward: get on the directories. Claim Yelp, claim Angi, claim BBB, keep HomeAdvisor current, and accept that the aggregators would keep winning the answer box because they had the domain authority and you did not.

That advice is going stale faster than anyone expected.

Between July 7 and September 30, 2026, our tracking system ran **2,100 automated checks** of real buyer prompts against one Dallas–Fort Worth plumbing market, across six surfaces — ChatGPT, Claude, Perplexity, Gemini, Google AI Overviews, and Google

Maps. For each check we stored the answer and the full list of sources the engine consulted on the way to it. This paper is about what happened to those source lists over twelve weeks.

The short version: the commercial directory layer is losing ground, quickly, and what is taking its place cannot be bought.

TRACKED CHECKS

2,100

42 buyer prompts · 6 surfaces · Jul 7 – Sep 30, 2026

YELP CITATIONS

−53%

1.03 → 0.48 per check, July to September

ANGI CITATIONS

−55%

1.27 → 0.57 per check

CITY GOVERNMENT PAGE

+79%

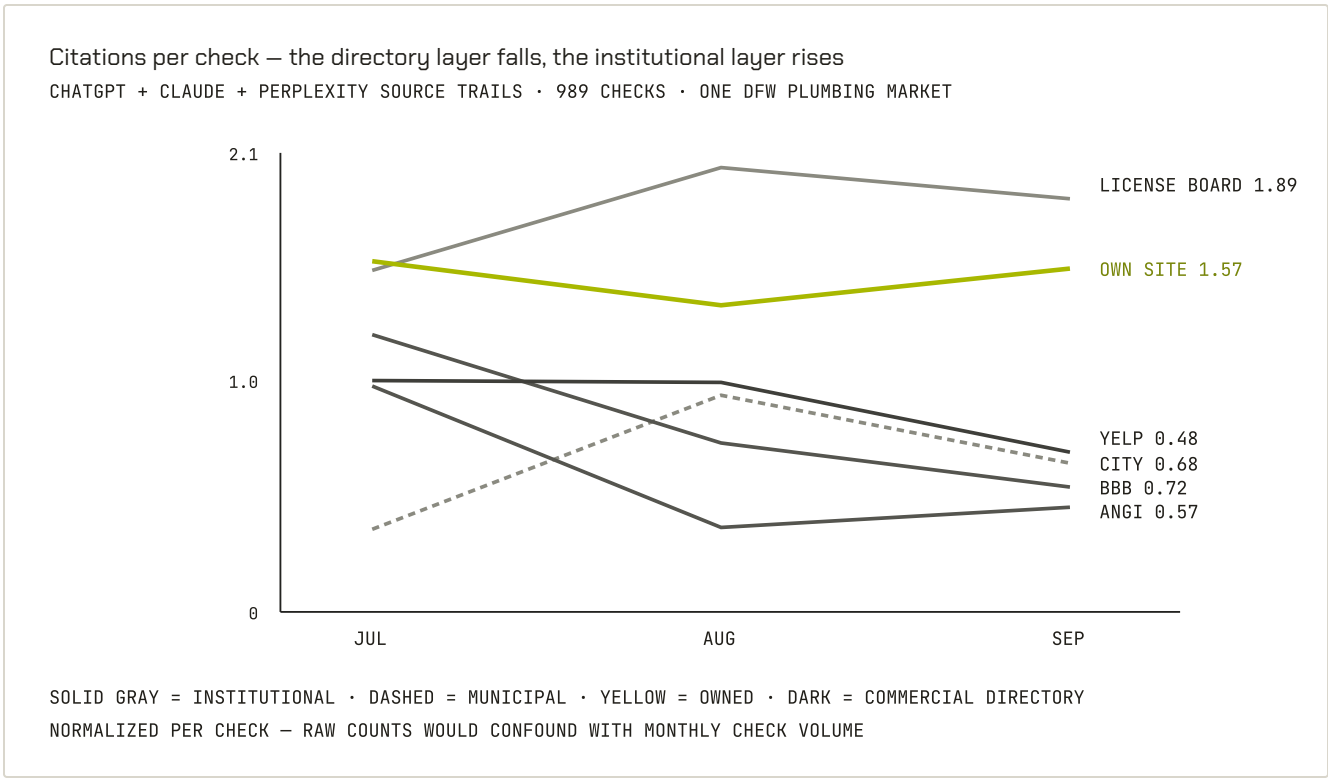
0.38 → 0.68 per check — the fastest riser

Part I — The measurement, and why it is normalized

A raw citation count is a misleading number. We did not run the same volume of checks every month — July produced 629 checks, August 615, September 856 — so a domain could appear more often in September purely because we looked more often. Every figure in this paper is therefore expressed as **citations per check**: the number of times a domain appeared in a source trail, divided by the number of checks run in that month against the engines being counted.

The source-trail figures count only the three research engines that return a readable list of consulted sources — **ChatGPT, Claude, and Perplexity** — which between them account for 989 of the 2,100 checks. Gemini returns its grounding URLs through a redirect service that cannot be resolved back to a publisher domain, so its trails are excluded from source counts throughout, exactly as in our first published study. Google AI Overviews and Google Maps are excluded from this section for a different and more uncomfortable reason, which Part IV deals with directly.

With that stated, here is the twelve-week movement.



Every directory line ends lower than it started. Every institutional line ends higher. The owned line barely moves — and stays on top.

The table, for readers who would rather have the numbers than the shape:

DOMAIN	CLASS	JULY	AUGUST	SEPTEMBER	CHANGE
x-actplumbing.com	Owned	1.60	1.40	1.57	−2%
tsbpe.texas.gov	Institutional	1.56	2.03	1.89	**+21%**
redoaktx.org	Municipal	0.38	0.99	0.68	**+79%**
bbb.org	Commercial	1.06	1.05	0.72	**−32%**
angi.com	Commercial	1.27	0.77	0.57	**−55%**
yelp.com	Commercial	1.03	0.39	0.48	**−53%**

Part II — What the decline is not

Three explanations present themselves before the interesting one, and each needs disposing of.

It is not that the engines got quieter. If the engines had simply started consulting fewer sources, every line would fall together. They did not. ChatGPT's trails averaged 78.8 sources per check in July and 61.4 in September — a real narrowing — but Claude held flat at 10.8 then 9.0, and Perplexity went *up*, from 10.6 to 19.4. Against that mixed background, the license board rose and the directories fell. The engines did not stop looking. They changed where they looked.

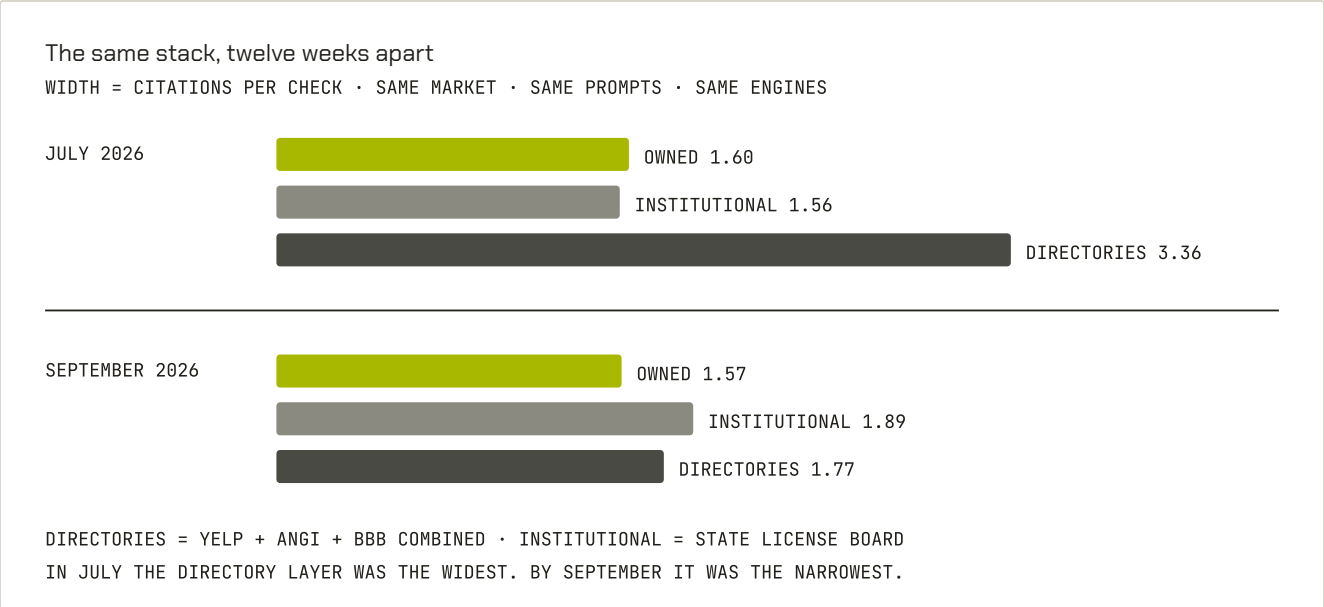
It is not that the company changed. The tracked business did not delist from Yelp, lose its BBB accreditation, or alter its directory presence during the window. Its own site stayed the most-cited domain in all three months. What moved was the engines' weighting, not the business's footprint.

It is not a single engine's quirk. Every figure above comes from domains present in **all three** research engines' trails. A change driven by one vendor's index refresh would not survive that filter.

What remains is the uncomfortable reading: over twelve weeks, three independently operated AI systems each reduced how much they leaned on commercial review platforms when assembling an answer about a local contractor, and each increased how much they leaned on government records.

Part III — The trust stack, re-weighted

The Supplier Signal described a four-layer trust stack — owned, institutional, commercial, community — and argued that engines cross-reference the layers rather than ranking them. That model survives this data. What has changed is the weighting between layers.



Combined, Yelp, Angi, and BBB went from 3.36 citations per check to 1.77 — a 47% fall in twelve weeks. The state license board alone now outweighs all three together.

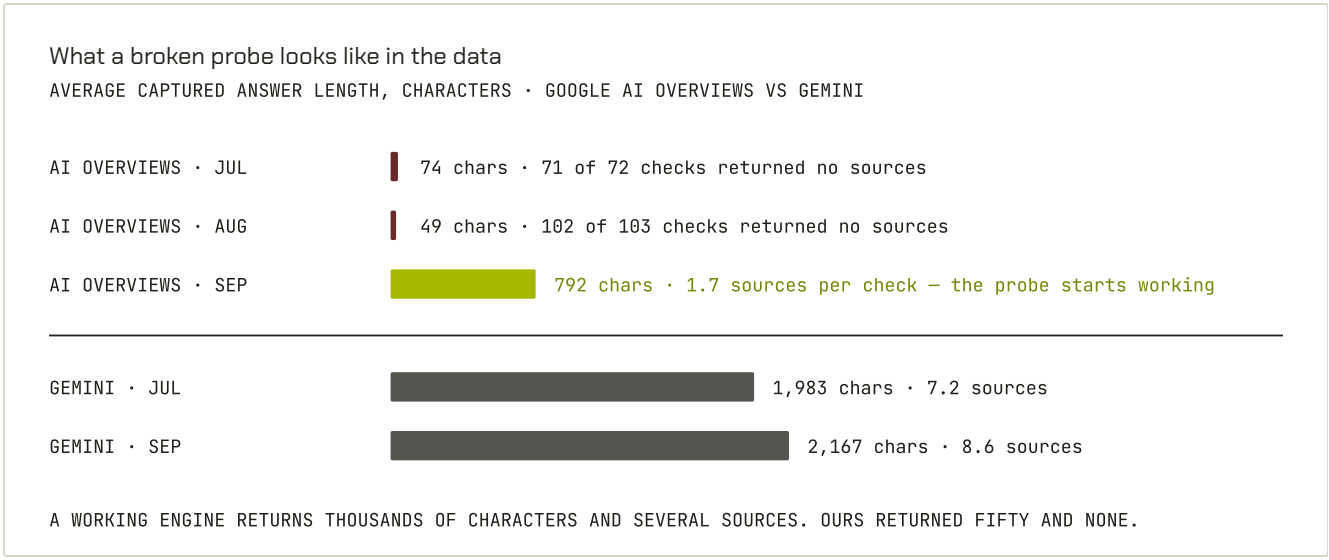
That last sentence is the paper. In July, a contractor's three big commercial profiles were collectively the heaviest thing in the engines' research. By the end of September, a single government database the contractor has never thought about outweighed all three.

There is a plausible mechanism, and it is worth stating as a hypothesis rather than a finding. Through 2026 the major engines have been under sustained pressure — commercial, legal, and reputational — over the reliability of what they assert. An answer that names a contractor is a recommendation, and a recommendation that turns out to be wrong about licensure or good standing is a liability of a different class than a wrong restaurant suggestion. Under that pressure, the rational move for an engine is to shift weight toward primary records: the regulator, the municipality, the registry. We cannot see inside the systems, and we are not going to pretend otherwise. But the behavior is consistent with verification displacing popularity.

Part IV — A correction to our own published work

In August we published a field note on the first 700 checks and made a point of highlighting one number: Google AI Overviews had mentioned the tracked business **zero times in 84 checks**. We wrote that the zero was the finding, and we published it deliberately rather than quietly dropping it.

Extending the dataset to 2,100 checks has given us enough signal to say that the zero was ours, not Google's.



An engine that is genuinely silent about a business still returns an answer about the market. Ours was returning almost nothing at all — the signature of a capture failure, not a verdict.

In July and August, our AI Overviews probe captured an average of **74 and 49 characters** per check, with no sources attached in 71 of 72 and 102 of 103 checks respectively. That is not an AI Overview. A functioning Overview in this dataset runs to thousands of characters and carries seven or more sources, as every Gemini capture did in the same window. In September the Overviews probe began returning real content — 792 characters on average, 1.7 sources per check, and the first 12 mentions.

So the honest correction reads: **we do not know what Google AI Overviews said about this business in July and August, because we were not successfully reading it.** The published zero measured our instrument. We have left the original field note online with a correction notice rather than editing the number out of it, because a research practice that silently revises its own numbers is worth less than one that does not.

The same caution applies to Google Maps, which recorded 0 mentions across 223 checks in July and August and then 85 in September, with average captured answer length rising from roughly 330 characters to 407 and then 662. We are treating the

Maps series before September as unreliable for the same reason, and have excluded both surfaces from every trend in this paper. Our earlier reporting of a 25-of-25 Maps grid presence came from a separate grid scan, not from this probe, and is unaffected.

This is the second time publishing the raw data has cost us a headline. It will not be the last. The alternative — reporting numbers we have not stress-tested — is how the AI visibility field filled up with statistics nobody can source.

Part V — What a contractor should actually do

The temptation is to read "directories are declining" as "cancel the directory listings." That is the wrong conclusion and the data does not support it. A 47% fall from a high base still leaves Yelp, Angi, BBB, HomeAdvisor, and BuildZoom appearing hundreds of times across the window, all five still inside the three-engine consensus set. They are losing weight, not disappearing.

The right conclusion is about **where the next hour of effort goes.**

01 **Pull your state license record and read it like an auditor.** For Texas plumbers that is the Texas State Board of Plumbing Examiners; for most other trades, the Texas Department of Licensing and Regulation. It was the single heaviest-cited third-party domain in our data by September. Confirm it is current, confirm the business name matches your website character for character, and confirm the responsible party is the person you expect.

02 **Claim your city and chamber listings.** The Red Oak municipal site was the fastest-rising domain we measured, up 79% per check in twelve weeks. Municipal business directories and chamber pages are usually free, frequently out of date, and almost

never maintained by the business they describe.

03 **Treat your own website as the primary record, because the engines do.** It was the most-cited domain about this company in every single month, at a rate the nearest third-party source only approached in August. Services in plain language, cities actually served, license number visible, credentials named. If a crawler cannot read it, none of the rest of this matters.

04 **Keep the directory profiles accurate and stop buying placement on the assumption it drives AI answers.** Claim them, correct the name-address-phone, and move on. Whatever a paid upgrade does for human traffic, our data gives no reason to expect it moves an AI recommendation — and the trend line is going the wrong way for that bet.

05 **Make every layer agree.** The engines cross-reference. A business whose license record says one name, whose website says a second, and whose Yelp page lists an old address is not a business with three listings. It is, to a verification system, three weak claims about possibly-different companies. Entity consistency is the whole game and it is boring work, which is why it is available.

06 **Re-measure in ninety days.** This paper describes twelve weeks. Twelve weeks from now the weighting will have moved again. Run the checks yourself with the measurement playbook, or let the $27 AI Visibility Audit score all six signals for your trade and city in one pass.

Method, and what this paper does not claim

The data comes from our internal tracking system: 2,100 automated checks of 42 buyer prompts for one Dallas–Fort Worth plumbing market, July 7 – September 30, 2026, run against ChatGPT, Claude, Perplexity, Gemini, Google AI Overviews, and Google Maps, with each engine's answer and source list stored at check time. Source-trail figures use only ChatGPT, Claude, and Perplexity (989 checks), and count only domains present in all three engines' trails. All citation figures are normalized per check. Gemini's grounding URLs resolve to a Google redirect service rather than publisher domains and are excluded from source counts. Google AI Overviews and Google Maps are excluded from all trend figures for the capture failure documented in Part IV.

This is **one market, one trade, one business, over twelve weeks.** It is not a cross-market study and we are not presenting it as one. A three-month window cannot distinguish a structural re-weighting from a long seasonal swing, and the mechanism proposed in Part III is a hypothesis about systems we cannot inspect. What we can say precisely is what we measured: in this market, against these prompts, across these three engines, commercial directory citations fell by roughly half while government-record citations rose, and the company's own website remained the most-consulted source throughout. We will publish the next window whether or not it agrees with this one.

The Directory Collapse is part of 6 Signal's visibility research. For the trust-stack model this paper updates, read The Supplier Signal; for the original baseline and the field note now carrying a correction, read We Ran 700 AI Visibility Checks in One Texas Market; for

the entity-consistency work the playbook depends on, read The Local Entity Gap. To see where your business stands today, start with the free AI visibility check or the $27 AI Visibility Audit.

Sources

- 6 Signal internal tracking system — 2,100 probe records with stored answers and source trails, July 7 – September 30, 2026 (method above)
- Texas State Board of Plumbing Examiners — tsbpe.texas.gov
- Texas Department of Licensing and Regulation — tdlr.texas.gov
- City of Red Oak, Texas — redoaktx.org

NEXT STEP

See where you *actually stand.*

The AI Visibility Audit runs your company through all six layers — GEO, AEO, LEO, VEO, PEO, IEO — and delivers instant results. $27. Specific to your business, your trade, and your market.

GET THE AUDIT

6signal.co/visibility-check