Scott Tischler

FOUNDER & CHAIRMAN · AIRECOMMEND.AI

HOW AI WORKS

Why AI Recommends You to One Person and Not Another

By Scott Tischler · October 16, 2026 · 8 min read

Here's a conversation I have constantly. An owner tells me, "I asked ChatGPT and it recommended us — we're winning at this AI thing." Two weeks later the same owner is deflated: "My biggest prospect asked the exact same question and we weren't even mentioned." Both things are true. Same question, same week, two different people, two different answers. If you don't understand why that happens, you'll either celebrate a mirage or panic over a ghost. So let me give you the direct answer first: AI answers differ between people because the model isn't answering a question in a vacuum — it's answering *your* question, shaped by your account, your location, your history, which model version you happened to hit, and a dose of built-in randomness. Some of that you can influence. Some of it you can't. Knowing which is which is the whole game.

The five reasons the same question gets different answers

Account history and memory. If you've talked to an assistant about your industry, your competitors, or your own business before, that history can bleed into later answers. When you ask "who's the best provider in my category," the model may be quietly weighting what it already knows you care about. This is why testing from your own logged-in account is the single most misleading thing you can do — you're not seeing what a stranger sees, you're seeing a version of the answer that's been warmed up by your own footprint.

Location and context. Many AI surfaces fold in where the user appears to be, what device they're on, and sometimes the broader context of the session. A question that reads as local gets a local answer, and "local" is defined by the asker, not by you. Two people in two cities asking an identical question are, functionally, asking two different questions.

Which model actually answered. The brand name on the box — ChatGPT, Gemini, Perplexity — is not one fixed thing. Providers route traffic across model versions and tiers, run free users and paid users on different configurations, and ship updates constantly. Your prospect on a free tier and you on

a paid one may be talking to meaningfully different systems. The label is the same; the machine behind it isn't.

Retrieval variation. When an engine searches the live web before answering — as Perplexity and the AI-search modes routinely do — it may pull a slightly different set of sources each time, depending on what's fresh, what's reachable in that moment, and how the query got expanded behind the scenes. Different sources in, different names out. The underlying question of how AI retrieves and blends sources is doing a lot of the work here, and it's not deterministic.

Plain randomness. Finally, these models are probabilistic by design. Ask the identical question twice, in two fresh sessions, and you can get two different answers for no reason other than the sampling that makes the output feel natural rather than robotic. This isn't a bug to be fixed; it's how the thing works. It means a single test is a single draw from a distribution, not a measurement.

What you can control, and what you can't

Lay those five out and a useful line appears between them. Account history, the specific model version someone hits, and raw randomness are not things you can optimize — they're conditions of the user, not properties of your business. Chasing them is wasted effort. What you *can* move is the thing underneath all of it: the strength, consistency, and breadth of the evidence about you that any of those systems draws on.

Think of it as a signal-to-noise problem. The user-side variables are noise — they jostle the answer around. Your authority is the signal. When your evidence is thin, you live or die on the noise: a lucky draw names you, an unlucky one doesn't, and your "AI visibility" swings wildly between people. When your evidence is strong and corroborated across many independent sources, you show up *across* the variations — different accounts, different cities, different model versions, different random draws. The goal was never to win one person's answer. It's to become the name that survives the noise, whoever's asking.

That reframes a lot of AEO work. Owning a prompt isn't getting the model to say your name once; it's getting it to say your name reliably across the spread of people and conditions that will actually ask it. Corroboration and co-citation matter precisely because they're what let you clear the noise floor consistently rather than occasionally.

How to test without fooling yourself

Almost every owner tests wrong, and the errors all point the same direction — toward flattering yourself. Here's how to do it honestly.

Never trust your own logged-in account. It's the most contaminated sample you have. Test signed out, and if you can, from a connection and device that aren't your business's. You're trying to see what a stranger sees, and you are not a stranger to your own tools.

Treat one answer as one data point, never a verdict. Because of randomness and retrieval variation, a single run tells you almost nothing. Ask the same question several times, in fresh sessions, and look at the *rate* — named four times out of five is a genuinely different reality than one out of five, even though a single lucky test can't tell them apart. If you take nothing else from this piece, take this: measure frequency, not incidents.

Vary the phrasing on purpose. Real people don't all ask the same way. Test the natural variants — "best," "most trusted," "affordable," "near me," "for [specific use case]." You'll often find you own one phrasing and vanish on a neighboring one, which tells you exactly where your evidence is thin.

Test the models your customers actually use, on the tiers they actually use. If your buyers live in one engine's free tier, your paid-tier results in a different engine are a vanity metric. Match the test to the customer.

Keep a simple log over time. Any given week is noisy. The trend across weeks — is my named-rate climbing as my authority work compounds? — is the thing that's actually real and actually worth watching.

Where the variance hits hardest

Not every question swings equally, and knowing which ones move tells you where you're actually exposed. Branded questions — someone typing your exact name — are relatively stable. The model either has a clear record of you or it doesn't, and user-side variables don't change that much. If you only ever test your own name, you'll badly overestimate how solid your position is, because you're testing the one kind of query that barely moves.

The variance explodes on the unbranded, category-level questions — "who's the best [category] in [city]," "what's a trustworthy [provider] for [need]." These are the discovery questions, the ones asked by people who don't yet know you exist, and they are the entire point of AI visibility. They're also where thin authority gets punished hardest: with a weak evidence base, your appearance on these queries is almost pure coin-flip, lurching between people and sessions. The cruel part is that the queries worth the most to you are exactly the ones where the noise dominates when your signal is weak. A business can genuinely own its own name across every engine and still be invisible on the category queries that bring new customers — and never notice, because it only ever tested the easy question.

So weight your testing toward the queries that matter: the unbranded discovery prompts a stranger would use, not the branded reassurance of seeing your own name come back. That's where you'll find the honest picture of where you stand.

Stop optimizing for your own screen

The trap in all of this is emotional as much as technical. The answer on your own screen feels like the truth because it's the one you can see. It isn't the truth — it's one draw, warmed by your own history, from a system that will answer your prospect differently for reasons that have nothing to do with how good you are. Celebrating it is how owners talk themselves into believing they've won when they've only gotten a friendly roll of the dice.

The healthier mental model: you're not trying to win an answer, you're trying to raise a floor. Every piece of real authority you build — consistent facts, corroborating mentions, fresh evidence across independent sources — lifts the baseline probability that you get named, across all the variation you can't control. Do that long enough and the difference between what you see and what your prospect sees shrinks, not because you optimized the noise, but because you drowned it out. That's the only version of "winning at AI" that holds up when someone other than you is asking.

Key takeaways

- The same AI question yields different answers to different people because the model answers a user-shaped query — colored by their account history, location, which model version they hit, live-retrieval variation, and built-in randomness.
- Three of those five factors (history, model tier, randomness) are conditions of the user, not properties of your business — you can't optimize them, so don't try.
- The one thing you can move is the signal underneath: the strength, consistency, and breadth of evidence about you. Strong corroborated authority is what lets you survive the noise across all those variations.
- Testing from your own logged-in account is the most misleading thing you can do — it's your most contaminated sample. Test signed out, ideally off your own network.
- One answer is one data point, never a verdict. Measure your named-rate across several fresh runs and varied phrasings; frequency is the signal, individual incidents are noise.

- Stop optimizing for the answer on your own screen. You're not winning a single answer, you're raising the baseline probability of being named for whoever asks.

Frequently asked questions

ChatGPT recommends my business when I ask — does that mean my AI visibility is good?

Not necessarily, and testing from your own account is the least reliable way to find out. Your logged-in history can warm the answer toward what you already care about, so you may be seeing a flattering version a stranger never gets. Test signed out, run the question several times in fresh sessions, and look at how often you're named rather than whether you were named once. A single friendly answer is one lucky draw, not a measurement.

If I can't control a user's account, location, or which model they hit, what's actually worth my effort?

The evidence underneath all of it. Account history, model version, and randomness are conditions of the user — they jostle the answer around but you can't optimize them. What you can build is strong, consistent, corroborated authority across many independent sources. That's the signal that lets you get named across the variation instead of only on a lucky roll. Raise that floor and the gap between what you see and what your prospect sees shrinks.

How many times should I test the same question before trusting the result?

Enough to see a rate, not an incident — several fresh, signed-out sessions at minimum, and ideally repeated over weeks. Because these models are probabilistic and live retrieval varies run to run, one answer tells you almost nothing. Being named four out of five times is a genuinely different reality than one out of five, and only repetition reveals which one you're in. Keep a simple log so you can watch the trend as your authority work compounds.

About the author. Scott Tischler is the Founder & Chairman of AIrecommend.ai and a practitioner-authority on AI search and Answer Engine Optimization. With 20+ years in marketing technology — including American Express, MetLife, and UBS — and executive and professional study at Wharton, Harvard, and Oxford, he helps businesses become the ones AI recommends.

© 2026 Scott Tischler · scotttischler.com
scott@airecommend.ai · AIrecommend.ai