Choosing the Right Prompts to Track

How to design a prompt panel that measures your AI visibility accurately without blowing your budget.

INSIDE THIS GUIDE

1. Why AI visibility is a poll, not a ranking
2. The emerging field of prompt tracking
3. How to design a prompt panel
4. Locations and personalization
5. Brand prompts vs. discovery prompts
6. Prompt tracking guidance, in table form

1. Why AI Visibility Is a Poll, Not a Ranking

Buyers are asking AI before they ask anyone else. It happens on small day-to-day purchases, and it happens on six-figure B2B deals where a committee wants a shortlist before the first call. In both cases the assistant names a handful of brands, and the ones it names get the meeting.

That creates a measurement problem. In traditional search you have first-party data; Search Console tells you the exact queries people typed to reach you. Nothing like that exists for AI assistants. Those conversations happen inside a private chat, and getting access to them would mean reading people's private conversations. No responsible platform is going to hand that over, and no responsible vendor should ask for it.

There's a second complication. AI answers are nondeterministic. Ask the same question twice and you can get two different lists of brands. That isn't a flaw in the measurement; it's how the models work.

So AI visibility measurement behaves much more like audience polling than like rank tracking. You aren't looking up a fixed position. You're sampling a population of possible answers and estimating what share of them mention you. Everything else in this guide follows from that. A poll gets more accurate as you ask more people. A prompt panel gets more accurate as you track more prompts.

2. The Emerging Field of Prompt Tracking

Every organization tracking AI visibility is solving the same equation. More prompts, more platforms, and more locations produce a tighter estimate, and they all cost more to run. The real question is never "what is the perfect panel"; it's "what is the most accurate panel I can afford".

You can put numbers on that trade-off before you commit. The Spyglasses reliability calculator shows the margin of error you would get at a given prompt count, so you can see what an extra ten prompts actually buys you.

> **TRY IT**
>
> Reliability calculator: spyglasses.io/en/is-your-ai-visibility-data-reliable

The industry is settling on the same view. The IAB's **Measuring Visibility in the AI Era** guidelines treat AI visibility as a sampled measurement with a stated confidence level rather than a fixed position. AMEC's **GEO hub** makes a similar case from the communications measurement side, tying AI visibility back to established standards for evaluating earned media.

Spyglasses research points the same way. Rewording the same question changes which brands appear, so a single phrasing is a single data point and not the answer. Identical runs of the same prompt move the brand list around. And a generated panel only stands in for the real distribution of buyer questions when it is built deliberately.

Every organization balances accuracy against cost. The rest of this guide is about doing that on purpose instead of by accident.

3. How to Design a Prompt Panel

Start with your brand snapshot

The best starting point is an AI Visibility Report, and at minimum a Brand Snapshot. The snapshot captures the three facts every good prompt is built from: your **category**, your **target customer segments**, and the **problems you solve**. If those three fields are vague, your prompts will be vague. Fix the snapshot first.

Build prompts like Mad Libs

Once the snapshot is right, writing prompts is mostly a fill-in-the-blank exercise. You take a template and drop in the snapshot facts. A "best in category" prompt looks like this:

> *What is the best [description of category] for [description of target customer]?*

Fill it in and you get something a real buyer would type, like "What is the best appointment scheduling software for independent dental practices?" Every prompt type has a template like this, and Spyglasses generates the first set for you.

The minimum viable panel

Start with at least one prompt from each of these five types.

Prompt type	What it tests
Category Best	Whether you show up when someone asks for the best option in your category
Use Case Solution	Whether you show up when someone describes a problem instead of a category
Segment Focused	Whether you show up for the specific kind of buyer you sell to
Solution Comparison	Whether you make the shortlist when AI is asked to compare options
Budget Constrained	Whether you show up when price is the leading concern

If you don't sell to budget-conscious buyers, swap Budget Constrained for a type that fits your market better. A premium brand is usually better served by a second Segment Focused prompt.

How many prompts is enough

Five prompts gives you directional data. It tells you roughly where you stand and it catches large moves. More is better, and the improvement is steepest at the low end.

Even a small panel detects drift. You're watching your own share of voice and your competitors' share of voice over time, and when those lines cross or diverge, something changed in the market. That signal shows up long before the absolute numbers get precise.

Spyglasses puts the statistics in front of you rather than making you estimate them. When you assign prompts to a Project, the **Data Accuracy** card shows the margin of error at 90% confidence for your selected set, plus the smallest change you would be able to detect.

Rating	Margin of error	What it's good for
Reliable	Under 8%	Reporting numbers to a client or an executive team
Directional	8% to 15%	Watching trends and spotting real movement
Exploratory	Above 15%	Getting a first read on where you stand

Use the Coverage Matrix as your guide

Once you're past the first five prompts, the question becomes what to add next. The Coverage Matrix answers that by scoring your library across three dimensions:

1. **Buyer segment and subsegment.** Every distinct kind of buyer should have at least one prompt written from their point of view.

2. **Problems solved.** Every problem in your snapshot should appear in a prompt that describes the problem without naming a category.
3. **Buyer-journey stage.** Awareness, consideration, and decision, following the standard buyer's journey framework.

The matrix shows you the empty cells and prioritizes which to fill first. That's a better use of your next ten prompts than adding more variations of the ones you already have.

A note on awareness-stage prompts

Awareness prompts feel wrong the first time you read the answers. A buyer at that stage often doesn't know the market well enough to ask for a comparison. They ask about the problem, not the product, so the answer is an explanation rather than a list of brands.

Don't worry if your brand isn't in those answers. That's expected. Track whether your **key messages** appear instead. If AI explains the problem using the framing your marketing put into the world, your differentiators are already shaping the conversation at the earliest possible stage, which is worth more than a name-check. Spyglasses separates awareness prompts out of share of voice and reports them as **Share of Influence**.

Consideration and decision prompts

Consideration is where AI builds the shortlist. These are the "best X for Y" and "compare A and B" prompts, and share of voice matters most here because the answer is literally a list of who gets considered.

Decision prompts are narrower. The buyer has mostly made up their mind and is asking specific questions about your solution and their own circumstances. Does it integrate with the tools they already run? Does it meet their compliance requirements? Can it handle their volume? These prompts often surface a different set of competitors than the consideration prompts do, which is exactly why they're worth tracking separately.

Turn your set into a Project

Once your panel fits both your accuracy target and your budget, set it up as a Project. Projects run your prompts on a schedule with Daily Prompt Tracking, hold your baseline, and give you the trend lines that make the whole exercise useful.

4. Locations and Personalization

Tracking visibility in multiple locations

AI assistants personalize by location. Ask for the best accounting software in Berlin and you'll get a different answer than you would in Chicago, even with identical wording.

In Spyglasses you set a location for the brand, either a city or a country, and every platform is then tested with that location as context. If you sell into multiple markets, or you're launching in a new one, you can add more locations and track them side by side. Be deliberate here, because locations multiply.

> **WATCH YOUR RUN COUNT**
>
> Every prompt in a project runs once per location. 5 prompts across 4 locations is **20 prompt runs**, not 5. A panel that looked affordable in one market can quadruple when you add three more.

If budget is tight, it's usually better to run a smaller panel in each market you actually sell into than a large panel in the one market you happen to be headquartered in.

Tracking personalization in AI responses

AI assistants also personalize on memory. If someone has told an assistant they run a small nonprofit, later answers lean toward that context, and two people asking an identical question can get meaningfully different lists.

We can't simulate an individual's memory, and for the same privacy reason we can't see real user prompts. The best available approximation is to write prompts aimed at the different customer segments in your brand snapshot. A prompt that says "for a two-person law firm" and a prompt that says "for a 400-attorney firm" produce answers shaped by different assumed contexts, and comparing them tells you which segment you're winning.

5. Brand Prompts vs. Discovery Prompts

Keep your brand out of discovery prompts

Do not put your brand name, product names, or proprietary terms in prompts you intend to use for share of voice. A prompt that names you will almost always return you, and it tells you nothing about whether a buyer who has never heard of you would find you.

The test is simple. A discovery prompt should read like it was typed by a prospect researching a category with no preconception about who's in it.

When you do want to ask about your brand

There's a legitimate reason to ask AI about your brand directly, which is to check how it describes you. Spyglasses has two dedicated prompt types for exactly that.

- **Brand Identity** asks what your brand is, what it does, and whether it's any good.
- **Brand Comparison** puts you head to head with a named competitor.

Both are excluded from share of voice, mention rate, citation rate, and property-wide key message matching, so using them doesn't distort your discovery numbers. Both also carry their own expected messages, defined per prompt before the prompt ever runs. Positive messages are things that should appear in the answer; negative messages are things that should not. Spyglasses checks each daily answer against both lists and tracks how far the answer's themes have drifted from the trailing 28-day baseline.

Brand Comparison prompts add per-competitor verdicts: which brand the answer favors and for which segments, the strengths and weaknesses AI attributes to each brand, and whether the answer named a clear winner, gave segmented winners, merely listed the options, or reached no verdict at all.

Recommendation: at least one Brand Identity prompt, plus one Brand Comparison prompt for each key competitor.

6. Prompt Tracking Guidance, in Table Form

Use this as a starting checklist. Adjust the core count up as your accuracy target tightens.

Prompt group	How many	What to use
Core prompts	5 to 50, depending on the accuracy you need	Any of the Standard AI Visibility types; Local prompts (Local Services, Local Comparison, Local Reviews); Buyer's Journey; Jobs to Be Done; Stakeholder Perspectives; Category Entry Points
Brand Identity prompts	At least 1	Add more to test specific features, a launch, a rebrand, or M&A activity
Brand Comparison prompts	One per key competitor	Pick the competitors you actually lose deals to, not the whole category
Project-specific prompts	As many as the campaign needs	Prompts tied to a campaign, launch, or PR push. More prompts improve accuracy, but watch for overlapping prompts that don't test anything unique

On that last point, Spyglasses will tell you when you're duplicating coverage. The overlap checker compares a new prompt against everything already in your library before you save it, and redundant-cluster detection flags groups of prompts that all measure the same thing. Ten prompts that ask the same question in ten ways cost ten times as much and give you roughly one prompt's worth of information.

Where to Go Next

Full documentation for everything referenced in this guide:

- **Choosing the right prompts to track** (the web version of this guide)
  https://www.spyglasses.io/en/docs/getting-started/choosing-prompts
- **Brand Snapshot**
  https://www.spyglasses.io/en/docs/dashboards/brand-snapshot
- **Prompts dashboard**, including prompt types, framework generation, the Coverage Matrix, and the overlap checker
  https://www.spyglasses.io/en/docs/dashboards/discovery-queries
- **Brand Identity and Brand Comparison prompts**
  https://www.spyglasses.io/en/docs/dashboards/discovery-queries#brand-identity-and-brand-comparison-prompts
- **Projects**
  https://www.spyglasses.io/en/docs/dashboards/projects
- **Locations**
  https://www.spyglasses.io/en/docs/dashboards/locations
- **Key Messages**
  https://www.spyglasses.io/en/docs/dashboards/key-messages
- **Awareness tracking and Share of Influence**
  https://www.spyglasses.io/en/docs/ai-visibility-features/awareness-tracking
- **Brand Dashboard**
  https://www.spyglasses.io/en/docs/dashboards/brand-dashboard
- **Reliability calculator**
  https://www.spyglasses.io/en/is-your-ai-visibility-data-reliable

Industry guidance

- IAB, **Measuring Visibility in the AI Era**
  https://www.iab.com/guidelines/measuring-visibility-in-the-ai-era/
- AMEC, **GEO hub**
  https://amecorg.com/amec-geo-hub/
- Harvard Business School Online, **the buying journey**
  https://online.hbs.edu/blog/post/buying-journey

Spyglasses research

- Prompt phrasing consistency
  https://research.spyglasses.io/articles/prompt-phrasing-consistency

- Rank stability in ChatGPT
  https://research.spyglasses.io/articles/rank-stability-chatgpt
- Synthetic prompt coverage
  https://research.spyglasses.io/articles/synthetic-prompt-coverage

SPYGLASSES

Spyglasses is an AI visibility platform. It tracks how AI assistants describe and recommend your brand, tells you which sources they consult, and shows you what to change.

spyglasses.io