Brandella Journal

FREE SEO AND AEO GROWTH ASSET 19 OF 36

AEO Prompt Library Kit: Prompt Library, Test Log, Share of Voice and AI Prompts

Free AEO prompt library in Excel: test how ChatGPT, Gemini, Perplexity and AI Mode name your brand, track mention rates and share of voice, fix wrong facts.

Updated 26 November 2026
Online version and downloads: https://brandella.in/resources/seo-aeo-assets/aeo-prompt-library/
Free for personal and client work. By Santaji Gadepatil, Founder of Brandella Journal.

What is in this kit

Item	What it is
Excel workbook, 11 tabs	Prompt Library (500 ready rows with 19 automatic columns), Test Log (1,000 rows that find your brand, position and competitors by themselves), Share of Voice, Fix Plan, Prompt Builder, Dashboard with charts, Settings, AI Prompts, Checklist and Sources. Opens in Excel and Google Sheets.
Worked demo inside	20 prompts, 207 test runs across 4 AI engines and 6 fixes for a fictional help desk app and 5 fictional competitors, so you see every rate, status and result working.
Prompt Builder	20 editable prompt templates across 8 intent types and 6 personas that write ready prompts from your category, audience, location and competitor.
Full guide (PDF)	This complete guide as a printable PDF, to share with your team or clients.
22 AI prompts	Prompts to build the library, write persona variants, read answers into the log, find the sources behind wrong facts and plan fixes, including a 5 prompt chained workflow.
Printable checklist	The whole testing cycle on one page, to tick off every month.

Asking ChatGPT "what is the best tool for X?" once tells you almost nothing. Ask again and you get a different list. This free **AEO prompt library** turns that guesswork into a method. An Excel workbook holds the prompts real buyers ask, logs every run in every AI assistant you track, and works out how often you are named, how often you are named first or cited, whether the answers get your facts right, and your share of voice against 5 competitors.

It is built on what the research says about AI answers. A 2026 SparkToro study found less than a 1 in 100 chance that ChatGPT or Google's AI gives the same list of brands twice for the same prompt, and recommends measuring how often a brand appears across many runs instead of a rank. So every number in this kit is a rate over many runs, with a check that tells you when you do not have enough runs to trust it yet.

How to use this AEO prompt library kit

1. **Set up your brand and rivals.** On Settings, type your brand, the other ways people write it and up to 5 competitors, as AI answers write them. Set today's date.
2. **Build the prompt library.** Write 30 to 100 prompts from sales calls, support tickets, Search Console, Bing grounding queries and the Prompt Builder. Cover every intent type and funnel stage, branded and unbranded.
3. **Run each prompt the same way.** Use a fresh chat, the same wording and the same mode every time. Run each prompt at least 3 times in every engine you track, so 4 engines give 12 runs.

4. **Log every run.** On the Test Log, paste the brands each answer names, in order, separated by commas. Choose whether your URL was cited, the tone and whether your facts were right.
5. **Read the rates, not single answers.** The Prompt Library shows mention, named first and cited rates, tone, wrong answers, the top competitor and a status for every prompt. Share of Voice compares you with competitors by engine and intent.
6. **Fix the problems.** Wrong facts come first, then absent prompts with high value. Add each one to the Fix Plan with the source to fix, within your weekly hours.
7. **Retest every month.** Run the library again every 30 days. After a fix, add the mention rate before and after, or the number of wrong answers in the retest. The Result column reads it.

Why one AI answer tells you nothing

For a study published in January 2026, Rand Fishkin of SparkToro and Patrick O'Donnell of Gumshoe.ai asked 600 volunteers to run 12 brand recommendation prompts through ChatGPT, Claude and Google's AI, 2,961 runs in all. They found less than a 1 in 100 chance that ChatGPT or Google's AI gives the same list of brands in any two answers, and closer to 1 in 1,000 for the same list in the same order. Lists also changed length, from 2 or 3 names to 10 or more.

Their conclusion was not that tracking is pointless. A ranking position in AI is close to meaningless, they wrote, but how often a brand appears across many runs held up well. In one example a cancer hospital appeared in 97 percent of ChatGPT answers but was the first name in only 25 of them.

The same study had 142 people write their own prompt for the same need. Their prompts had a semantic similarity of just 0.081: almost no two people asked the same way. That is why this kit uses personas and many prompt variants, not one perfect prompt.

A later Conductor study of 14,000 API calls across ChatGPT, Perplexity, Claude and Gemini found consistency depends on the intent type. Brand overlap between runs was 63 percent for comparison prompts but 40 percent for purchase prompts. Conductor recommends tracking by intent type and engine separately, with unbranded recommendation prompts as a baseline. The workbook does exactly that.

What is inside the AEO prompt library workbook

Tab	What it does	What you type
Prompt Library	Spots branded prompts and rolls up every run: runs, engines, mention, named first and cited rates, tone, wrong answers, top competitor, confidence, status, next action, priority and retest date.	ID, prompt, topic, intent, funnel stage, persona, origin, value, target page, active
Test Log	Finds your brand and its position in the list, flags each competitor, numbers the runs and checks every row for typing mistakes.	Date, prompt ID, engine, mode, brands named in order, cited, tone, facts, what it got wrong
Share of Voice	Mention rates for you and 5 competitors by engine and by intent type, and share of voice with a chart. Unbranded prompts only.	Nothing
Fix Plan	Adds the prompt, current rate, suggested fix and hours, sets the measure date, reads the result and checks hours per week.	Prompt ID, problem, what the AI said, the correct fact, source to fix, owner, week, status, before and after
Prompt Builder	Writes 20 prompts from templates and 6 persona prompts, and flags which are already in the library.	Category, audience, location, competitor, use case; templates and personas if you like
Dashboard	20 key numbers, the top 10 prompts to work on, prompts by status and your mention rate by engine.	Nothing

Yellow cells are inputs. Grey cells under green headers are formulas. The library holds 500 prompts and the log 1,000 runs, all with dropdowns.

How to build a prompt library that matches real buyers

The best prompts come from your customers' own words, not from a brainstorm. Good places to find them: sales call notes, support tickets, the Search Console Performance report (long questions), People also ask, Reddit and forums, and the grounding queries in Bing Webmaster Tools. Then fill the gaps with the Prompt Builder.

Intent type	What it tests	Example (demo)	Starter share of the library
Recommendation	Are you suggested when someone asks what to choose?	What is the best help desk software for small teams?	About 30 percent
Comparison	How are you described next to rivals?	Compare the top shared inbox tools for email support.	About 10 percent
Alternatives	Do you appear as an alternative to a rival, and who appears as yours?	What are the best alternatives to HelpHive?	About 10 percent
Problem or how to	Are you suggested as the fix for a problem?	How can a small team manage support emails better?	About 15 percent
Pricing	Are prices and free plans right?	Is there a free help desk for small teams?	About 10 percent
Brand facts	Does the AI know what you are, do and cost?	How much does CalmDesk cost?	About 10 percent
Reviews and trust	Is the tone about you fair?	Is CalmDesk good? What do users say?	About 10 percent
Local	Are you named for your city or country?	Best help desk software companies in India	About 5 percent

The shares are a starting point, not a rule. A local business will want far more local prompts. Keep each prompt in plain words, the way a person would type or say it, and give every prompt the page on your site that should answer it.

How to run tests you can compare month to month

- **Fresh chat every run.** Memory and earlier messages change answers. Use a new chat, logged out where the engine allows it, and record the mode on the Test Log.
- **Same wording every time.** Copy the prompt from the library; never retype it.
- **Several runs per engine.** The kit asks for at least 10 runs per prompt before it trusts a rate (for example 3 runs in each of 4 engines). This is an assumption you can change. SparkToro suggests 60 to 100 runs for solid numbers, which is what paid tracking tools automate; a manual library gives you a useful direction, not a precise measurement.
- **Log the brands in order.** Paste the names exactly as the answer writes them, separated by commas. The sheet does the rest.
- **Same day for all engines.** Answers change over time, so test all engines for a prompt on the same day.

How the rates and statuses work (worked example)

In the demo, CalmDesk (a fictional help desk app in India) tests **"What is the best help desk software for small teams?"**, business value 3, unbranded. It was run 3 times in each of ChatGPT, Gemini, Perplexity and Google AI Mode in November: 12 runs.

- **Mention rate:** CalmDesk is named in 6 of 12 runs, so 50 percent.
- **Named first rate:** it is the first name in 1 run, so 8 percent.
- **Cited rate:** an answer links to calmdesk.example in 2 runs, so 17 percent.
- **Tone:** 2 positive and 4 neutral answers among the 6 that name it, so +0.33.
- **Top competitor:** HelpHive and TicketNest are named in all 12 runs.
- **Confidence:** 12 runs is above the 10 run minimum, so Enough runs.

50 percent sits between the 20 percent and 60 percent thresholds, so the status is **Mentioned sometimes: build proof**. The priority score has four parts: business value 35 x 3/3 = 35, visibility gap 30 x (1 minus 0.50) = 15, unbranded 15 x 1 = 15 and wrong answers 20 x 0 = 0. That is **65, P1: now**.

Second example. "How much does CalmDesk cost?" names CalmDesk in every run, but 5 of 12 answers quote an old price that still sits in an FAQ. Any wrong or partly wrong answer makes the status **Wrong facts: fix the source** and the priority **P1: fix facts**, so it tops the Dashboard. The FAQ was fixed on 12 November; the Fix Plan reads "Wait: 16 days left" until the 30 day retest.

Third example. "Is CalmDesk good? What do users say?" gets 4 negative, 1 positive and 7 neutral answers, a tone of minus 0.25, so the status is **Branded: negative tone**.

Fourth example. For "Best help desk software companies in India" CalmDesk got listed in two India software directories in October. The mention rate went from 8 percent to 33 percent, so the result reads **Up 25 points: keep going**.

Across all unbranded runs CalmDesk is named in 31 percent and holds a **15 percent share of voice**, behind TicketNest (21 percent) and HelpHive (20 percent). Perplexity names it most often (44 percent of unbranded runs) and Google AI Mode least (21 percent).

Column by column guide to the Test Log

Column	What goes in it	Tip
Date, Prompt ID, Engine, Mode	When, which prompt, which assistant and how you ran it.	Keep the mode the same every month.
Brands named in the answer	Every brand the answer names, in order, separated by commas.	Write names as on Settings; aliases are found too.
Your URL cited?, Cited URL or source	Yes when the answer links to one of your pages, and which.	Note competitor and third party sources too; they show where to get mentioned.
Tone about you	Positive, Neutral, Negative or Not mentioned.	Judge only what it says about you.
Facts about you, What it got wrong	Correct, Partly wrong, Wrong or Not mentioned, and the wrong claim.	Wrong facts drive the Fix Plan.
Prompt, Intent type, Branded prompt?	Automatic from the library.	Branded prompts are left out of share of voice.
You named?, Your position, Brands named	Automatic from the brands column.	Position is shown, but trust the rate more.
Competitor 1 to 5 named, Competitors named	Automatic 1 or 0 per competitor.	Feeds Share of Voice.
Run number, Data check	Automatic.	Fix any row where Data check is not OK.

Statuses, thresholds and what to do

Status	Rule	What to do
Too few runs: test more	Under 10 runs	Log more runs before acting
Wrong facts: fix the source	Any wrong or partly wrong answer	Fix the fact on your site and profiles, then retest
Branded: negative tone	Branded prompt, tone below 0	Find the source of complaints, fix the cause, publish proof
Branded: accurate, protect it	Branded prompt, no problems	Keep pages and profiles current
Strong: protect it	Named in 60 percent of runs or more	Keep the target page fresh
Mentioned sometimes: build proof	Named in 20 to 59 percent of runs	Strengthen the page and earn mentions on cited sources
Absent: competitors win	Under 20 percent, rivals named	Get on the sources the engines cite for this prompt
Absent: no brand named	Under 20 percent, no rivals named	Publish the clearest answer and become the source

The 10 run minimum and the 60 and 20 percent thresholds are kit assumptions on Settings. Priority P1 starts at a score of 65, P2 at 45 and P3 at 25. Wrong facts always come first.

Free platform data to use alongside the library

- **Bing Webmaster Tools.** The AI Performance report (February 2026) shows how often your pages are cited in Microsoft Copilot, Bing AI summaries and select partners, which pages, and a sample of the grounding queries the AI used. In June 2026 Bing added intents, topics, citation share and period comparison. Grounding queries make excellent prompts for your library.
- **Google Search Console.** The generative AI performance report shows impressions of links to your site in AI Overviews and AI Mode by page, country, device and date. Google says it reached all websites on 31 August 2026.
- **ChatGPT referrals.** OpenAI's publisher FAQ says any public website can appear in ChatGPT search if it does not block OAI-SearchBot, and that ChatGPT adds `utm_source=chatgpt.com` to referral links, so you can see that traffic in your analytics. GPTBot is a separate crawler for training.

None of these shows how assistants describe you in an answer, or which competitors they name instead. That is the gap the prompt library fills. Our guides on how ChatGPT chooses websites and how Perplexity finds and ranks sources explain where the answers come from.

How to fix what the tests find

Problem	Where to look	Typical fix
Wrong fact	Your own pages first (old FAQs, blog posts, PDFs), then profiles, listings and old articles	Correct every copy, show an updated date, retest the same prompt
Negative tone	Review sites, forums and old press the answers echo	Fix the cause, reply in public, publish proof such as numbers and case studies
Absent: competitors win	The sources the engines cite for that prompt	Earn a fair mention on those lists, directories and comparisons; publish your own honest comparison
Absent: no brand named	Nobody owns the answer yet	Write the clearest answer on the topic with your own data
Mentioned, not cited	Your target page and crawler access	Make the page crawlable and the most complete answer
Outdated information	Old pages that still rank	Update or redirect them

The hours for each fix on Settings are assumptions; change them to your team's pace. For the content side, see our guide on how to increase AI search citations and brand mentions.

Your first 30 days with this kit

- **Week 1: library.** Set up Settings, then write 30 to 50 prompts from real customer language and the Prompt Builder. Cover all 8 intent types.
- **Week 2: first test.** Run every prompt 3 times in each engine you track, in fresh chats, and log every run. Fix any Data check warnings.
- **Week 3: read and plan.** Read the statuses, Share of Voice and the Dashboard. Put every wrong fact and the top absent prompts on the Fix Plan.
- **Week 4: fix.** Correct wrong facts at the source first, then start on the highest value absent prompts.

After that, retest the whole library every 30 days and compare the rates.

The chained prompt workflow (from a buyer question to a logged run)

Prompts 18 to 22 below are designed to run in order in one AI chat, in a different assistant from the one you are testing:

1. **Brief the assistant** on your brand, competitors and rules.
2. **Turn customer language into prompts** across the 8 intent types.

3. **Write persona variants** so the library sounds like real people.
4. **Read a test answer** you paste in and pull out the brands in order, citations, tone and wrong facts.
5. **Log it** as a tab separated row for the Test Log.

Always check the extracted row against the answer itself before you paste it in.

AEO testing mistakes to avoid

- **Trusting one answer.** One run is a sample of one. Use rates over many runs.
- **Reporting a rank.** Position in an AI list changes almost every run. Report mention rate and share of voice.
- **Only branded prompts.** Prompts that already name you always mention you. Your real visibility is on unbranded prompts.
- **Changing the method.** New wording, logged in one month and out the next, or memory on: the numbers stop being comparable.
- **Ignoring wrong facts.** An AI quoting an old price to buyers costs sales. Fix it at the source.
- **Buying fake mentions.** Earn mentions on the sources engines already trust; do not plant them.
- **Treating tools as truth.** SparkToro advises asking any tracking vendor to show its method and math.

AI prompts

Replace the parts in [square brackets]. Always check the output against real data.

1. Turn customer words into prompts

```
Here are real messages from my customers, sales calls or support tickets: [paste]. Turn
them into 20 prompts a person
would type into ChatGPT or Gemini when looking for a product like mine. Keep their
words. Do not name my brand.
```

2. Fill gaps by intent type

```
Here is my prompt library with intent types: [paste]. My category is [category] for
[audience] in [location].
Write 3 more prompts for each intent type that has fewer than 3: recommendation,
comparison, alternatives,
problem or how to, pricing, brand facts, reviews and trust, local.
```

3. Persona variants

```
Rewrite this prompt in the voice of 5 different buyers, each with one line of real
context (team size, budget, tool they use now,
country): "[prompt]". Keep the core need the same. No brand names unless the original
has one.
```

4. Long conversational versions

```
People ask AI assistants long questions with constraints. Turn these short prompts into realistic 20 to 40 word versions
with a budget, a deadline or a must have feature: [paste prompts].
```

5. Find prompts in Bing grounding queries

```
Here are grounding queries from the AI Performance report in Bing Webmaster Tools: [paste]. Group them by intent type and
turn the 15 most useful into natural prompts for my library.
```

6. Read an answer into the Test Log

```
Here is an AI answer to the prompt "[prompt]": [paste answer]. List every brand it names, in order, separated by commas.
Then say whether it links to [my domain], its tone about [brand] (positive, neutral, negative or not mentioned) and any
claim about [brand] that I should check. Do not judge accuracy yourself.
```

7. List the sources an answer relies on

```
Here is an AI answer with its cited links: [paste]. List each source, what type it is (brand site, review site, directory,
news, forum, comparison article) and which brand it supports.
```

8. Why a competitor wins a prompt

```
For the prompt "[prompt]", answers name [competitor] far more often than [brand]. Here are the cited sources and what they
say: [paste]. Based only on this, give the likely reasons and 3 things [brand] could do about it.
```

9. Hunt down the source of a wrong fact

```
AI answers say this about [brand]: "[wrong claim]". The truth is: [correct fact]. List the kinds of places this old or wrong
claim usually comes from and a search I can run for each one to find copies of it.
```

10. Write a correction page section

```
Write a short, factual section for our [page] that states the correct [fact type]: [correct facts]. Make it easy to quote:
a clear heading, one direct sentence first, then the details and the date it changed.
```

11. Answer a negative review theme

```
AI answers say users complain that [complaint]. Here is what we changed and our proof: [paste]. Write a calm, specific
public reply and a short proof section for our site. No hype.
```

12. Honest comparison page outline

```
Outline an honest comparison page: [brand] vs [competitor] for [use case]. Include a summary table, who each one is best for,
pricing as of [date] and where the competitor is stronger. Mark every fact I must verify.
```

13. Get on the lists the engines cite

```
These list articles and directories are cited when AI assistants answer "[prompt]": [paste]. For each one, suggest a fair
way to be considered (update request, review, data, expert quote) and what proof I should send.
```

14. Unbranded answer page

```
No brand is named for the prompt "[prompt]". Outline the clearest page on the web for this question: a 40 to 60 word answer
first, then sections with our own data and examples from [brand]. Mark facts to verify.
```

15. Monthly change summary

```
Here are last month's and this month's mention rates by prompt and engine: [paste]. List the biggest gains and losses,
the prompts that changed status and what to check first.
```

16. Share of voice report for a manager

```
Write a short update for my manager. Data: [paste Dashboard and Share of Voice numbers]. Explain our mention rate, share of
voice against competitors, wrong facts found and fixed, and the plan for next month. No jargon.
```

17. Review the Fix Plan results

```
Here is my Fix Plan export with problem, source fixed, done date and rates before and after: [paste]. Which fixes moved
the rates? What should we do more of and what should we stop?
```

18. Workflow 1 of 5: brief the assistant

```
You are my AI visibility testing assistant. My brand: [brand], also written as [alias]. Competitors: [list].
Category: [category]. Audience: [audience]. Country: [country]. Rules: never invent facts, keep customer wording,
no brand names in unbranded prompts. Confirm, then wait for the next step.
```

19. Workflow 2 of 5: prompts from customer language

```
Here is real customer language: [paste]. Write 24 prompts, 3 for each intent type
(recommendation, comparison, alternatives,
problem or how to, pricing, brand facts, reviews and trust, local), as a table: ID,
Prompt, Intent type, Funnel stage.
```

20. Workflow 3 of 5: persona variants

```
For the 8 recommendation and problem prompts above, write one persona version each with
one line of real context,
the way a busy person would actually ask. Add them to the table with new IDs.
```

21. Workflow 4 of 5: read a test answer

```
I ran prompt [ID] in [engine]. Here is the answer: [paste]. Extract the brands named in
order, whether it cites [domain],
the tone about [brand] and any claim about [brand] I should check.
```

22. Workflow 5 of 5: log it

```
Turn that into one tab separated row for my Test Log: Date, Prompt ID, Engine, Mode,
Brands named in the answer,
Your URL cited?, Cited URL or source, Tone about you, Facts about you, What it got
wrong. Use [today] and "Logged out, fresh chat".
```

Checklist

- [ ] Typed your brand, its other spellings and up to 5 competitors on Settings
- [ ] Collected prompts from sales calls, support tickets and Search Console
- [ ] Added grounding queries from the Bing Webmaster Tools AI Performance report
- [ ] Covered all 8 intent types and every funnel stage
- [ ] Mixed branded and unbranded prompts
- [ ] Gave every prompt a business value and a target page
- [ ] Used a fresh chat, the same wording and the same mode for every run
- [ ] Ran each prompt at least 3 times in every engine you track
- [ ] Logged the brands named in order for every run
- [ ] Fixed every row where the Data check is not OK
- [ ] Read mention rates and share of voice on unbranded prompts only
- [ ] Added every wrong fact to the Fix Plan and fixed it at the source
- [ ] Planned fixes for the highest value absent prompts within your weekly hours
- [ ] Checked that OAI-SearchBot, Bingbot and Googlebot can reach your target pages
- [ ] Retested the library 30 days later and compared the rates

Frequently asked questions

Is this AEO prompt library really free?

Yes. The Excel workbook, PDF guide, prompts and checklist are free to download and use for your own brand and for clients, with no sign up.

What is an AEO prompt library?

A list of the questions real buyers ask AI assistants about your category and brand, run on a schedule in each assistant, with every answer logged. It shows how often you are named, how you are described and who is named instead.

How many prompts and runs do I need?

Start with 30 to 100 prompts. The kit asks for at least 10 runs per prompt, for example 3 in each of 4 engines, before it trusts a rate. Research suggests many more runs for precise numbers, so treat a manual library as a direction finder.

Why not just track my position in ChatGPT?

AI answers change almost every run. A 2026 SparkToro study found less than a 1 in 100 chance of the same list of brands twice. How often you are named across many runs is far more stable than your position in any one answer.

Which AI assistants should I test?

The ones your buyers use. The demo uses ChatGPT, Gemini, Perplexity and Google AI Mode, and the workbook also lists Google AI Overviews, Microsoft Copilot and Claude. Test the same engines every month.

Does the workbook work in Google Sheets?

Yes. Upload the .xlsx file to Google Drive and open it with Google Sheets. Formulas, dropdowns and colours carry over. The charts may need a quick refresh.

Sources

- SparkToro: AIs are highly inconsistent when recommending brands or products (January 2026)
  https://sparktoro.com/blog/new-research-ais-are-highly-inconsistent-when-recommending-brands-or-products-marketers-should-take-care-when-tracking-ai-visibility/
- Conductor: AI brand recommendation study, why intent type predicts AI output consistency
  https://www.conductor.com/academy/ai-recommendation-consistency-analysis/
- Bing Webmaster Blog: Introducing AI Performance in Bing Webmaster Tools (public preview)
  https://blogs.bing.com/webmaster/February-2026/Introducing-AI-Performance-in-Bing-Webmaster-Tools-Public-Preview
- Bing Search Blog: New AI visibility insights in Bing Webmaster Tools (intents, topics, citation share, compare)

https://blogs.bing.com/search/June-2026/New-AI-Visibility-Insights-in-Bing-Webmaster-Tools-Intents-Topics-Citation-Share-Compare

- OpenAI Help Center: Publishers and developers FAQ
  https://help.openai.com/en/articles/12627856-publishers-and-developers-faq
- Search Console Help: Generative AI performance report
  https://support.google.com/webmasters/answer/16984139

More free templates, checklists and prompts: https://brandella.in/resources/seo-aeo-assets/