The AI Search Visibility Audit Framework

About this framework

I designed this framework for the way I audit sites now, after running organic growth in-house for 18 years across fintech, iGaming, publishing, marketplace and hosting. It came out of a simple problem: The SEO audits I was running told me whether a page could rank and told me nothing about whether the brand turned up when an assistant answered a buying question.

So I built a second audit for that second question and this document is how I run it. Eight stages, with the tools, the checks and the output for each one, written so you can work through it with the document open beside you.

It extends the methodology I published in AI Search Optimisation 101, specifically the persona by journey prompt matrix, the source gap diagnosis and the layered approach to measurement. The walkthrough version, with the free Claude skill that automates part of it, is at oritmutznik.com.

What this AI search visibility audit does

Most SEO audits tell you whether a page can rank, while this one tells you whether your brand turns up when somebody asks an assistant a buying question, which brand gets named first, whether the description of you is accurate and which sources the answer was built from.

It works through 8 stages: You pick the commercial territories that matter, write the questions your buyers would genuinely type, check that AI crawlers can actually fetch your pages, run those questions across ChatGPT, Gemini, Perplexity, Claude, Copilot and Google's AI surfaces, then look at who the answers already trust. At the end you match every gap to one of 7 causes, hand each one to a named team and score it, so the whole thing lands as a roadmap instead of a report.

The free skill takes the mechanical half off your hands: It scopes the clusters, drafts the prompt set, runs the accessibility checks, pulls the cited domains and tests how assistants describe your brand, then writes the lot up as a Word document with a tracking spreadsheet. What it cannot do is query the assistants for you, so it builds your run sheet for that part and tells you plainly where it stops. Skip to the downloads if you would rather start there.

In short:

1. **What it is:** An 8-stage AI search visibility audit, worked through cluster by cluster
2. **The 8 stages:** Scope, prompt set, AI bot accessibility, visibility baseline, citation sources, entity clarity, diagnosis, then scorecard and roadmap
3. **The 4 movements:** Map, measure, diagnose, prioritise

4. **What makes it an audit:** The diagnosis stage, where every gap gets matched to a cause and an owner
5. **What you end up with:** A score per cluster, one number for leadership and a 90-day roadmap
6. **What it costs:** The core stages run free and Ahrefs covers the keyword and citation work from $29 a month
7. **How long it takes:** 2 to 3 days the first time, about a day after that

Why AI search visibility has to be audited per cluster

Two findings from the last year make this case better than I can.

1. **Platforms disagree with each other:** Ahrefs analysed the 50 most-mentioned websites across 76.7 million AI Overviews, 957,000 ChatGPT prompts and 953,500 Perplexity prompts. Seven sites appeared in all three lists, which is a 14% overlap. In a separate study, Ahrefs found 13.7% citation overlap between AI Mode and AI Overviews, which are two surfaces of the same Google product
2. **They disagree with the SERP too:** Another Ahrefs study found around 11% of citations across four assistants overlapped with Google and Bing's top-ten results for the same queries, while Semrush found Perplexity aligned most closely with Google's top ten and ChatGPT least of all

Put those together and one blended AI visibility number stops meaning anything, because you can be the default recommendation in Perplexity for one cluster while being absent from ChatGPT for the cluster next door. Report that number with the clusters visible underneath it and you have an audit. Report it alone and you have a dashboard.

AI search visibility audit vs SEO audit: what changes and what stays

The two audits share a foundation and diverge everywhere above it, so it is worth being precise about which is which, because the overlap is what makes people assume one covers the other.

	SEO audit	AI search visibility audit
Unit of analysis	The domain and its pages	The cluster, persona and platform
The question it answers	Can this page rank for this query?	Does this brand appear in this answer, in what order, described how?
Primary evidence	Crawl data, index coverage, rankings, backlinks	Prompt runs across platforms, cited sources, bot logs
Sources that decide the outcome	Mostly your own pages and the links to them	Mostly third-party pages you do not control
Who has to be able to read the page	Googlebot and Bingbot	Around 19 AI agents doing 3 different jobs
JavaScript	Google renders it, with delay	Most AI crawlers never run it

	SEO audit	AI search visibility audit
What success looks like	A position	Inclusion, order, an accurate description and an owned citation
Volatility to expect	Core updates, a few times a year	Citation drift of 40.5% to 59.3% in a single month
Attribution	Clicks and impressions in Search Console	Incomplete, with around 1% of traffic measurably attributable
What "fixed" looks like	The page ranks	The brand shows up correctly across several platforms

What genuinely carries over:

1. **Crawlability and indexation:** If nothing can fetch the page, neither audit has anything to work with
2. **Server-rendered content:** It mattered for SEO and it matters more here, since most AI crawlers skip JavaScript entirely
3. **Entity clarity and structured data:** Consistent naming, schema and third-party agreement help both
4. **Genuinely useful content with evidence in it:** Original data and first-hand experience earn rankings and citations alike
5. **Internal linking:** Both systems follow links to find what else you have

What has no equivalent in an SEO audit:

1. **An inaccurate mention counts as a finding:** Ranking has no concept of being described wrongly
2. **Third-party pages become your primary battleground** on commercial questions
3. **Platform-by-platform divergence**, where citation overlap between assistants runs as low as 14%, so one platform tells you very little about another
4. **Bot job separation**, since training, indexing and real-time fetching are 3 different activities with 3 different consequences when you block one

What you need to run an AI search visibility audit

Six things, 2 of which cost money: The core of this audit runs on the free four, so sort those first and treat the paid rows as an upgrade.

Worth saying plainly, because the free-tools angle gets oversold. The more you spend here, the better your output gets. Paid platforms run more prompts, across more platforms, far more often and they hold the history you cannot reconstruct after the fact. A manual monthly run tells you the direction of travel. A daily automated one tells you the week something moved and gives you the citations to work out why. Both are honest pieces of work and one of them has considerably more data behind it, which is simply how it is.

Before you commit to anything, run the trials: Profound offers a free trial on its self-serve plans, Peec AI has one on every tier and Rankscale gives you 7 days of its Pro plan with no charge until day 7. AccuRanker's AI tracking is demo-led at the time of writing, so ask them directly. A fortnight of trials across two of these will

also show you exactly how differently they each count a mention, which is the point I made above and much easier to see than to explain.

Tool	What it does	Cost	Worth knowing
The AI assistants	ChatGPT, Google AI Overviews, Google AI Mode, Gemini, Perplexity, Claude, Copilot. You run your prompts in each one	Free tiers work	Run them logged out or in a private window, because personalisation and memory contaminate the baseline
A spreadsheet	Holds the prompt set, the 6 recording fields and the scorecard	Free	The prompt sheet is the real deliverable of this audit. Everything else derives from it
Screaming Frog	Crawls the site, compares raw HTML against rendered HTML, checks status codes and schema	Free to 500 URLs	The JavaScript rendering comparison does most of the accessibility stage in one crawl
Ahrefs, Lite or above	Keyword expansion for the prompt set, `serp_features` for AI Overview presence, Brand Radar for cited domains	From $29 a month	Brand Radar sits on higher tiers. Without it, the citation source stage runs off the citations you logged by hand
Server or CDN logs	Splits AI bot activity into training, indexing and real-time fetches	Free if you can get an export	The hardest thing to obtain and the most valuable thing in the audit. Ask for it in week one
Profound, Peec AI, Rankscale or AccuRanker AccuLLM	Automates the visibility baseline across platforms, daily, with stored history	Paid, most with a free trial	They give you cadence and coverage. What you record stays the same either way

One warning about that last row: Every AI visibility platform defines a mention, a citation and an owned link slightly differently, so two tools will hand you two numbers for the same brand in the same week and both will be right inside their own definition. Ask a vendor which of the three they count before you compare their figure to anyone else's.

The 8 stages of an AI search visibility audit

Here are the 8 stages and what each one produces.

1. **Scope:** Choose the clusters, personas, markets and competitors the audit covers
2. **Prompt set:** Write the 25 to 50 questions your buyers would actually type, per market
3. **AI bot accessibility:** Check that assistants can fetch and read the pages in the first place
4. **Visibility baseline:** Run every prompt on every platform and record what comes back
5. **Citation sources:** Work out which domains and formats the answers already trust
6. **Entity clarity:** Check that the systems understand who you are and what you sell
7. **Diagnosis:** Match every gap to one of 7 causes and give it an owner
8. **Scorecard and roadmap:** Score each cluster, roll it up to one number and sequence the fixes

Those 8 group into 4 movements and each movement answers one question.

1. **Map**, covering scope and the prompt set, answers where you should appear
2. **Measure**, covering accessibility, the baseline, citation sources and entity clarity, answers where you actually appear and whether you can be retrieved at all
3. **Diagnose**, which is the diagnosis stage on its own, answers why the gap exists in each cluster and who owns the fix
4. **Prioritise**, which is the scorecard, answers what to do first and what to report upwards

Accessibility sits at stage 3 for a reason: A page an assistant cannot fetch will never be cited, whatever else you fix.

Stages 1 to 6 collect information and the diagnosis stage turns that information into findings, which is where most published AEO audits stop short: They present the baseline as though it were the conclusion, when a baseline is only the raw material a conclusion gets built from.

Two rules hold across all 8 stages:

1. **Work cluster by cluster:** Pick 3 to 5 clusters and carry each one through every stage, so every finding stays attached to a commercial territory
2. **Keep the prompt set fixed between runs:** Profound calls the movement citation drift and in a study of around 80,000 prompts per platform it found 40.5% to 59.3% of cited domains changed over a single month, rising to 70% to 90% comparing January with July 2025. Ahrefs saw similar volatility tracking 43,000 AI Overview keywords for a month. Change your prompts between runs and you lose the ability to separate your own work from that background movement

---

Stage 1: How to scope an AI search visibility audit

Tools: Ahrefs Site Explorer for the top pages and keywords, Google Search Console for what already converts, plus whatever holds your sales conversations: HubSpot, Salesforce, Gong or the support inbox.

Decide 4 things before you write a single prompt:

1. **Clusters:** Pick 3 to 5 commercial territories and name them for what they are, so "payroll software" earns a place where "cheap payroll tool" does not. Pull your top pages by traffic and revenue in Ahrefs or Search Console, then group them into territories the business would recognise. If only 3 matter commercially, audit 3
2. **Personas:** Choose 2 to 4, grounded in CRM records, recorded sales calls, support tickets, site search logs or win-loss notes. The test I apply: If the persona changes none of the questions, it is a demographic. A technical evaluator and a budget owner ask genuinely different things about the same product, while a "35-44 urban professional" asks nothing in particular
3. **Markets:** List the countries and languages in scope, since each one is a separate audit surface, because assistants behave differently from a localised SERP

4. **Competitors:** Build 2 lists deliberately: Commercial competitors, meaning who the business loses deals to and answer competitors, meaning whoever shows up in the answers. Stage 5 will tell you how far apart those lists sit and in my experience the distance between them is where the interesting work is

> **Output:** A one-page scope note listing clusters, personas, markets, both competitor lists and the single business question this audit has to answer.

Stage 2: How to build an AI search prompt set

Tools: Ahrefs Keywords Explorer for matching terms and questions, G-Trendalyser for up to 250 rising and related queries at a time, AlsoAsked for People Also Ask chains and FanoutFox for the fan-out itself.

FanoutFox deserves a line of its own: It is a free Chrome extension built by Suganthan Mohanadasan that shows the search queries ChatGPT actually runs behind an answer, which pages it fetched and whether your page was cited, mentioned or only fetched. That last distinction takes field 4 of the visibility baseline from an educated guess to something you can read off the screen and it costs nothing.

This stage is the keyword research of AI search and the spine of everything downstream: Get it wrong and the next 6 stages will measure the wrong thing very precisely.

Build a persona by journey matrix for each cluster, covering 5 journey stages:

1. Problem discovery
2. Category education
3. Brand and product comparison
4. Objections, limitations and risk
5. Buying and post-purchase

Then apply 4 rules as you write them:

1. **Use the persona's own language and their real constraints:** "Best payroll software" is a keyword. "Which payroll software works for a 40-person agency paying contractors in 3 countries on under £200 a month" is a prompt and it is the version that reveals whether your positioning survives contact with a constraint
2. **Write the same cluster for each persona:** A budget owner asks which options fit their company size and budget and what evidence backs the recommendation. A technical evaluator asks you to compare options on integration effort and limitations
3. **Add a follow-up question to every prompt**, because assistants are conversational and the second question is usually where the shortlist forms
4. **Hold the volume at 25 to 50 core prompts per market:** Below 25 no pattern is visible and above 50 nobody keeps running it, which makes an abandoned prompt set worth less than a small one you maintain

Then expand for fan-out, since assistants break one question into many sub-queries and consistency across that surface matters more than perfecting a single page. Pull from 4 places:

1. Rising and related queries per head term, in G-Trendalyser or Google Trends
2. Matching terms and questions in Ahrefs Keywords Explorer
3. People Also Ask chains in AlsoAsked
4. Your 3 most commercial prompts run through FanoutFox, which shows the real sub-queries in place of your guess at them

The query fan-out entry in my AI search glossary has the longer definition if you want it.

One caution, which Lily Ray has made well: Resist turning every fan-out query into its own page. Cluster the repeated themes and strengthen the best destination you already own. Treat the fan-out list as research input.

> **Output:** The prompt sheet. One row per prompt, with columns for cluster, persona, journey stage, market, prompt text, follow-up, date added and whether it is active or retired.

Stage 3: How to run an AI bot accessibility audit

Tools: Screaming Frog for the crawl and the rendering comparison, Terminal or Command Prompt for the per-bot fetch test, Ahrefs Site Audit or Sitebulb for the wider technical sweep and JetOctopus, Botify or Lumar if you can get a log export.

Two words to define before this section makes sense:

1. **A user agent** is the name a visitor gives when it requests a page. Your browser says it is Chrome, OpenAI's crawler says it is GPTBot and servers read that name and can treat each one differently
2. **A WAF** is a web application firewall, the security layer sitting in front of your site that blocks traffic it judges suspicious. It can block AI crawlers without anyone telling the SEO team, which is why this stage exists

Accessibility comes before measurement because it caps everything downstream: A page an assistant cannot fetch will never be cited, whatever the content says.

Work through the 4 checks below.

3a. Check what robots.txt says, then check what your server actually does

First, read the file: Go to `yourdomain.com/robots.txt` in a browser and look for rules naming these agents:

`GPTBot`, `OAI-SearchBot`, `ChatGPT-User`, `ClaudeBot`, `Claude-User`, `PerplexityBot`, `Google-Extended`, `Bytespider`, `Amazonbot`, `meta-externalagent`, `Bingbot`

Write down whether each one is allowed, disallowed or unmentioned and treat unmentioned as allowed.

Second, test what actually happens, because the file states an intention and the server states the truth. Here is how to run that test with no prior command line experience.

On a Mac: Open Spotlight with Command and Space, type Terminal, then press Enter. **On Windows:** Press the Windows key, type PowerShell, then press Enter.

Paste this line, swapping in your own URL, then press Enter:

```
curl -A "GPTBot" -o /dev/null -s -w "%{http_code}\n" https://example.com/pricing
```

Reading that command left to right: `curl` fetches a web page the way a bot does, with no browser and no JavaScript. `-A "GPTBot"` tells the server you are GPTBot, `-o /dev/null` throws away the page content, since you only want the response code and `-s` hides the progress bar. `-w "%{http_code}"` prints the response code on its own.

You get back a 3-digit number:

1. **200** means the page loaded: That bot can reach it
2. **403** means forbidden: Something is blocking that bot, usually a WAF or a CDN rule and this is a critical finding
3. **429** means too many requests: The server is rate limiting that bot, which slows retrieval down
4. **301 or 302** means a redirect: Follow it and check where it lands
5. **404** means the page is missing for that agent even when it loads in your browser

Run it for 5 to 10 important URLs per cluster, swapping the agent name each time. Anything returning 403 for a bot you meant to allow goes straight to the top of your roadmap.

No terminal, no problem: Install the User-Agent Switcher extension in Chrome, add GPTBot and the others as custom agents, then load your pages while pretending to be each one. Screaming Frog can do the same thing under Configuration, then User-Agent, where you pick or type a custom agent and crawl as that bot.

3b. Check whether your pages work without JavaScript

Most AI crawlers do not run JavaScript, so anything that only appears after a script executes is invisible to them. You are checking this per template rather than per page, meaning one product page, one category page, one blog post, one pricing page. Pages built from the same template all behave the same way.

The 2-minute browser check, which needs no tools at all:

1. Open one of your pages in Chrome and press Command-Option-U on a Mac or Control-U on Windows. That shows the raw HTML the server sent
2. Press Command-F or Control-F and search for a sentence you can see on the live page

3. If you find it, that content is server-rendered and crawlers can read it. If you cannot find it, JavaScript is adding it after load and most AI crawlers will never see it

The Screaming Frog version, which covers the whole site:

1. Open Screaming Frog, go to Configuration, then Spider, then the Rendering tab
2. Choose **Text Only**, crawl your site, then export the Internal HTML report
3. Go back to the same setting, choose **JavaScript**, crawl again, then export that crawl too
4. Compare the Word Count column between the 2 exports, template by template

Where the JavaScript crawl shows far more words than the text-only crawl, that gap is content AI crawlers cannot see. A page showing 120 words in text-only mode and 1,400 with JavaScript enabled has a serious problem.

In one bot log audit I ran, a developer documentation subdomain recorded 69 AI citations across 30 days against 220,707 for the main domain, a ratio of 0.03%. The docs were a client-side rendered single-page app behind a hash fragment URL, so every assistant that fetched it received an empty shell. The content itself was excellent and nothing could read it.

3c. Split your bot logs into the 3 jobs they represent

Server logs are the most valuable dataset in this audit and the hardest to get, so ask for them in week one.

How to ask: Email whoever runs your hosting, infrastructure or DevOps and request the raw access logs for the last 30 days, including user agent, requested URL, response code and timestamp. If your site sits behind Cloudflare, the same data lives under Analytics, then Security, then Bots. If nobody can produce a log file, say so in the audit and move on, since the other 3 checks still stand.

What to do with them: Upload the export to JetOctopus, Botify or Lumar, all of which segment by user agent for you. For a smaller site, a spreadsheet pivot by user agent and URL does the job.

Then split the AI bot traffic into 3 activities, because most dashboards add them into one "AI bots" line:

1. **Training crawls**, building the model's background knowledge
2. **Indexing visits**, building a corpus that can be retrieved later
3. **Real-time fetches**, answering a question somebody is asking right now

Verify each bot by reverse DNS, because the user-agent string is trivial to spoof. Any decent log tool has a verification setting for this.

The split matters because providers behave very differently: In the same log export, OpenAI's bots generated 2.6 million real-time fetches while Amazon and ByteDance generated none at all, with ByteDance contributing 1.1 million training visits and zero indexing. Bing sent a million indexing visits, which is what feeds Copilot. So blocking one provider has an entirely different consequence depending on which of the 3 jobs it was doing for you.

Then use the real-time numbers as a demand signal: Assistant fetches landing on a cluster's pages are direct evidence that prompts in that cluster are triggering retrieval right now.

3d. Find the crawl budget going nowhere

The same log export also shows where AI bots waste their requests and these are the 4 patterns I find most often, all of which turned up in that same export:

1. **Tracking and affiliate redirect URLs:** One set of internal redirect paths pulled roughly 861,000 bot visits, so bots were training on redirect infrastructure with no content in it. The fix is a robots.txt disallow
2. **Discontinued product pages:** A retired product page ranked third most-crawled on the entire site at 592,105 visits. The fix is a 301 to the current equivalent, which also recovers the link equity a disallow would strand
3. **Dynamic search results pages:** One returned 161,165 assistant citations and 160,235 training visits against only 871 indexing visits, which is the signature of a page bots can store nothing useful from. The fix is a disallow plus a noindex
4. **Pagination duplicates splitting a signal:** A pricing page and its `/pricing/1` variant were crawled separately at 35,691 and 22,327 citations, dividing one signal in two. The fix is a canonical tag

What about llms.txt?

Check whether you have one at `yourdomain.com/llms.txt`, then keep it in proportion. I went through the evidence in AI Search Optimisation 101 and the short version is that Google says it ignores the file outright, so it neither helps nor harms you in Google Search and in the bot logs I have looked at the file barely gets requested at all.

If you do not have one, leave it: Nothing in this audit depends on it and the 4 checks above will move your visibility considerably further than adding one.

If you already have one, check its size: On one site the file returned a healthy 200 response, ran to 297,000 characters, far past the point where most models truncate and never appeared in the top 200 crawled paths across the period. A file that exists is not a file that gets read.

> **Output:** An accessibility table per cluster, showing each key URL, its robots policy, the response code returned per bot, whether the content survives without JavaScript and what the logs say about it. Every failure here goes to the top of the roadmap.

Stage 4: How to measure AI search visibility across platforms

Tools: The assistants themselves, run by hand against your prompt sheet. Ahrefs Keywords Explorer with `serp_features` selected shows AI Overview presence on terms you already rank for and Profound, Peec AI, AccuRanker's AccuLLM or Rankscale automate the whole stage.

Run every core prompt on every platform in scope, logged out or in a private window and date every run.

Record 6 fields for each prompt on each platform:

1. Whether an AI answer appears at all
2. Which brands are included and in what order
3. How each brand is described and whether that description is accurate
4. Which pages and domains are cited
5. Whether each cited source is owned, earned, community-led or commercial
6. Whether anything in the answer is wrong or out of date

Field 3 gets skipped most often, though it is where the money frequently sits. Being mentioned inaccurately does you no good: "A budget option for small sites" attached to an enterprise product is a positioning failure created by a retrieval system and no volume of new content corrects it while the description comes from a third-party page nobody on your side has read.

For cadence, run the full core set monthly and a 10-prompt sentinel subset weekly. By hand that costs roughly 2 hours a month, while the paid platforms take it to daily and store the history, which is their real value.

Two shortcuts are worth taking whatever your stack looks like:

1. Pull your top organic keywords in Ahrefs with `serp_features` included, since every keyword returning `ai_overview` shows where Google's AI layer already sits on terms you rank for, at no extra cost on a call you were making anyway. Be clear on its limit: It tells you an AI Overview appears, which differs from telling you that you are cited inside it
2. Run your 3 most commercial prompts through FanoutFox, which shows the sub-queries ChatGPT ran and whether your page was cited, mentioned or only fetched

> **Output:** A visibility matrix per cluster, covering presence rate on each platform, your average position among the brands mentioned, description accuracy and the share of prompts citing at least one page you own.

Stage 5: How to audit AI citation sources

Tools: The citations you logged in the visibility baseline, Ahrefs Brand Radar for cited domains and cited pages where your plan includes it and Profound's Answer Engine Insights if you have it.

Turn the question round here, because who the answer trusts is usually more actionable than whether it mentions you.

For each cluster, aggregate every cited domain from your baseline runs, rank them by how often they appear, then label each one as owned, earned, community, commercial or competitor-owned.

The pattern is the finding: Profound's analysis of 680 million citations found ChatGPT cited Wikipedia most often across its August 2024 to June 2025 dataset, while Reddit led for both Google AI Overviews and Perplexity. Those preferences shift over time, so treat the durable conclusion as platforms differing from each other, then read your own cluster data for the local version.

Study the winning formats next, since the format is often the barrier. Are the cited pages comparisons, original research, documentation, tools, videos or third-party lists? A cluster that consistently cites research will keep citing research however many product pages you publish, which is obvious written down and still the most common recommendation in the AEO audits I read.

Then audit the shortlists you do not own: For commercial prompts, the answer is usually assembled out of third-party "best of" and comparison pages, which makes 3 questions worth answering per cluster.

1. Which of those pages get cited most often?
2. Does your brand appear on them?
3. Is what they say about you still true?

I have spent years on both sides of this: I have run the comparison pages that assistants now quote and I have been on the brand side trying to get into them. The thing that surprises in-house teams is how much of their AI visibility problem sits on somebody else's website, in a table last updated eighteen months ago, next to a price that has changed twice since.

> **Output:** For each cluster, the top 10 cited domains with their type, the dominant winning format and the 3 third-party pages where getting listed or corrected would move the most.

Stage 6: How to check entity clarity in AI search

Tools: The assistants in clean sessions, Google's Rich Results Test, the Schema.org validator, Wikidata and a manual brand SERP review.

When a system misunderstands who you are, every other signal lands on the wrong entity, so run this before spending anything on content.

Work through 4 checks:

1. **Ask each assistant cold**, in a fresh session with memory off: Who is [brand], what does [brand] do, what is it known for and how does it compare to [main competitor]. Record what it says, what it gets wrong and what it cites, because that citation column is the actionable one. A wrong description almost always traces back to a specific source you can go and correct
2. **Validate your schema** with the Rich Results Test and the Schema.org validator. Check that Organization, Product and Person markup exists, resolves and agrees with your copy
3. **Check consistency across the web:** Your name, category description and positioning should match on your site, LinkedIn, Crunchbase, Wikidata and the main industry directories. Assistants reconcile these, so contradictions between them become hedged or wrong answers

4. **Read your brand SERP:** It is the public record of the entity and anything contradictory sitting on page one contradicts you everywhere at once

> **Output:** An entity accuracy table listing each claim, what each assistant says about it, whether that is correct, the likely source and where the fix has to be made.

Stage 7: How to diagnose why you are missing from AI answers

Tools: None needed, because this stage runs on judgement and it is what separates an audit from a report.

Take every gap from stages 3 to 6, the four measurement stages and match it to one of 7 causes. The list comes from my AI search optimisation guide, where I set it out as a diagnosis step. Each one below includes how to confirm it is yours.

1. **Access:** The useful page cannot be retrieved. *Diagnose it:* Stage 3a returned a 403, 429 or a redirect chain for at least one AI user agent, or stage 3b showed the content appearing only after JavaScript runs. Owner: engineering
2. **Relevance:** Content exists and answers a different question from the one being asked. *Diagnose it:* Your page is cited for informational prompts in a cluster and absent from the commercial ones, or the prompts it wins do not match its stated purpose. Owner: content and SEO
3. **Evidence:** Claims carry nothing worth quoting. *Diagnose it:* Compare your page against the cited pages from the citation source stage and count original numbers, named sources, first-hand testing and dates. If the cited pages have them and yours does not, that is your cause. Owner: content and product marketing
4. **Entity clarity:** The system misunderstands the brand, product or category. *Diagnose it:* Stage 6 produced inconsistent or wrong answers to "who is", or your schema, LinkedIn and Wikidata descriptions disagree with each other. Owner: SEO and brand
5. **Format:** Competing sources present the same information more usefully. *Diagnose it:* Stage 5 shows one dominant format in the cluster, comparison tables for instance and your best page is a different format entirely. Owner: content and design
6. **External context:** Publishers, affiliates, reviews and communities tell an incomplete story. *Diagnose it:* Stage 5's cited domains are mostly third party and you are either missing from those pages or described inaccurately on them. Owner: partnerships, PR and community
7. **Freshness:** Prices, specs or claims are out of date, on your pages or on theirs. *Diagnose it:* Compare the numbers in the AI answer against your current pricing and specification. Where they differ, trace the citation back to the page carrying the stale figure. Owner: content operations

Three rules govern this stage:

1. **Give every gap a cause and an owner:** A finding with no owner is a complaint with a table around it
2. **Notice that one cause out of 7 is solved by publishing more content.** Publishing more content gets commissioned almost every time, so this diagnosis mostly exists to interrupt that reflex

3. **Rescope if everything fires:** All 7 causes appearing on every cluster means the scope was too broad, so return to the scope stage before pressing on

> **Output:** A diagnosis table with one row per gap: Cluster, gap, cause, owner, the fix, effort as small, medium or large and which visibility baseline metric it should move.

Stage 8: How to score AI search visibility and build the roadmap

Tools: A spreadsheet is all this stage needs.

Score each cluster out of 100 using 5 weighted components:

1. **Presence, 30%:** The share of core prompts where your brand appears, from stage 4
2. **Prominence and accuracy, 20%:** Your position among the brands mentioned and whether the description is correct, from stage 4
3. **Owned citations, 20%:** The share of prompts citing at least one page you own, from stage 4
4. **Third-party strength, 20%:** Your presence and accuracy in the cited shortlists, from stage 5
5. **Access, 10%:** How retrievable the cluster's key pages are, from stage 3

One design decision matters more than the weights themselves: Cap every component at 100% of its own target. Without that cap, a component that massively overshoots hides one that is failing and the single number stops being honest. I built a weighted AI visibility model on exactly this principle for a leadership team and the cap is the part that made it survive scrutiny. Adjust the weights to your business, write down the reasoning once, then leave them alone.

Roll the clusters up to one number per market and one overall and that becomes the figure for the board slide.

Put the caveat in writing, so it survives the meeting: Prompts are sampled, citations fluctuate and attribution is incomplete. Conductor's 2026 benchmark put measurable AI referrals at 1.08% of total website traffic across 10 industries, ranging from 2.80% in Information Technology to 0.25% in Communication Services, while Ahrefs' June 2026 panel saw 0.33% of total visits. Treat both figures as floors, because SearchPilot's review found the same platform landing in Referral, Organic, Unassigned or Direct depending on the device and interface. Aleyda Solis makes the same point about measured referrals being the floor of AI's contribution, then layers presence, retrieval and wider influence on top.

An honest directional number will serve you better than a precise imaginary one, so say exactly that in the deck.

Then sequence the roadmap by impact over effort, with 2 overrides:

1. **Access fixes go first**, whatever they score, because they cap everything else
2. **Freshness corrections on heavily-cited third-party pages go second**, since they deliver the most per hour spent in the whole framework and usually take an email rather than a project

> **Output:** One page carrying the per-cluster scores, the roll-up number, the 90-day roadmap and the caveat.

How often to run an AI search visibility audit

Activity	Frequency
Sentinel subset, 10 prompts	Weekly
Full core prompt set	Monthly
Scorecard refresh	Monthly
Full audit, stages 3 to 8	Quarterly
Scope review	Quarterly, after the audit

One practical note: When a major model ships, mark the date on every trend line. Baselines reset and the step change that follows belongs to the model release rather than to your work, in either direction.

Why the diagnosis stage decides whether the audit was worth running

> Stages 1 to 6 produce information and stage 8 produces a number. Stage 7 is the only stage that produces a decision and it is the one that gets skipped.

Diagnosis is what separates an audit from a report, so protect the time for it. A report tells a business what is happening, while an audit tells it what to do, in what order and which team has to do it. Most AI search work stalls on the first because the second requires somebody to hold a view and put their name against it.

Two things make the diagnosis stage straightforward to run: Every cause in the list has a written test attached, so you are checking evidence you already collected in stages 3 to 6 instead of forming an opinion from scratch. And every cause has a default owner, so the conversation about who does the work happens while the finding is fresh.

If your diagnosis lands 5 of the 7 causes on content, go back through the accessibility stage before you believe it. In every audit I have run, access and external context came out larger than anyone expected and content came out smaller.

Appendix: prompt sheet fields

One row per prompt, per platform, per run.

Field	Values
Cluster	From the scope stage
Persona	From the scope stage
Journey stage	Discovery, education, comparison, objections, or purchase and after
Market	Country and language
Prompt	The exact text, in the persona's words
Follow-up	The second question a real user would ask
Platform	AI Overviews, AI Mode, ChatGPT, Gemini, Perplexity, Claude or Copilot
Answer present	Yes or no
Brand present	Yes or no, plus position among the brands mentioned
Description	Verbatim, plus whether it is accurate
Citations	The domains and URLs cited
Source types	Owned, earned, community, commercial or competitor-owned
Errors	Anything wrong or out of date in the answer
Run date	The date of the run
Model note	Any known model or interface change since the last run

---

This framework is free to use and share, with attribution. If you want a second pair of eyes on the AI search side of a domain, or a hand turning a process you repeat into something you can hand to a machine, get in touch at oritmutznik.com.