

Tau Field Guide · September 2026

The AEO Playbook

What makes an answer engine cite your page — and the nine-skill procedure for finding out whether yours qualifies.

Retrievable → Quotable → Corroborated · nine skills · one output contract

Almost nothing about how assistants choose their citations is documented. This guide separates what providers publish from what is only inferred, labels every claim, and lists every source it used. Where the honest answer is *nobody knows*, it says so.

CONTENTS

What is in this guide

01	How an answer gets assembled Two paths, and which one you can influence	04
02	The three conditions Why the order is not negotiable	05
03	The question set What the page is for, before anything else	06
04	Condition 1 — retrievable Access, discovery, and the fetch that tells the truth	07
05	Crawlers, by provider Three jobs, separately controlled	08
06	Preview and use controls How much of the page may be used	11
07	Condition 2 — quotable The seven rules for a passage that survives alone	12
08	Claims and evidence Four labels, and the ones that get deleted	14
09	Condition 3 — corroborated The one you cannot write your way into	15
10	Where this departs from ranking Five places the SEO instinct misleads	16
11	The pre-publish checklist Six checks with answers, not judgements	17
12	Measuring without rank tracking Four sources, and what each is blind to	18
13	The observation panel Sampling a non-deterministic system, honestly	19
14	Uncertainty register and output contract What this guide will not claim	20
15	The nine skills What is in the pack and how it chains	21
16	Sources Every document this guide was built from	22

The honest framing, first

Search ranking is not published either, but it has decades of public guidance and outside study behind it. Citation selection inside an assistant is younger, thinner on documentation, and changes with every model release. Almost everything here that is not a mechanism of how these systems are built is labelled as an observation — and observations in this field expire faster than they do in search.

HOW TO READ THIS

Four labels, and they travel with the claim

Every material statement in this guide carries one of four labels. The label is not decoration. It tells you how long the statement is likely to remain true and how much weight it will bear in an argument.

LABEL	WHAT IT MEANS	EXAMPLE FROM THIS GUIDE
Documented	A provider publishes it. Verifiable today, and still liable to change with the next release.	A page must be indexed and eligible to be shown with a snippet to appear as a supporting link in Google's AI Overviews or AI Mode.
Structural	It follows from how the systems are built rather than from any announcement.	Retrieval operates over passages, and a source that was never retrieved cannot be cited.
Observed	Consistent with practice and with the mechanism; not confirmed by any provider.	Clear, self-contained, well-sourced passages tend to be quoted more readily.
Unverified	Any share, multiplier or ranking factor about a system whose internals are private. Nobody outside these companies can measure it.	Anything of the form "AI search now drives <i>n</i> % of discovery in your category".

A number attached to a private ranking system is a sales device rather than a finding

The rule this guide applies to itself, and the one the pack's claim-and-evidence-check skill applies to your pages, is simple: **unverified claims are removed, not softened**. "Some studies suggest" attached to an untraceable figure is worse than the figure alone, because it performs rigour without doing any.

What this guide cannot do

Nothing here can make an assistant cite you. What it can do is remove the reasons not to, in the order that matters, and give you a record of what actually happened instead of a number that sounds like coverage. That is the whole of the offer. Anyone promising more is selling something.

Sources for this page. Google, *AI features and your website*; the labelling scheme is Tau's, and matches the uncertainty register in the AEO skill spec at get-tau.com/resources/aeo-skill-spec.

01 · MECHANICS

How an answer gets assembled

An assistant answers from two different places, and the difference decides what you can influence.

PATH	WHAT IT IS
What the model learned	Everything absorbed during training, held as parameters rather than as documents. There are no citations here and no live updates: the model produces what its training made likely, up to its cutoff. What the web said about you before that point is part of what it knows.
What the model fetched	Text pulled in at the moment of the question — by searching an index, by crawling, or by opening a URL the user supplied. This is where citations come from, and it is the path a published page can enter within hours rather than at the next training run.

The retrieval sequence

On the fetched path the sequence is roughly the same everywhere, even though the implementations are not public. A query is derived from the conversation, often rewritten and often split into several. Candidate documents come back. Those documents are broken into passages, and **passages, not pages, are what gets ranked and passed to the model**. The model then writes an answer over the passages it was given and attaches citations to the sources it used.

Google describes its own version of the first step openly, as a *query fan-out* technique — "issuing multiple related searches across subtopics and data sources — to develop a response". That is a documented statement about one engine, and it is consistent with what the structure of the problem requires everywhere.

The unit of value is a passage rather than a page, and a source that was never retrieved cannot be cited no matter how good it is

Two consequences follow directly, and they drive everything else in this guide. The first is that writing is a per-passage discipline, not a per-page one: a brilliant page whose third section depends on its second has one usable passage. The second is that access is not a technicality to be delegated to whoever owns the CDN. It is the gate, and it is upstream of every word you write.

What follows for freshness

Freshness splits in two, and the split is the most useful practical consequence of the two-path model. The retrieval path can surface a page published today. The trained path cannot know about it at all until a later model. A new page therefore needs retrieval to work for it; a long-standing reputation is what the trained path uses. Teams that conflate the two end up either expecting a launch page to appear in a model's parameters, or expecting years of brand-building to show up in this week's citations.

Sources. Google, *AI features and your website* (query fan-out, eligibility). The passage-level mechanism is **structural**: no provider publishes its chunking, and nothing here claims to know the chunk size.

02 · STRUCTURE

The three conditions

A page has to clear three separate bars, in order. Failing any one of them makes work on the other two worthless, which is why the pack checks them in this sequence and stops at the first failure.

1 · Retrievable

The assistant is allowed and able to fetch the page, and something in the index connects it to the question. Permission, delivery and discovery are three separate things and all three must hold.

2 · Quotable

A passage on the page answers the question on its own, in a form a model can lift without repairing it. Read out of context, next to passages from other sites, by something that will paraphrase it.

3 · Corroborated

The claim is supported elsewhere, so a system that weighs agreement has a reason to trust it. The only condition you cannot satisfy from your own page.

The order

There is no point rewriting passages on a page an agent cannot fetch, and no point measuring a page whose passages cannot be quoted. Weeks are lost to this every quarter.

Why teams get the order wrong

Because condition two is the fun one. Rewriting is visible work with a satisfying output, and it can be started in an afternoon by someone who does not need access to the CDN. Condition one requires finding out who owns the WAF. Condition three requires a quarter of unglamorous corrections to profiles nobody has looked at since the company was founded.

The most common single finding in a first audit is a page that renders perfectly in a browser and returns a loading shell to an anonymous fetch. No amount of writing helps it. The second most common is a blanket Disallow added years ago for an unrelated reason, quietly covering every agent nobody thought to name.

Condition three is a programme, not a release gate

Corroboration moves on a scale of months and sits mostly off your own site. Blocking a release on it stops the shipping and does not speed the corroboration. Record it, assign it an owner and a date, and ship — but record it, because "corroboration is slow" becomes "corroboration never happened" inside two quarters.

Sources. Structural throughout, derived from the retrieval mechanism in section 01. Google's eligibility statement — indexed and eligible to be shown with a snippet, with no additional technical requirements — is the documented anchor under condition one.

03 · BEFORE THE PAGE

The question set

Write the questions the page must be able to answer, phrased the way a person asks an assistant. Everything downstream is scored against this list.

A passage review with no question set is a style review. An observation panel with no question set is somebody typing their brand name into an assistant and feeling something about the result.

A keyword is what somebody typed into a box. A question is what somebody wants to know.

Assistants rewrite what the user said before they search, carry context from earlier turns, and often issue several searches for one question. Optimising for one exact phrase is optimising for the one thing that no longer reliably happens. Google's own guidance notes that users in its AI experiences are "asking longer and more specific questions — as well as follow-up questions to dig even deeper".

The procedure

01 Name the decision the page serves

One sentence: *a person who reads this page is trying to decide* _____. If the sentence is hard to write, the page has no job, and no question set will rescue it.

02 Write the primary question

In the words a person would use out loud. Not *enterprise pricing SaaS* but "How much does this cost for a team of twenty?"

03 Write the neighbours

Six to twelve more across the four families below, plus the follow-up each invites. Covering the question and its surroundings beats matching one exact phrase.

04 Assign each question an owner page

This page, another page, or nothing. The third outcome is the gap list, and it is the most valuable output of the step.

05 Write the answer you would accept

Two or three sentences per question — not marketing copy, but the acceptance test. If the page does not contain that answer as one contiguous passage, the passage does not exist yet.

06 Freeze and date the set

A panel whose questions drift between runs measures nothing. Change the set in a new dated version, never in place.

The four families

FAMILY	ASKS	EXAMPLE
Definition	What is this thing, in one sentence?	"What is answer engine optimisation?"
Comparison	How does it differ from the adjacent thing?	"Is AEO different from SEO?"
Procedure	How do I actually do it?	"How do I check whether an assistant can read my site?"
Objection	Why might this be wrong or not worth it?	"Does llms.txt actually do anything?"

04 · CONDITION 1

Retrievable

Two questions, answered with evidence rather than assumption: **is this agent allowed in**, and **what does it get when it comes**.

Start with the fetch, not the file

Before reading a single line of `robots.txt`, look at what an agent would actually receive. Fetch the page as an anonymous client, with scripts disabled and nothing signed in, then strip the tags and read what is left.

```
curl -sSL -A "Mozilla/5.0 (compatible; retrievability-audit)" \
  https://example.com/page -o fetch.html -w '%{http_code} %{size_download}\n'
```

Three outcomes, and only one is fine.

- **The answer to the question is in what came back.** Pass.
- **A shell came back and the content arrives by script.** The content may not reach the model at all. This is the single most common silent failure, and it is invisible in a browser.
- **A challenge page, a consent wall or a 403 came back.** A fetch that succeeded technically and carried no content.

Record the status, the byte count and the first 500 words of extracted text. A 4KB pricing page is a finding, not a statistic.

Then read `robots.txt` per agent, not as a mood

Resolve the file separately for each token and record, for each: permitted, disallowed, or not mentioned — and which rule block decided it. Two rules of the protocol cause most of the mistakes.

The two rules everybody gets wrong

A group matching a **specific** user agent token wins over the `*` group entirely: an agent named in its own block inherits nothing from `*`. And an agent named **nowhere** falls into `*` — which is why a blanket `Disallow: /` added years ago for a different reason silently covers every agent nobody thought to name.

Then ask the question that matters: **was each decision deliberate?** A disallow on a training crawler is a decision. A disallow on the search or answering agent, when the goal is to be cited, is almost always an accident.

The layer `robots.txt` cannot see

Server and edge rules are enforced where `robots.txt` is only a request. Check the WAF and bot-management logs for the agent names rather than for blocks; check rate limits, which make a crawler functionally absent on a large site; and check IP rules against each provider's published ranges, because a rule that allows a user-agent string but blocks its address range permits nothing. If your CDN serves or appends a managed `robots.txt`, the file at your domain is not the file in your repository — fetch it.

Sources. Provider crawler documentation from OpenAI, Anthropic, Perplexity and Google (see section 05 and the source list); Cloudflare, *Content Signals Policy*, for managed `robots.txt` and the limits of preference signals.

05 · REFERENCE

Crawlers, by provider

A provider does not have one crawler. It has several with different jobs, they carry different names, and they are controlled separately.

JOB	WHAT IT DOES	BLOCKING IT MEANS
Training	Gathers text that may contribute to model training	Your text is not used to train. It does <i>not</i> stop you being cited.
Search / index	Builds the index the assistant searches	You are not in the index the assistant searches. This is the one that stops citation.
User-initiated fetch	Fetches a page because a person asked about it a second ago	Varies. Several providers state robots.txt rules may not apply, because a user asked.

Tokens, as published at the time of writing

TOKEN	JOB	PUBLISHED NOTE
Googlebot	Search	Eligibility for AI Overviews and AI Mode runs through ordinary Search indexing.
Google-Extended	Training	Training future Gemini models; grounding in Gemini Apps. Does not affect Search ranking or indexing.
GoogleOther	Other	Various product teams; research and development crawls.
OAI-SearchBot	Search	Surfaces websites in ChatGPT's search features. Sites opted out "will not be shown in ChatGPT search answers, though can still appear as navigational links".
GPTBot	Training	Crawls content for training generative AI foundation models.
ChatGPT-User	User fetch	"Because these actions are initiated by a user, robots.txt rules may not apply."
Claude-SearchBot	Search	Analyses online content for search result relevance and accuracy.
ClaudeBot	Training	Collects web content that may contribute to training. Supports Crawl-delay.
Claude-User	User fetch	Supports users accessing websites during queries. Honours robots.txt.
PerplexityBot	Search	"Designed to surface and link websites in search results on Perplexity. It is not used to crawl content for AI foundation models."
Perplexity-User	User fetch	"Since a user requested the fetch, this fetcher generally ignores robots.txt rules."
bingbot	Search	Builds the index behind Microsoft Copilot and Bing's AI summaries.

Sources. OpenAI, *Crawlers*; Anthropic, *Does Anthropic crawl data from the web...*; Perplexity, *Perplexity Crawlers*; Google, *Google's common crawlers*; Microsoft Bing Webmaster Blog, *Introducing AI Performance in Bing Webmaster Tools*, 10 February 2026. All **documented**.

05 · REFERENCE, CONTINUED

A robots.txt written on purpose

Verify, do not copy

The table opposite was verified against provider documentation on 3 September 2026. These lists change, and at least one control is a robots.txt token rather than a crawler, so read the current documentation rather than copying a list out of an article — including this one. Every provider publishes verified IP ranges as JSON; agent strings are trivially forged, so verify identity against the ranges rather than the string.

The file below permits search and answering agents, declines training, and is explicit rather than inherited. Every line is a decision somebody can defend.

```
# Search and answering agents – permitted, deliberately.
User-agent: Googlebot
Allow: /
User-agent: bingbot
Allow: /
User-agent: OAI-SearchBot
Allow: /
User-agent: Claude-SearchBot
Allow: /
User-agent: PerplexityBot
Allow: /

# Training crawlers – declined, deliberately. This does not affect citation.
User-agent: GPTBot
Disallow: /
User-agent: ClaudeBot
Disallow: /
User-agent: Google-Extended
Disallow: /

# Everything else.
User-agent: *
Allow: /
Disallow: /admin/
Disallow: /internal/

Sitemap: https://example.com/sitemap.xml
```

Two things this file cannot do. It cannot bind a user-initiated fetcher whose provider states that robots.txt may not apply to it. And it is a request, not an enforcement — the WAF is where enforcement lives.

Sources. Provider crawler documentation as listed opposite. The two robots.txt precedence rules are from the protocol itself.

05 · REFERENCE, CONTINUED

Discovery, push and the speculative

Access is permission. Discovery is whether anything connects the page to the question — and they fail independently.

Assistants that search reach pages through an index, so the ordinary search work applies unchanged: HTTP 200 and not noindex; a canonical URL that is itself and is the URL you want cited; links from pages already crawled, because an orphan page is reachable and undiscovered; a current sitemap. This is the prerequisite that AEO advice tends to skip because it is unglamorous.

Where the retrieval path matters, push rather than wait. IndexNow accepts a single URL by GET or up to 10,000 in one POST, with a key file of 8 to 128 hexadecimal characters at the site root, and participating engines share submissions. A page entering an index within hours is the one place where the fetched path beats the trained path outright.

The speculative item, labelled as such

llms.txt is a real proposal with real adoption, offering a markdown index of a site for language models. No major assistant has committed to consuming it, and Google states that you do not need new machine-readable files or markup to appear in its AI features. Publishing one is cheap and speculative — write it down as speculative, so nobody later reports it as a win.

Sources. IndexNow documentation; llmstxt.org; Google, *AI features and your website*. IETF's AIPREF working group is standardising a vocabulary for AI usage preferences and its carriage in robots.txt and an HTTP header — direction of travel, not finished, and nothing here depends on it.

06 · CONDITION 1, CONTINUED

Preview and use controls

Access and use are separate decisions, and both are set deliberately or set by accident. A page can be perfectly crawlable and still be excluded from an answer because a snippet control says so.

CONTROL	WHERE	EFFECT
<code>noindex</code>	robots meta / X-Robots-Tag	Page not shown in search results at all. Every other control becomes irrelevant.
<code>nosnippet</code>	robots meta	No snippet or video preview. Google states this also prevents the content being used as a direct input for AI Overviews and AI Mode.
<code>max-snippet: [n]</code>	robots meta	Limits snippet length; 0 equals <code>nosnippet</code> , -1 lets Google choose. Also limits how much may be used as a direct input for AI features.
<code>data-nosnippet</code>	attribute on span, div, section	Excludes specific elements from snippet use, at element level rather than page level.
Search generative AI control	Search Console → Settings	Property-level include or exclude for AI Overviews, AI Mode and generative AI in Discover. Does not affect AI training.
Google-Extended	robots.txt	Training and Gemini Apps grounding only. Does not affect Search ranking or indexing.
Content-Signal	robots.txt	A CDN-originated convention expressing search, ai-input and ai-train preferences. Explicitly a preference, not an enforcement.

Google's own summary of the trade-off is one sentence: **"More restrictive permissions will limit how your content is featured in our AI experiences."**

Separate the training decision from the citation decision

They use different controls, so a team that wants "not trained on, but cited" can express exactly that. A team that blanket-disallows because of a training concern has usually switched off citation without meaning to. Write both decisions down, with the name of who made them — and keep the deliberate restrictions, with the reason in a comment beside them so the next person does not delete it as an obstacle.

There is no AEO schema

Google states plainly that no new machine-readable files, AI text files or markup are needed to appear in its AI features, and that no special schema.org structured data is required. Ordinary structured data still earns its place, under two constraints people break constantly: **everything in the markup must also be visible on the page**, and the markup must validate. An empty `aggregateRating` left in by a template is a violation that helps nothing.

The three most common inherited settings

A `noindex` carried over from a staging template; a `max-snippet` set low for a legal reason years ago and never revisited; a blanket AI opt-out added to solve a training concern, which also removed the site from answers. All three are invisible in a browser, and one fetch away from being found.

07 · CONDITION 2

Quotable

A retrieved passage is read out of context, next to passages from other sites, by a model that will paraphrase it. Writing for that is a real constraint, and it is mostly the discipline of never depending on the paragraph above.

The seven rules

Each rule is a repair, not a preference, and each has a test that returns yes or no.

RULE	WHAT IT REQUIRES	THE TEST
1 · Name the subject	Every paragraph states what it is about rather than opening with <i>it</i> , <i>this</i> or <i>the above</i> . A passage whose subject lives two paragraphs earlier is unusable the moment it is separated from them.	Delete the paragraph before it. Does this one still say what it is about?
2 · Answer, then qualify	The first sentence under a heading answers the heading. Conditions, exceptions and history follow. A passage that spends four sentences arriving at the answer is often cut before it gets there.	Read only the heading and the first sentence. Is the question answered?
3 · Make claims checkable	Give the figure, the date it applies to and where it came from, in the same sentence. A number with no date and no source is exactly the kind of claim a careful system declines to repeat.	Does every number carry a date and an attribution within the same sentence?
4 · Define terms in place	The first use of any term the audience might not share carries a one-sentence definition. The passage may be the reader's only contact with your vocabulary.	Could a competent stranger to your category read this without opening another tab?
5 · Use question-shaped headings	Headings written as the question a person would ask make the match between passage and query more direct, and force the section beneath to actually answer something.	Is this heading a question from the question set, or a phrase a person would say?
6 · Keep structure in the text	Comparisons, steps and criteria belong in prose, lists and headings that survive conversion to plain text. Anything that only exists as layout, or only appears after a script runs, may not reach the model at all.	Run the plain-text extraction. Is the comparison still legible?
7 · Say what is uncertain	Marking the limits of a claim makes a passage safer to quote, not weaker. It also stops a system finding the contradiction elsewhere and choosing the other source.	Does the passage overclaim anything a reader could falsify in one search?

Self-contained does not mean voiceless. The rules are about dependency, not about tone.

Sources. The rules are **structural** consequences of passage-level retrieval, and they align with Microsoft's published site-owner guidance in the Bing AI Performance announcement, which advises "clear headings, tables, and FAQ sections", supporting claims with "examples, data, and cited sources", and content that is understandable in isolation. That a well-formed passage is quoted more readily is **observed**, not confirmed by any provider.

07 · CONDITION 2, CONTINUED

Taking a page apart

The procedure is short. The discipline is in step two, which almost nobody does properly.

01 Split the page into passages

Use the headings. Where a section runs past roughly 300 words with no heading, treat each paragraph cluster as its own candidate passage.

02 Read each passage in isolation — literally

Copy it into an empty file with no title and nothing around it. You cannot un-know the paragraph above, and reading it on the page is not a fair test. This one step finds most of what there is to find.

03 Score it against the seven rules

Yes or no per rule. Not a rating out of ten.

04 Write the specific repair for every no

"Not self-contained" is not a repair. "Opens with *this*; the subject is the re-crawl schedule, three paragraphs up" is a repair.

05 Check coverage against the question set

For each question the page owns, is there one *contiguous* passage that answers it? If the answer is assembled from three places on the page, it does not exist yet. This is the single most productive test in the pack.

06 Rewrite, then re-read in isolation

The second read is where the leftover pronouns are.

Three repairs, before and after

Naming the subject

BEFORE *"It also handles the second case, which is where most teams get stuck."*

AFTER *"Scheduled re-crawling also handles the second case — pages that changed after their last crawl — which is where most teams get stuck."*

Answering first

BEFORE *"There has been a lot of debate about this, and the answer has changed several times since 2023. Some argue... Ultimately, most sites should use a 301."*

AFTER *"Use a 301 redirect. The debate since 2023 has been about edge cases, and the edge cases are below."*

Marking uncertainty

BEFORE *"AI search now drives the majority of discovery."*

AFTER *"How much discovery moves through assistants is not measurable from outside these companies. What is measurable is our own referral log, and it is in the appendix."*

Failure modes worth naming

Front-loading the brand instead of the answer — "At Tau, we believe..." is four wasted sentences at the top of a passage. Turning every heading into a question, including the one over a price table. Adding a new unsourced claim during a rewrite, which makes the page worse than it was.

08 · CONDITION 2, CONTINUED

Claims and evidence

Make every claim on the page either checkable or gone. A number with no date and no source is also the kind of claim that ages into a liability.

The procedure

1. **Extract every claim.** Every sentence containing a number, a superlative, a comparison, or an assertion about how a third party's system behaves. A page usually has more than the author remembers.
2. **Trace each to a primary source.** Not the article that cited the article that cited it. Three or four hops ending at a vendor blog quoting an unnamed "recent study" is the ordinary outcome, and it is a finding.
3. **Rewrite each survivor into the checkable form.** Figure, date it applies to, source — in the same sentence.
4. **Carry the source's caveat across.** Inheriting a source's authority without inheriting its limits is how a careful claim becomes a wrong one.
5. **Date the page.** "Reviewed 3 September 2026; crawler names and console features verified against provider documentation on that date" is worth more than a redesign.
6. **Diarise the re-check.** Every documented claim has an expiry. Quarterly is a reasonable default for provider behaviour.

A worked rewrite

BEFORE "Most AI answers cite multiple sources."

AFTER "In Pew Research Center's analysis of 68,879 Google searches by 900 US adults in March 2025, 88% of AI summaries cited three or more sources and 1% cited a single source. Pew notes it could not analyse AI summaries on non-Google engines."

The second version is longer, duller, and the one a model can quote without exposing itself.

The rule for unverified claims is that they do not go on the page. Not hedged — removed.

What gets deleted

- **Citation laundering.** Vendor A cites vendor B citing vendor A. Follow the chain; it usually loops.
- **The plausible round number.** "60% of buyers now start in an assistant" has never been measured by anyone who published a method.
- **Undated first-party numbers.** "Our customers see 3x" with no period, no cohort and no metric definition is unverified even though you own the data. Replace it with something you can actually show: "Across accounts active in August 2026, the median time from opening a bank feed to marking a month reconciled was 4 days, down from 9 before automatic matching shipped in March 2026. Measured in our own product; it excludes work done elsewhere."

Sources. Pew Research Center, *Google users are less likely to click on links when an AI summary appears in the results*, 22 July 2025 — 900 US adults, 68,879 searches, March 2025. The only click-behaviour figures in this guide, and usable precisely because Pew publishes the method.

09 · CONDITION 3

Corroborated

The third condition is the one you cannot write your way into on your own page. Both paths reward agreement across independent sources.

A model trained on text that repeatedly associates your name with a category learns that association. A retrieval system assembling an answer from several documents has an easier time using a claim that its other sources do not contradict. A product described three different ways in three places gives a model no stable fact to learn.

The procedure

01 Write the canonical description

Three sentences: what it is, in the category noun a stranger would use; who it is for, specifically enough to exclude people; what is distinctive and not true of the obvious alternative. Then a one-sentence and a six-word version. These are the only descriptions anyone writes from, and they live in the repository, not in a deck.

02 Inventory every surface

Owned pages, profiles and directories, third-party coverage, and the structured data. Most teams find between twenty and sixty.

03 Diff each against the canonical description

Matches, stale, wrong category, or not yours to edit. *Stale* is the highest-value fix: a stale description elsewhere outweighs a fresh one on your site when everything else agrees with the stale version. *Wrong category* is the most urgent, because category is the fact a model learns first.

04 Make the machine-readable version agree

Organization markup on the home page or the page that describes the organisation — Google says you do not need it on every page — with the canonical description and sameAs links to the profiles that are genuinely yours. Everything in the markup must be visible on the page.

05 Handle the surfaces you do not own

Correct the record where correction is invited. Publish the things other people need to cite you for — the definition, the method, the comparison, the pricing — not only the pages that sell. Treat mentions without links as real: training text and retrieved passages carry names whether or not anything is hyperlinked.

06 Re-run quarterly

Positioning drifts and a rename takes months to propagate.

The honest caveat, stated plainly

How much any of this weighs is not documented by any provider. The corroboration effect is inferred from how these systems are built rather than confirmed by them. It is a sound bet, not a measured one, and this guide does not pretend otherwise. The nearest thing to provider endorsement is Microsoft's advice to publishers to "align text, images, and video so they consistently represent the same entities" — a provider recommending consistency, not a provider quantifying its weight.

What this is not

Seeding identical copy across low-quality directories is a well-known pattern with a well-known outcome. Fabricated consensus is detectable and is the opposite of this work. So is chasing links while ignoring mentions: the unlinked mention counts on both paths, and no backlink audit will show it to you.

10 · CONTRAST

Where this departs from ranking

The overlap with search work is large, which is why teams assume the two are the same job. These are the five places where that assumption produces wrong decisions.

RANKING ON A SEARCH ENGINE	BEING CITED IN AN ANSWER
A results page shows many links and rewards being anywhere on it.	An answer shows one composed response with a handful of citations, so being fourth-best is much closer to being invisible .
The click is the outcome, and a visit is what the work is measured by.	The user often has what they needed inside the conversation, so being the cited source is worth something even without a visit .
The query is the phrase the user typed.	The assistant rewrites what the user said, carries context from earlier turns, and may issue several searches, so covering the question and its neighbours beats one exact phrase .
There is a position, and it can be looked up and compared over time.	Ask twice and you can get different sources, because model versions, retrieval variance and personalisation all move it .
Freshness is one path: publish the page, get it crawled, become eligible.	Freshness splits in two — the retrieval path can surface a page published today, and the trained path cannot know about it at all until a later model.

A single observation proves very little. A pattern across many is the smallest useful unit of evidence.

The one piece of outside evidence worth carrying

Pew Research Center's March 2025 study is the only click-behaviour data in this guide, because it is the only one whose method is published. Across 68,879 Google searches by 900 US adults, about 18% produced an AI summary. Users who saw one clicked a search result in 8% of visits, against 15% for those who did not, and clicked a link inside the summary in 1% of visits.

What that supports is narrow and important: **click-based measurement was already measuring a small and shrinking slice**, on one engine, in one month, in one country. What it does not support is any claim about your category, any other engine, or any trend since. Pew itself notes it could not analyse AI summaries on non-Google engines.

The academic framing is worth knowing too. The KDD 2024 paper that named the field, *GEO: Generative Engine Optimization*, reports visibility improvements of "up to 40%" from content-side changes. Read the caveat before quoting the number: it is measured on the authors' own benchmark against their own generative-engine implementation, not against production assistants, and "up to" is doing real work in the sentence. This guide cites the paper for its framing — that content-side changes move visibility at all, and that the effect varies by domain — and carries the figure into no procedure.

Sources. Pew Research Center, 22 July 2025; Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan and Deshpande, *GEO: Generative Engine Optimization*, KDD 2024 — filed as **observed**, not documented.

11 · THE GATE

The pre-publish checklist

Six checks. Each is a check with an answer, not a judgement call. A review that returns "looks good" has not been run.

#	CHECK	DO THIS	PASS CONDITION
1	Fetch the page as an anonymous client	Scripts disabled, nothing signed in	The answer is in what comes back
2	Read robots.txt and any bot filtering	Resolve per agent; check the WAF and the live file	A permission list that was deliberate rather than inherited
3	Read each section on its own	Copy it into an empty file	It names its subject, answers its heading and defines its terms
4	Check every figure	Trace to a primary source	A date and a source, or the figure removed
5	Compare the description with your other surfaces	Diff against the canonical description	The same description everywhere, not three
6	Record the questions this page answers	From the frozen question set	Something for the review to test against

Three verdicts, and no scores

Ship

All six pass. Corroboration work, if outstanding, is recorded with an owner and a date.

Ship with follow-ups

Conditions 1 and 2 pass; specific, dated follow-ups are attached to named people.

Blocked

A condition-1 or condition-2 failure, named, with the single next action. Nothing else is worth doing until it clears.

Not a verdict

"7 out of 10." A score is not a decision anybody can act on, and it lets a blocking failure hide inside an average.

The order the gate runs in

question-set	what is this page for	
↓		
retrievability-audit	can an agent fetch it	CONDITION 1
structured-data-and-preview-controls	how much of it may be used	
↓		
passage-rewrite	does a passage answer alone	CONDITION 2
claim-and-evidence-check	is every claim checkable	
↓		
entity-consistency	do other sources agree	CONDITION 3
↓		
assistant-observation-panel	what do assistants actually say	
citation-measurement	what do platforms and logs count	

Run the baseline observation panel *before* publishing a new or substantially rewritten page, where the questions are answerable without it. A rewrite with no baseline can only ever be reported as a feeling.

12 · MEASUREMENT

Measuring without rank tracking

There is no position to track and no results page you can fetch and score. Everything available is either a sample you took, a partial count somebody else made, or a referral that happened to carry attribution.

SOURCE	COVERS	BLIND TO
Google Search Console Generative AI performance reports	Impressions in generative AI features on Search — AI Overviews and AI Mode — and in Discover. Dimensions: pages, countries, devices, dates from hourly to monthly. Announced 3 June 2026; rolled out to all websites worldwide as of 31 August 2026.	Clicks, position and click-through rate are not in this report. Search Labs experiments are excluded. Do not construct a CTR from it.
Bing Webmaster Tools AI Performance	Citations across Microsoft Copilot, AI-generated summaries in Bing, and select partner integrations. Total citations, average cited pages per day, grounding queries , page-level activity and trends. Public preview from February 2026.	Every non-Microsoft surface.
Referral traffic Analytics and logs	Sessions arriving from assistants that pass a referrer or a tagged URL.	Every session that arrives carrying nothing, and every reader who never clicks. A floor, not a total.
Server logs Verified agent fetches	That a named agent fetched a page, and when — verified against the provider's published IP ranges rather than trusted on the user-agent string.	Whether it used the page. Whether it cited it.

The grounding queries in Bing's report deserve particular attention: it is the only place that tells you *which question* caused a citation, which makes it the best available input to a question set.

Six rules for the report

1. **Never add the sources together.** An impression in one engine, a citation in another, a referral session and an agent fetch are four different events. A single "AI visibility score" is arithmetic on incompatible units.
2. **Write each source's blind spot before the numbers.** This is the step that keeps the report from becoming a claim.
3. **Pair every number with its period or its denominator.** "412 impressions, 1–31 March" survives scrutiny. "Up 40%" does not.
4. **Report changes with their absolutes.** "Citations up 50%" over a base of four.
5. **Join to the observation panel where they corroborate.** Independent sources agreeing about one page is the strongest statement this discipline supports.
6. **State the three unmeasurable things.** Use without citation; your share of all answers in the category; any ranking factor. Putting the unmeasurable in the report is what makes the measurable credible.

Sources. Google Search Central, *Introducing Search Generative AI performance reports in Search Console*, 3 June 2026, and Search Console Help, *Generative AI performance report (Search)*; Bing Webmaster Blog, 10 February 2026. That ChatGPT tags referral URLs with `utm_source=chatgpt.com` is reported by OpenAI's help centre, which we could not fetch directly — verify before relying on it.

13 · MEASUREMENT, CONTINUED

The observation panel

A panel of your own questions, run on a schedule across the assistants you care about, with the answers and the cited sources kept. A sampling exercise, not a rank tracker.

Four things must not change between runs

FIXED	WHY
The question set	Verbatim, with its version and date. Rephrasing a question because the answer disappointed you makes it a different question.
The assistants	The same roster every run, or the change is recorded.
The conditions	Signed out; personalisation and memory off where a switch exists; a consistent locale; one question per fresh conversation; no follow-ups. Signed in, you measure your own history.
The repeats	The same number per question per assistant. Three is a workable minimum. One is not a sample.

What is recorded, per observation

Date and time; assistant and model version if shown; the question verbatim; the repeat number; the answer **verbatim**; every cited source in order with its URL; whether you were cited and which URL; whether you were named without a citation; the competitors cited; whether the answer was correct. A summary is not evidence — the wording tells you which passage was used.

Three honest aggregate measures, and no others

Citation frequency

Observations in which any page of yours was cited, out of the total. Reported as a fraction with its denominator, never as a bare percentage.

Mention without citation

The same count for being named with no link. Real, common, and invisible to every log-based method.

Source concentration

Which domains appear most often, and whether the cited set is stable between repeats.

Not a measure

"Share of voice." There is no denominator that makes it mean anything: the population of questions people ask assistants in your category is unknowable from outside.

Reading the diff conservatively, and setting the cadence

The honest reading of most changes between two runs is: *this moved, and a model release, retrieval variance or our rewrite could each explain it*. Attribute a change to your work only when it is large, persists into a later run, and lines up with a page you altered. Run the panel monthly, plus once around any significant rewrite or rename and when a provider ships a change large enough to be news — not daily, or the variance will swamp the signal and the team will start reacting to noise.

The clearest result the method can produce

You publish a genuinely new, self-contained page answering a question you were never cited on, and in the next run you are cited on it with that page as the cited URL every time. A URL cannot be cited before it exists. That is as close to causal as this discipline gets — and almost nothing else in a panel diff reaches that bar.

14 · HONESTY

Uncertainty register and output contract

Anything written about this subject should carry a register like this one, and this guide is not exempt.

CLASS	WHAT SITS IN IT
Documented	Crawler names and robots.txt behaviour, the existence of a training cutoff, the preview and generative-AI controls, the reports each console offers, and each provider's stated policies. Verifiable today and still liable to change.
Structural	That retrieval operates over passages; that models answer from parameters when they do not fetch; that a source not retrieved cannot be cited; that agreement across sources is easier to use than contradiction. These follow from how the systems are built rather than from any announcement.
Observed	That clear, self-contained, well-sourced passages tend to be quoted more readily. Consistent with practice and with how the systems work, and not confirmed by any provider.
Unverified	Anything expressed as a share, a multiplier or a ranking factor. Nobody outside these companies can measure it, and a number attached to it is a sales device rather than a finding.

The output contract

Every full run of the pack returns the same four artefacts per page. Anything else is commentary.

The question set

The questions this page should be able to answer, phrased as a person would ask an assistant rather than as keywords — with each question's owner page and its acceptance answer.

The access report

Which agents are permitted, what the anonymous fetch returned, and any bot filtering that interfered — with the deciding rule named for each verdict.

The passage review

Each section marked as self-contained or not, with the specific repair where it is not, and coverage checked against the question set.

The observation log

The dated panel runs, the answers verbatim, the sources cited, and what changed since the previous run — with a noise reading beside every signal reading.

What this guide deliberately does not contain

No vendor "AI visibility study", no share-of-voice figure, and no ranking-factor list. None of those can be measured from outside the providers, and the rule in section 08 would delete them from a client's page. Applying it to ourselves is the minimum.

Report both as what they are: samples and partial counts, not a percentage that sounds like coverage

15 · THE PACK

The nine skills

This guide is the reading edition. The pack is the working one: nine `SKILL.md` files with frontmatter, procedures, checklists, prompts and output templates, plus a full source list and two worked examples. Download it at get-tau.com/resources/aeo-skill-spec.

SKILL	WHAT IT DOES
question-set	Writes the questions the page must answer, assigns each an owner page, and freezes the set so a panel can use it.
retrievability-audit	Establishes, per agent, whether the page can be fetched and what comes back. Separates training from search from user-initiated fetch, and robots.txt from the WAF.
structured-data-and-preview-controls	Sets the preview and generative-AI controls deliberately, and makes markup match the visible page.
passage-rewrite	Takes each section apart against the seven rules and rebuilds it to survive alone. Checks coverage against the question set.
claim-and-evidence-check	Traces every figure to a primary source, labels it, and removes what cannot be traced.
entity-consistency	Makes the description of the organisation the same everywhere it appears, on your surfaces and off them.
assistant-observation-panel	Runs the fixed question set across the assistants on a schedule and keeps the raw record.
citation-measurement	Assembles what the consoles and logs actually report, with each source's blind spot written first.
aeo-release-review	The gate. Runs the six checks in order, stops at the first blocking failure, and returns the four artefacts and one of three verdicts.

Also in the pack

- **references/references.md** — every source used, what each contributed, and whether it was read directly or only through a summary.
- **references/crawlers-and-controls.md** — the agent tables, a robots.txt written on purpose, and the preference conventions that are not enforcement.
- **examples/01** — a pricing page through the full release review, including the stop condition and the fix.
- **examples/02** — two observation panel runs around a rewrite, and exactly how much may honestly be concluded from the difference.

Where Tau fits

Tau is the marketing autopilot for new products: it runs the visibility work across search, AI answers and social so a small team does not have to. The free diagnostic at platform.get-tau.com/diagnostics runs the access and discovery half of condition one against a URL you give it, and tells you what an anonymous fetch actually returns. The skills above are yours either way — they are free, and they do not depend on the product.

16 · SOURCES

Every document this guide was built from

Verified on 3 September 2026. The subject moves; re-read before relying on any of it.

Google

- **AI features and your website** — developers.google.com/search/docs/appearance/ai-features. Eligibility, query fan-out, and the statement that no new machine-readable files or special schema are needed.
- **Top ways to ensure your content performs well in Google's AI experiences on Search**, 21 May 2025 — developers.google.com/search/blog/2025/05/succeeding-in-ai-search. Preview controls, markup matching visible content, longer and follow-up questions.
- **Introducing Search Generative AI performance reports in Search Console**, 3 June 2026 — developers.google.com/search/blog/2026/06/gen-ai-performance-reports. Dimensions and the 31 August 2026 worldwide rollout.
- **Generative AI performance report (Search)**, Search Console Help — support.google.com/webmasters/answer/16984139. The impression definition; impressions only.
- **Search generative AI control**, Search Console Help — support.google.com/webmasters/answer/16908024. Property-level include/exclude; does not affect training.
- **Robots meta tag, data-nosnippet, and X-Robots-Tag specifications** — developers.google.com/search/docs/crawling-indexing/robots-meta-tag.
- **Google's common crawlers** — developers.google.com/search/docs/crawling-indexing/google-common-crawlers.
- **Organization structured data** — developers.google.com/search/docs/appearance/structured-data/organization. sameAs and placement guidance.

OpenAI, Anthropic, Perplexity, Microsoft

- **OpenAI crawlers** — developers.openai.com/api/docs/bots. OAI-SearchBot, GPTBot, ChatGPT-User, OAI-AdsBot, and the published IP-range endpoints.
- **Does Anthropic crawl data from the web, and how can site owners block the crawler?** — [support.claude.com](https://support.claude.com/article/8896518), article 8896518. ClaudeBot, Claude-SearchBot, Claude-User; Crawl-delay; verified IPs.
- **Perplexity Crawlers** — docs.perplexity.ai/docs/resources/perplexity-crawlers. PerplexityBot and Perplexity-User, including the user-initiated robots.txt exception.
- **Introducing AI Performance in Bing Webmaster Tools (public preview)**, 10 February 2026 — blogs.bing.com/webmaster. Surfaces, metrics, grounding queries, and Microsoft's site-owner guidance.
- **Searching the web with ChatGPT** and the publishers FAQ — help.openai.com. *Not read directly; both returned HTTP 403 to our fetch*. Consulted through search summaries only, and flagged wherever used.

Continued overleaf. Protocols, conventions and independent evidence.

16 · SOURCES, CONTINUED

Protocols, conventions and independent evidence

- **IndexNow documentation** — indexnow.org/documentation. Key file, GET and POST submission, the 10,000-URL batch limit, and engine sharing.
- **Cloudflare Content Signals Policy** — blog.cloudflare.com/content-signals-policy. The search, ai-input and ai-train definitions, and the enforceability caveat.
- **The llms.txt proposal** — llmstxt.org. Jeremy Howard, September 2024. Filed as speculative.
- **IETF AI Preferences (AIPREF) working group** — ietf.org/blog/aipref-wg and the vocabulary draft on datatracker. Consulted via summaries; direction of travel only.
- **Pew Research Center**, *Google users are less likely to click on links when an AI summary appears in the results*, 22 July 2025 — pewresearch.org. 900 US adults, 68,879 searches, March 2025.
- **Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan, Deshpande**, *GEO: Generative Engine Optimization*, KDD 2024 — arxiv.org/abs/2311.09735. Framing only; the headline figure is benchmark-internal.

Read the date, then read the source

Every document listed here was read directly on 3 September 2026 with two exceptions: the OpenAI help-centre articles returned HTTP 403 to our fetch, and the IETF drafts were consulted through summaries. Both are flagged wherever they are used, and neither carries a procedure in this guide. If you are reading this months from now, treat every **documented** claim as a hypothesis until you have re-opened its source.

Take the pack, not the promise.

Nine skills, a full source list and two worked examples. Free,
and they do not depend on the product.

The skill pack and this guide · get-tau.com/resources/aeo-skill-spec

The free visibility diagnostic · platform.get-tau.com/diagnostics

Tau is the marketing autopilot for new products — it runs the visibility work across search, AI answers and social so a small team does not have to.

This guide is draft work product about systems whose internals are private. Every claim is labelled, every source is listed, and the honest answer to a good many questions in this field remains *nobody outside these companies knows*. Verify before you rely on it.