GEO & MEASUREMENT

From visibility to demand.

GEO: From AI Visibility to Qualified Demand

Field guide / Edition 1.0 2026-09-09

ranketize.com/resources

About this edition

Public research, practical methods and clearly labelled examples. Prepared with AI assistance; this edition reports no original Ranketize experiment results.

Read the current web edition and follow its source links for changing platform guidance.

https://ranketize.com/resources/geo-measurement-guide

Contents

Using the worksheets

Copy the tables from the web edition or download its Markdown text. Record your own observations in the blank fields; any figures in worked examples are illustrative.

Use this guide to choose the buyer questions worth monitoring, check whether AI answers represent your business accurately, and connect observable discovery to qualified enquiries. In 30 days, a small team can establish a repeatable baseline, improve a few important source pages and decide what deserves further investment. The deliverable is a documented decision, supported by a measurement workbook.

This first edition synthesizes public documentation and research, checked on 9 September 2026. Its worksheets and operating method are Ranketize's editorial recommendations. The worked example is fictional. No original Ranketize study or client performance results are presented.

1. Start with a buying decision

A useful GEO program begins with a decision your customer needs to make. For a SaaS business, that might be whether its product supports a required integration, suits a particular team or can replace an existing workflow. Write down the decision before choosing prompts or commissioning content.

Use one sentence: “Help [specific buyer] assess [specific choice] using [evidence they need], then offer [appropriate next step].” A project management vendor might focus on operations managers evaluating migration from spreadsheets. Its evidence could include an import walkthrough, limits on field mapping and a sample implementation plan. A relevant next step is a migration assessment.

This produces a narrower, more useful research question than “How visible are we in AI?” Ask whether answers to migration questions describe the product correctly, identify its limitations and connect buyers with evidence they can inspect.

Keep five stages separate: technical eligibility, appearance in an answer, accurate representation, a website visit and a qualified enquiry. Our recommendation to separate them follows from evidence with different units of measurement: platform documentation describes access and reporting; citation research evaluates support; behavioral research observes clicks. None supplies a complete business attribution model.

Choose a commercial outcome your team can assess consistently. For example, a qualified enquiry might require a relevant company, a stated migration need and an agreed follow-up. Define exclusions, such as recruitment, suppliers and spam. A download can be useful engagement without satisfying that definition.

Before collecting anything, name the person who will use the findings and the decision they will make. If nobody can explain what changes when the metric moves, remove it from the first dashboard.

2. Establish eligibility and a clear source of truth

Google says its generative search features retrieve information from its Search index and can expand a question into related searches. Its current guidance requires indexing, snippet eligibility and inclusion in Search generative AI features. Meeting those conditions does not guarantee selection. The same guidance says Google ignores **llms.txt** and does not require special AI schema or forced content chunking. These are Google-specific statements. Google's optimization guidance

OpenAI documents separate controls for OAI-SearchBot, which supports search discovery, and GPTBot, which concerns potential training use. ChatGPT-User performs certain user-triggered visits. A log entry from one agent does not establish what another system indexed or cited. Record the actual agent and request rather than calling every automated visit “AI visibility.” OpenAI crawler documentation

For each priority page, inspect the response status, intended canonical URL, rendered main text, indexability, crawler access and internal links. Check that important product facts survive mobile rendering and do not exist solely inside an inaccessible interaction. Record evidence and the inspection date. A technical checklist should identify a fault someone can fix.

Then choose a source of truth for each changeable fact: pricing, availability, integrations, implementation requirements or service scope. Give it a responsible owner and a revision date. Link explanatory articles to that source instead of maintaining conflicting copies of the same specification.

An answer may consult material without citing it. OpenAI's web search API distinguishes inline citations from a broader list of consulted sources. That API documentation does not describe every consumer interface or reveal every step behind an observed answer. Consequently, record a linked citation as a linked citation; avoid inferring training inclusion or exclusive influence. OpenAI web search documentation

Your output from this stage is a short fault list and an evidence map. Resolve missing or contradictory product facts before expanding the publishing schedule.

3. Build a baseline with explicit denominators

Construct a fixed question set from consented customer questions, sales objections or documented support needs. If your team invents a question, label it as constructed. A plausible prompt is not evidence of search volume.

Group questions by decision: understanding the problem, comparing approaches and evaluating a supplier. Keep branded questions separate from non-branded discovery. Record the complete wording, product surface, search mode, date, locale and account context. Use fresh conversations under a documented procedure, and preserve unsuccessful attempts.

Google's current Generative AI performance report covers impressions in AI Overviews and AI Mode, with page, country, date and device dimensions. Property and page aggregation differ, and Search Labs experiments are excluded. Treat it as a native impression report with documented limitations. Its exports can represent unavailable values as zeros, so retain availability notes when collecting data. Search Console report documentation

Use that report alongside your constructed sample and website analytics. They describe different populations. Do not divide sampled citations by native impressions or interpret a prompt sample as your share of all buyer searches.

Metric dictionary

Adopt these definitions before the baseline. Report counts alongside rates.

Metric	Definition for this workbook	Boundary
Collection completion	Attempts made / attempts planned	Missing collection must remain visible.
Answer availability	Inspectable substantive answers / all attempts	Record no-answer, refusal and collection failure separately.
Brand presence	Answers naming the target brand / valid answers	A mention can be unfavorable or irrelevant.
Owned-source citation	Answers linking an owned domain / valid answers	Count each answer once, regardless of link count.
All-attempt presence	Attempts yielding a brand mention / all attempts	Shows the effect of unavailable answers on coverage.
Citation support	Fully supported claim/source pairs / inspectable pairs	Report partial and unsupported pairs separately.
Audit coverage	Inspectable pairs / all pairs selected for audit	An inaccessible source is unverifiable.
Observed AI-referred sessions	Sessions classified by a documented source rule	Includes only traffic identifiable under that rule.
Qualified enquiries	Distinct enquiries meeting the written qualification rule	Deduplicate and record the review date.

Keep results separated by surface and question group before presenting a total. Repeated runs on the same question show variability; they are not independent people. Preserve the question-level records so a rise caused by one frequently mentioned brand query cannot masquerade as broader discovery.

4. Audit the answer, then improve the evidence

A mention is worth inspecting when it can influence a buying decision. Prioritize claims about capability, suitability, price, implementation and comparisons. Split compound statements into checkable claims, associate each with its cited URLs, and retain uncited claims in the audit.

Research by Liu, Zhang and Liang distinguishes whether generated statements have citation support from whether individual citations support their associated statements. We adapt that distinction here. Their 2023 systems and historical error rates should not be treated as current product performance. Evaluating Verifiability in Generative Search Engines

For each claim/source pair, record full support, partial support, no support or unverifiable. Preserve the source passage and date where permitted. Our pair-level tally is an operational simplification, not a reproduction of the paper's complete metric. Where several citations jointly support a claim, record that combined judgment separately.

Support and truth require separate checks. A citation may accurately reproduce an outdated price. In that case, the answer is supported by its source but wrong about the current offer. Correct the owned source and record the factual error. If an external source is wrong, document the discrepancy and pursue an appropriate correction without assuming the publisher will accept it.

Turn each important problem into an evidence brief:

Observed problem	Useful source improvement	Acceptance check
Integration described too broadly	List supported objects, direction and limitations	Product owner verifies each capability.
Migration effort unclear	Publish a scoped walkthrough and required inputs	A reader can identify prerequisites and exclusions.
Comparison overlooks a tradeoff	Explain where each approach fits	Criteria and supporting evidence are inspectable.

The original GEO paper tested source modifications against defined visibility metrics. Its main experiments used GPT-3.5-turbo; a Perplexity extension supplied source files. Those results motivate testable hypotheses about evidence presentation, but do not establish a traffic uplift from editing a live website. GEO research paper

Improve material because it resolves a real uncertainty. Adding a statistic without checking its population or relevance can make the page look documented while making the advice worse.

5. Connect discovery to commercial outcomes

Pew examined 68,879 Google searches from 900 US adults, using March 2025 browsing and search results reconstructed in April. Traditional-result clicks occurred on 8% of visits with an AI summary and 15% without one; summary-source clicks occurred on 1% of visits with summaries. This observational comparison does not isolate a causal effect or establish present-day B2B conversion rates. It does show why answer exposure and website visits need separate measurement. Pew's study and method

Define how your analytics identifies AI referrals and retain that rule's version. Test a real visit where feasible. Keep the landing page, observed source, consent state where available and subsequent meaningful actions within your existing analytics design. Do not collect sensitive prompt text to make a dashboard more detailed.

In Google Analytics, first-user, session and event traffic-source dimensions answer different attribution questions. Choose the scope that matches the report and document it. Google Analytics scope documentation

For enquiries, retain both observed acquisition data and an optional response to “Where did you first hear about us?” A buyer's recollection is useful additional evidence. It should not overwrite the recorded source. Direct traffic can reflect unavailable referral information, so an unexplained rise is insufficient to reclassify sessions as AI-generated demand. Google Analytics traffic-source processing

Worked example: fictional migration software business

The figures below are invented to demonstrate interpretation. They are not Ranketize or client results. The team runs 12 questions across two surfaces on three occasions per period: 72 attempts before and 72 after improving its migration documentation. Each period contains the same number of days.

Observation	Before	After
Attempts / planned attempts	72 / 72	72 / 72
Valid answers	60	63
Answers mentioning the brand	15	21
Answers citing an owned page	8	12
Observed AI-referred sessions	24	37
Distinct enquiries from those sessions	2	3
Enquiries qualified by review date	1	1

Brand presence among valid answers moves from 25% to 33.3%. Against all attempts, presence moves from 20.8% to 29.2%. Both belong in the record because availability changed. The team must inspect surface-level results and failure reasons before explaining the movement.

The defensible conclusion is that this sample contained more mentions and owned citations, while observed referrals increased. Qualified enquiry counts remained unchanged at the review date. Documentation changes, existing demand, other marketing and platform behavior remain plausible contributors. The next decision is to inspect source accuracy and enquiry quality, then continue observation. These counts do not justify claiming that GEO increased revenue.

6. Run a 30-day improvement cycle

Days 1-5: define and instrument

Agree on the buyer decision, qualification rule and collection procedure. Select the question set and priority source pages. Check technical eligibility, analytics classification and contact-form completion. Make one test submission identifiable as a test and exclude it from business reporting. Save the initial content versions and document other planned marketing activity.

Days 6-10: collect and diagnose

Run the frozen questions on separate days using the same procedure. Inspect the material claims and linked sources. Export the available native reports with date and aggregation notes. Rank issues by their effect on a customer's decision, the strength of available evidence and the effort required to correct them.

Choose one to three interventions. For example, replace an ambiguous integration paragraph, add a verified migration walkthrough and remove an obsolete price from a comparison. Assign an owner and acceptance check to each. Avoid changing every page simultaneously if you want an interpretable record.

Days 11-20: publish and verify

Publish accepted changes through the normal release process. Record the exact URLs, deployment date and changed claims. Check the live pages, internal links and the next step offered to readers. Record observed crawler visits or indexing evidence when available; deployment does not establish that a search system has incorporated the revision.

Keep unrelated product launches, campaigns and site changes in the same intervention log. They may matter when interpreting the next measurement.

Days 21-30: repeat and decide

Repeat the question procedure and inspect whether answers cite the revised material. Compare counts, availability, representation errors, referrals and qualified enquiries. Align reporting windows and annotate incomplete recent data. If the revised content has not been observed in the system, mark the exposure interval unknown rather than calling the intervention unsuccessful.

Choose a concrete outcome: correct another documented evidence gap, repair measurement, extend observation, or stop a low-value activity. Thirty days is an operating cadence, not a promised time to traffic or sales. Preserve useful page improvements even when their AI contribution remains uncertain.

7. Copy the worksheets and document the limits

These templates can be copied into a shared document or spreadsheet. Keep the observation table separate from the claim audit so an answer with several claims does not inflate your answer count.

Buyer question and evidence brief

Field	Your entry
Buyer and decision	
Exact question and question ID	
Origin: customer record or constructed	
Decision stage and brand/non-brand group	
Existing source URL and responsible owner	
Missing evidence and acceptance check	
Intended next step for the reader	

Observation and citation record

Field	Your entry
Observation ID, question ID and run	
Product surface, mode, locale and account context	
Date, time and output reference	
Valid answer, no answer, refusal or collection failure	
Brand mentioned, context and owned-domain citation	
Claim ID and exact factual claim	
Cited URL, source passage and inspection date	
Pair verdict and combined-source verdict if relevant	
Independently verified factual error, if any	
Actual reviewer, disagreement and resolution	

Intervention and decision log

Field	Your entry
Problem and evidence supporting it	
Changed page, claim and version	
Acceptance check, owner and publication date	
Observed indexing or retrieval evidence	
Baseline and follow-up windows	
Other changes and measurement gaps	
Outcome, next action and decision owner	

For a larger descriptive study, an optional protocol is 30 questions, three surfaces and three runs: 270 planned observations. This study has not been conducted for this edition. Freeze the questions and coding instructions, retain failures, and publish actual coverage. If a second reviewer is available, independently code a stratified sample and document disagreements; never claim independent review without that work.

This design describes a selected question set under recorded conditions. It cannot estimate consumer market share, isolate the causal effect of content changes or represent every personalized answer. Multiple runs help expose inconsistency, but do not transform a convenience sample into a representative survey.

Before circulating findings, check every headline against its denominator, dates and evidence class. Label a hypothesis as a hypothesis, an illustrative example as illustrative and unavailable data as unavailable. Preserve unsuccessful interventions. The resulting record should let another person understand both the recommendation and the circumstances that would change it.

8. Sources and edition notes

Public sources checked on 9 September 2026. Living documentation can change; verify the relevant sections when repeating this workflow.

1. Google Search Central: optimizing for generative AI features - technical eligibility, source retrieval and Google-specific guidance.
2. Google Search Console: Generative AI performance report - impression definitions, dimensions and reporting limitations.
3. OpenAI: crawler documentation - search, training and user-triggered access distinctions.
4. OpenAI: web search documentation - API citations and consulted sources.
5. Liu, Zhang and Liang: Evaluating Verifiability in Generative Search Engines - citation evaluation framework, 2023 revision.
6. Aggarwal and colleagues: GEO, Generative Engine Optimization - experimental visibility metrics, 2024 revision.
7. Pew Research Center: Google users and AI-summary clicks - historical observational behavior study, July 2025.
8. Google Analytics: scopes of traffic-source dimensions - attribution scope definitions.
9. Google Analytics: campaigns and traffic sources - source processing and direct traffic.