Evidence, editorial refresh and citation of answer pages

Protocol version 2, frozen for registration November 2, 2026. Prospective, three-arm controlled study. No baseline answers have been collected. Registration records a plan, not peer review.

Question and scope

Does an evidence treatment change how often live answer pages are retrieved, cited or associated with Broadcastwell, beyond the effect of a similarly sized editorial refresh? We compare evidence, placebo and unchanged pages on our own public documentation. The test estimates the effect of the complete evidence package in this selected set. It cannot isolate the contribution of a statistic, quotation, source link or opening paragraph separately.

We do not claim priority as the first experiment in the field. A bounded literature search cannot establish that no prior controlled experiment exists. The public title will describe the design without a priority claim.

Sampling and randomization

The revised frame contains 77 general pages across measurement (30), GEO basics (19) and content fixes (28), all in the current 300-page served library. Client pages, named-provider procurement, local services, platform-specific traffic and topics outside the added evidence are excluded before assignment. The revision uses the historical page-impression file received before registration; it is not based on answer outcomes. No baseline answer has been requested.

The sample is 45 pages, 15 per arm and 15 per cluster. The only eligible pages with observed positive Web impressions are one in each cluster: 5, 2 and 4 impressions over August 1 through October 31, 2026. The export is capped and supplies no complete retrieval-potential measure. Missing rows remain UNKNOWN, never zero. All three observed-positive pages are included; within each cluster, shuffle sorted known and unknown lists with the SHA-256-derived integer seed Broadcastwell evidence field test 2026-11-02 protocol 2, take the known page followed by the first 14 unknown pages and divide in that order into five blocks of three. Selection is randomized within this constrained frame, rather than a representative sample of all pages.

Randomly permute treated, placebo and control within each of the 15 blocks. Redraw the full assignment until the three observed-positive pages occupy different arms. Thus each arm contains one observed-positive and 14 UNKNOWN pages, and each block contains all three arms. This is limited balance on the available historical impression evidence, not exact matching on baseline retrieval. Source SHA-256 values use UTF-8 text without a byte-order mark and LF newlines, so checkout line endings do not change identity. The source frame, selection, source hashes and assignment are frozen before any answer measurement. No substitutions after registration; retain an unavailable page and flag it under intention to treat.

The 45 questions are the pages' current public question headings, copied exactly except for typographic apostrophe normalization. Each question is paired with one exact canonical page URL. The same list is used in both waves. Requests must not name Broadcastwell, supply the target URL, ask for a citation to us, or disclose the assigned arm. Those additions would condition retrieval on the desired outcome.

Arms and treatment specification

The direct answer is already the first paragraph on every selected page and is retained in all arms. The treatment therefore does not identify an answer-placement effect. Treated pages add two paragraphs, 90 to 140 words, before the existing What this means for you explanation. Each addition supplies a relevant, additional statistic with count, base, observation period and linked primary table, one exact short quotation from our CC BY 4.0 study, and the limits needed to interpret the statistic. No existing factual text is removed. This is the full evidence package being tested, not an isolated quotation or statistic effect.

Placebo pages add two paragraphs of the same scale, expanding and reflowing the existing explanation without a new factual proposition, figure, quotation or source. Both arms therefore add text; a treatment-length increase is not compared with a zero-length placebo reflow. The exact additions are 118 to 127 words; within each block, the treated and placebo additions differ by at most four words. The original explanation remains below the addition in both arms. Both modified arms receive the same modification-date convention at the single release. Control pages receive no source, date, navigation, metadata or content changes.

The appendix supplies all 30 exact proposed additions, their placement, existing source hashes and exact proposed text. The new evidence uses the two completed descriptive studies: October 2026 page retrieval of September citation URLs and the September 2026 B2B software answer archive. Their results do not establish an effectiveness claim. The relevance review checks the fit between each page's question and the newly stated measurement boundary; source-choice counts do not become accuracy or repeat-stability estimates.

Exact proposed content and before/after hashes will be included in the registered appendix. The complete protocol and appendix must be accepted and frozen before any baseline or edit. No page is edited before baseline validation.

The freeze runs from accepted registration until the post wave completes. Other than the registered edits, all sampled pages and site-wide presentation affecting them are held fixed. Necessary security or truth corrections override the freeze, are dated, and are analyzed as deviations. Changes to indexing submissions, links, promotion, domain configuration or the research pages can cause spillovers; record them. Do not intentionally promote individual sample pages or submit different arms for recrawling at different times. The report, claim audit, academic access page, protocol page and their discovery links are planned context publications before baseline. Record their completed release times; baseline must not begin until those publications finish. After baseline starts, hold this context fixed through post collection except for necessary truth or security corrections, which remain deviations. Publishing the common protocol does not authorize any sampled-page edit.

Engines, modes, location and repetitions

The engines are ChatGPT in web-search mode, Claude in web-search mode, Perplexity in web-search mode, Google AI Overviews in its normal search surface, and Google AI Mode. These are five distinct products/modes. No sixth-engine label is used.

Use a declared United States location where supported and a United States network country. Record the actual network country, declared location granularity, language (English), account/search state and any unsupported location setting. Do not invent an unexposed setting. A mismatch is a protocol deviation, not a silent correction. Hold the configuration constant across waves and record service changes.

Three independently initiated repeats per question, engine and wave are the fixed minimum: 675 scheduled answer attempts per wave, 1,350 across baseline and post. Repeats are interleaved with randomized question order, and arms are mixed within short time blocks rather than measured sequentially by arm. Fix the order before each wave. A failure is a retained scheduled attempt; it is not replaced with a favorable return. A single documented service retry may be recorded as a linked diagnostic record but never replaces the scheduled outcome in the primary analysis.

Seven repeats are permitted only by a prospective amendment accepted before baseline starts, with unchanged page sample and both waves funded. Do not use a baseline estimate to decide to add repeats. There is no optional second post wave in the confirmatory design. A later exploratory wave needs its own dated amendment and is labeled exploratory.

Timeline

Registration acceptance and archival completion precede baseline. Baseline spans no more than 72 hours. After validation, all treated and placebo edits deploy in one release, defining time zero as the completed public deployment in UTC. The post window begins 21 full days after time zero and completes within 72 hours, with the same daily time bands as baseline where practical. Twenty-one days allows recrawl and indexing lag; it does not guarantee either. Late or incomplete post capture is disclosed and retained. The fixed edit date is November 9, 2026, with a release window from 14:00 to 15:00 UTC (09:00 to 10:00 Eastern standard time). Baseline opens November 6 at 14:00 UTC and must finish and validate before that edit time. Post opens exactly 21 full days after actual deployment completion, therefore November 30 between 14:00 and 15:00 UTC, and closes 72 hours after that opening. Record the actual deployment completion and derived post opening; never measure before 21 full days have elapsed. If the fixed edit window cannot be met, do not edit late under this plan: accept a dated prospective amendment with a new calendar before any baseline capture or edit. If a baseline already exists, retain it, disclose the timing failure and register a new confirmatory schedule before a new wave.

Outcomes and denominators

Report these separately, overall and for every engine:

1. Retrieval: the exact target page appears among exposed consulted sources. When the engine does not expose consulted sources, use null with not_exposed, never zero. Its eligible denominator is attempts for which retrieval exposure exists, with the full scheduled base printed beside it.

2. Citation: the answer explicitly cites the exact target page URL or a verified canonical equivalent. Primary citation probability uses all scheduled attempts, including no search, no overview, no citation, refusal and service failures as observed unsuccessful citation outcomes. Report a valid-return sensitivity separately.

3. Mention: the answer text contains the name Broadcastwell under a frozen case-insensitive exact-name rule. A source URL alone is not a textual mention. Retain all scheduled attempts in this base.

4. Position: first appearance of the target page among ordered cited sources, if order is exposed. Use reciprocal rank with zero for uncited attempts as the unconditional prominence outcome; give conditional rank separately among cited returns. Missing exposed order is null, not an invented rank.

The exact-page URL rule lowercases host, strips fragment and standard tracking parameters, and normalizes a terminal slash only where the server proves equivalence. It does not merge different paths, subdomains, language versions, translated pages or query-dependent documents. Record original URL, resolved URL, canonical decision and reason. Redirect checks are ordinary document reads, never new answer requests.

No overview, no search, empty answer, refusal, blocked retrieval and service error are distinct outcome codes. Exclude only a demonstrably duplicated diagnostic record or a malformed record with no recoverable scheduled identity; print all excluded counts and reasons. The scheduled attempt still appears as missing/failure in the base. No data point disappears because it is unfavorable.

Analysis fixed before data

For each page and engine, compute the mean of the three scheduled binary outcomes in each wave and the post-minus-baseline difference. For citation and mention, give each of the five engines equal weight within page, then each page equal weight within arm. For retrieval and reciprocal rank, a page-engine change requires an exposed, non-null outcome in both waves. Average the available paired engines equally within page and retain pages with at least one such pair; print the usable page, engine and attempt bases. If no paired exposure exists, report the contrast as not estimable, not zero. This exposure-conditional secondary estimate is distinct from the full-scheduled-base citation outcome. The confirmatory effect is the treated-minus-placebo difference in these page changes. The placebo-minus-control contrast is the freshness check. Also print treated-minus-control and the three within-arm changes, explicitly as secondary descriptions. Do not pool retrieval-exposure denominators with citation bases.

Uncertainty uses a stratified block bootstrap with 10,000 resamples of five blocks within each of the three clusters, for 15 resampled blocks per draw, retaining all pages, engines, repeats and paired waves within each selected block. Give percentile 95% intervals and the number of usable blocks. This preserves page clustering and blocking. A randomization-inference sensitivity permutes the three arm labels within each block and rejects assignments that do not retain one observed-positive page per arm, respecting the original conditional allocation. It never permutes repeats as independent units. Engine-specific estimates resample the same blocks. Source-position sensitivity is descriptive and cannot replace the citation outcome.

Citation is the sole primary confirmatory outcome and its treated-minus-placebo contrast the primary test. Retrieval, mention and reciprocal-rank results retain their own estimates and intervals and are labeled secondary. Use Holm adjustment across the four treated-minus-placebo outcome tests in any joint claim; never choose the significant engine or outcome after inspection. Raw per-engine estimates remain visible even when their uncertainty crosses zero.

The main analysis is intention to treat by frozen arm. A per-protocol sensitivity may flag unauthorized page changes, availability problems or location deviations, while showing all excluded blocks and retaining the main estimate. No regression adjustment is selected after seeing outcomes. No engine is dropped because its observed result differs.

Detectable effects and precision

With 15 pages per arm, five engines and three repeats, each arm contributes 225 scheduled attempts per wave. These are not 225 independent pages. A planning approximation at two-sided 5% and 80% power is 2.80 * sqrt(4*p*(1-p)*DE/(15*15)), where DE=1+14*rho and rho is within-page correlation. This deliberately assumes zero beneficial baseline/post correlation and is an approximation, not a power guarantee for the block-bootstrap test.

At p=0.05, the approximate minimum detectable absolute contrast is 8.1 percentage points for rho=0 and 12.6 points for rho=0.10. At p=0.50 it is 18.7 points for rho=0 and 28.9 points for rho=0.10. At rho=1 the p=0.50 approximation is 72.3 points. A very small true effect will not be resolved by this design. Print these assumptions with the results. Estimate the realized uncertainty from the registered analysis, never call a wide null interval proof of no effect. Planning arithmetic checked November 2, 2026.

Claim tiers before outcomes

• If the citation treated-minus-placebo interval is above zero, with a consistent intention-to-treat result and no decisive fidelity/implementation defect: supports an effect of this registered evidence package on citation for the sampled pages and dates. It suggests replication elsewhere. It does not support a universal tactic, mechanism, client promise or revenue gain.

• If placebo-minus-control is above zero while treated-minus-placebo is not: supports an observed editorial-refresh contrast in this setting; evidence-specific benefit remains unresolved. Do not relabel freshness as evidence success.

• If the primary interval crosses zero: supports the recorded estimate and uncertainty; the effect is unresolved at the measured precision. No guarantee of absence or equivalence.

• If the primary interval is below zero: supports a reduction under this package in this setting. Publish it with the same title prominence and tables.

• If retrieval changes without citation, or mention changes without retrieval: report the separate stage-specific pattern. It does not establish a total benefit or substitute for the primary outcome.

• Severe implementation, missingness or configuration failure: supports a failed or compromised field test, with the actual observations and deviations. Publish the failure; do not claim a clean causal result.

Every result page carries: Observed answers on stated dates, not a ranking.

Integrity, publication and data

No hidden instructions, fabricated figures, review solicitation, client pages or third-party page edits. Verify all factual treatment additions against public primary sources and preserve semantic qualifications. The quote is brief and attributed. An evidence review before baseline checks exact diffs, word counts and unchanged commercial terms; it is not an outcome review.

Publish the protocol, accepted registration DOI, frozen page and arm list, source and treatment hashes, outcome data, analysis specification, exclusions, deviations and complete results regardless of sign. The report distinguishes acceptance timestamp, baseline dates, edit date and post dates. Registration records a plan and is not peer review. Protocol amendments are new dated records beside the frozen registration. The study uses the public question heading of each listed page, unchanged between waves.

Appendix A. Frozen assignment

The historical Web-impression period is August 1 through October 31, 2026. UNKNOWN means no matching retained export row, not zero impressions. All pages were served at selection. The downloadable assignment contains the full source hashes.

Block	Arm	Page	Historical Web impressions
measurement-1	treated	Which AI visibility tools can a buyer verify?	5
measurement-1	placebo	What can Search Console show about AI visibility queries?	UNKNOWN
measurement-1	control	Why can a page rank in Google without being cited by ChatGPT or Perplexity?	UNKNOWN
measurement-2	control	What counts as an own-domain citation in an AI answer?	UNKNOWN
measurement-2	treated	What should a client AI visibility report contain?	UNKNOWN
measurement-2	placebo	How can I check a tool's historical AI visibility data?	UNKNOWN
measurement-3	treated	How should I check an AI baseline before changing content?	UNKNOWN
measurement-3	control	Who has documented results from GEO work in ChatGPT?	UNKNOWN
measurement-3	placebo	How can a buyer verify the conclusion?	UNKNOWN
measurement-4	placebo	How do we compare citation metrics when changing AI tools?	UNKNOWN
measurement-4	control	How can I check AI search visibility?	UNKNOWN
measurement-4	treated	How many repeated runs make an AI visibility rate useful?	UNKNOWN
measurement-5	control	Which AI visibility platforms offer historical data and time-series reporting?	UNKNOWN
measurement-5	placebo	How does an AEO or GEO audit differ from an SEO audit?	UNKNOWN

Block	Arm	Page	Historical Web impressions
measurement-5	treated	Can AI answers get company facts wrong?	UNKNOWN
geo-basics-1	control	What does unverifiable mean in an AI product answer?	2
geo-basics-1	placebo	Can I use AI to do my SEO?	UNKNOWN
geo-basics-1	treated	Do brand-name prompts measure discovery?	UNKNOWN
geo-basics-2	treated	Does AI search replace SEO?	UNKNOWN
geo-basics-2	control	Is Perplexity good for SEO?	UNKNOWN
geo-basics-2	placebo	What does LLM in AI mean?	UNKNOWN
geo-basics-3	treated	What is an AI software shortlist?	UNKNOWN
geo-basics-3	control	What does GEO mean in AI search?	UNKNOWN
geo-basics-3	placebo	Does being named by AI prove product quality?	UNKNOWN
geo-basics-4	treated	Are failed AI answers the same as a brand absence?	UNKNOWN
geo-basics-4	placebo	Does an AI shortlist show what buyers actually purchase?	UNKNOWN
geo-basics-4	control	Can AI search use information from the live web?	UNKNOWN
geo-basics-5	control	What does AI visibility mean?	UNKNOWN
geo-basics-5	placebo	Why does the buyer question matter to an AI shortlist?	UNKNOWN
geo-basics-5	treated	Why can the same AI question produce different names?	UNKNOWN
fixes-1	placebo	How should a comparison page handle unknown product features?	4
fixes-1	treated	What does publishing llms.txt prove about AI visibility?	UNKNOWN
fixes-1	control	How should I audit third-party sources before requesting a correction?	UNKNOWN
fixes-2	control	What does Google's guidance say about llms.txt?	UNKNOWN
fixes-2	placebo	What is an llms.txt file?	UNKNOWN
fixes-2	treated	Is llms.txt actually used?	UNKNOWN
fixes-3	placebo	How should I check an llms.txt file?	UNKNOWN
fixes-3	treated	How should a page distinguish native features from integrations?	UNKNOWN
fixes-3	control	How do AI search crawlers differ from AI training crawlers?	UNKNOWN

Block	Arm	Page	Historical Web impressions
fixes-4	treated	How can I generate llms.txt from existing pages?	UNKNOWN
fixes-4	control	What is an llms.txt checker?	UNKNOWN
fixes-4	placebo	How should I choose content for SEO, AEO and GEO?	UNKNOWN
fixes-5	treated	How should I request an AI fact correction with a source receipt?	UNKNOWN
fixes-5	control	Can Google AI Overview answers be changed?	UNKNOWN
fixes-5	placebo	What should I include in an llms.txt file?	UNKNOWN

Appendix B. Exact proposed additions

These additions are part of the prospective plan. Insert each pair of paragraphs before the existing explanatory section and retain the original text. The direct answer already appears first and remains there in all arms. Both modified arms use the same modification-date convention; controls have no edits.

measurement-1: treated

Which AI visibility tools can a buyer verify?

Added text (118 words):

Check the event behind a visibility number before treating it as evidence about a source. Our registered citation study found qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. The base is named pairs, rather than every attempted answer or every citation.

The study states, “A citation describes a source choice.” A third-party source can accompany a vendor name without qualifying as that vendor's own domain. Ask whether the tool preserves that distinction and the underlying observation. This descriptive result does not compare monitoring products or establish that any particular tool is accurate. It illustrates why an export's event definition belongs beside its displayed rate.

measurement-1: placebo

What can Search Console show about AI visibility queries?

Added text (122 words):

Use the query table for the question it can answer: which reported searches suggest a topic for investigation. Keep the reported wording, dates and metrics together so that a reader can identify the demand signal being discussed. If the question concerns an engine's answer, preserve that observation separately rather than substituting a search-table entry.

Apply the same care when describing a missing event. A filtered query-table result should retain its own scope instead of becoming a statement about the whole site's traffic. Where tables use different populations or aggregation rules, keep that difference explicit. Share only an aggregate that preserves the accounting when a raw query contains personal or commercial detail; the person's identity is not needed to explain the reported table.

measurement-2: treated

What should a client AI visibility report contain?

Added text (120 words):

Keep an answer-level base beside a citation-level base in the report. Our registered citation study retained 2,095 eligible target answers from the September 2026 B2B software archive; 39 had empty citation arrays, or 1.9%. Those answers contribute to the answer denominator but supply no entries to a citation-occurrence denominator.

The study states, “A citation describes a source choice.” A source count therefore needs its own definition, alongside the answer count and the exclusions. These observations describe the archived questions across five engines. They do not measure client visits, sales or the effect of content work. Preserve those boundaries when writing a result sentence so that an empty source array does not disappear merely because another table counts citations.

measurement-2: placebo

How can I check a tool's historical AI visibility data?

Added text (123 words):

Read each historical point as a dated record with a definition. Establish whether its underlying answer was collected at that time, calculated later from retained answers or reconstructed under a changed method. Preserve those distinctions when the earlier and later periods appear beside one another. A line on a chart does not resolve an unavailable underlying record.

Keep the original file when a calculation changes, and describe the new calculation without silently changing what the earlier report meant. Check whether the same buyer questions and engine set support the comparison. If they do not, present separate dated series or state the unresolved boundary. Ask which records remain available for inspection after access ends so that a later reader can understand the historical point.

measurement-3: treated

How should I check an AI baseline before changing content?

Added text (120 words):

Fix the citation rule as well as the question list before changing content. Our registered citation study found 39 empty citation arrays among 2,095 eligible target answers, or 1.9%, in the September 2026 B2B software archive. A rule that silently drops those answers changes the base before any later comparison begins.

The study states, “A citation describes a source choice.” Keep source selection separate from support for a sentence and from a textual brand mention. These descriptive counts do not establish baseline stability or the required number of repeats; the repeat evidence above addresses a different question. They show a concrete denominator decision to freeze and retain when the same questions are measured again after a page change.

measurement-3: placebo

How can a buyer verify the conclusion?

Added text (124 words):

Read the conclusion together with the dates and the retained classification. The answer quotation and the later source check describe different observations. Preserve the distinction between an incorrect historical statement and one classified as stale rather than presenting both as a single current-price conclusion. A later review date does not change the capture date of the answer.

Keep each flagged statement identifiable when several observations support it. A question, a repeat and a supporting quotation remain evidence for the statement; they need not become additional flagged statements. The reference field should lead to the page relevant to that particular claim. Where a statement remains unverifiable, retain that label beside the other outcomes instead of converting uncertainty into either a confirmed error or a zero.

measurement-4: placebo

How do we compare citation metrics when changing AI tools?

Added text (123 words):

Start the comparison by matching the event and the base, then check how the number is stored. A naming measure and a cited-domain measure describe different observations even when a display uses the same general label. Preserve the definitions before joining the earlier and later reports so that the comparison does not depend on a familiar column name.

Keep an unmatched field unresolved when the necessary definition or supporting records are unavailable. A prominence measure cannot be reconstructed merely by renaming another column. The collection conditions also belong beside the comparison: questions, engines, dates, repeats and exclusions define what the record covers. If those conditions change, describe the new series without presenting the definitional change itself as a client's observed gain or loss.

measurement-4: treated

How many repeated runs make an AI visibility rate useful?

Added text (120 words):

Retain a no-citation answer when calculating a citation rate across repeats. Our registered citation study recorded 39 empty citation arrays among 2,095 eligible target answers, or 1.9%, in the September 2026 B2B software archive. Removing these answers would make the denominator depend on whether a source was returned.

The study states, “A citation describes a source choice.” The archived count is not an estimate of repeat agreement and supplies no universal repeat threshold. Its relevance is the outcome rule: repeated observations need the same eligibility definition, including valid returns with no source. Keep the repeat count, observation dates and exclusions beside any later rate. A different collection can have a different empty-array share without demonstrating a content effect.

measurement-5: placebo

How does an AEO or GEO audit differ from an SEO audit?

Added text (125 words):

Match the audit task to the observation being discussed. If the question concerns a buyer criterion in an answer, preserve the passage that states it and where it appears in the response. A record of whether a supplier name appears would leave that criterion unexamined. The relevant evidence is the captured wording rather than a broader audit label.

Keep passages and answers as separate counting units. Several retained quotations can come from a single returned answer, so they should not silently become additional answers. The technical page checks address a different part of the work: whether a page can supply accessible, supported information. Keep those checks beside the answer evidence while preserving the distinction between the page's technical state and the buyer criterion actually observed.

measurement-5: treated

Can AI answers get company facts wrong?

Added text (122 words):

A cited source still needs to be checked against the particular company claim. Our registered citation study counted qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. That classification records naming and domain matching, rather than whether a feature sentence is true.

The study states, “A citation describes a source choice.” It follows that the source-choice count cannot certify the description attached to a name. Inspect the dated passage and preserve any unresolved mismatch. These data do not estimate a company-fact error rate or show that own-domain citations are more accurate than third-party references. The point is to keep the observed link and the factual verification as separate records.

geo-basics-1: placebo

Can I use AI to do my SEO?

Added text (126 words):

Verify the input claim before using a generated draft as the basis for an editorial decision. A mistaken allowance can change a comparison before any decision about its title, headings or links. Follow the cited reference and check the precise proposition rather than treating the presence of a citation as a completed factual review.

Keep the answer capture date separate from the later product-page check. A historical reference can explain why a study classified a statement without establishing today's terms. Preserve the distinction between a numerical ceiling and an unlimited allowance when that is the disagreement under review. A broad description of a generous plan would lose the detail needed to understand the source comparison. The next editorial step follows the checked claim and its scope.

geo-basics-1: treated

Do brand-name prompts measure discovery?

Added text (122 words):

Keep the question condition attached when reporting whether a vendor was named or cited. Our registered citation study found qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. The base is already-named pairs under the archived matching rules.

The study states, “A citation describes a source choice.” That conditional percentage cannot be read as a generic discovery rate, and supplying a brand in a question changes the condition being investigated. Preserve the question text beside the result. The study is descriptive and does not test the causal effect of adding a brand to a request. It provides a concrete example of why a conditioned denominator needs a conditioned interpretation.

geo-basics-2: treated

Does AI search replace SEO?

Added text (126 words):

Separate visible page characteristics from evidence that an optimization caused selection. Our registered page study classified 3,913 of 4,420 successfully captured target URLs as definitely optimized, or 88.5%, using its stated precedence rules in October 2026. The URLs came from citations in the September B2B software archive.

The study states, “Presence does not prove consumption.” Its categories describe observed signatures on later-captured pages; they are not treatment assignments or a comparison with pages that were never cited. The result therefore cannot establish that AI search replaces SEO or that adding a signature increases citation probability. Keep technical access, factual page quality and the later answer observation as distinct checks. This source supplies descriptive context for that separation, rather than a universal publishing prescription.

geo-basics-2: placebo

What does LLM in AI mean?

Added text (125 words):

Keep the explanation of a language model separate from the verification of a company claim. Tokens, context and attention describe how text is represented and generated. They do not by themselves establish that a current product page was consulted or that a statement about an account is supported by the relevant source.

When a returned answer provides a link, inspect the dated passage for the particular proposition. Retain the answer's capture date and the reference check date as separate observations. That distinction matters when a historical company allowance differs from what the retained vendor page states. The abbreviation LLM does not resolve the disagreement. The source review concerns the actual claim and the available evidence, while the architectural explanation answers how the model produces text.

geo-basics-3: treated

What is an AI software shortlist?

Added text (124 words):

Keep a named option separate from the source attached to it. Our registered citation study found qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. A pair could be named without a qualifying citation to that vendor's own domain.

The study states, “A citation describes a source choice.” The percentage is conditional on naming, not a measure of suitability, product quality or purchase. A third-party page may still be cited when the vendor's domain is absent. Read the returned list with those separate observations intact, then inspect the claim and the buyer's requirements. These archived questions do not represent every buying situation or establish that a particular shortlist option deserves selection.

geo-basics-3: placebo

Does being named by AI prove product quality?

Added text (125 words):

Evaluate an appearance and the description attached to it as separate observations. First establish that the product was named; then inspect whether the important claims are supported by dated evidence. A stated feature can justify a narrower description without settling whether the product performs well for the buyer's actual requirements.

Keep both records when a named product is described incorrectly. Correcting the description is a factual task, not a request to certify quality or a conclusion that the product is poor. A comparison should preserve the distinction between stated features, tested functionality and unknown evidence. Use the returned name as a reason to investigate where relevant, then follow the requirement and the available evidence. The frequency of the name alone does not complete that evaluation.

geo-basics-4: treated

Are failed AI answers the same as a brand absence?

Added text (124 words):

Preserve the distinction between an excluded attempt and an eligible answer with no citation. Our registered citation study retained 2,095 eligible target answers in the September 2026 B2B software archive after five target exclusions. Among eligible answers, 39 had empty citation arrays, or 1.9%; the excluded attempts were retained in a separate audit.

The study states, “A citation describes a source choice.” An empty array records a source outcome for an eligible answer, whereas a failed attempt does not establish a brand absence. Keep both states visible with their own bases. These counts describe this archive and its eligibility rules. They do not make every service failure a valid answer or establish that missing citations and missing brand names are equivalent events.

geo-basics-4: placebo

Does an AI shortlist show what buyers actually purchase?

Added text (124 words):

Describe a returned shortlist as a record of the options presented in an answer. Keep later buyer actions in their own records, with the dates and definitions that make those events identifiable. The person may investigate an option, reject the list or make no purchase. The list alone does not distinguish those possibilities.

Apply the same boundary when reporting a business outcome. A suitable recommendation still leaves price, requirements, procurement and the person's evaluation unresolved. Preserve visits, self-reports, opportunities and sales as separate events rather than renaming an appearance measure. A later comparison can connect records only through a stated, supported link. If that link is absent, the contribution remains unknown. The answer sample can remain useful while retaining that limit on its interpretation.

geo-basics-5: placebo

Why does the buyer question matter to an AI shortlist?

Added text (125 words):

Read the buyer condition before applying a category-wide result to a particular decision. A question about residential builders retains that work type as part of what was measured. Team size, budget and other needs can define a different question, even when both questions sit within the same broad software category.

Keep that wording stable when comparing dates. Replacing the question with an easier or broader one changes the observation being compared. Preserve a missing name as a record of the captured answers under the stated condition, with the uncertainty that accompanies the repeated runs. It does not establish whether the product would suit the particular buyer. The useful interpretation stays attached to the question and the retained observations rather than to the category label alone.

geo-basics-5: treated

Why can the same AI question produce different names?

Added text (127 words):

Fix what counts as a name and what counts as a source when comparing repeated answers. Our registered citation study counted qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. Naming and source selection were separate recorded events.

The study states, “A citation describes a source choice.” A change in sources therefore need not be a change in the returned vendor set, and neither observation alone identifies why an answer changed. This descriptive analysis does not measure repeat stability or prescribe a sufficient repeat count. Retain the repeated question's wording, dates and separate outcome definitions. The earlier repeat evidence remains the evidence for variation; this later count clarifies an additional boundary in interpreting it.

fixes-1: placebo

How should a comparison page handle unknown product features?

Added text (124 words):

Keep a comparison cell tied to the precise scope of its supporting page. A statement about a suite, connected tool or different plan may leave the buyer's narrower question unresolved. Review the product, module, region and date before choosing a source-state label, and keep the qualification near the comparison rather than hiding it elsewhere.

Use source silence carefully. If the reviewed page does not state a capability, record that limit instead of treating it as a universal test of the software. An unknown remains a question for further evidence. A later suitable primary source can support an update when the claim and review date change together. Retaining the earlier evidence makes the correction traceable without turning the unresolved cell into an invented feature conclusion.

fixes-1: treated

What does publishing llms.txt prove about AI visibility?

Added text (125 words):

A published overview file records availability at a particular address, rather than demonstrated use in an answer. Our registered page study found llms.txt present at 774 of 1,600 target host roots, or 48.4%, during October 2026 retrieval of hosts cited in the September B2B software archive. Another 85 root-file states remained unknown.

The study states, “Presence does not prove consumption.” This is a descriptive host observation, not a controlled comparison of publishing the file. It cannot establish that a particular service requested the file, used its links or cited a page because of it. Keep the served-file check separate from later request and answer evidence. A successful publication can be verified directly while the file's contribution to a generated answer remains unresolved.

fixes-2: placebo

What is an llms.txt file?

Added text (125 words):

Give the overview a clear scope before shortening the material it describes. A documentation file should identify the covered part of the site, and a link description should lead to the relevant detailed page. The concise overview and its destinations need to retain the same product boundaries and dates so that shortening does not alter the claim.

Keep the access check separate from that editorial check. An overview does not grant permission to fetch restricted content, and its presence does not establish that a linked page can be retrieved successfully. Follow the destinations and compare the descriptions with what a person reads there. A file can be concise while still making the covered path and the purpose of its selected links clear to the reader.

fixes-2: treated

Is llms.txt actually used?

Added text (124 words):

Distinguish finding a file from observing its use. Our registered page study found llms.txt at 774 of 1,600 target host roots, or 48.4%, during October 2026 retrieval of hosts cited in the September B2B software archive. The root-file states also included 741 absent and 85 unknown cases under the stated response rules.

The study states, “Presence does not prove consumption.” These data describe file availability on the captured hosts. They do not record whether an answer engine requested an overview or relied on its contents. A request log and an answer observation would address different events and need their own dates and definitions. Preserve the unknown state instead of treating a blocked or ambiguous response as proof that no file exists.

fixes-3: placebo

How should I check an llms.txt file?

Added text (122 words):

Record what was checked before calling the file valid. Begin with the public address and its actual response, then inspect the intended text and follow the destinations. A successful retrieval and accepted syntax do not establish that a description matches the current page, the relevant plan or the study the link promises to explain.

Keep those outcomes distinct in the receipt. If a destination is unresolved, leave that state visible rather than hiding it in an overall score. After a deployment, compare the served text with the approved version and revisit changed links. The check then identifies which static facts were confirmed and which remain unresolved. Any later statement about use, citation or recommendation still requires its own observation and stated method.

fixes-3: treated

How should a page distinguish native features from integrations?

Added text (124 words):

Check the cited passage for the exact product boundary, even when the answer supplies a link. Our registered citation study found qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. The matching rule identifies a name and its qualifying domain, not the delivery method of a feature.

The study states, “A citation describes a source choice.” A source link alone cannot establish whether a capability is native, supplied through an integration or unresolved. Inspect the passage's module, plan and date before using it to fill that comparison. This descriptive count is not a native-feature accuracy test and does not turn an unmatched own-domain citation into proof that a capability is missing.

fixes-4: treated

How can I generate llms.txt from existing pages?

Added text (127 words):

Generate an overview that accurately describes the linked pages, then verify what is publicly served. Our registered page study found llms.txt at 774 of 1,600 target host roots, or 48.4%, during October 2026 retrieval of hosts cited in the September B2B software archive. Eighty-five root-file states remained unknown under the response rules.

The study states, “Presence does not prove consumption.” The observation does not compare file generators, validate their summaries or establish a citation benefit from creating a file. Your generation check therefore still needs to compare each description with its destination and retain the page's scope and date. File availability, summary accuracy and later use in an answer remain separate questions. This count supplies context for that boundary, rather than an effectiveness promise.

fixes-4: placebo

How should I choose content for SEO, AEO and GEO?

Added text (125 words):

Let the missing buyer detail define the content task. If the question concerns an accounting integration, preserve the named system, the information that moves and the documented limits as the matters to investigate. A broader article can provide context without resolving that narrower product question. The format should follow the evidence the reader needs.

Keep the underlying question consistent across the SEO, AEO and GEO brief when they concern the same decision. Repeating the same answer under different acronym labels does not supply distinct information. Also preserve the distinction between eligibility and content: a technical access state and an unanswered accounting detail are different problems to establish. The next page change should follow the actual missing fact or technical condition identified in the retained evidence.

fixes-5: treated

How should I request an AI fact correction with a source receipt?

Added text (125 words):

Submit evidence for the particular disputed statement, rather than treating a source link as the conclusion. Our registered citation study found qualifying own-domain citations in 2,608 of 8,085 named vendor-answer pairs, or 32.3%, in the September 2026 B2B software archive. This count records source selection under domain-matching rules, not successful factual corrections.

The study states, “A citation describes a source choice.” A correction receipt must still identify the dated passage, the claim it supports and the remaining uncertainty. These descriptive data do not test whether a correction request will be accepted or change an answer. Keep the original observation and the later response as separate records. A documented request is evidence of the request; any claimed correction needs a subsequent, retained observation.

fixes-5: placebo

What should I include in an llms.txt file?

Added text (123 words):

Choose each link for the task it helps the reader complete. A pricing destination, method page and product page serve different purposes, and the description should make that scope clear. Preserve the relevant date or historical label when shortening an approved description so that the overview does not become more confident than its destination.

Keep the selected set small enough to maintain accurately. Group the essential company, product and research destinations by their purpose rather than reproducing the full navigation. Supporting detail can stay behind the appropriate link when a short heading would conceal its qualifications. Recheck a changed summary against the human-facing page. The file is another publication of the same facts and needs to retain their meaning, scope and permission boundaries.

Sources and rights

• Broadcastwell research team, Priya Nair, Ananya Rao and Vikram Patel. (2026). AI search visibility: source evidence and two registered analyses. Version DOI: https://doi.org/10.5281/zenodo.23804758. CC BY 4.0.

• Broadcastwell. (2026). Absence Index, release September 2026. https://doi.org/10.5281/zenodo.22907695.

• Study 1 frozen plan: https://doi.org/10.17605/OSF.IO/DVYFU.

• Study 2 frozen plan: https://doi.org/10.17605/OSF.IO/VCZKT.

Prepared by the Broadcastwell research team. Lead: Vikram Patel. License: CC BY 4.0.