seogirl

Playbook, November 2026

Claim to Source Verification

A practical method for checking whether the evidence actually supports the words on your page, or in an AI answer.

Working draft · v0.1 · research-informed method

ONE SENTENCE, THREE SEPARATE CLAIMS

ON THE PAGE

"This change increased bookings by 25%."

Cited source: booking dashboard, two periods

- Counts: 100 to 125 — Supported
- Cause: no control group — Unsupported
- Period: not stated — Unresolved

A citation shows where a claim came from. This playbook checks whether the source actually **backs every part** of the sentence.

06	05	07
review steps	support decisions	real claims checked

Ameneh Saeednia
Senior SEO and Generative Search Strategist

Research synthesis

A citation is the start of the check

A link makes a statement traceable. To judge support, open the source and compare the **exact claim, subject, timeframe and conditions**. A credible source can still be irrelevant to a particular sentence.

Evidence from primary sources	How this playbook uses it
The ALCE research separates answer correctness and citation quality. [1]	Assess the source-to-claim link separately from whether the statement is independently true.
Google's self-assessment guidance asks about reliable sourcing and original value. [2]	Record evidence and contribution. Do not treat a bibliography alone as original research.
Google's structured-data policies require representative, non-misleading information. [3]	Keep the evidence review upstream of markup. Machine-readable repetition does not verify a claim.

Where the idea comes from

ALCE, an **EMNLP 2023 benchmark**, scored AI answers on correctness and citation quality as two separate things. I borrow that split. The benchmark is about older models, so it says nothing about today's failure rates.

What this playbook adds

A **claim register** you can run on your own pages and on captured AI answers. It's the claim-level companion to *Proof on the Page*, and I tested it on my own blog first (page 4).

The short version

Coverage asks whether each material claim **has** a source. Support asks whether that source **backs** that claim. A page can score high on the first and poorly on the second.

The six review steps

Extract, compare, classify, repair

Steps 1 to 3 make the claim and the source visible before anyone judges them. Steps 4 to 6 turn the comparison into a decision, a fix and a dated readback.

1

Split compound claims

Break a sentence into independently checkable statements. Separate a **measured outcome**, its **cause** and a claim about the wider market.

Evidence: Claim ID, exact words, subject, metric, timeframe and conditions.

2

Open the actual cited source

Record the source version and relevant passage. Distinguish **original research** from commentary, summaries and repeated copies.

Evidence: Source URL, title, author or publisher, date and precise locator. An inaccessible source remains unresolved.

3

Compare the complete meaning

Check whether the evidence supports the **subject, amount, direction, timeframe and limits** of the claim. Do not fill gaps with plausible assumptions.

Evidence: A brief supporting or conflicting passage, plus an explanation of the match.

4

Classify the source-to-claim link

Choose **Supported, Partly supported, Unsupported by cited source, Contradicted** or **Unresolved**. Explain the decision rather than assigning a decorative score.

Evidence: The exact missing condition or conflict, not a general statement that the source is weak.

5

Repair the statement or the evidence

Narrow the wording to the supported scope, obtain the missing source, remove an unsupported claim or **label analysis as analysis**.

Evidence: Approved revised wording, source and reviewer. Never invent a client result or first-hand activity.

6

Recheck the published version

Confirm the final words and their linked source after implementation. For an AI answer, **preserve the original capture** and record a separate later observation.

Evidence: Dated final URL or capture, remaining limits and owner.

Useful distinction

A supported quotation can still describe a disputed or outdated assertion. Keep **citation support** and **overall factual confidence** as separate notes.

Real example: I ran it on my own site

Six claims passed. One didn’t.

Before publishing this playbook, I checked my October 2026 post on **Bing's AI Performance report** against the two Microsoft announcements it cites, then checked a results claim on my own homepage and About page. Checked 3 November 2026.

Claim on my page	Source passage	Decision
AI Performance covers Copilot, AI summaries in Bing and select partner integrations.	“across Microsoft Copilot, AI-generated summaries in Bing, and select partner integrations” [5]	**Supported**
Grounding queries are phrases used to retrieve cited content, and the data shown is a sample.	“The data shown represents a sample of overall citation activity.” [5]	**Supported**
Citation Share is distinct from traffic share, ranking and content quality.	“It does not… represent traffic share, or assign quality scores to content.” [6]	**Supported**
The June 2026 update added Intents, Topics, Citation Share and Compare in preview; labels can be broad.	“four new capabilities available in preview…”; “some labels may still be broad” [6]	**Supported** (3 claims)
Homepage: “90% traffic recovered.” About page: “about 100% of lost visibility” restored.	No source linked. Two different figures and **two different metrics** for what reads as the same project.	**Unresolved**

What the Bing check showed

All six source-backed claims matched what Microsoft actually wrote. That's the expected result for a careful post, and it's the point: **a passing review is evidence too**. It tells a reader the citations were actually opened.

Repair for the results claim

Choose **one metric** (organic traffic or visibility), state the **period and data source**, and use the same number on both pages. Until then the claim stays unresolved, not wrong.

Two fictional examples

How wording outruns evidence

The most common gap: the evidence supports something **smaller** than the sentence claims.

Claim	Available evidence	Decision and repair
"This change increased bookings by 25%."	A dashboard shows 100 bookings in one period and 125 in another. No control or attribution analysis is provided.	The change in counts is supported. **Causation is not established.** State the periods and observed change; describe the intervention separately.
"Every guest returns to the ship on time."	One review says the guest returned early. The itinerary lists a planned return time.	The universal guarantee is **unsupported**. State the scheduled return and actual policy, if verified. Do not turn one review into a guarantee.

Handling numerical evidence

Check the **denominator and period** before reporting a rate. A change from 100 to 125 is a 25% relative increase in counts. A conversion rate moving from 4% to 5% rises by 1 percentage point, or 25% relatively. Neither comparison identifies a cause by itself.

Handling research summaries

Keep **sample size, selection, geography, model versions, observation dates and exclusions** attached to the finding. Do not rewrite an observed sample result as a universal rule. If a study measures citations, do not report it as bookings or revenue.

Both scenarios are invented for teaching.

Template

The claim register

One material claim per row. The register is the deliverable: it makes every decision traceable.

Field	What to capture
Claim ID / exact words	One material claim per row. Preserve the original text.
Origin	Your page, draft, captured AI answer or third-party statement.
Claim type	Identity, product fact, measured result, causal assertion, recommendation or analysis.
Source and locator	Original URL, date or version, relevant section and short passage.
Support decision	Supported / Partly supported / Unsupported by cited source / Contradicted / Unresolved.
Factual confidence note	Independent evidence, source reliability concerns or remaining disputes. Do not confuse this with entailment.
Scope / uncertainty	Who, when, where, sample, denominator, conditions and missing access.
Revision / owner / recheck	Concrete repair, person responsible and dated final verification.

Define your denominator

For a declared sample of 20 material claims, report how many were assessed and how many remain unresolved. If reporting a proportion, state whether unresolved claims are included. **Preserve counts** so a high percentage cannot hide excluded claims.

Sensitive or consequential claims

Route **health, legal, financial and safety** statements to an appropriately qualified reviewer. This method does not certify compliance or replace specialist judgment.

Pilot and limitations

Build a review you can repeat

Before turning this into a Skill, run it by hand on a small, declared sample.

Pilot step	Deliverable
Choose a bounded sample	Ten material claims from one page or one captured answer. State how they were selected.
Review independently	Two reviewers classify source support before reconciling differences.
Investigate disagreements	Clarify compound claims, implicit assumptions and inaccessible sources. Record the agreed rule.
Implement approved repairs	Revised wording and a source trail. Unresolved items stay visible.
Recheck	Compare the final wording to the same source version and capture any new evidence.

A small Skill test set

Include a **compound claim** supported only in part, a relevant but **inaccessible source**, a source that **contradicts** the sentence, an accurate quotation from a **weak source** and a **correlation rewritten as causation**. The expected output identifies the precise gap and preserves uncertainty.

What to measure

Reviewer disagreement, unresolved items and repaired claims within the declared sample. Fewer unsupported claims is an **editorial improvement**, not proof of better rankings or conversion.

Sources, checked 2 and 3 November 2026

Evidence and boundaries

01 **Enabling Large Language Models to Generate Text with Citations. Gao, Yen, Yu and Chen, EMNLP 2023**

ALCE evaluates fluency, correctness and citation quality. Historical benchmark context, not current platform performance.

https://arxiv.org/abs/2305.14627

02 **Creating helpful, reliable, people-first content. Google Search Central**

Primary self-assessment guidance on sourcing, expertise, original information and reader value. It does not provide this playbook's classification scale.

https://developers.google.com/search/docs/fundamentals/creating-helpful-content

03 **General structured data guidelines. Google Search Central**

Supports the boundary between accurate page representation and syntax validation. Markup is not independent evidence for a factual statement.

https://developers.google.com/search/docs/appearance/structured-data/sd-policies

04 **Proof on the Page. SEO Girl**

The broader proof-placement playbook this one builds on.

https://www.seogirl.ca/resources/files/proof-on-the-page.pdf

05 **Introducing AI Performance in Bing Webmaster Tools (Public Preview). Microsoft Bing, February 2026**

Source for the real example on page 4.

https://blogs.bing.com/webmaster/2026/2/Introducing-AI-Performance-in-Bing-Webmaster-Tools-Public-Preview/

06 **New AI Visibility Insights in Bing Webmaster Tools. Microsoft Bing, June 2026**

Source for the real example on page 4.

https://blogs.bing.com/search/2026/6/New-AI-Visibility-Insights-in-Bing-Webmaster-Tools-Intents-Topics-Citation-Share-Compare/

Scope

A method playbook by Ameneh Saeednia, prepared with AI assistance. The real example covers seven claims on seogirl.ca; everything else is a proposed workflow, not a study or a ranking score.