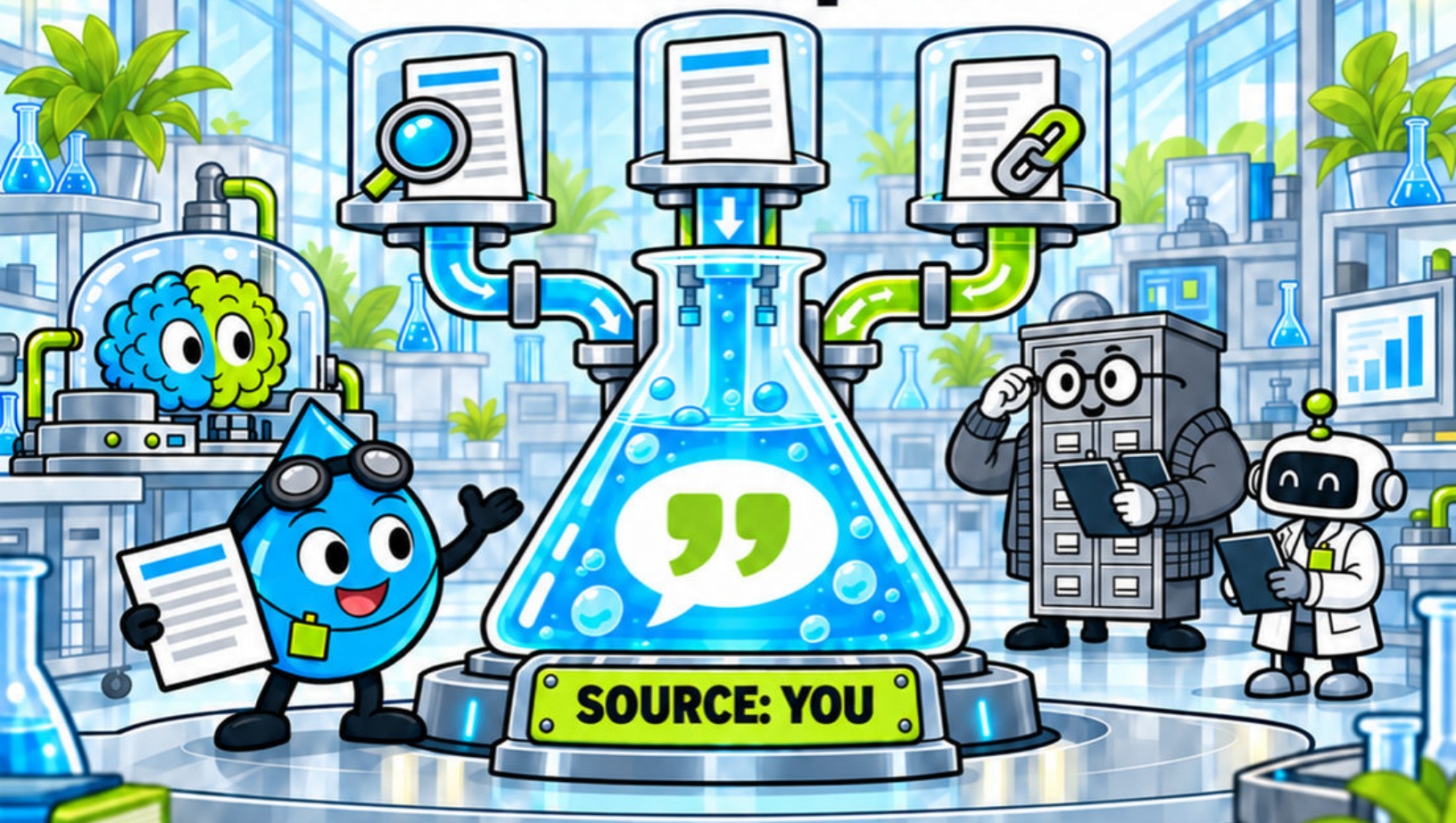


2026 FIELD MANUAL

THE COMPLETE GUIDE TO GEO

**Generative Engine Optimization
for SEO Experts**



**Inside GEO: How AI Finds,
Understands and Cites Content**

JOURDAN ROMBOUGH
OPTIMIZATION THEORY

VERSION 2.5.1

CONTENTS

START HERE

Why I Built This Lab	3
How to Use This Guide	4
Meet the Lab Crew	5
GEO in 60 Seconds	6
How an AI Answer Gets Made	7
Anatomy of a Citable Chunk	8
Lab Map	9

HOW GEO WORKS

● Room 01: The Library Floods	10
The Click Drain	11
Ways Back to the Web	12
The New Engine Room	13
SEO vs. GEO: What Changes	14
SEO Is the Power Plant	15
● Room 02: Inside the Answer Machine	16
Anatomy of an AI Engine	17
The Answer Pipeline: Part 1	18
The Answer Pipeline: Part 2	19
Query Fan-Out	20
Passages Need Context	21
The Crawl-to-Click Gap	22
Engines Consult Different Shelves	23
● Room 03: The Periodic Table of GEO	24
The Experiment That Started It All	25
From Vague to Specific	26
Do AI Answers Prefer Fresh Pages?	27
Busted Experiments	28

DOING GEO

● Room 04: Brand DNA	29
The Entity Molecule	30
The DNA Match Test	31
Repair Wrong Answers	32
Fill Real Decision Gaps	33
● Room 05: The Technical Airlock	34
The Badge Wall	35
Clear the JavaScript Fog	36
Specimen Labels	37
Different Files, Different Jobs	38
Publisher Controls: Choose the Right Lever	39
Publisher Controls: Read the Panel	40
Markdown for Agents	41
Markdown: What the Evidence Says	42
Who Pays for the Crawl	43
Who Pays for the Crawl: Your Settings	44
When the Agent Buys	45
When the Agent Buys: The Timeline	46
Can an Agent Complete the Task?	47
● Room 06: Growing Cultures	48
Web Mentions and AI Visibility	49

Help the Community	50
Useful Video	51
Search With a Camera	52
Evidence to Coverage	53
GEO for the Local Shop	54
● Room 07: Hazard Zone	55
Why Formulas Don't Travel	56
The Clean Room Pledge	57

MEASURING GEO

● Room 08: Under the Microscope	58
The Prompt Test Rack	59
Check the Answer, Then the Source	60
Scoring Answers and Sources	61
The Gauge Panel	62
The Tracker Bench	63
Built in This Lab	64
● Room 09: The 90-Day Protocol	65
Audit, Then Build	66
Grow, Measure, Repeat	67

THE TACTICS BOARD

The Tactics Board	68
Every Tactic at a Glance	69
Files, Formats and Access	70
Writing Tactics	71
Off-Site and Paid Tactics	72
The Black-Hat File	73
The Black-Hat File: Case by Case	74

FROM MY OWN LAB

90 Days of AI Traffic on vuss.io	75
----------------------------------	----

TOOLS AND REFERENCE

The Perfect GEO Page	76
The GEO Lab Master Checklist	77
Field Kit: Robots Policy	78
Field Kit: Entity Graph	79
Field Kit: Prompt Panel	80
Field Kit: GA4 AI Traffic	81
Field Kit: Markdown Alternates	82
Field Kit: Test Your Markdown Setup	83
Glossary and Lab References	84
Glossary: Newer Terms	85
Chunk Made It. Will Yours?	86
What Changed in v2.5	87
What Changed in v2.5.1	88
Source Drawer: 1 of 6	89
Source Drawer: 2 of 6	90
Source Drawer: 3 of 6	91
Source Drawer: 4 of 6	92
Source Drawer: 5 of 6	93
Source Drawer: 6 of 6	94
Back Cover	95

WHY I BUILT THIS LAB



Hi, I'm Jourdan.

Technical SEO for more than 15 years

Founder, Optimization Theory, Cooperstown, NY

Builder of AI Citation Trace, PageRank Simulator and Black Hat SEO Terminal

linkedin.com/in/jourdanrombough

I've worked in technical SEO for more than 15 years. For most of that time the goal was easy to explain: help the right page rank, earn the click, and give the visitor what they came for.

Then the answers started arriving before the links. People now ask ChatGPT, Gemini, Perplexity or Google's AI Mode and get a written reply, sometimes with a source attached and sometimes without one. I wanted to know how those replies are built: where the facts come from, why one page is quoted and another is passed over, and what a site owner can honestly do about it.

So I read the research, built small tools to test ideas, and checked my own sites. This guide is the result: GEO explained one piece at a time, with pictures, plain language, and a named source behind every number.

If you already know SEO, much of this will feel familiar, and the differences will stand out. Start at the front and go room by room. By the end you'll know how AI answers choose their sources, and what makes a page worth choosing.

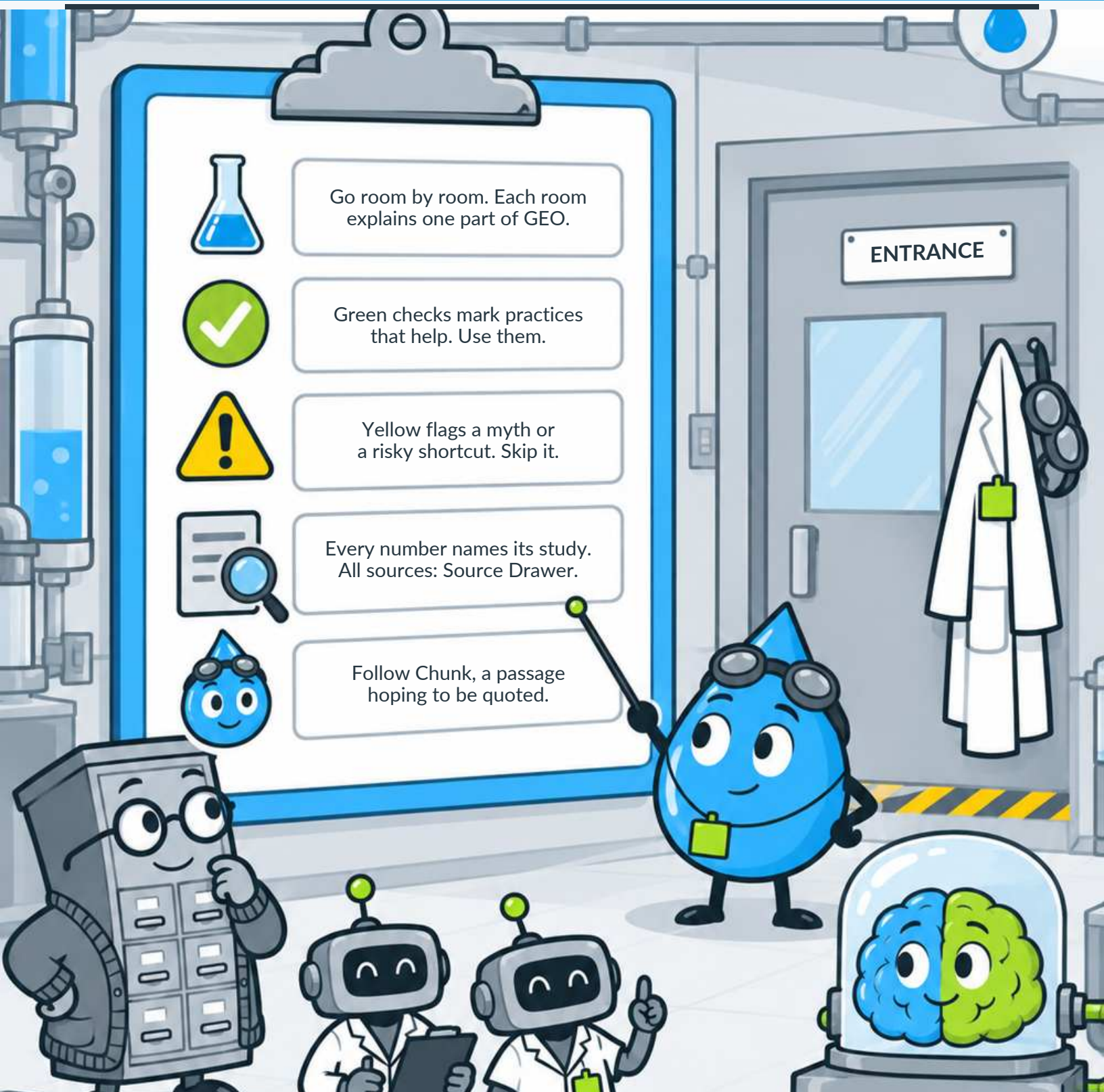
Jourdan Rombough

Version 2.5.1 | October 31, 2026



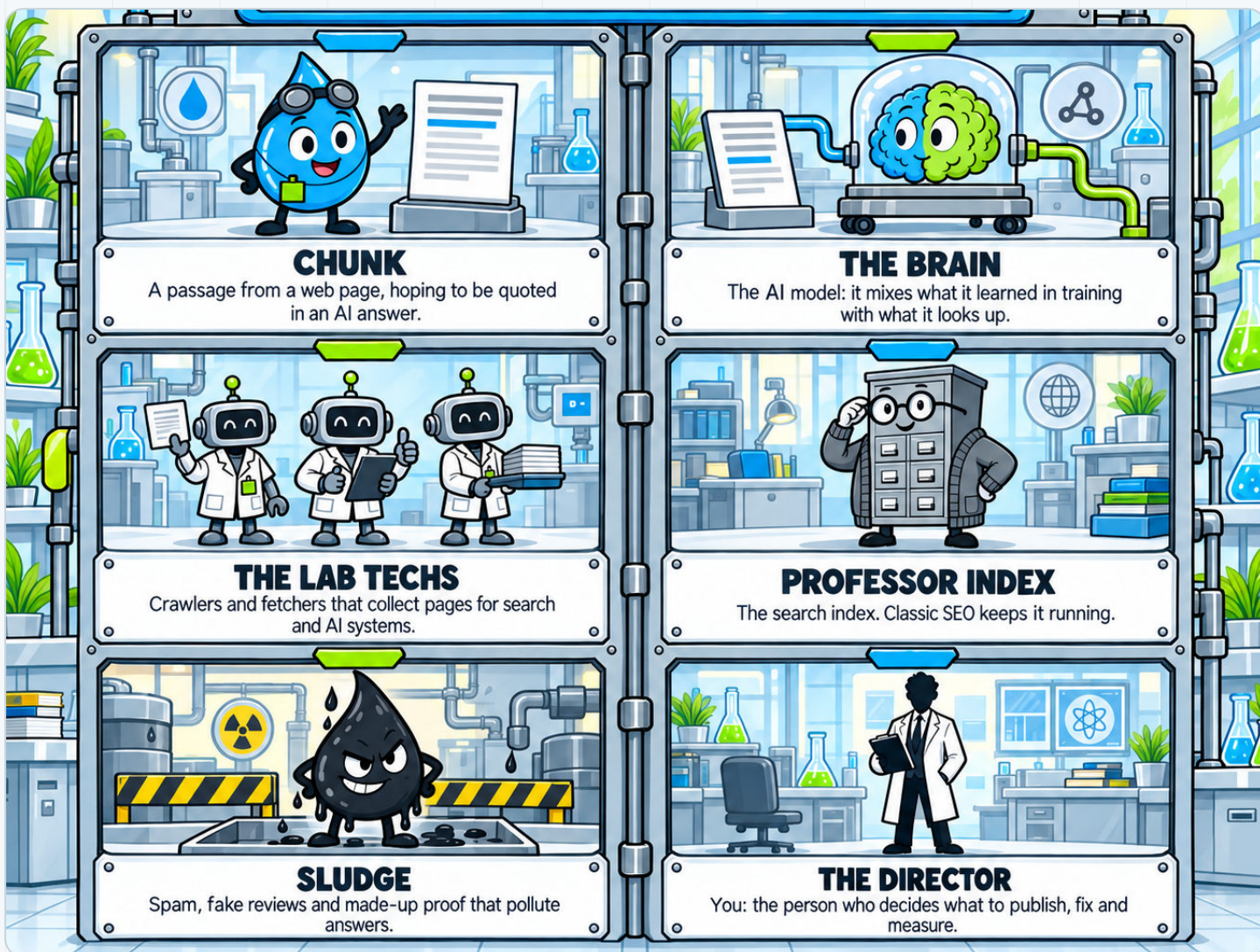
**JOURDAN BUILT THE LAB. I'M THE
PASSAGE YOU'LL FOLLOW THROUGH IT.**

HOW TO USE THIS GUIDE



This edition concentrates on publicly accessible English-language web content and a defined set of AI search surfaces. For other languages or markets, build a separate prompt panel, verify local product availability and test the relevant content and sources. Findings from one locale do not establish performance in another.

MEET THE LAB CREW



WHO THEY ARE IN REAL LIFE

- **CHUNK** is a passage on your page: a paragraph, list item or table row an engine can lift and quote.
- **THE LAB TECHS** are crawlers and fetchers such as Googlebot, Bingbot, OAI-SearchBot, ChatGPT-User, Claude-SearchBot and PerplexityBot.
- **SLUDGE** is spam, fake reviews, hidden prompts and cloaking. Its case file is the Black-Hat File on page 73.
- **THE BRAIN** is the language model behind ChatGPT, Gemini, Claude or Copilot, answering from training and from what it looks up.
- **PROFESSOR INDEX** is the search indexes answers draw on, such as Google's and Bing's. A page that isn't indexed is hard to retrieve.
- **THE DIRECTOR** is you: choose the questions, fix the pages and run the tests. Room 09 is the plan.

GEO IN 60 SECONDS

GEO (generative engine optimization) is the work of making your content easy for AI answer engines to find, understand, trust and cite. Six ideas cover most of it:

01 Start with the task

People ask AI tools to help them decide things: which drone to buy, which software to choose, which agency to hire. Find the questions your audience really asks. A page that directly helps with one gives an answer engine something to use.

[Room 01: The Library Floods, page 10](#)

02 Keep the route open

Before an AI system can quote a page, something has to fetch and read it. Check that the crawlers you want are allowed, pages load without errors, and key facts sit in the HTML. A blocked page can't be cited from a fresh visit, however good it is.

[Room 05: The Technical Airlock, page 34](#)

03 Make the evidence useful

Answers are stitched together from passages. Write ones that make sense alone: name the subject, answer plainly, add the number or example that proves it, and say where it came from. In the GEO benchmark, adding quotes, statistics and citations raised a simulated engine's visibility score.

[Room 03: The Periodic Table of GEO, page 24](#)

04 Keep identity coherent

Engines piece together who you are from your site, profiles, directories and press. When names, addresses or services disagree, answers can blur you with someone else or repeat old facts. Keep one approved set of facts and make every profile match.

[Room 04: Brand DNA, page 29](#)

05 Earn corroboration

AI answers often cite sources you don't control: reviews, Reddit, YouTube, news and research. You can't buy that trust. You can earn it by publishing findings people can check and by helping where your audience already talks.

[Room 06: Growing Cultures, page 48](#)

06 Measure the right outcome

One screenshot of an AI answer proves little; answers change from run to run. Track a fixed set of questions over time and count how often you're named, cited, accurate and visited. A mention isn't a click, and a click isn't a sale.

[Room 08: Under the Microscope, page 58](#)

First step: write down ten questions your audience asks and record how AI answers handle them today, before you change anything.

HOW AN AI ANSWER GETS MADE

Ask ChatGPT, Gemini, Perplexity or Google's AI Mode a question with search turned on, and several steps happen before the reply appears. Products differ, but most search-backed answers follow a pattern like this one.



1 ASK

A person asks a question.

2 SEARCH

The system runs several related searches.

3 RETRIEVE

It gathers candidate pages and passages.

4 SELECT

It keeps the passages that fit best.

5 WRITE

The model writes an answer from them.

6 CITE

Some sources get a visible link.

WHERE GEO FITS

BE FINDABLE

Stages 2 and 3. Pages that crawlers can reach and search can index, matching the sub-questions people ask. Rooms 02 and 05.

BE THE BEST PASSAGE

Stage 4. Self-contained passages with a named subject, a direct answer, evidence and a source. Room 03.

BE DESCRIBED RIGHT

Stages 5 and 6. Consistent facts about you across the web, and independent sources that back them up. Rooms 04 and 06.

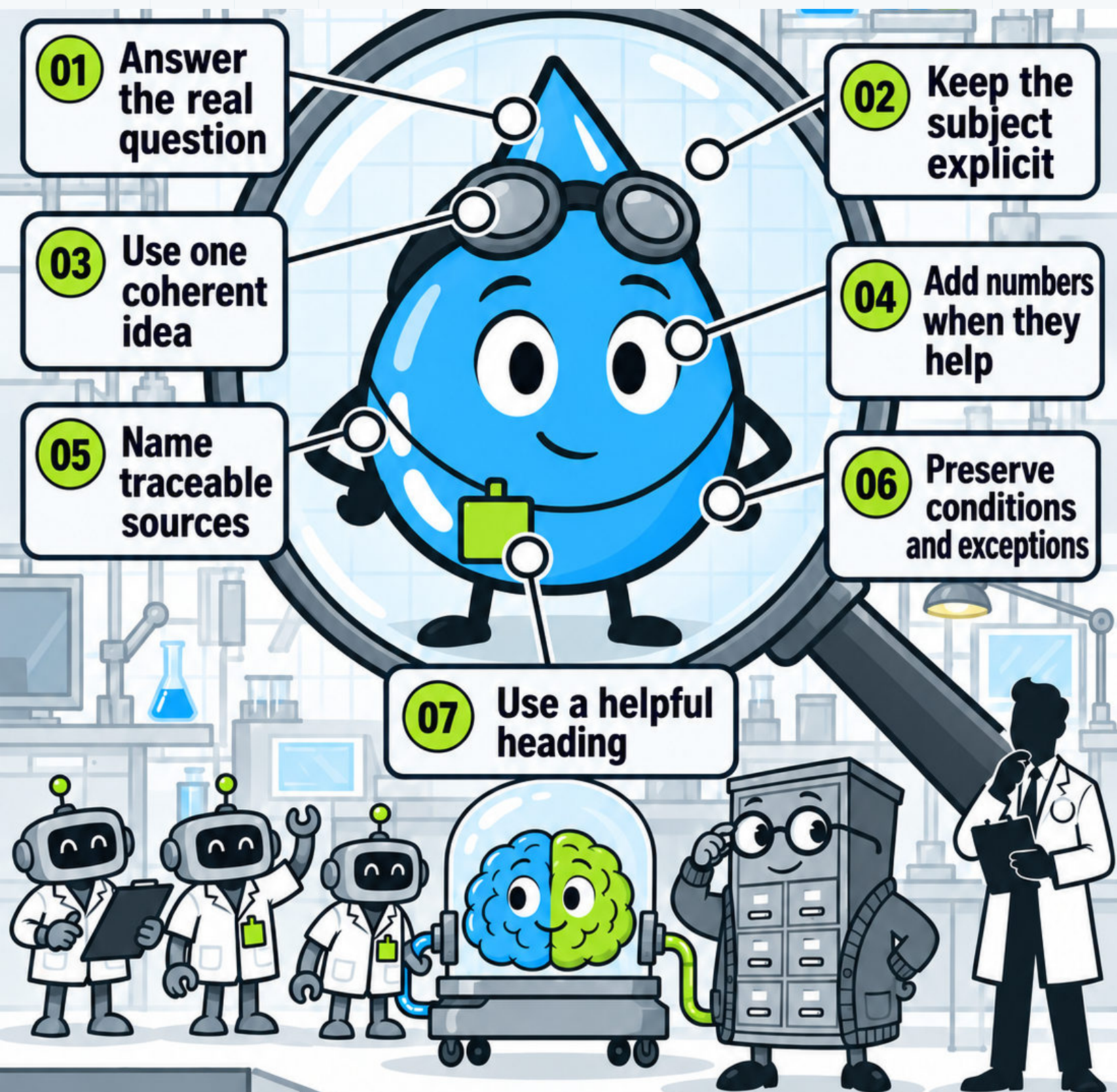


**I CAN'T WRITE THE ANSWER.
I CAN BE THE BEST EVIDENCE IN IT.**

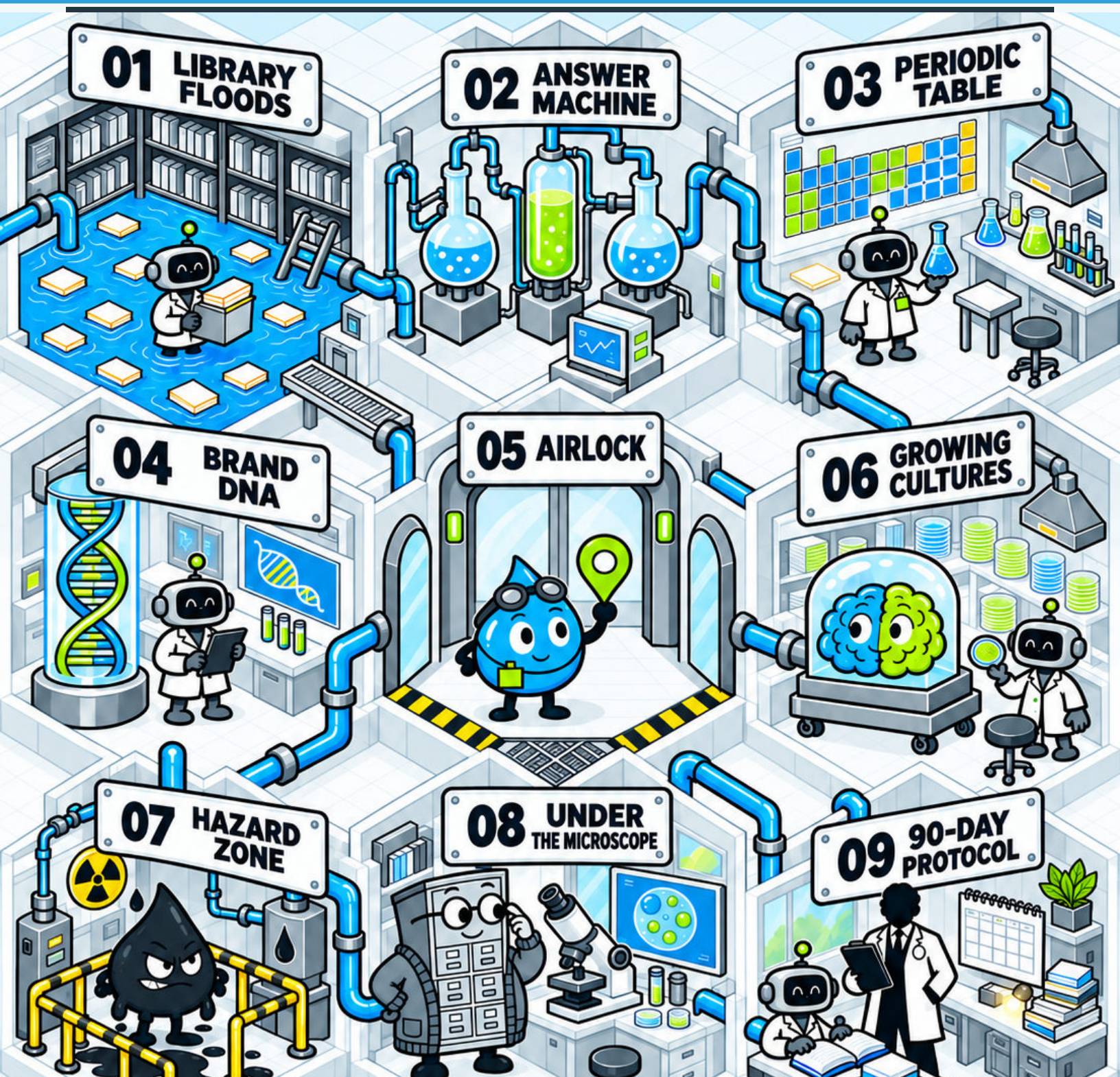
A simplified pattern. Some products answer from training alone, repeat steps, or show no links. Room 02 (page 16) walks through each stage.

ANATOMY OF A CITABLE CHUNK

AI answers rarely quote a whole page. They lift a passage, a "chunk", that makes sense on its own. These seven traits make a chunk easy to understand, check and cite. Google says its AI features need no content written or "chunked" specially for AI: these are simply traits of clear writing, and there is no required passage length.



LAB MAP



01 The Library Floods > p. 10

04 Brand DNA > p. 29

07 Hazard Zone > p. 55

02 Inside the Answer Machine > p. 16

05 The Technical Airlock > p. 34

08 Under the Microscope > p. 58

03 The Periodic Table of GEO > p. 24

06 Growing Cultures > p. 48

09 The 90-Day Protocol > p. 65

Every popular GEO trick, rated: page 68. Copy-ready field kits: page 78. Real traffic example: page 75.

THE LIBRARY FLOODS

01

WHAT'S HAPPENING

Search used to hand people a list of links. Now an AI answer often comes first: Google's AI Overviews and AI Mode, ChatGPT, Gemini, Perplexity. It summarizes several sources in one reply, and the reader may never scroll to the links.

WHY IT MATTERS

Ranking is no longer the only way to be seen, or a guarantee of a visit. To reach people now, a page has to be findable, useful enough to quote, and described accurately inside the answer.

JOURDAN'S TAKE

Rankings still matter. They just aren't the finish line anymore. Start with the decision your reader is trying to make, then check who's answering it for them.



ANSWERS NOW ARRIVE BEFORE THE LINKS. I WANT TO BE PART OF THEM.

IN THIS ROOM

11 The Click Drain

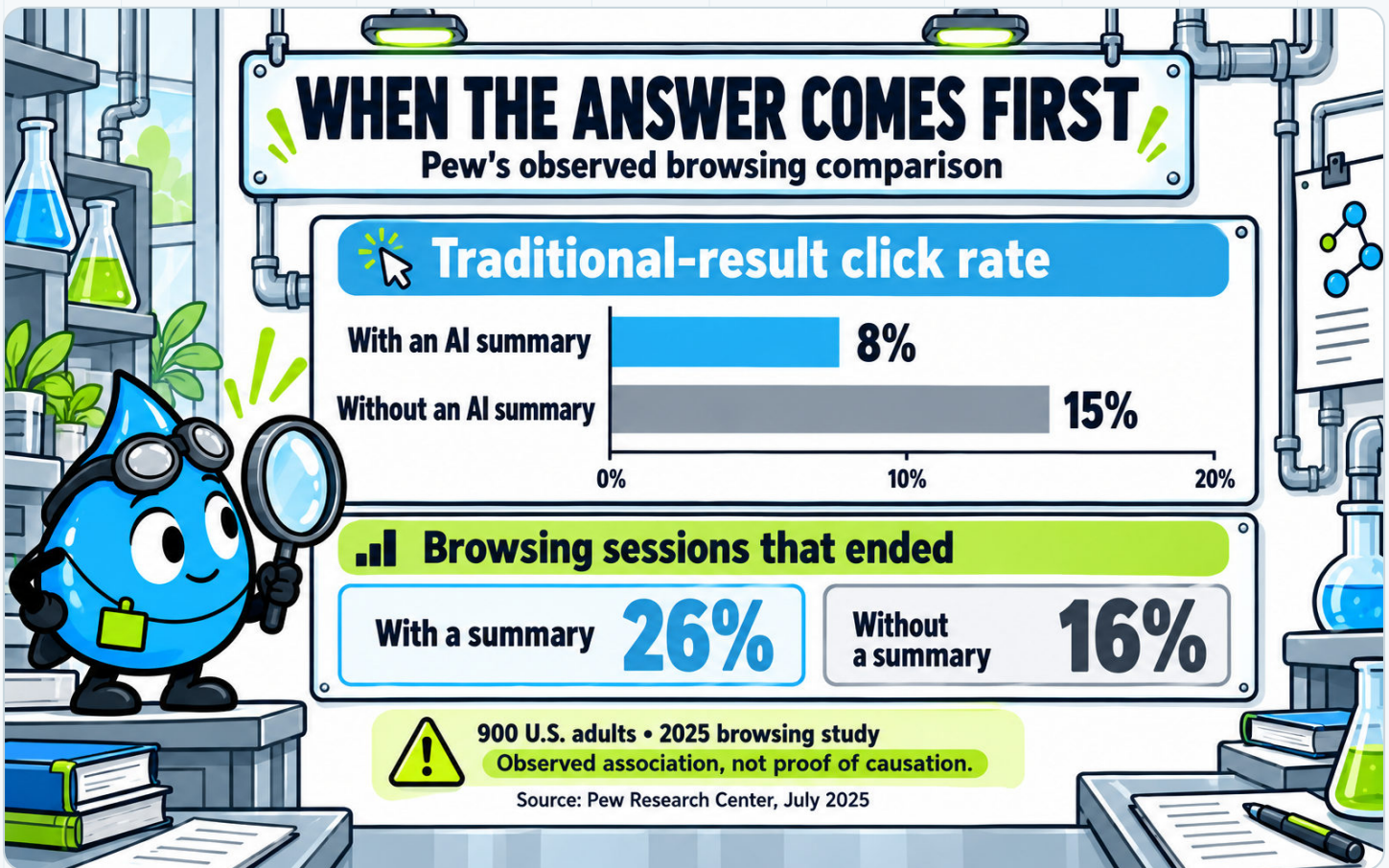
12 Ways Back to the Web

13 The New Engine Room

14 SEO vs. GEO: What Changes

15 SEO Is the Power Plant

THE CLICK DRAIN



58% LOWER DESKTOP CTR

Ahrefs: position-one CTR with an AI Overview versus a modeled counterfactual. Not a measured causal effect.

1% CLICKED A SUMMARY SOURCE

Pew: visits to summary-bearing results pages that clicked a link inside the summary itself.



A CITATION CAN BE VISIBLE EVEN WHEN THE CLICK NEVER COMES.

Ahrefs: data Dec. 2025, published Feb. 4, 2026; 300K keywords; modeled desktop comparison. Pew: data March 2025, published July 22, 2025; 900 U.S. adults. Both observational. Oct. 15, 2026: Google added more inline links to AI Mode and AI Overviews and began testing a section for official sites; it gave no click numbers. See *Ways Back to the Web* (p. 12).

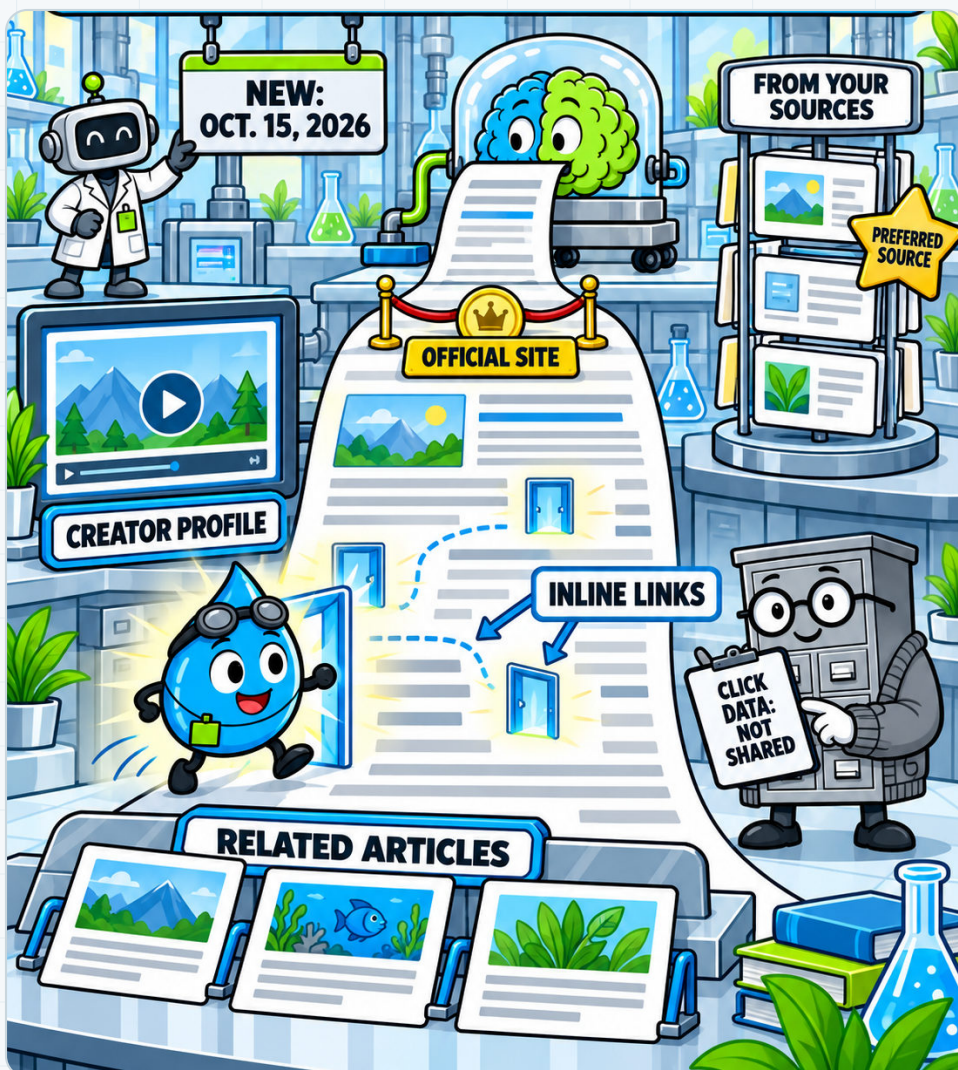
WAYS BACK TO THE WEB

7x

People using Preferred Sources have grown more than sevenfold since the start of 2026 (Google, Oct. 15, 2026).

WHAT GOOGLE CHANGED, OCT. 15, 2026

More inline links, placed higher in answers. A test section for official sites on task queries. News carousels in more places that highlight Preferred Sources. Related-article links with thumbnails. Clickable creator and publisher names on video results (test).



DO THIS

Write passages specific enough to earn the inline link. Make the one official page for each task (booking, support, pricing) unmistakable. If you publish timely articles, add Google's "Add as a preferred source on Google" button. Use one creator name across your videos and your site.



**A LINK INSIDE THE ANSWER IS A DOOR.
MAKE THE ROOM BEHIND IT WORTH ENTERING.**

Google shared no click data with the announcement. Compare your own Search Console clicks before and after.

THE NEW ENGINE ROOM

GOOGLE AI MODE

1B+

monthly active users

Company reported, July 2026

GEMINI APP

950M

monthly active users

Company reported, July 2026

CHATGPT

1.2B

weekly users

OpenAI DevDay 2026 recap, Sept. 29, 2026

ALSO INVESTIGATE

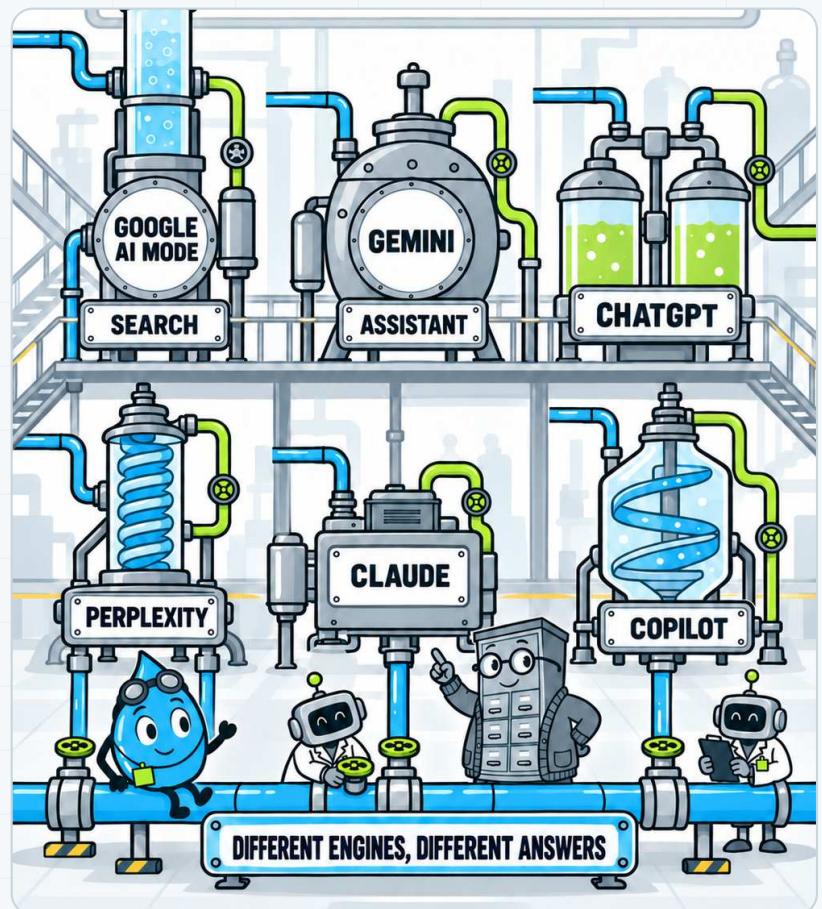
Claude, Perplexity and Copilot have their own audiences and search modes.

WHY IT MATTERS

Each product searches, selects and cites differently. Test the product and search mode your audience actually uses.

WHY THE NUMBERS DON'T ADD UP

Monthly users counts anyone active once in a month; weekly users counts a single week. Many people use several assistants, so the figures overlap. Read them as scale, not market share.



WHICH ENGINE IS MY AUDIENCE ACTUALLY USING?

Different metrics (monthly vs. weekly users) and time windows, and the audiences overlap, so these numbers cannot be added together. ChatGPT figure: OpenAI DevDay 2026 recap, Sept. 29, 2026.

SEO VS. GEO: WHAT CHANGES



SEO ASKS

Crawl and index access

Search bots can reach and store the page.

Useful pages and passages

Content that answers what the searcher wants.

Queries and customer tasks

Keywords mapped to what people need.

Relevant links and reputation

Other sites vouch for you with links.

Search clicks and outcomes

Rankings, clicks and conversions.

Performance over time

Rank tracking on a fixed keyword set.

GEO ALSO ASKS

Can this AI surface consult the source?

AI fetchers and search modes have their own access rules.

Does the answer keep the meaning?

A passage quoted on its own must still be accurate.

Which details need evidence?

Conversational questions add conditions (budget, place, use) to address.

Who else backs up the claim?

Answers lean on independent sources: reviews, communities, press.

Were you mentioned, cited, or visited?

Three different outcomes. Count each one separately.

How much do repeated answers vary?

The same question can return different sources. Sample it repeatedly.

These overlap: SEO foundations power GEO, and GEO adds questions on top. Neither replaces the other.

SEO IS THE POWER PLANT

When an answer engine searches the web, it needs pages it can find, read and trust. Google's own guide says optimizing for its AI features is still SEO: the same foundations apply.

CRAWLABLE

Bots can reach the page: robots.txt allows it, the server answers 200, and no login or script wall hides the text.

INDEXABLE

Search can store and show it: no stray noindex, a sensible canonical, and the URL in your sitemap.

USEFUL

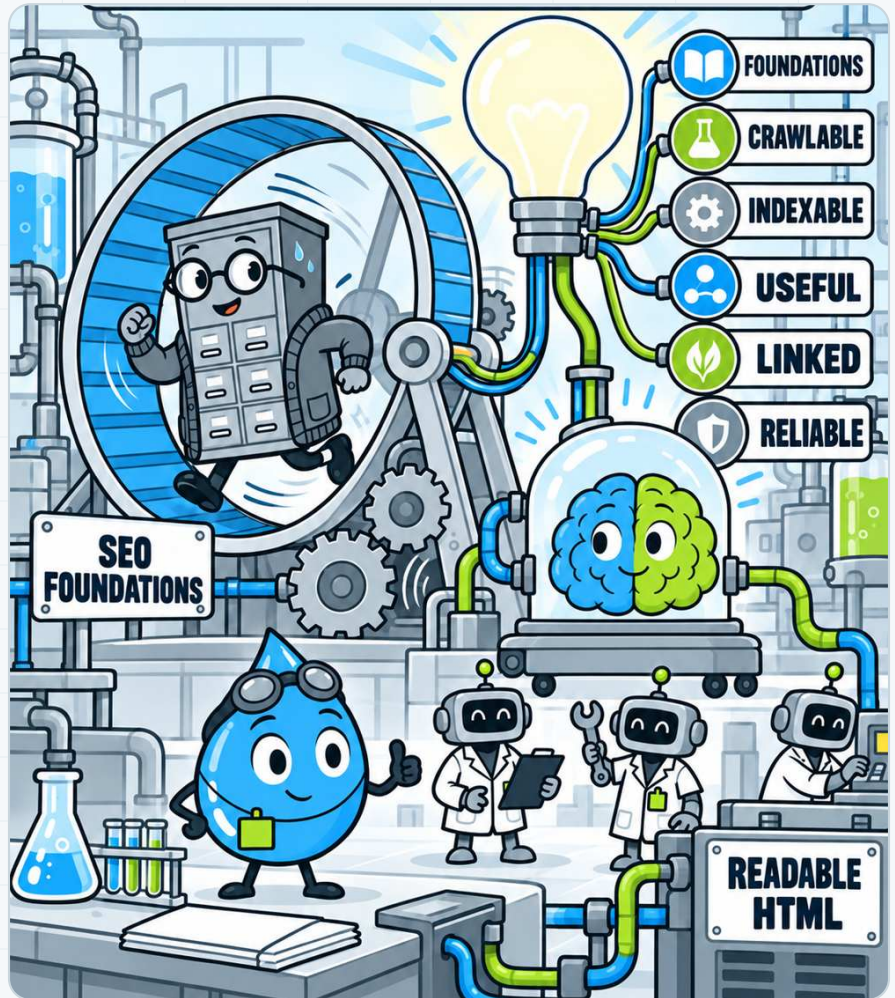
It answers a real question well, with the facts a person needs to decide or act.

LINKED

Other pages point to it, from your own site and from independent sources.

RELIABLE

Accurate, current, and consistent with your other pages and profiles.



THE BASICS THAT HELP SEARCH FIND ME ALSO HELP AI SEARCH FIND ME.

The same Google guide says AI-only files such as llms.txt aren't needed to appear in Google Search. Other answer engines document their own crawlers and rules (Room 05).

INSIDE THE ANSWER MACHINE

02

WHAT'S HAPPENING

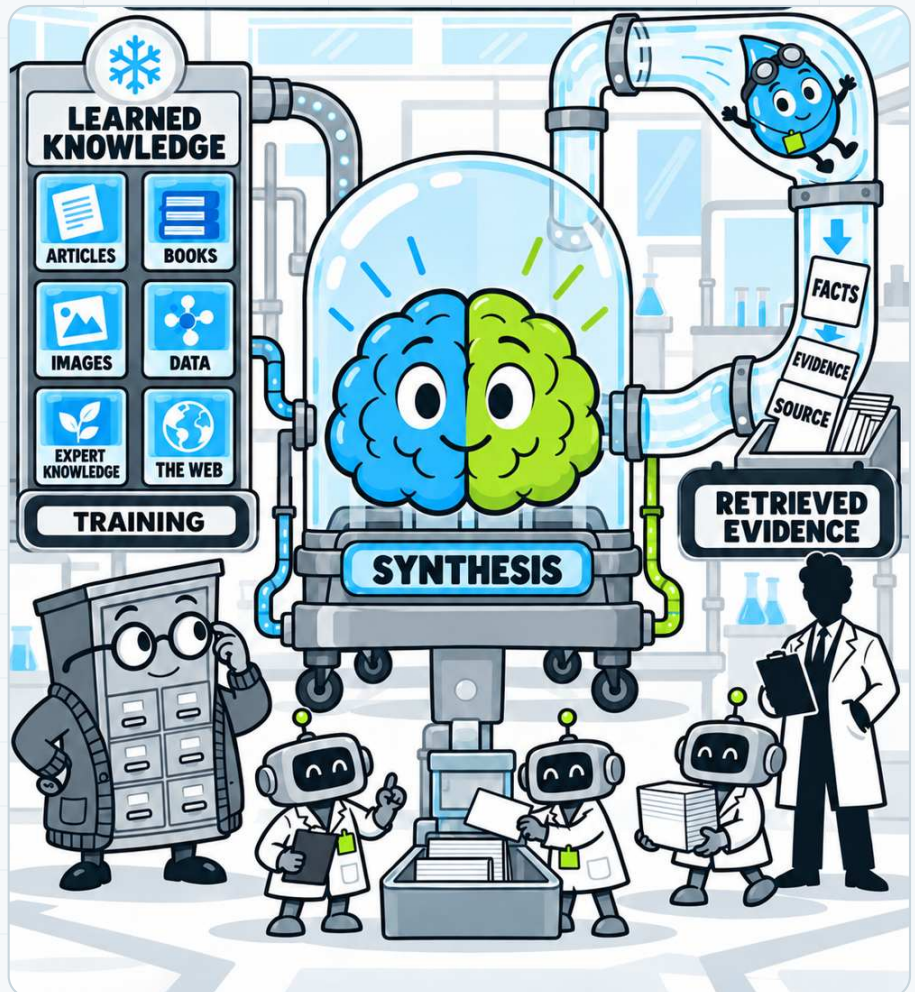
An AI model answers from two places: what it learned in training, and what it looks up for this question. With search on, it retrieves pages, picks passages, and writes a reply from them.

WHY IT MATTERS

You can't edit what a model already learned. You can work on what it finds when it searches: whether your page can be retrieved, and whether your passage is the clearest one available.

JOURDAN'S TAKE

A key claim can lose its meaning once it leaves the paragraph around it. I write the subject, the conditions, and the source into the claim itself.



THE BRAIN CAN KNOW THINGS AND STILL NEED TO LOOK THINGS UP.

IN THIS ROOM

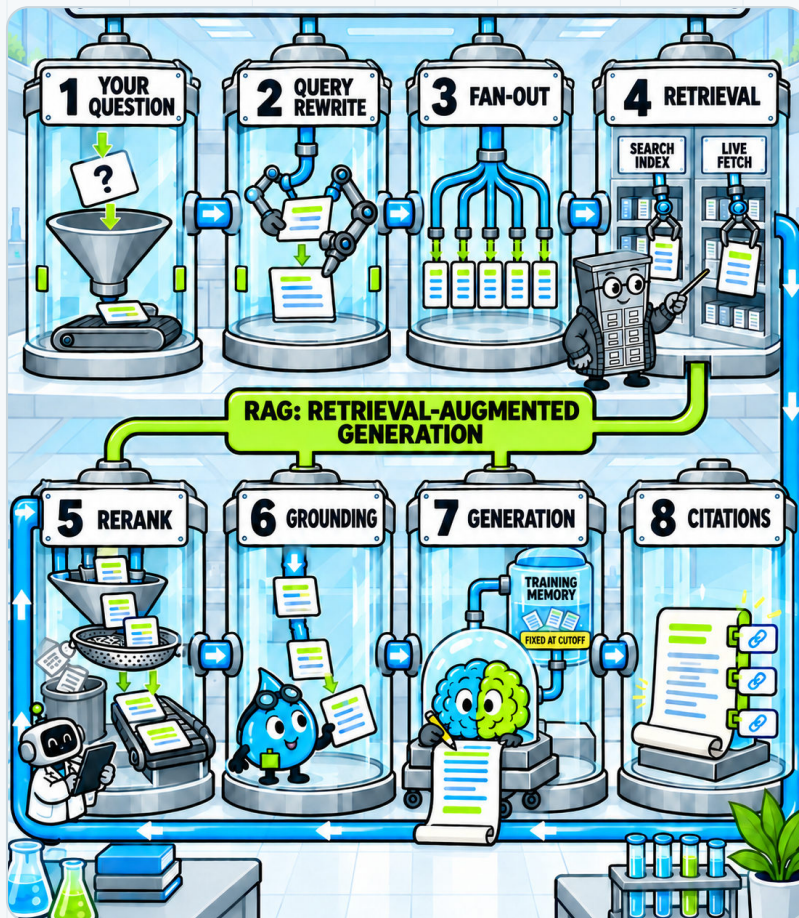
17 Anatomy of an AI Engine
18 Answer Pipeline: Part 1
19 Answer Pipeline: Part 2

20 Query Fan-Out
21 Passages Need Context
22 The Crawl-to-Click Gap

23 Engines Consult Different Shelves

ANATOMY OF AN AI ENGINE

Most AI products that search the web follow a pattern like this one. The numbers match the chambers: what each part does, and where your page can win or lose.



- 1 YOUR QUESTION.** The prompt, plus context the product adds, such as location or chat history.
Your move: Learn the questions your audience really asks.
- 2 QUERY REWRITE.** The system turns the prompt into search queries.
Your move: Use the words people search with.
- 3 FAN-OUT.** One question becomes several searches, one per angle.
Your move: Cover the sub-questions (p. 20).
- 4 RETRIEVAL.** Candidates come from a search index or a live fetch.
Your move: Be crawlable, indexed and fast (Room 05).
- 5 RERANK.** Passages are scored against the question; a few survive.
Your move: Write passages that stand alone (p. 21).
- 6 GROUNDING.** The survivors go into the model's context as evidence.
Your move: Keep the fact, the number and the source together.
- 7 GENERATION.** The model writes from that evidence and its training memory.
Your move: Say the same thing everywhere (Room 04).
- 8 CITATIONS.** Some sources get a visible link; most do not.
Your move: Count mentions, citations and visits separately (Room 08).

HOW TO READ THE MACHINE

Stages 1 to 3 turn a question into searches; 4 to 6 find and sort evidence; 7 and 8 write and link. A printed link alone does not show that a page was retrieved or that it supports the claim.

RAG, IN PLAIN WORDS

Retrieval-augmented generation: the model looks up evidence (stages 4 to 6) and writes from it, instead of answering from memory alone. That lookup is the door your page comes through.

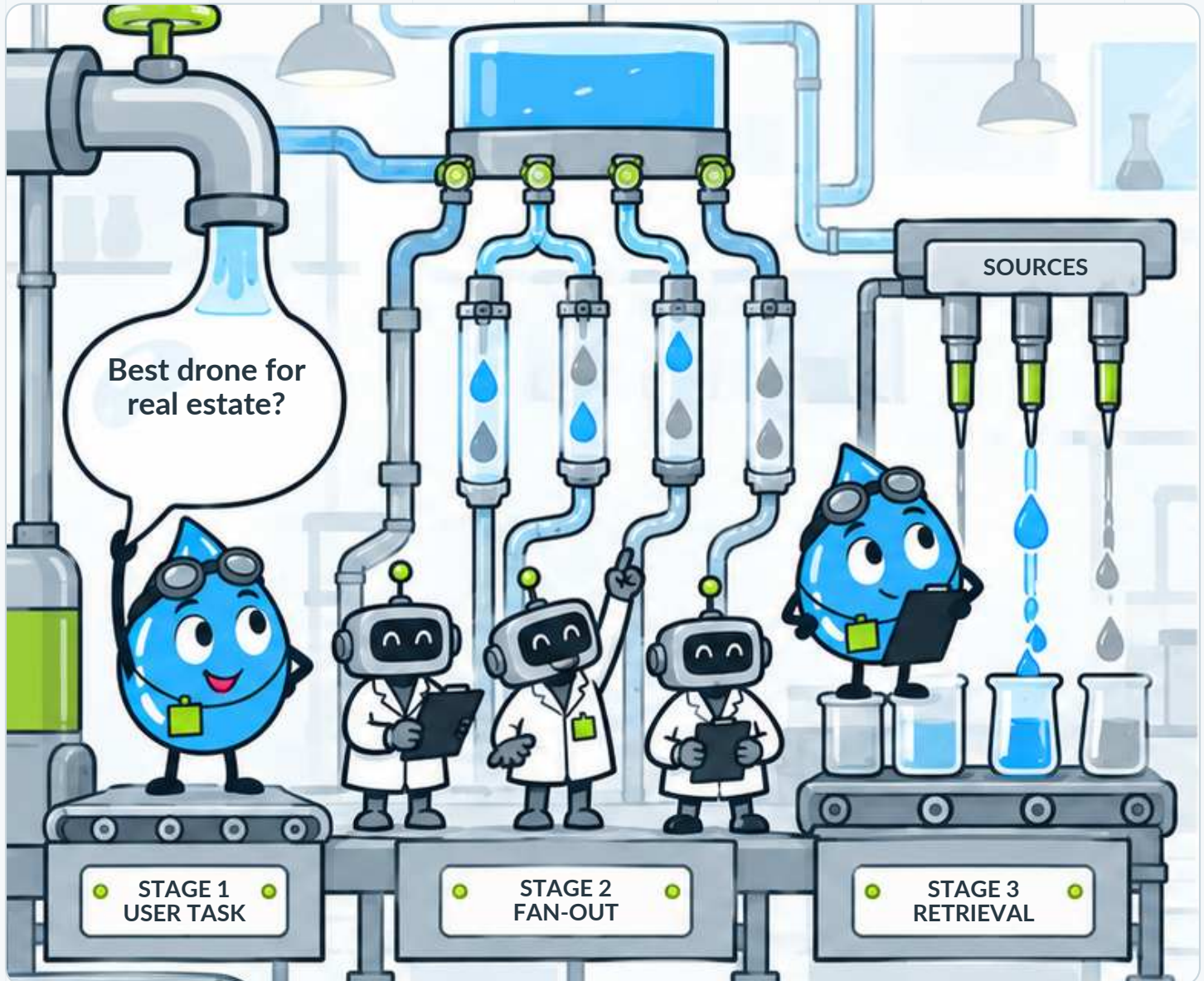
TRAINING MEMORY

What the model learned in training is fixed at a cutoff, and you cannot edit it. New pages, prices and facts reach answers through retrieval.

A simplified model, not any one vendor's system. This pipeline explains how retrieved evidence can support an answer. A model can also emit an unsupported or mistaken URL; a visible link does not prove that the page was retrieved or that it supports the claim.

THE ANSWER PIPELINE

PART 1 OF 2

[Continue: stages 4 to 6 >](#)

1. USER TASK

A user asks a question.

2. FAN-OUT

Possible related searches explore useful sub-questions.

3. RETRIEVAL

Retrieval gathers candidate information.

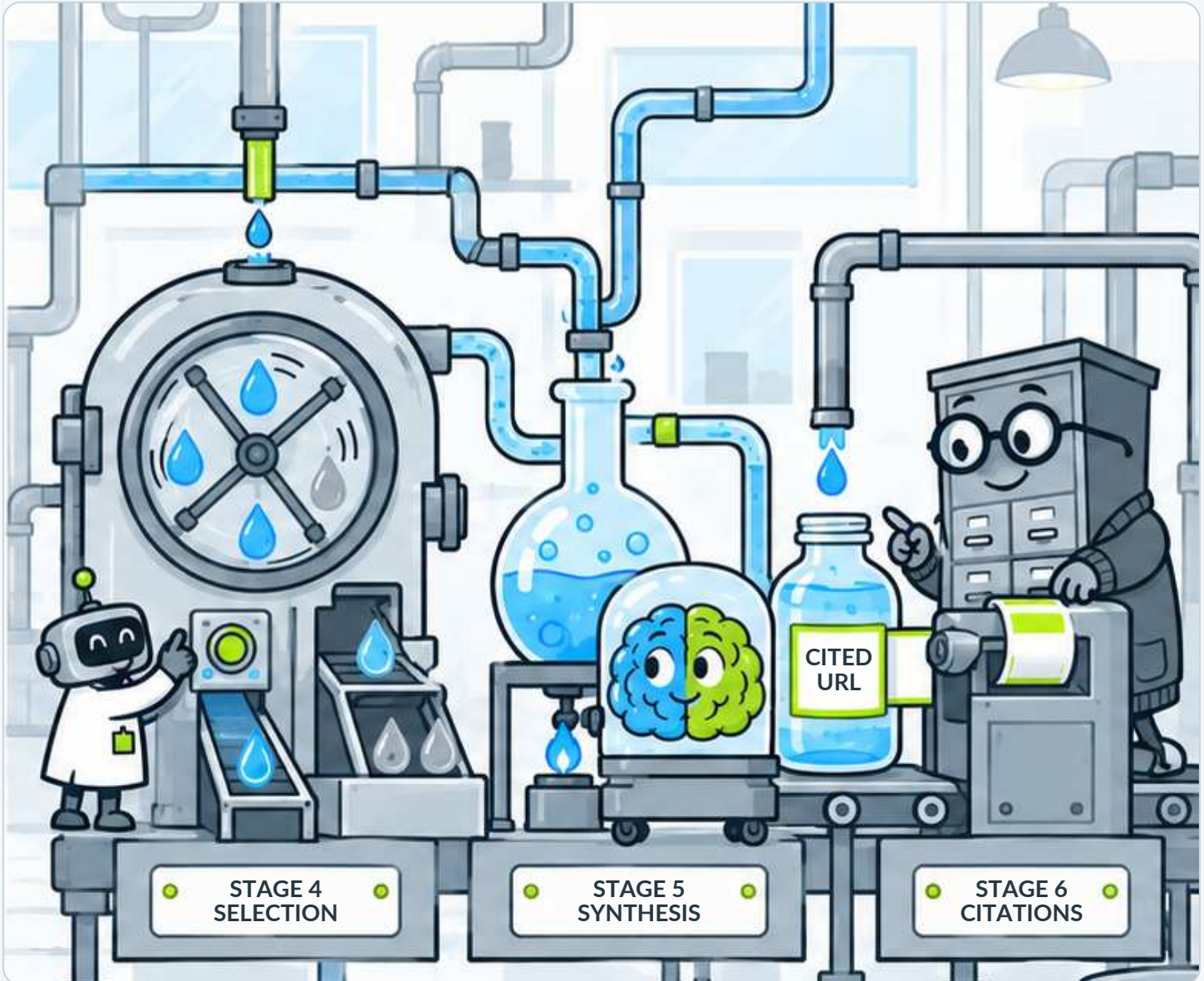


**EVERY ANSWER STARTS AS A QUESTION,
THEN A SEARCH FOR EVIDENCE.**

A simplified workflow: search-enabled products can combine or repeat these stages.

THE ANSWER PIPELINE

PART 2 OF 2

[< Back to stages 1 to 3](#)

4. SELECTION

Candidate selection may reorder the available evidence.

5. SYNTHESIS

The model synthesizes an answer.

6. CITATIONS

Some supporting sources receive visible links.



***I HAVE TO SURVIVE THE JOURNEY
WITH MY MEANING INTACT.***

QUERY FAN-OUT

Illustrative task: best drone for real estate photography

WHERE AIO CITATIONS RANK ORGANICALLY

Ahrefs | Data: collection not stated | Published: March 2026 | 863K keywords, 4M cited URLs

37.1%

26.2%

36.7%

Top 10

11 to 100

Outside top 100

ONE QUESTION, MANY SEARCHES

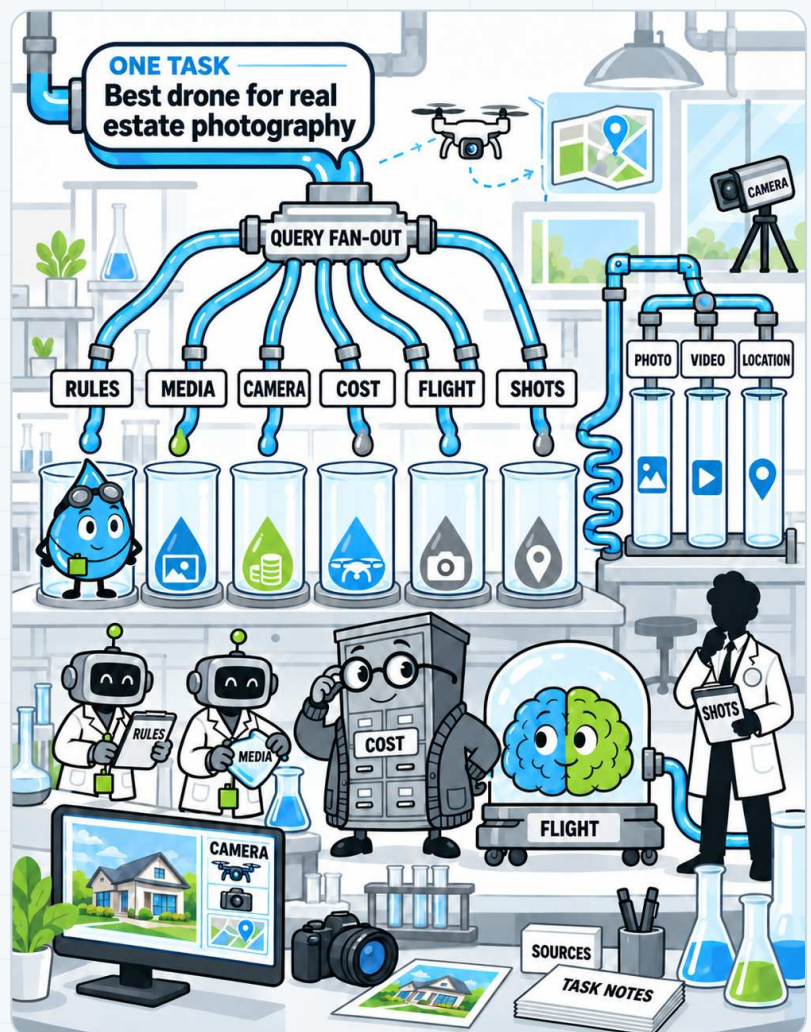
An engine can split one task into sub-questions: rules, media, camera, cost, flight limits and shots. Each one is a separate search.

WHAT IT MEANS

A cited page does not have to rank first. Pages that answer one sub-question well can earn a place in the answer.

TRY IT ON YOUR TOPIC

List the smaller questions inside one customer question. For a drone buyer: which camera handles interiors, what the flight rules allow near homes, what a full kit costs. Each can be a focused section, and a separate chance to be the source an answer uses.



ONE QUESTION CAN OPEN SEVERAL EVIDENCE TRAILS.

Data: collection dates not stated / Published: March 2, 2026. Organic-only rank buckets for the same query; 863,000 keyword SERPs and 4 million cited URLs.

PASSAGES NEED CONTEXT

WRITE IT TO STAND ALONE

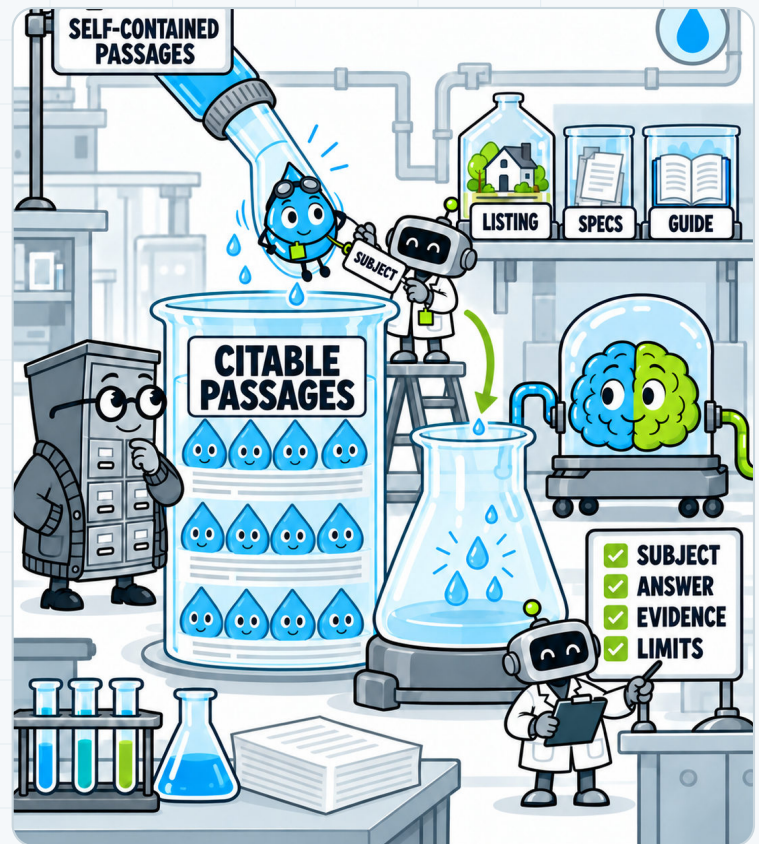
Write self-contained passages inside coherent pages.

STATE THE FOUR PARTS

State the subject, answer, evidence, and limits clearly enough to survive being read alone.

KEEP THE PAGE COHERENT

Build deep source pages for specific tasks; keep the whole page coherent.



THE CHECKLIST ON THE PLATE

Subject, answer, evidence, limits. If a passage lifted out of the page still has all four, it can be quoted accurately.

TRY THIS TEST

Copy one paragraph from your page into a blank document and read it cold. If you can't tell what "it" refers to, which product or place is meant, or where a number came from, put those words back into the paragraph itself. A passage that keeps its subject and its source can be quoted without being misread.

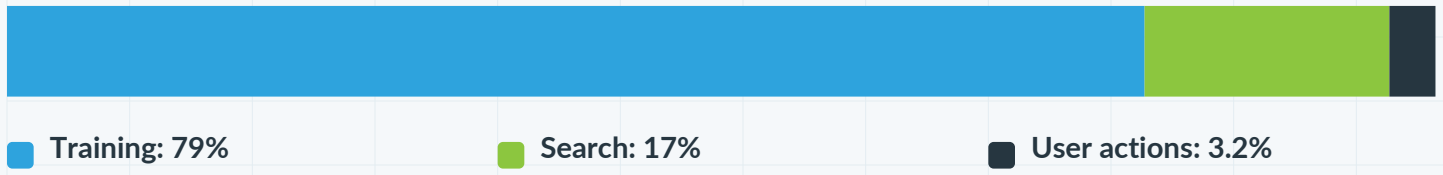


READ ME ALONE. DO I STILL MAKE SENSE?

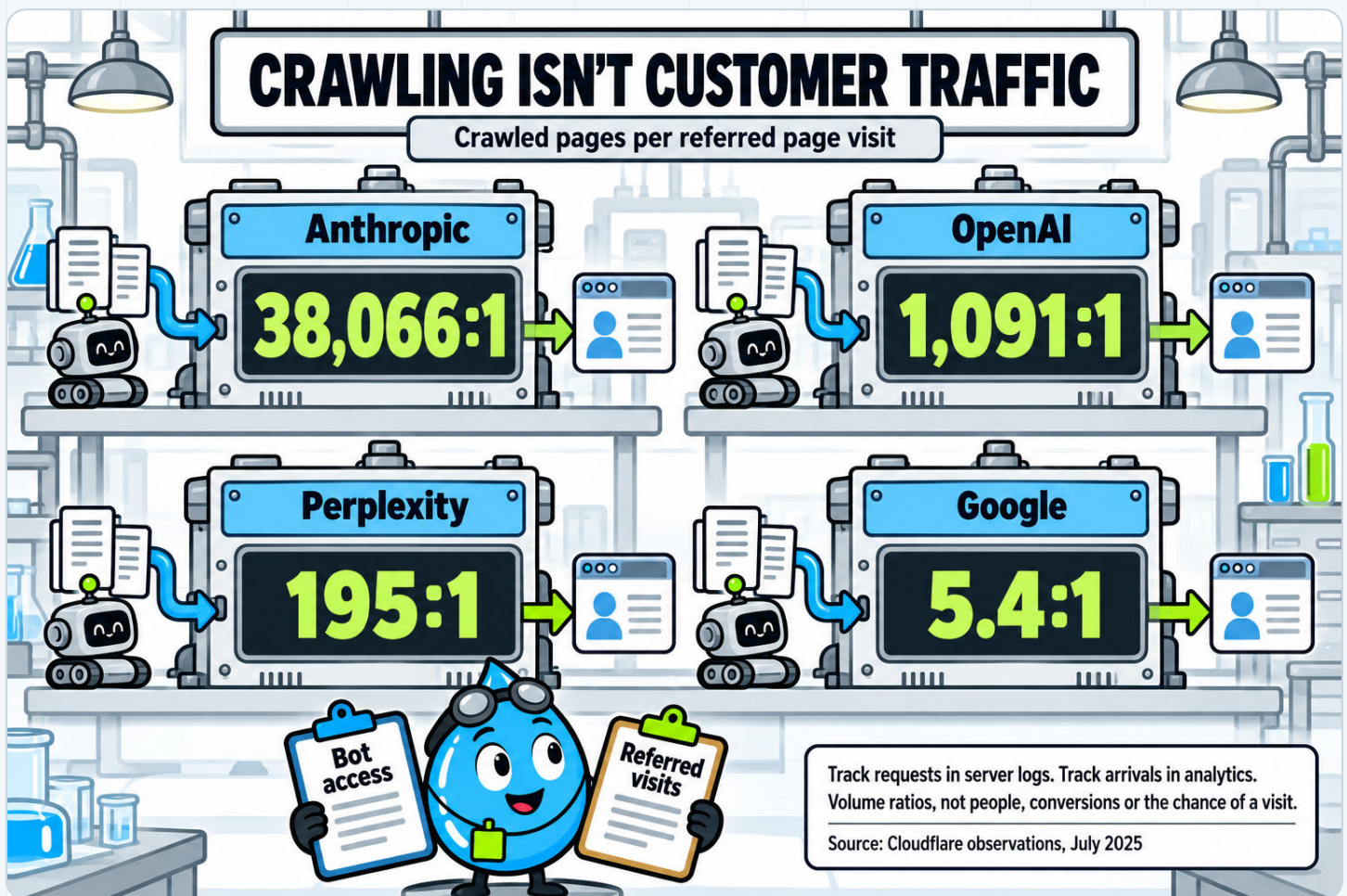
THE CRAWL-TO-CLICK GAP

CLASSIFIED CRAWL PURPOSE

Cloudflare | Data: July 2025 | Published: Aug. 29, 2025 | observed classified traffic



Rounded shares total 99.2%; remaining 0.8 percentage points are not assigned here.



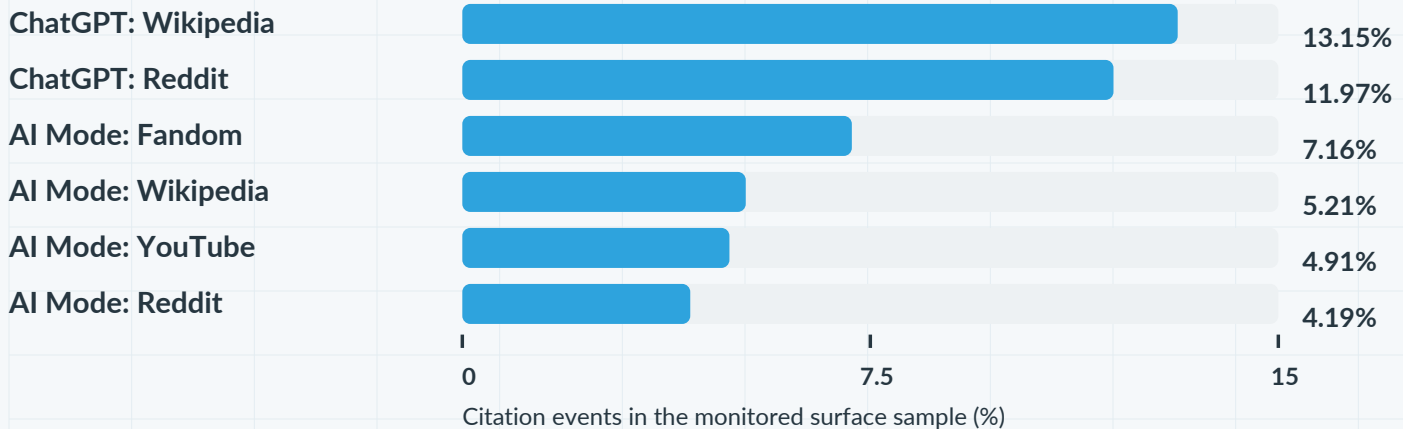
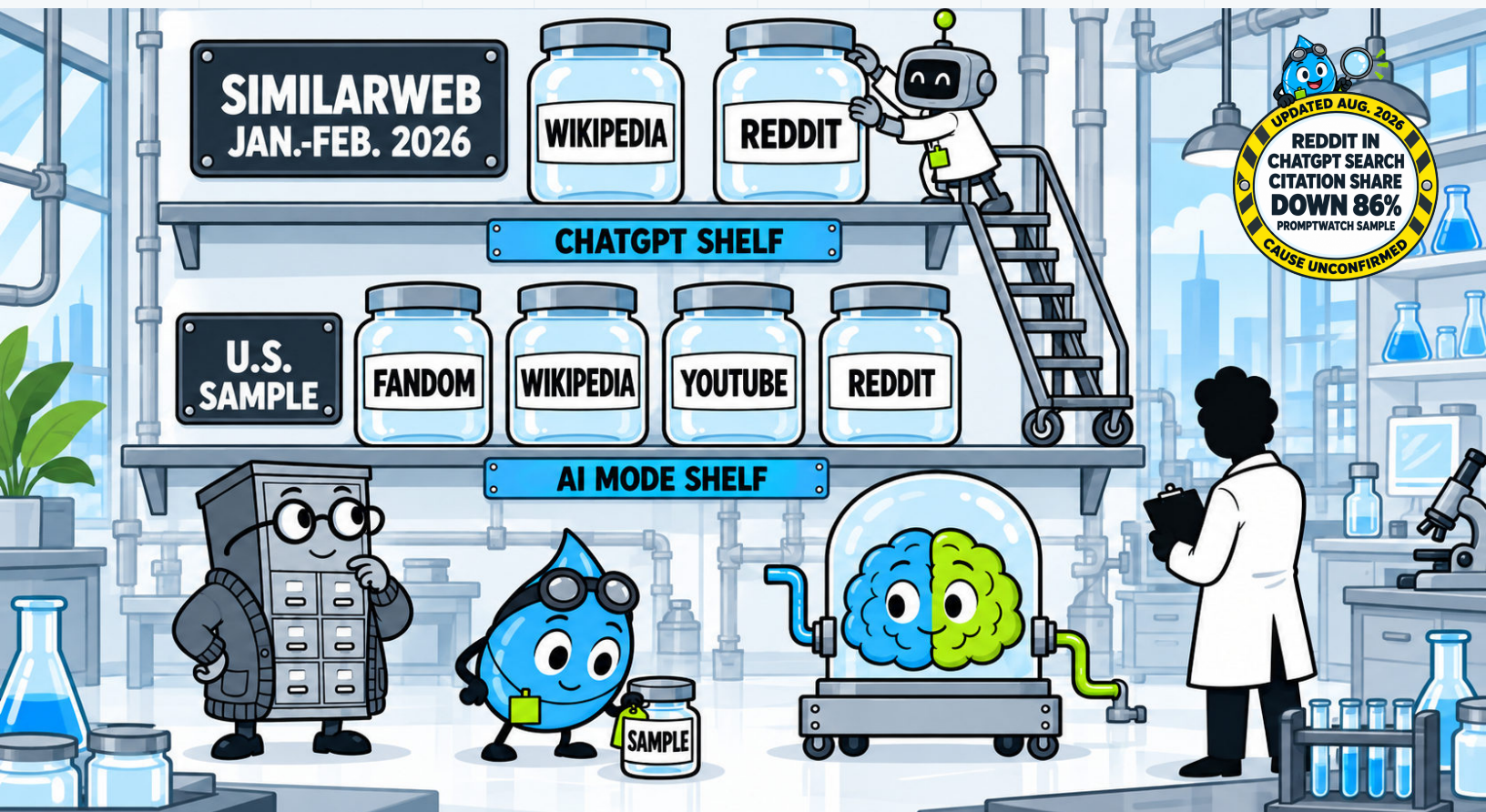
BEING FETCHED IS ONLY THE START OF MY JOURNEY.

Data: July 2025 / Published: Aug. 29, 2025. Rounded crawl shares and crawled-pages / referred-page-visit ratios. Denominators differ from users and conversions.

ENGINES CONSULT DIFFERENT SHELVES

SHARE OF TRACKED CITATION EVENTS

Similarweb | Data: Jan.-Feb. 2026 | U.S. monitored prompts, sample count not disclosed



**EACH ENGINE LEANS ON DIFFERENT SITES.
LEARN WHICH ONES YOURS USES.**

Similarweb U.S. monitored prompts, Jan. to Feb. 2026. Percentages count tracked citation events by surface. Update, Aug. 2026: Promptwatch reported Reddit's share of ChatGPT Search citations fell from a 3.83% average (July 18 to Aug. 7) to 0.52%, an 86.4% drop. Cause unconfirmed; Promptwatch has not ruled out a data issue. Different tracker and sample from the chart above.

THE PERIODIC TABLE OF GEO

03

Aq Answer clearly	Qt Authentic quotes	St Verified statistics	Cs Cite sources	Fl Fluency	Or Original research
Tb Useful tables	Fr Current facts	En Clear entities	Au Relevant expertise	Sa Valid sameAs	Ff Task coverage
Cr Crawl access	Ss Readable HTML	Sc Accurate schema	Il Internal links	Mn Corroboration	Rd Communities
Yt Useful video	Rv Genuine reviews	Pr Digital PR	Pt Prompt sampling	Lg Verified logs	Ai Referred visits
Ks Keyword stuffing	Pi Prompt injection				
Content	Entity	Technical		Hazard	
Off-Site	Measure				

HOW TO READ IT

Each tile is a practice you can work on, colored by type: content, entity, technical, off-site, measurement, and two hazards. Click a tile to jump to the room that explains it.

JOURDAN'S TAKE

Most GEO hacks are good writing with a new label. Clear answers, real numbers and named sources are where I'd spend the first hour, not on another keyword.

IN THIS ROOM

25 The Experiment That Started It All
26 From Vague to Specific

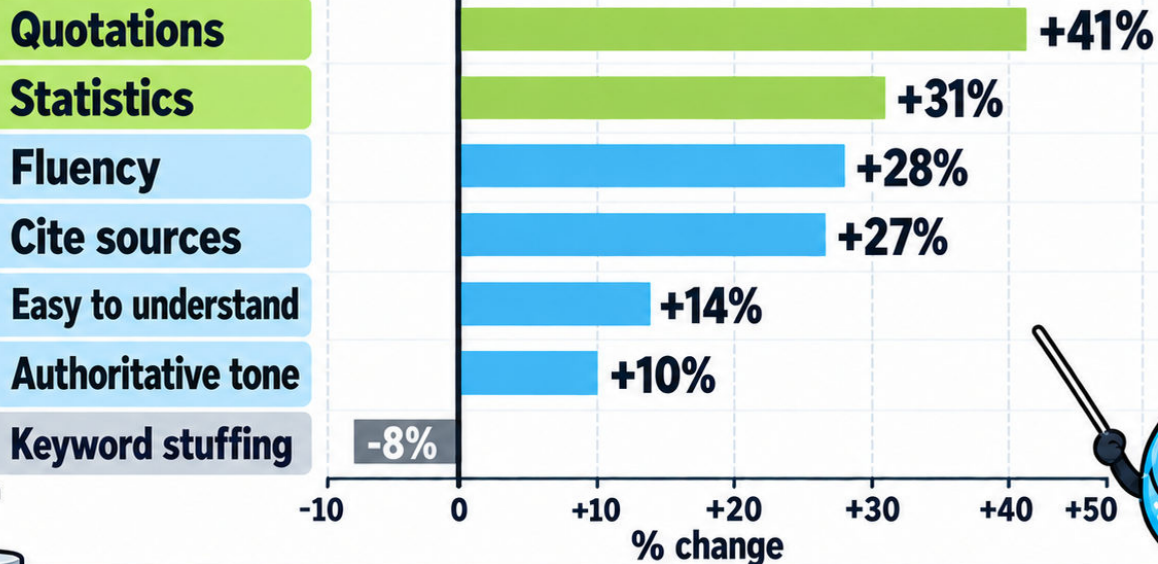
27 Do AI Answers Prefer Fresh Pages?
28 Busted Experiments

No tile is a known ranking weight or a guaranteed citation: the table is a menu for judgment. The next pages show what research measured.

THE EXPERIMENT THAT STARTED IT ALL

WHAT HELPED IN GEO-BENCH?

Relative change in answer visibility



Simulated engine. Fixed-source visibility, not live traffic.
Overall Position-Adjusted Word Count • Baseline: 19.3

Source: GEO research, Table 1 • Rounded derived changes



Use authentic quotes and statistics

WHAT IT FOUND

The 2024 paper that named GEO rewrote source texts nine ways. Adding quotations, statistics and citations raised a simulated engine's visibility metric; keyword stuffing lowered it.

WHAT IT DOES NOT SHOW

Its published code allowed synthetic quotations, invented sources and hypothetical statistics, so it does not show an effect from verified evidence. Nor does it measure live citations or traffic.

WHAT TO DO

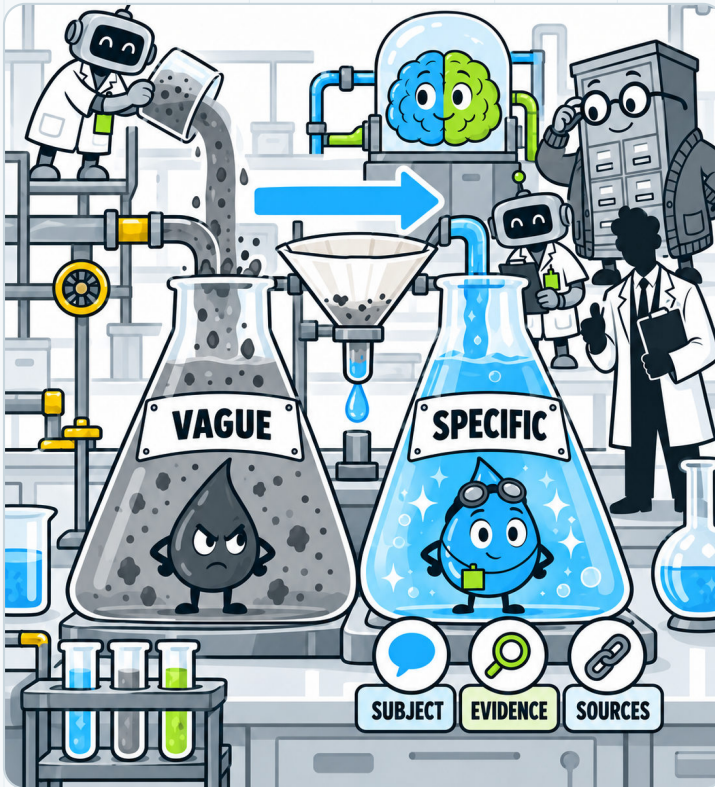
Publish authentic, traceable evidence: real numbers with dates, named people you quoted, links to the original source. It is honest and checkable, whatever a benchmark scores.



REAL EVIDENCE BEATS A LOUDER ADJECTIVE.

FROM VAGUE TO SPECIFIC

Most passages fail because they are vague, not because they are short. Here is the same advice before and after a rewrite.



VAGUE

Drones are popular, there are lots of options, and it depends on what you need.

SPECIFIC

For real estate listings, compare camera quality, obstacle sensing, flight time and the flight rules that apply where you shoot. Show sample images, the test conditions and the date of the specs.

WHY THE REWRITE WORKS

The vague version could describe any drone for anyone. The specific one says who it is for, what to compare and what proof to show, so a reader can act on it and an answer can quote it without guessing.

SUBJECT

Names who it is for and what it covers: drones for real estate listings.

EVIDENCE

Says what to compare and shows sample images and test conditions.

SOURCES

Dates the specs, so a reader or an engine can check them.



NAME THE SUBJECT. SHOW THE PROOF. SAY WHERE IT COMES FROM.

Illustrative rewrite, not a product recommendation or legal advice. The specific version is easier to check; neither version is guaranteed a citation.

DO AI ANSWERS PREFER FRESH PAGES?

MEAN AGE OF URLS

Ahrefs | Published: July 2025 | Collection date / sample count not stated; first-seen age proxy

ChatGPT citations

958

Organic results

1,416

0

750

1500

Mean age in days (first-seen proxy)

WHAT THE DATA SHOWS

In Ahrefs' study, pages cited by ChatGPT were younger on average than pages in organic results: about 958 days old versus 1,416.

WHAT IT DOES NOT SHOW

Age was estimated from when a crawler first saw each URL, and the collection dates are not stated. Older pages still get cited when they stay accurate.

UPDATE FACTS, NOT TIMESTAMPS

Revise the substance and keep its evidence date. A new timestamp alone is not stronger proof (page 71).



NEW PAINT IS NOT THE SAME AS NEW EVIDENCE.

Data: citation snapshot, collection dates not stated / Published: July 28, 2025. Mean age uses crawler first-seen as a publication proxy.

BUSTED EXPERIMENTS



SCALED MANIPULATION

Low-value mass production is not a useful evidence strategy, and Google treats it as scaled content abuse.

KEYWORD STUFFING

Tested and failed: about 8% lower on the GEO study's selected visibility metric. Bing's guidelines now name it.

llms.txt

10.13% of nearly 300,000 domains have one, with no positive link to citation frequency (live source updated Oct. 15, 2026). Useful for agents, not a ranking lever.

SPECIAL AI SCHEMA

Google requires no special GEO schema. Keep accurate markup for rich results and identity.



**LET'S KEEP THE USEFUL WORK
AND DISCARD THE SHORTCUTS.**

The llms.txt result is observational. The GEO figure uses its study's selected visibility metric. Every popular trick, rated: page 68.

WHAT'S HAPPENING

AI answers describe a business with facts gathered from many places: its website, LinkedIn, Google Business Profile, directories and press. When those facts disagree, the answer has to guess.

WHY IT MATTERS

A wrong founding year, an old phone number or a discontinued product can be repeated as fact. One approved set of facts, published everywhere and connected with structured data, is harder to confuse.

JOURDAN'S TAKE

Brand facts are data, so treat them like data. One owner, one approved version, one review date. A slogan isn't a fact.



A FAMILIAR NAME NEEDS A CONSISTENT IDENTITY.

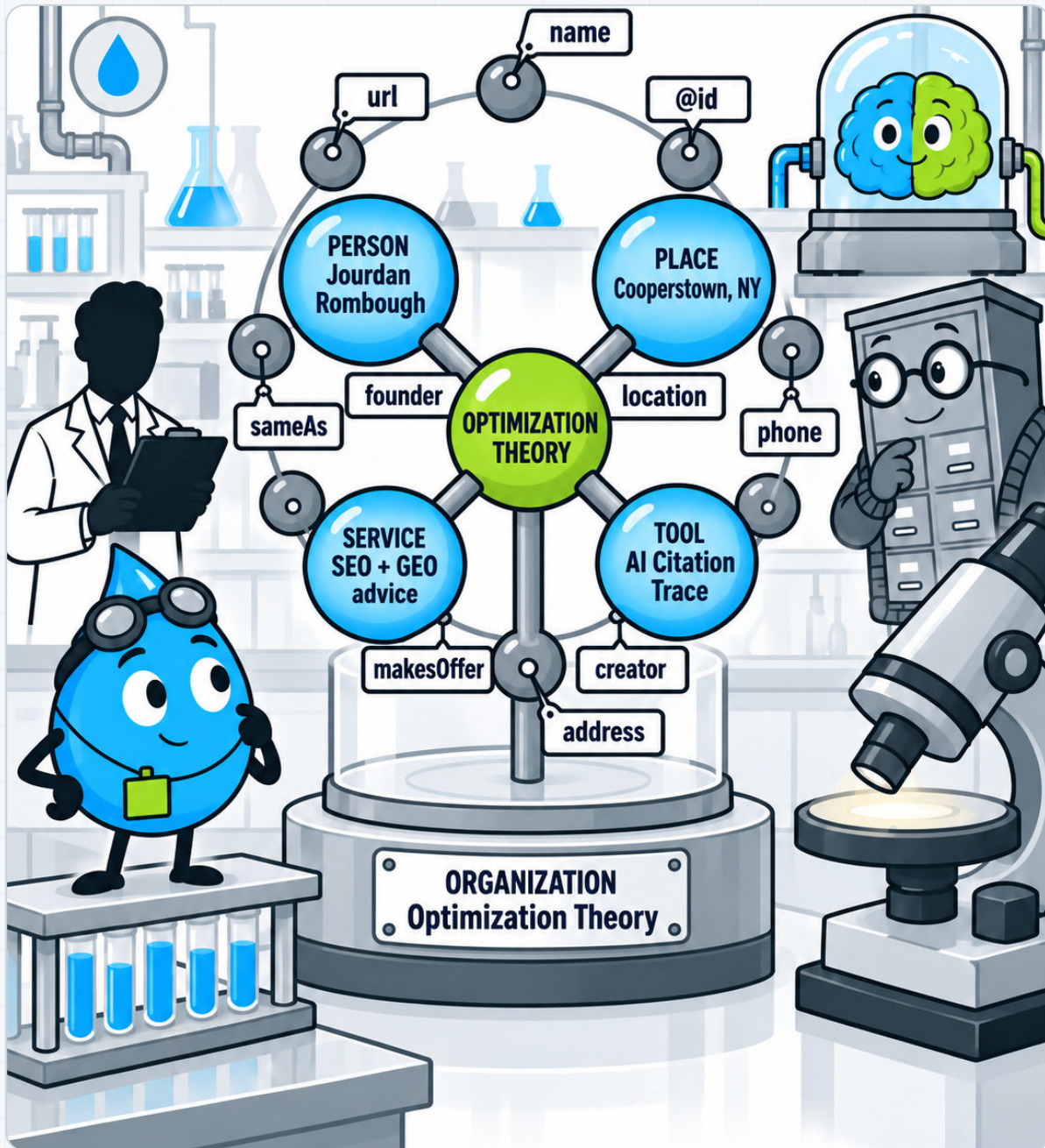
IN THIS ROOM

30 The Entity Molecule
31 The DNA Match Test

32 Repair Wrong Answers
33 Fill Real Decision Gaps

THE ENTITY MOLECULE

Structured data describes your organization as connected things: a person, a place, a service, a tool, each joined by a named relationship.



RELATIONSHIPS CONNECT US. THEY DON'T MAKE US THE SAME ENTITY.

Bond labels are schema.org properties read from the organization outward (creator reads from the tool). sameAs joins only equivalent identities, here the person's LinkedIn profile. name, url, @id, phone, and address are properties of a node, not separate entities.

THE DNA MATCH TEST

Illustrative mismatch check: current phone number

COLUMNS ON THE BOARD

WEB website, LI LinkedIn, GBP Google Business Profile, DIR directories, PR press. The warning sits on DIR: a directory still lists an old phone number.

REPAIR THE RECORD

Keep the approved value, its owner, where it is published and its review date in one record. Correct material errors, then retest the task.

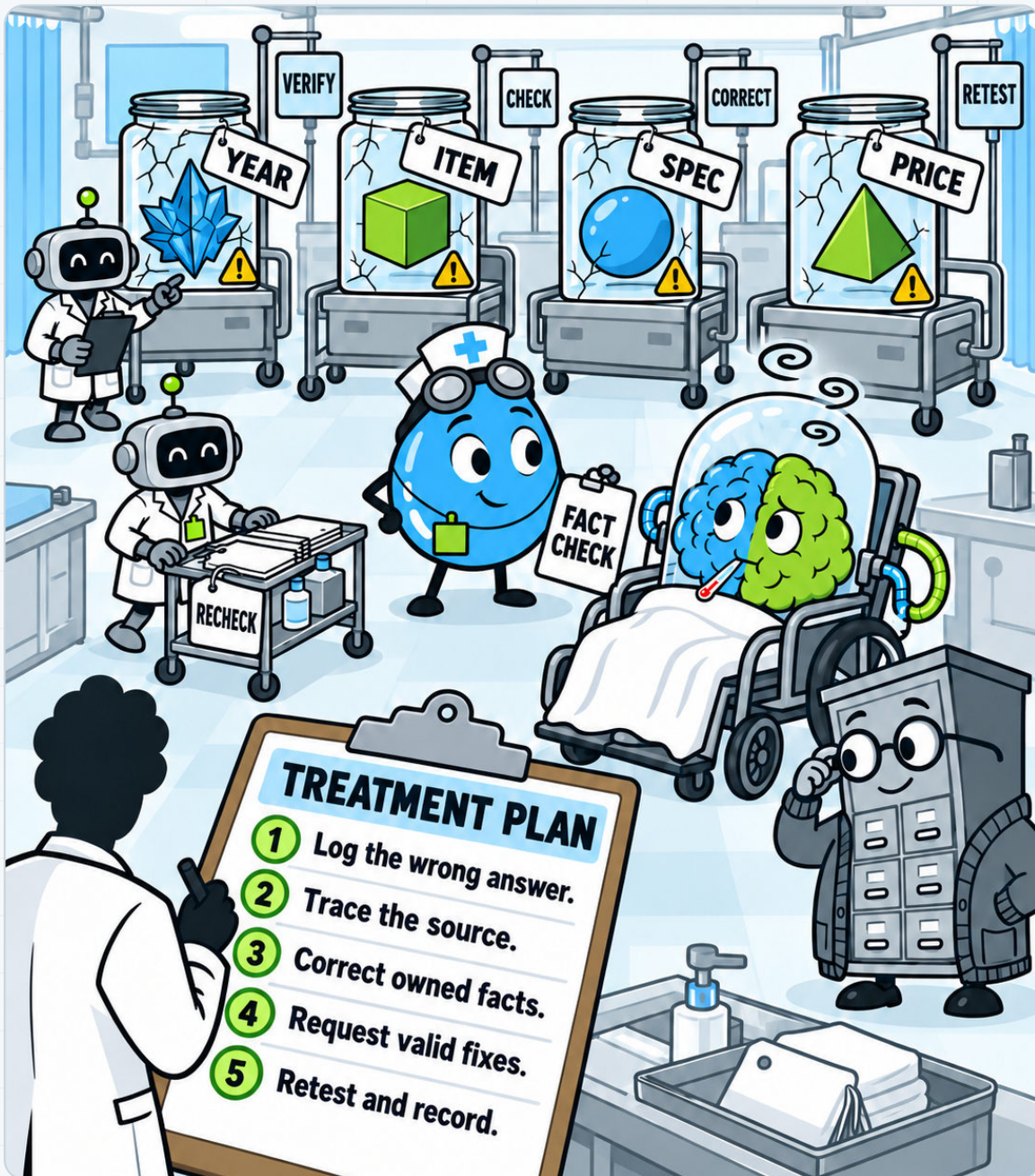
EXAMPLE STATUS

The board illustrates a discrepancy. It does not claim these live profiles currently publish conflicting numbers.



REPAIR WRONG ANSWERS

When an AI answer gets a fact about you wrong, treat it like a case: log it, trace where the claim came from, and fix the source.



IF AN ANSWER GETS US WRONG, TRACE THE CLAIM AND REPAIR THE SOURCE.

Jar tags: YEAR = wrong founding year, ITEM = discontinued product, SPEC = a competitor's feature, PRICE = old pricing. Visible citations do not reveal every causal input; a corrected source can improve the evidence without guaranteeing that every answer updates.

FILL REAL DECISION GAPS

Map the questions that affect a purchase or action. Filled tubes are questions your content already answers; the empty Gap tubes are questions people ask that it does not answer yet. Fill gaps with sourced explanations and merge pages that overlap.



THE EMPTY SLOTS ARE QUESTIONS WE HAVEN'T ANSWERED YET.

Tube tags: Define, How, Cost, Compare, Mistakes, Tools, Examples and FAQ are questions the content already answers; the two Gap tubes (Limits, Evidence) are questions it has not answered yet.

THE TECHNICAL AIRLOCK

05

WHAT'S HAPPENING

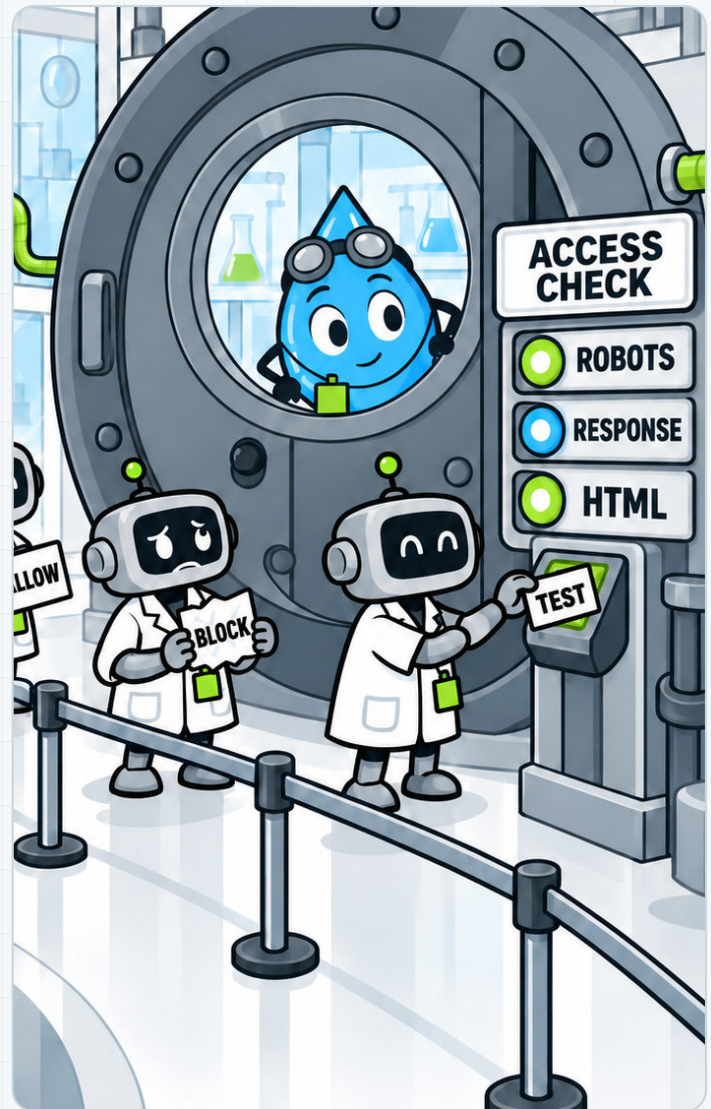
Different bots visit for different jobs: some collect training data, some build search indexes, some fetch a page because a user asked. Each can be allowed or blocked separately.

WHY IT MATTERS

A fetcher that is blocked, gets an error, or receives HTML where the key facts only appear after JavaScript runs may never see your best content, however good it is.

JOURDAN'S TAKE

For a direct citation to this page, check the response and content the target fetcher can access before testing content changes. Other routes exist (partner indexes, cached copies, learned knowledge); test the one you are measuring.



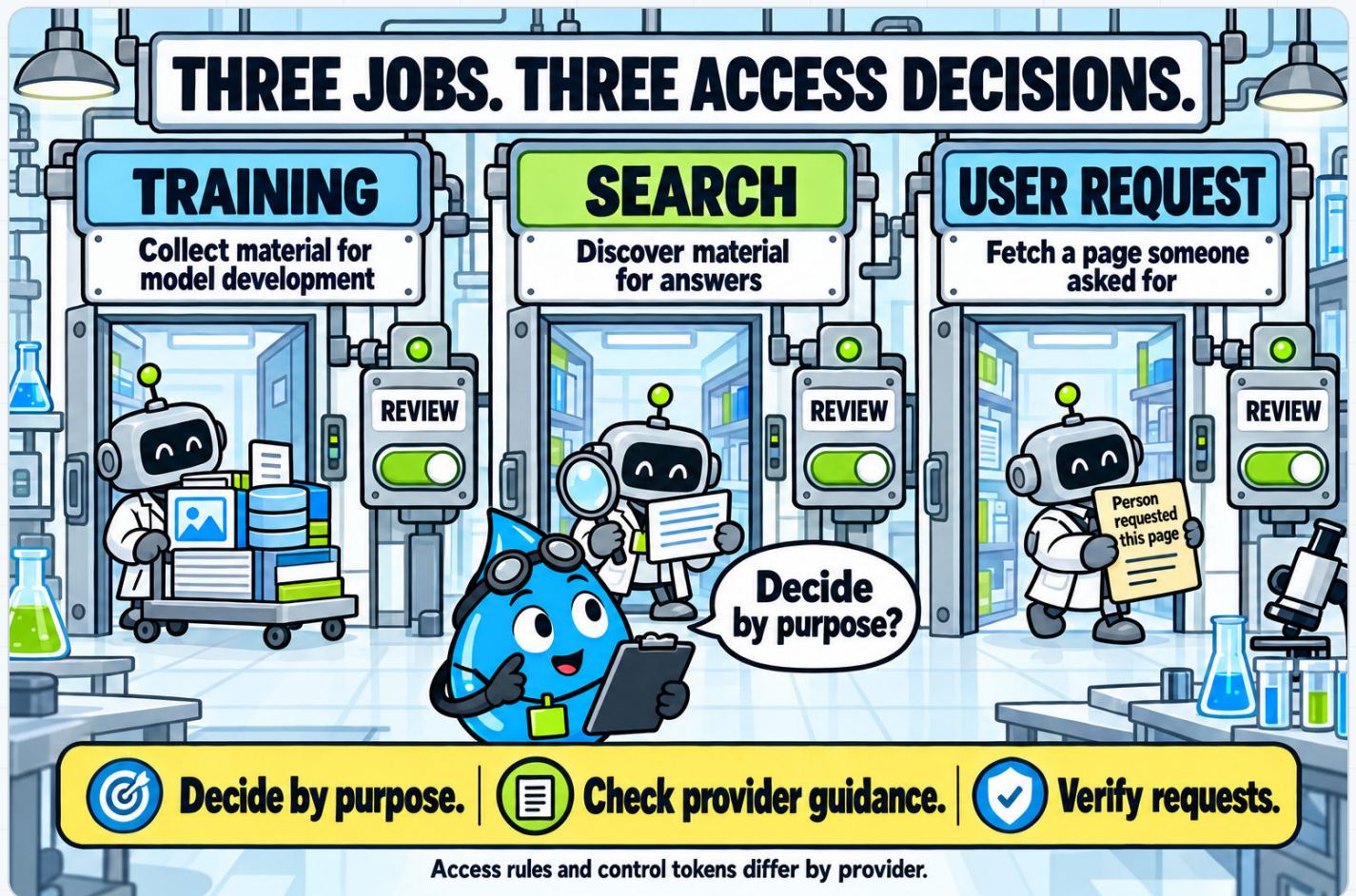
CAN THE RIGHT FETCHER GET THROUGH THIS AIRLOCK?

IN THIS ROOM

- | | | |
|-----------------------------|------------------------------|-------------------------|
| 35 The Badge Wall | 40 Publisher Controls: Panel | 45 When the Agent Buys |
| 36 Clear the JavaScript Fog | 41 Markdown for Agents | 46 Agent Buys: Timeline |
| 37 Specimen Labels | 42 Markdown: The Evidence | 47 Can an Agent Finish? |
| 38 Different Files, Jobs | 43 Who Pays for the Crawl | |
| 39 Publisher Controls | 44 Crawl: Your Settings | |

Access, indexation, and answer selection are separate checks.

THE BADGE WALL



TRAINING / COLLECTION

GPTBot

ClaudeBot

Google-Extended*

Applebot-Extended*

CCBot

SEARCH DISCOVERY

OAI-SearchBot

Claude-SearchBot

PerplexityBot

Googlebot / Bingbot

USER ACTIONS

ChatGPT-User

Claude-User

Perplexity-User

Provider-specific rules

* Google-Extended and Applebot-Extended are control tokens, not separate fetching identities. Merge policy choices with your existing robots rules.



GIVE EACH CRAWLER THE POLICY ITS JOB REQUIRES.

Extended names are control tokens, not separate fetchers. CCBot collects Common Crawl's public web corpus. Merge policies deliberately.

CLEAR THE JAVASCRIPT FOG

INSPECT THE DELIVERY

Keep essential facts in delivered HTML. Inspect redirects and error responses. Verify today's target fetcher.

THE SEVEN TAGS

Each tag is one check on the delivered page: 1 the URL, 2 its HTTP status, 3 the delivered HTML, 4 the key facts, 5 the named source, 6 the date, 7 the index status. The fogged jar is what appears only after scripts run.



HISTORICAL 404 SHARES

ChatGPT 34.82% / Claude 34.16% / Googlebot 8.22%. Vercel/MERJ, Dec. 2024; Applebot was the rendering exception.

WHY PAGES GO FOGGY, AND A TWO-MINUTE TEST

Many sites build pages in the browser: the server sends a thin shell, and scripts fill in the prices, specs and text a moment later. People never notice, but a fetcher that reads only the delivered HTML may miss the facts. To check, turn JavaScript off or view the page source, and search for one key fact, its source and its date. If they are missing, ask your developer to put them in the HTML the server sends.



MY ESSENTIAL FACTS SHOULD SURVIVE THE FOG.

SPECIMEN LABELS

SCHEMA TYPE	WHAT THE PAGE MUST SHOW	IN GOOGLE SEARCH, 2026
Organization	Your name, logo, contact details and sameAs profiles, on the home or about page.	● Supported: logo, legal name and contact details.
Person	A real person's name, role and profiles, on a page about them.	● No feature of its own; used for authors and Profile page markup.
Article	Headline, author, dates and images that match the page.	● Supported: title, image and date display.
Product	Name, price, availability and reviews shown on the page.	● Supported: product snippets and merchant listings.
LocalBusiness	Name, address, hours and phone that match your Business Profile.	● Supported: hours and details in the knowledge panel.
Review	A real review of a specific item, visible on the page.	● Supported, but reviews a business controls about itself get no stars.
HowTo	Steps a reader can follow, visible on the page.	● No rich result: removed from Search in 2023.
FAQPage	Questions and answers visible on the page.	● No rich result since May 7, 2026.

YOUR PAGE

Markup must describe content a person can see on the page. Valid schema.org vocabulary and Google eligibility are separate checks: test both.

THE AI ANGLE

Schema alone does not earn AI citations: Ahrefs found no significant lift (page 70). Keep it accurate for rich results and to say who you are.

In plain words. Markup is the label on a specimen jar. A page that shows a drone, its price and that it is in stock can say the same three things in Product markup. A label that claims a price or reviews nobody can see on the page breaks the first rule of this table.



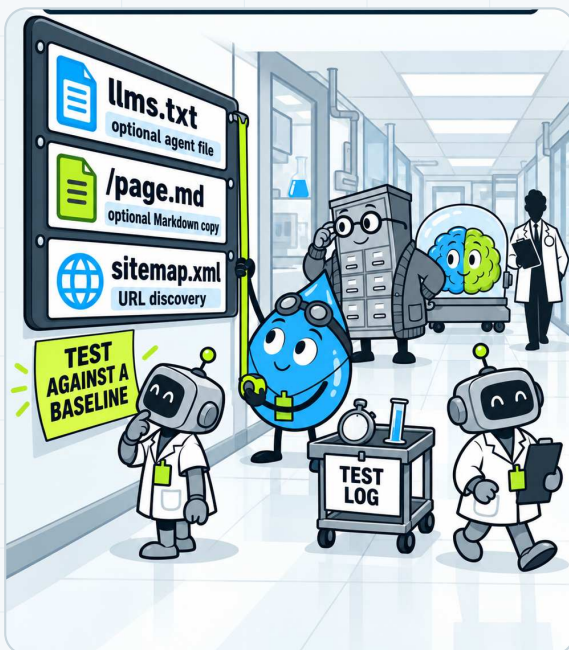
A SPECIMEN LABEL SHOULD DESCRIBE WHAT'S ACTUALLY IN THE JAR.

Google feature status checked Oct. 29, 2026 against its structured data search gallery, the FAQ deprecation notice and the review snippet guidelines. All three are in the Source Drawer.

DIFFERENT FILES, DIFFERENT JOBS

Several small files talk to machines about your site. Each has one job, and none of them is a shortcut into AI answers. What each one does, who reads it, and what the evidence says:

FILE	WHAT IT DOES	WHO READS IT	EFFECT ON AI ANSWERS
robots.txt	Tells crawlers which paths they may fetch, bot by bot.	Search and AI crawlers that honour it.	Access only: a blocked page cannot be fetched for a fresh answer (Room 05).
sitemap.xml	Lists your URLs and when each one changed.	Search engines, for discovery.	Indirect: helps pages get found and refreshed.
llms.txt	A curated list of key pages for agents.	Some AI tools; Google Search does not use it.	None shown: no citation link across about 300,000 domains.
page.md	A clean Markdown copy, linked with rel="alternate".	Agents and tools that ask for Markdown.	Helps agents read you; no citation lift shown (p. 41).
JSON-LD schema	Describes the page's entities and facts.	Search engines, for rich results and identity.	No significant AI citation lift in Ahrefs' test (p. 70).
IndexNow	Pings engines when a URL is added, changed or deleted.	Bing and other participating engines.	Keeps the index behind Copilot fresh (p. 70).



THE RULE OF THUMB

Use robots.txt and sitemaps for access and discovery. Treat llms.txt and Markdown copies as optional help for agents, never as ranking levers.

PROPOSED LAB TEST

Compare a maintained Markdown endpoint against a documented baseline. Preserve the same facts and watch for duplicate-version drift.



A DIRECTORY SIGN HELPS ONLY IF IT POINTS SOMEWHERE USEFUL.

Google says llms.txt does not affect Search ranking. The Markdown strategy, step by step: pages 41-42. The comparison on this page is proposed, not run.

PUBLISHER CONTROLS: CHOOSE THE RIGHT LEVER

Crawling, indexing, previews, training and AI-answer inclusion are separate choices. Pick the outcome first, then choose a control whose documented scope matches it. A publisher preference is not a security boundary.

CONTROL	WHAT IT CONTROLS	LIMIT TO REMEMBER
robots.txt, addressed to a crawler	Automated fetching of specified paths by crawlers that honor the rules	Does not reliably remove a known URL from an index or stop every user-directed fetch
Google Search generative AI control, in Search Console	Links and content, including grounding, in specified Search AI features: AI Overviews, AI Mode and Discover AI	Does not control training or other Search ranking/inclusion; check parent-property inheritance
Google-Extended, a robots.txt product token	Future Gemini model training and selected grounding in Gemini Apps and Vertex AI	No separate fetching identity; not the control for Google Search inclusion/ranking
noindex, in robots meta or X-Robots-Tag	Whether the resource can appear in Google Search results	Google must fetch the resource to read the rule
nosnippet	Text snippets/video previews and direct content input for Google's AI Overviews and AI Mode	A static image may still appear; not a general training preference
max-snippet	Maximum textual snippet length and amount of content available as direct input to Google's AI Overviews and AI Mode	Not an exact answer-length limit; separately granted uses can differ
data-nosnippet	Text sections excluded from Google's snippets	Not an access block; structured data has separate rules
CDN/WAF block or authentication	Whether a request can reach the content	Can stop intended search or agent visits too; inspect real responses and current settings

Next page: the panel, step by step

Where the Search Console setting lives, what X-Robots-Tag adds, and how other providers split their crawlers. Cloudflare's crawler presets: page 43.

PUBLISHER CONTROLS: READ THE PANEL



Switch positions are illustrative, not a recommended configuration.
Choose settings for your site's purpose and verify their documented scope.

WHERE TO LOOK IN SEARCH CONSOLE

Google's Search generative AI control was available worldwide by August 31, 2026, according to its documentation. Look under Settings > Search generative AI. A property without its own choice inherits one from its closest configured parent, so check the parent too. The default is to include your links and content.

X-ROBOTS-TAG

The same rules (noindex, nosnippet, max-snippet) can be sent as an X-Robots-Tag HTTP header, which also works for non-HTML files such as PDFs, images and video. Either way, Google has to fetch the resource to read the rule.

OTHER PROVIDERS, OTHER LEVERS

OpenAI separates OAI-SearchBot (ChatGPT search), GPTBot (training) and ChatGPT-User, which fetches a page when a person asks; OpenAI says robots.txt rules may not apply to those user-initiated visits.

DON'T FLIP EVERY SWITCH

Preview and indexing rules work only if the page can still be fetched. Record each change with its date, verify the result and repeat your prompt panel. Don't apply every restrictive control just because it appears in the table.



NAME THE OUTCOME BEFORE CHOOSING THE SWITCH.

MARKDOWN FOR AGENTS

Many AI agents and retrieval tools would rather read Markdown than a full web page: no menus, scripts or cookie banners, and far fewer tokens. You can offer a clean Markdown copy of an important page and point to it openly. There are three ways.

1. LINK IT IN THE PAGE HEAD

```
<link rel="alternate" type="text/markdown"
href="https://example.com/pricing.md">
```

The HTML page announces its Markdown twin. lms.txt v2 (Aug. 10, 2026) names this exact link.

2. OR SEND IT AS A HEADER

```
Link: <https://example.com/pricing.md>;
rel="alternate"; type="text/markdown"
```

Same signal as an HTTP header, set at the server or CDN. It works for any file type.

3. OR ANSWER WHEN ASKED

```
GET /pricing Accept: text/markdown
-> 200 Content-Type: text/markdown Vary: Accept
```

Same URL, Markdown on request. Cloudflare's Markdown for Agents (Feb. 12, 2026) converts pages this way for sites that turn it on.



THE STANDARD

lms.txt v2 puts Markdown copies at page.md, links them with `rel="alternate" type="text/markdown"`, and uses `rel="describedby"` to point a page at its lms.txt. Chrome's Lighthouse checks lms.txt in an experimental Agentic Browsing audit (May 2026).

DO IT RIGHT

Same facts and dates as the HTML. On the .md send `Link: <page URL>; rel="canonical"` (a strong signal, not a guarantee). Rebuild it when the page changes. Link it openly; never pick the version by user agent.



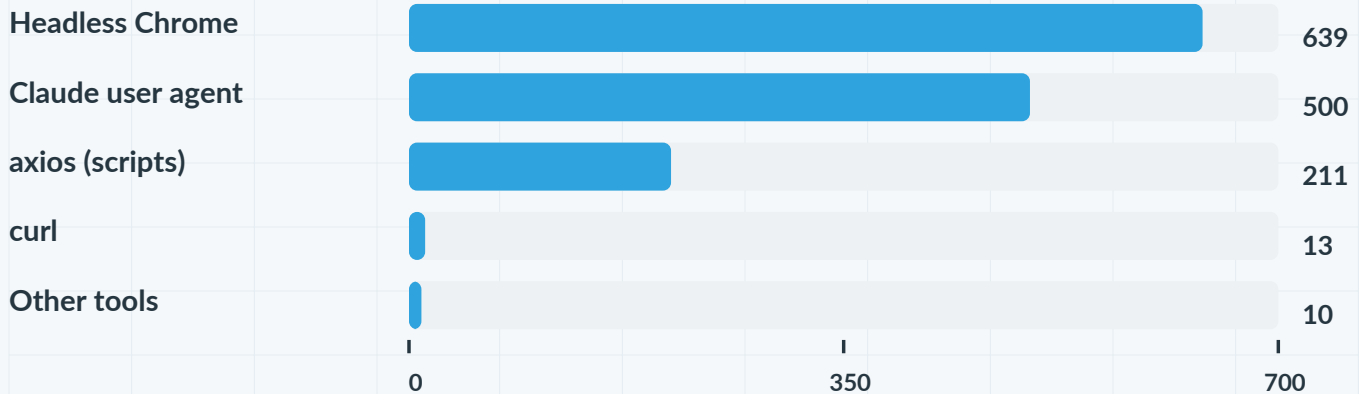
**GIVE AGENTS THE CLEAN COPY.
KEEP EVERY FACT IDENTICAL.**

lms.txt v2 is a proposal still open for feedback; Cloudflare converts pages only for sites that enable the feature. Copy-ready snippets: page 82.

MARKDOWN: WHAT THE EVIDENCE SAYS

WHO ASKED ONE SITE FOR MARKDOWN

Suganthan | Accept: text/markdown requests by user agent, March 7 to April 19, 2026 | one SEO blog on Cloudflare



WHAT GOOGLE SAYS

Search does not use it: "You don't need to create new machine readable files, AI text files, markup, or Markdown to appear in Google Search." On serving Markdown to bots, John Mueller (Feb. 2026): "Why would they want to see a page that no user sees?"

WHAT BING SAYS

Fabrice Canel (Feb. 2026): "We'll crawl anyway to check similarity," and "Non-user versions (crawlable AJAX and like) are often neglected, broken." A twin you don't maintain becomes a liability.

THE VERDICT: HELPS AGENTS, NOT RANKINGS

Worth doing for pages agents read to act: docs, pricing, specs, policies, product data. Cheap if your CDN converts pages. No study shows a citation lift, and it is not a Google ranking signal.

WHY WE STILL EMPHASIZE IT

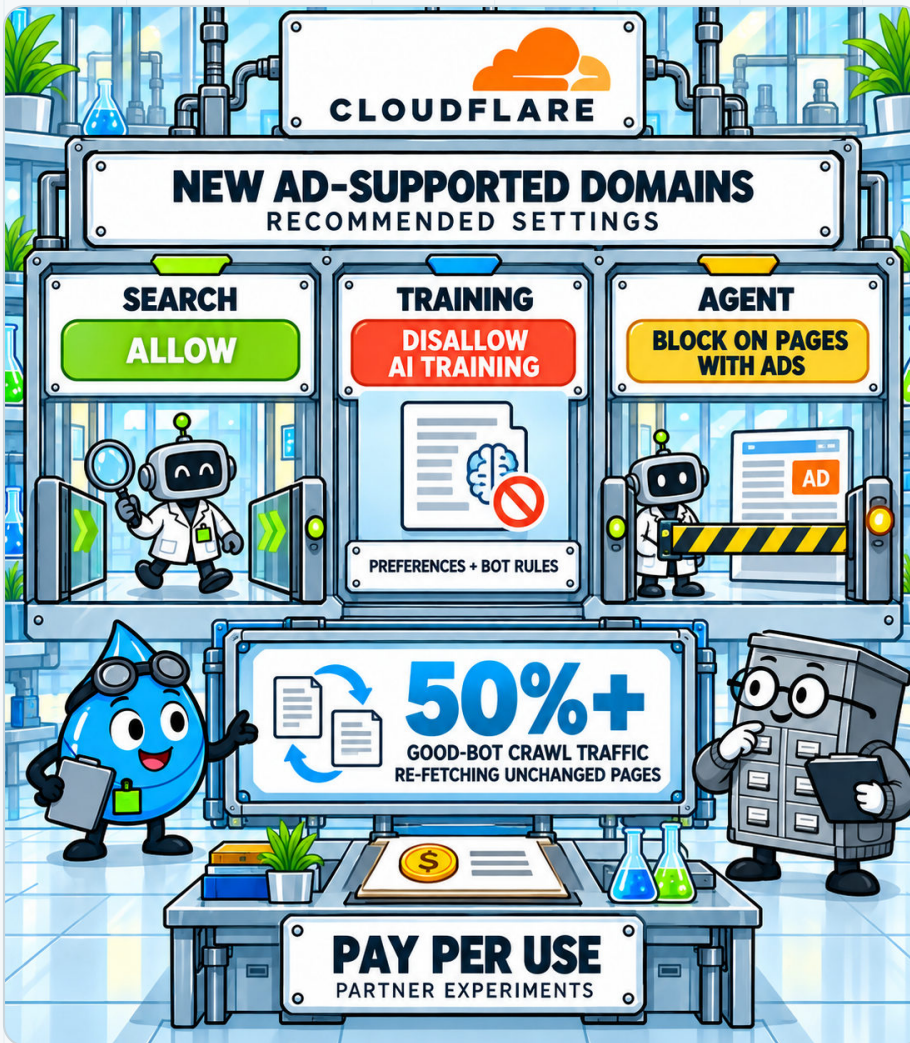
Agent traffic is real and growing on that site (about 15 a day in early March, about 30 in April), and agents are starting to browse and buy for people. A clean, accurate copy is a low-risk bet. Log .md and Accept: text/markdown requests to see who uses yours.



**REAL REQUESTS, UNPROVEN PAYOFF.
MEASURE IT ON YOUR OWN SITE.**

One site over 44 days; the post's own totals vary from 1,373 to 1,421 requests. Requests without the header are not counted, and no citation or traffic outcome was measured.

WHO PAYS FOR THE CRAWL



MORE THAN 50%

Share of good-bot crawl traffic spent re-fetching unchanged pages, according to Cloudflare's July 1, 2026 announcement.

PAY PER USE: A PILOT

Cloudflare is experimenting with payments tied to use, building on Pay Per Crawl. Ceramic.ai's model pays opted-in publishers when content appears in its search results; You.com can pay for access to premium content. These are partner pilots with different triggers, not a universal payment for every AI citation.

WHAT SHIPPED ON SEPTEMBER 15

Cloudflare's recommended preset for new ad-supported domains keeps Search allowed, sets Training to Disallow AI Training, and blocks Agents on pages with ads. Accountable mixed-use crawlers can still serve search. Existing settings migrate according to their previous configuration; inspect your actual domain settings (p. 44).



DECIDE BY PURPOSE. CHECK THE SETTING THAT ACTUALLY SHIPPED.

Cloudflare figures are company-reported. Settings described as of Cloudflare's Sept. 15, 2026 post; check your own dashboard.

WHO PAYS FOR THE CRAWL: YOUR SETTINGS

CLOUDFLARE PRESET FOR NEW AD-SUPPORTED DOMAINS (FROM SEPT. 15, 2026)

SETTING	PRESET	WHAT IT DOES
Search	Allow	Search crawlers can fetch pages. A Block setting would also stop Googlebot, Bingbot and Applebot, search included.
Training	Disallow AI Training	Publishes a no-training preference in robots.txt for accountable mixed-use crawlers, which stay allowed for search, and blocks other training crawlers. Bing is not covered yet, Cloudflare says.
Agent	Block on pages with ads	Agents are blocked on the pages that carry ads.

EXISTING DOMAINS

Settings carry over. If you never used the granular controls, the old Block AI setting maps across: Disabled becomes Allow everywhere; Block, or Block on pages with ads, becomes the preset above. If you did, your Search and Agent choices stay, and a Training block becomes Disallow AI Training.

THE TRADE-OFF

A hard block on Googlebot can also stop Search crawling. A training preference is a different choice. Google-Extended governs specified Gemini training and grounding uses; Search's generative AI inclusion control is separate (Publisher Controls, p. 39). Neither setting changes the need to verify what your CDN actually serves.

LICENSED SHELVES

Yelp announced its ChatGPT content integration in July 2026, bringing reviews, ratings, photos and business details into local answers.

DO THIS

Inspect Search, Training and Agent settings separately. Record the domain, previous choice, new choice and date. Verify crawler identity and the response it receives, then repeat your prompt panel.

IN PLAIN WORDS: A PREFERENCE OR A BLOCK

A preference is a published request that crawlers which honor it are expected to follow; a block refuses the request at the door. Cloudflare's training setting uses both: a no-training preference for accountable mixed-use crawlers, so they can keep serving search, and a block for other training crawlers. Choose by purpose, then check which of the two you actually have.

Settings as described in Cloudflare's Sept. 15, 2026 post; Yelp announcement as reported by Axios, July 23, 2026. Check your own dashboard.

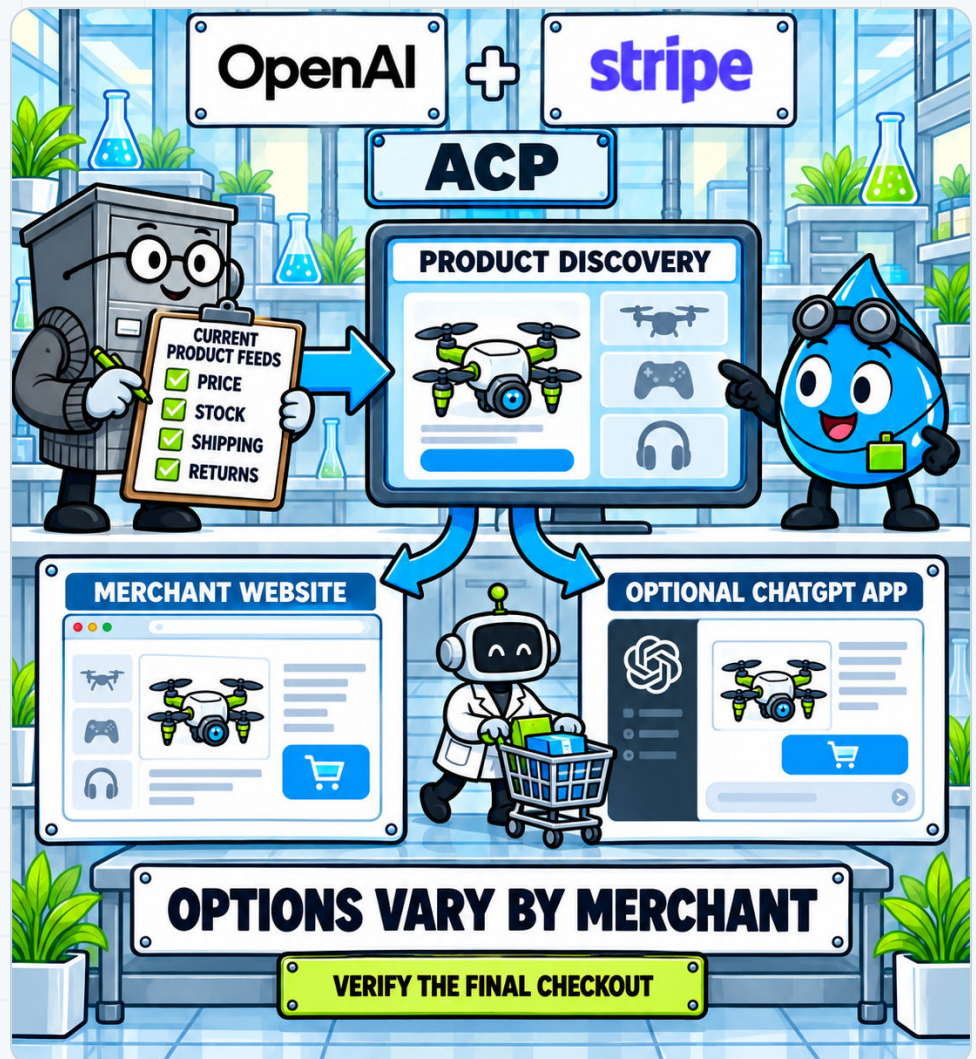
WHEN THE AGENT BUYS

TWO PROTOCOL EXAMPLES

ACP and UCP connect different parts of discovery and commerce. A protocol name does not mean every merchant supports checkout inside an assistant.

WHAT AGENTS READ

Keep price, stock, shipping and returns current across product pages and feeds. Answer buying questions such as compatibility, fit and what is included. Merchant Center's announced conversational attributes and AI insights are additional merchant tools, with their own eligibility and rollout.



Illustrative ACP path; UCP is discussed separately in the text. Options vary by merchant.



MAKE THE PRODUCT CLEAR, THEN TEST THE PATH TO PURCHASE.

Sources: OpenAI announcements of Sept. 29, 2025 and Mar. 24, 2026; Google announcements of May 20 and Oct. 15, 2026. All are in the Source Drawer.

WHEN THE AGENT BUYS: THE TIMELINE

WHAT HAPPENED, IN ORDER

DATE	WHO	WHAT HAPPENED
Sept. 29, 2025	OpenAI	ChatGPT launched Instant Checkout for U.S. Etsy sellers, built on the Agentic Commerce Protocol with Stripe. Shopify merchants were announced as coming soon.
Mar. 24, 2026	OpenAI	ACP expanded to product discovery and feeds, while merchants could use their own checkout. Shopify purchases could finish on merchant sites through an in-app browser. ChatGPT apps remained an optional deeper integration.
May 20, 2026	Google	Announced Universal Cart across Search and Gemini, with UCP supporting checkout on Google or a move to the merchant's site.
Oct. 15, 2026	Google	Said Universal Cart was rolling out in the U.S. in Search and the Gemini app. Participating stores offer checkout with Google Pay; shoppers can also move items to the merchant's cart.

ONE PRODUCT, SEVERAL PATHS

Discovery inside an assistant, checkout inside it, a handoff to the merchant's own site, or an app built for that assistant are different routes. Which ones exist depends on the merchant and the product.

SEPARATE OUTCOMES

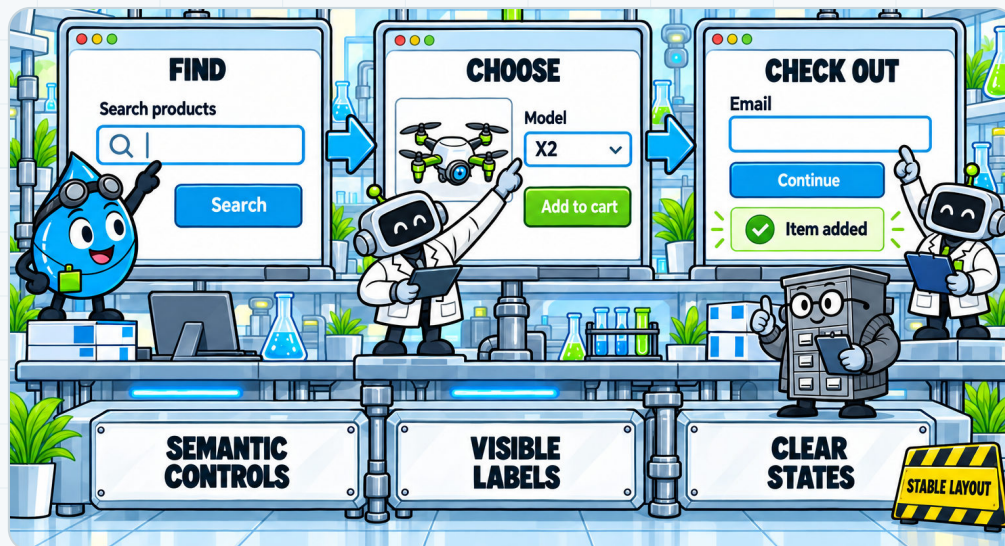
Discovery, a favorable recommendation, a visit and a completed purchase are separate outcomes. No protocol guarantees any of them, so count each one on its own.

DO THIS

Keep product data consistent with the transaction being offered. Test the actual discovery-to-checkout route available to your customers, including the merchant handoff (p. 47). Verify the final price, stock and confirmation state before treating the journey as complete.

Dates are the companies' own announcements. Availability varies by country, merchant and product; check before describing any path as available to your customers.

CAN AN AGENT COMPLETE THE TASK?



Getting cited and completing a task are different problems. Browser agents can inspect screenshots, the DOM and the accessibility tree, so the same page has to make sense to a person and to a machine reading its structure.

THE CHECKS

- Native buttons and links, visible field labels and understandable names.
- Clear selected, success, error and loading states.
- Stable layouts, with no overlays hiding controls.
- A tested route from finding a product to choosing an option to checking out, including the merchant handoff. Record where it fails.
- The right product, quantity and resulting state, with the user's required confirmation before a purchase or other consequential action.

WHAT THIS IS NOT

These are usability practices, not a citation boost or an accessibility certification. WebMCP is an optional emerging interface for supported agents, not a requirement for GEO.



CAN IT FIND THE CONTROL, UNDERSTAND THE CHOICE AND SEE WHAT HAPPENED?

Source: [web.dev](https://web.dev/build-agent-friendly-websites/), Build agent-friendly websites (last updated April 1, 2026). The guidance describes how browser agents operate; it makes no claim about citations or rankings.

WHAT'S HAPPENING

AI answers often cite sources you don't control: Wikipedia, Reddit, YouTube, reviews, news and research. What others say about you is part of the evidence an answer can draw on.

WHY IT MATTERS

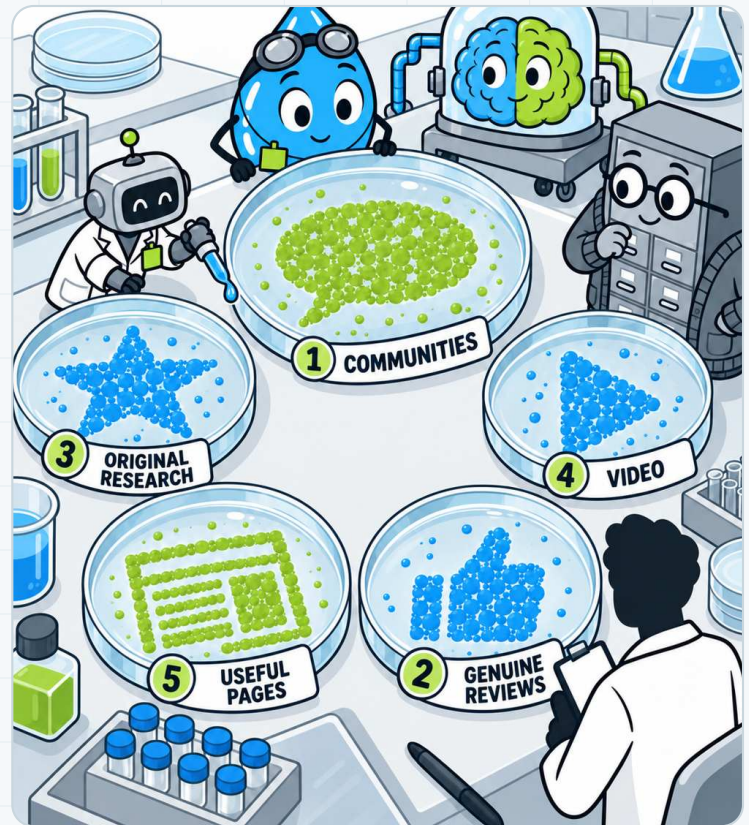
Brands mentioned more across the web tend to appear more in AI answers. That is a correlation, not a lever to buy, but it shows why genuine work that others can inspect and discuss is worth doing.

JOURDAN'S TAKE

Paid placement and independent corroboration are different things. Disclose sponsorship, publish a method people can inspect, and let the inspection be the reason they talk about it.

A PRACTICAL EXAMPLE

A drone shop publishes its flight-time test with the method, the conditions and the sample size, not just a boast. Reviewers and journalists can check it and point to it, and each page that does becomes evidence an AI answer can find.



OTHER PEOPLE CAN INSPECT AND CORROBORATE MY EVIDENCE.

IN THIS ROOM

49 Web Mentions and AI Visibility
50 Help the Community

51 Useful Video
52 Search With a Camera

53 Evidence to Coverage
54 GEO for the Local Shop

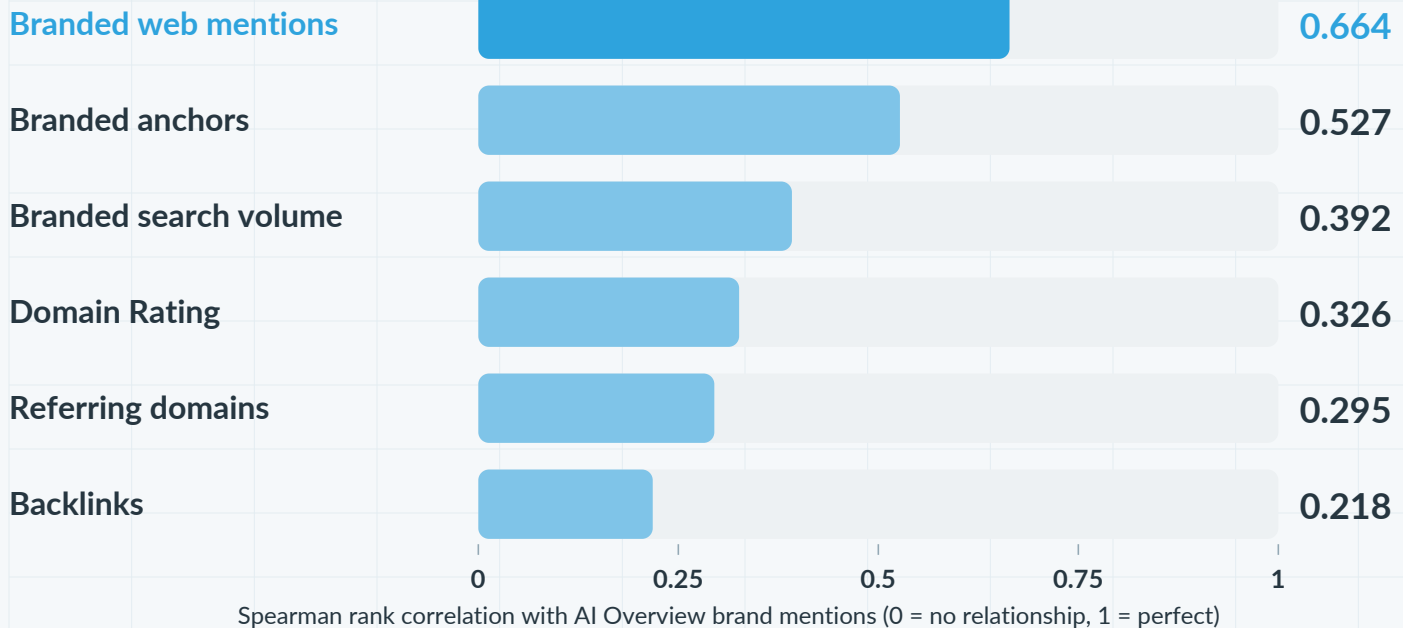
Inspect the sources actually cited for your tasks; do not read a universal trust weighting into a correlation.

WEB MENTIONS AND AI VISIBILITY

Which signals travel with being named in Google's AI Overviews? Ahrefs ranked six brand signals by how closely they moved with AI Overview brand mentions across a selected sample of brands.

ASSOCIATION WITH AI OVERVIEW BRAND MENTIONS

Ahrefs | Published: May 2025 | Selected sample; collection date / final n not stated | Spearman rank correlation, 0 to 1



MENTIONS LEAD

Brands talked about across the web were named more often in AI Overviews: 0.664 for branded web mentions, against 0.218 for backlinks.

169 VS 14

Brands in the top quarter for branded mentions averaged 169 AI Overview mentions; the bottom quarter averaged 14.

NOT A LEVER TO BUY

A correlation, not a cause. Earn mentions with work people want to cite (page 53); bought mentions are not the same thing.

How to read these numbers. Each score shows how closely two rankings moved together in this sample, from 0 (no relationship) to 1 (a perfect match). It does not show that mentions cause citations: well-known brands collect both. Use it to see where your brand is discussed, and where it is missing.



THAT IS AN ASSOCIATION TO INVESTIGATE, NOT A LEVER TO BUY.

Data: brand snapshot, collection dates not stated / Published: May 26, 2025. Spearman correlations; zero-AIO candidates excluded. Association, not a causal weight.

HELP THE COMMUNITY

SHARES SWING

Reddit was 11.97% of tracked ChatGPT browsing citations in Jan.-Feb. 2026 (Similarweb). In August 2026, Promptwatch saw Reddit's ChatGPT Search share fall to 0.52%.

PARTICIPATE DIRECTLY

Participate as a person.
Answer the actual question.
Disclose relevant affiliation.
Never fake a thread.

PRACTICAL EXAMPLE

Answer the actual question in the thread where it is asked, say plainly that the page you link is yours, and never manufacture a recommendation thread.



**ANSWER FOR THE PEOPLE IN THE THREAD.
THEY STAY WHEN THE CITATIONS MOVE.**

Similarweb U.S. monitored ChatGPT browsing citation events, Jan.-Feb. 2026; Promptwatch ChatGPT Search citation share, July-Aug. 2026. Different trackers and samples.

The two studies use different surfaces and denominators: Ahrefs, March 2026; Similarweb, Jan. to Feb. 2026.

SEARCH WITH A CAMERA



100B+

Visual searches with Google Lens in the first months of 2025, up 65% year over year (Sundar Pichai, Google I/O, May 2025).

NOW MEASURABLE

On Sept. 24, 2026, Search Console added a Web: Multimodal search type for results triggered by an image, photo or screenshot (Lens, Circle to Search, image uploads, Search This Image). It shows no queries.

WHAT A CAMERA SEARCH NEEDS

Original photos of the real product or place from several angles; the image beside the text that names it; descriptive file names and alt text; product data and Merchant Center images that match the page.

DO THIS

Photograph what you sell the way customers see it. Name and describe it next to the image. Compare Multimodal and Text-based clicks in Search Console each month.



SOMEONE WILL POINT A CAMERA AT YOUR PRODUCT. MAKE SURE THE PAGE THEY LAND ON NAMES IT.

EVIDENCE TO COVERAGE



WHY FINDINGS TRAVEL

Branded web mentions had the strongest association with AI Overview brand visibility (0.664) in Ahrefs' sample (page 49). Bing (Feb. 2026): "Examples, data, and cited sources help build trust when content is reused in AI-generated answers."

THE PATH ON THE PLATE

Raw data, then method, then study, then coverage: publish the numbers, show how you got them, package the finding, and only then pitch it.

PRACTICAL EXAMPLE

Publish the finding with its raw numbers, sample size and method, so journalists and researchers can check it before they cite it. Then record coverage, citations and visits as separate results.

WHAT A CITABLE STUDY SHOWS

The question, the sample size and dates, the method, the raw numbers or a downloadable table, the limits, and a named author.



A FINDING TRAVELS FARTHER WHEN THE METHOD TRAVELS WITH IT.

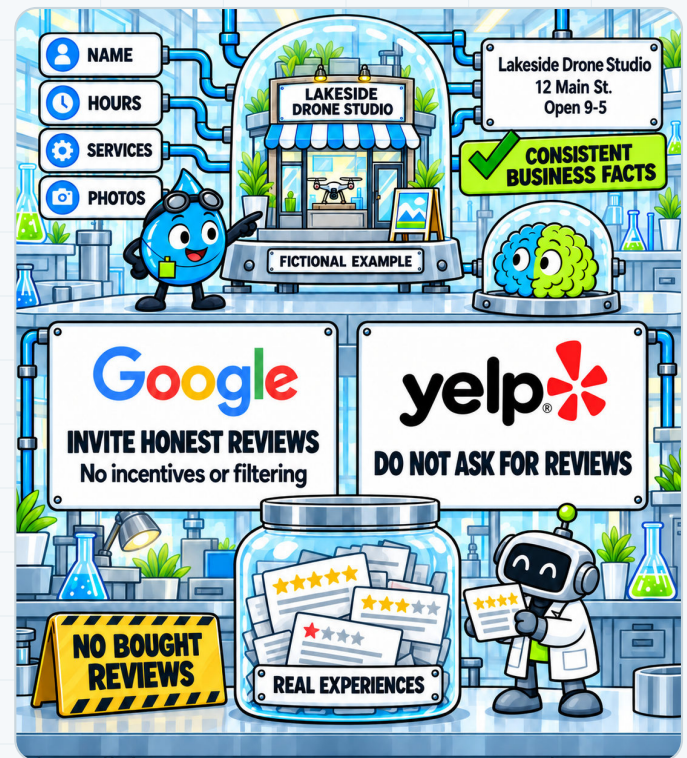
GEO FOR THE LOCAL SHOP

YELP IN CHATGPT

Reviews, ratings, photos and business details are licensed into ChatGPT's local answers (announced July 2026). Yelp's Request a Quote is announced for ChatGPT, with no date.

WHERE LOCAL ANSWERS COME FROM

Business profiles (Google Business Profile, Bing Places, Apple Business Connect), review sites (Yelp also feeds Apple Maps), your own site, local press and directories.



DO THIS

Keep names, contact details, services and hours accurate wherever the business appears. Use real photos and clear service pages with relevant service areas and prices where useful. Follow each review platform's rules: Google allows requests for genuine reviews without incentives or steering, but prohibits selectively seeking positive reviews. Yelp tells businesses not to solicit Yelp reviews. Respond to feedback without trying to manufacture consensus.

DON'T

Buy reviews, offer rewards for reviews, screen invitations by customer satisfaction, or publish interchangeable town pages without useful local information (The Black-Hat File, p. 73).



BE ACCURATE EVERYWHERE. RESPECT EACH PLATFORM'S REVIEW RULES.

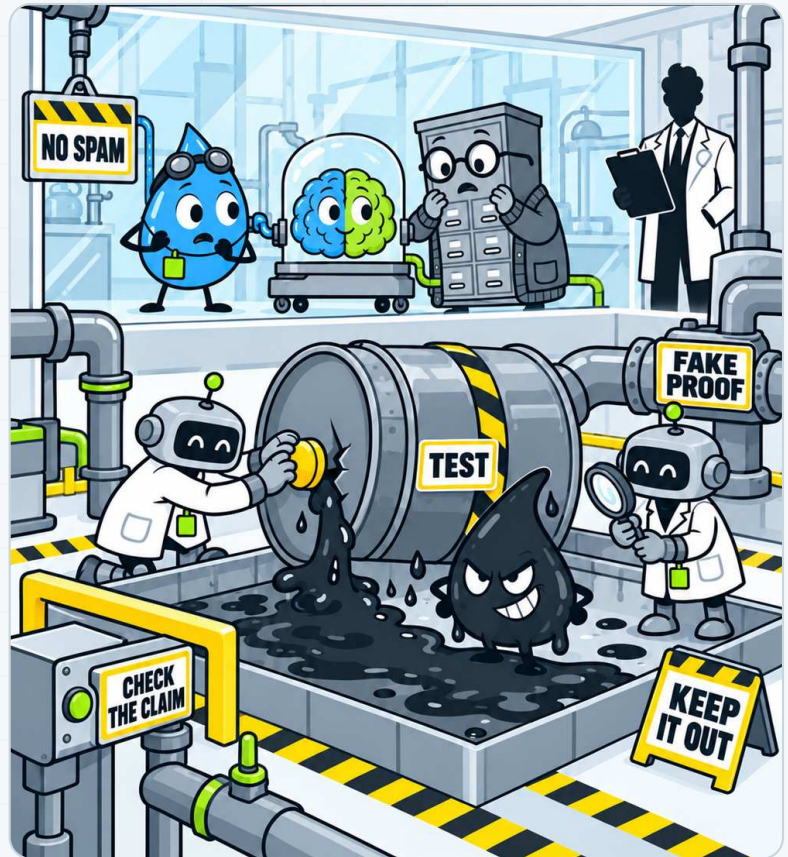
Lakeside Drone Studio is a fictional example. Reviews that show up as cited sources do not prove that collecting reviews causes citations.

WHAT'S HAPPENING

Some people try to game AI answers with hidden prompts, fake reviews, self-ranking 'best of' lists, pages mass-produced to manipulate rankings, or links planted on hacked sites.

WHY IT MATTERS

These tricks put false claims into the record that answers are built from, many break search engines' spam policies, and none of them teaches you anything you can reuse.



JOURDAN'S TAKE

I built a Black Hat SEO Terminal to study this stuff, and that's exactly why none of it belongs in this lab. Fabricated proof puts a false claim into the record, and the record is what we're trying to improve.

A PRACTICAL EXAMPLE

A rival keeps appearing in AI answers for your key question. The shortcut is a planted thread or a few fake reviews. The clean fix is slower and it lasts: improve the page, publish the proof, and reach out openly to people who might cover it.



**SLUDGE AND FABRICATED PROOF
AREN'T COMING INTO THIS LAB.**

IN THIS ROOM

56 Why Formulas Don't Travel

57 The Clean Room Pledge

WHY FORMULAS DON'T TRAVEL

LAB NOTE 01

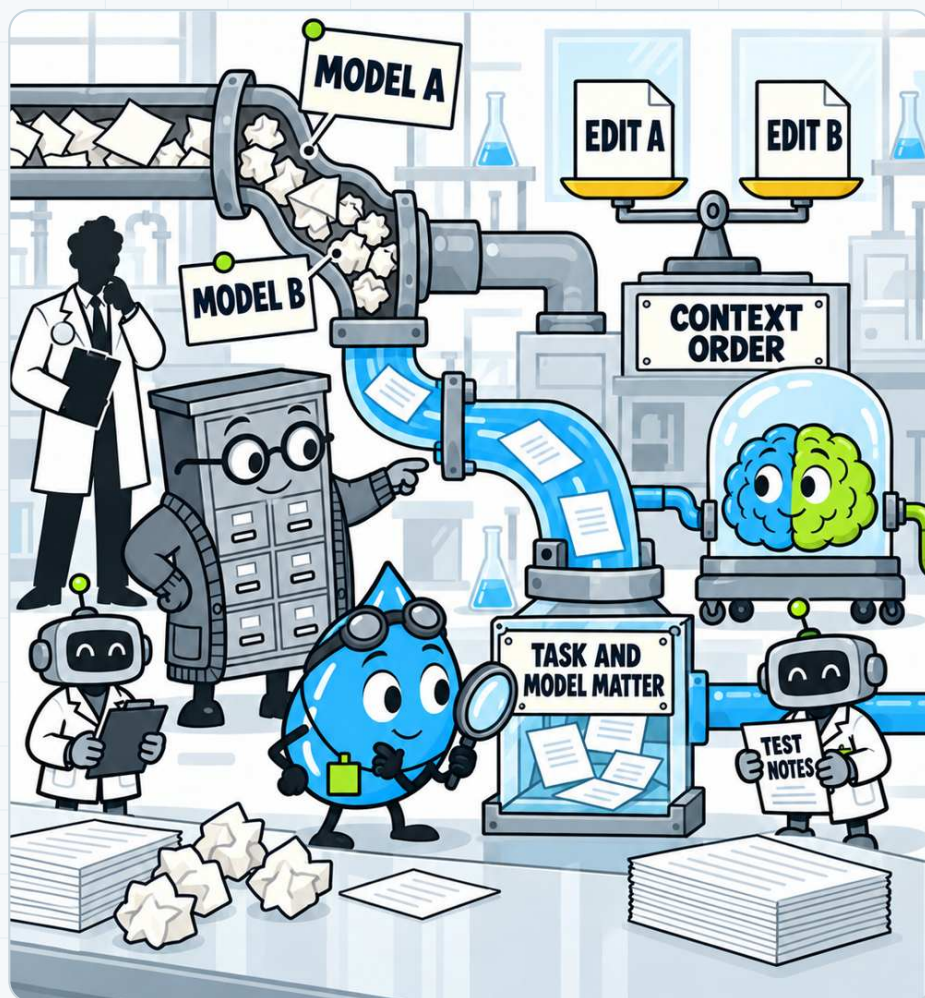
Most tested white-hat edits gave little citation-rank benefit in C-SEO Bench.

LAB NOTE 02

Effects varied by task and model. Gains per adopter shrank as competing sources used effective edits in the simulations.

LAB NOTE 03

Position in the supplied context mattered more than most edits.



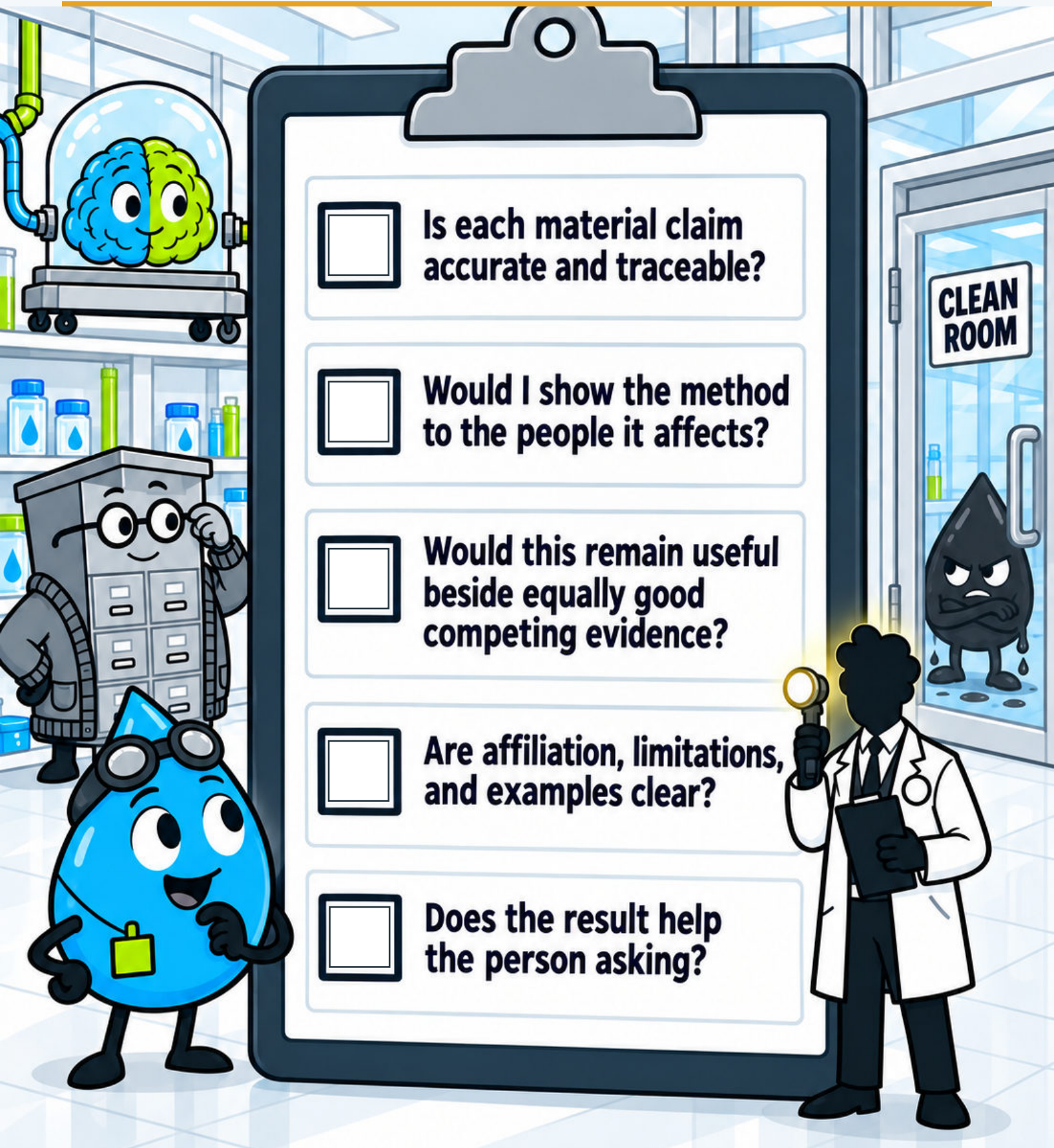
WHAT THIS MEANS FOR YOU

Every few weeks someone publishes a formula for AI citations. This benchmark is a reason to treat those formulas with care. The same edit helped in one setup and barely mattered in another, depending on the task, the model and the competing sources, and the gain shrank as more sources used it. It also tested controlled edits to supplied documents, not a live site. Treat a published result as a test to run: try it on a few pages, keep a comparison group, and measure it with a fixed prompt panel (Room 08, p. 58).



**A RESULT IN ONE SETUP IS
THE START OF ANOTHER TEST.**

THE CLEAN ROOM PLEDGE



WHAT'S HAPPENING

AI answers vary. Ask the same question twice and you may get different sources, so one screenshot can't tell you whether you're winning.

WHY IT MATTERS

Measuring GEO means sampling: a fixed set of questions, asked repeatedly and recorded the same way. Then you can count mentions, citations, accuracy and visits, and see whether a change made a difference.

JOURDAN'S TAKE

One screenshot of an AI naming your brand is one observation. It can't show consistency or whether a change made a difference. Run the same prompts again next week and count.



PUT ME UNDER THE MICROSCOPE MORE THAN ONCE.

IN THIS ROOM

59 The Prompt Test Rack
60 Check the Answer

61 Scoring Answers
62 The Gauge Panel

63 The Tracker Bench
64 Built in This Lab

THE PROMPT TEST RACK

THE PANEL, IN FIVE STEPS

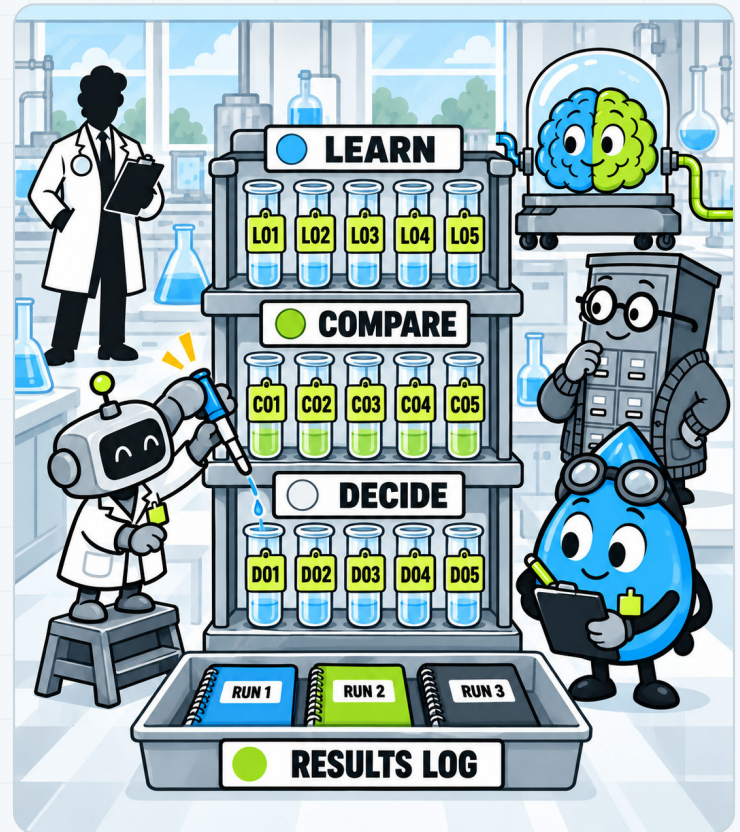
1. Define real audience tasks and constraints.
2. Group a fixed panel into learn, compare, and decide.
3. Repeat across runs and days; preserve surface, locale, and prior context.
4. Record mentions, citations, source URLs, accuracy, and fit.
5. Report counts, denominators, exclusions, uncertainty, and a comparison.

THE RACK

Each shelf holds one task group: LEARN, COMPARE and DECIDE, with five tasks each (L01-L05, C01-C05, D01-D05). The RUN 1-3 logbooks in the results tray hold the repeated runs.

WHAT A ROW MUST SAY

Who asked (audience, locale), where (surface, model, search on or off), with what context, which repeat, and what came back: named, cited, accurate, complete.



AN EXAMPLE ROW (ILLUSTRATIVE)

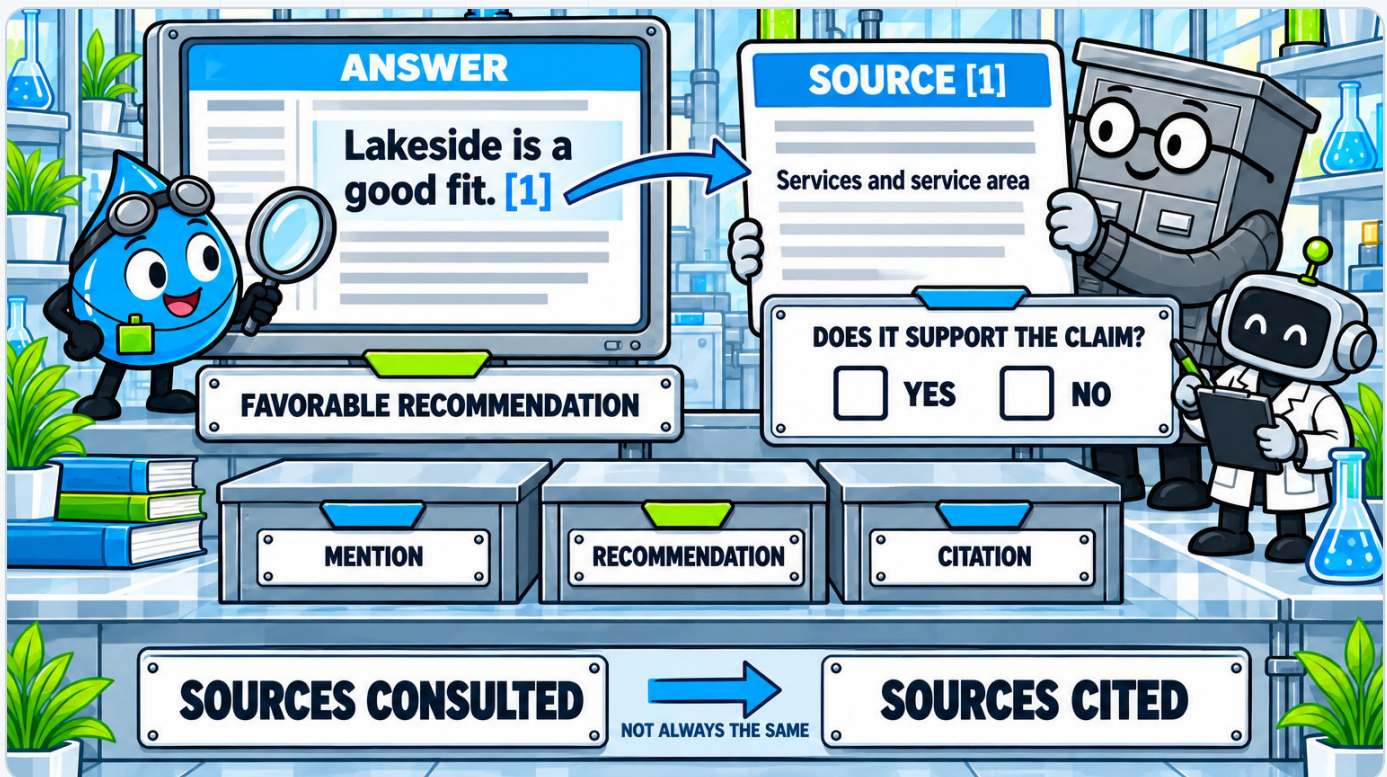
Task L03, "best drone for real estate photography", asked in Google AI Mode with search on, repeat 2 of 3. Brand mentioned: yes. Your domain cited: no. Accuracy: correct. One row like this tells you little. Fifteen tasks, each run three times, start to show a pattern you can compare next month.



**SAME TASKS, REPEATED RUNS,
A RECORD WE CAN INSPECT.**

Completed refusals and no-answer outcomes stay in the trial count; report instrument failures separately. The full 22-field panel is on page 80.

CHECK THE ANSWER, THEN THE SOURCE



Fictional example. Neither support box is selected: the picture asks the question, it does not answer it.

Count a mention, a recommendation and a citation separately. A favorable sentence can lack a source; a cited page can fail to support the claim; a correct source can be used in an inaccurate or unsuitable answer.

In the fictional illustration, a page describing Lakeside's services and service area does not by itself prove "Lakeside is a good fit." Check the user's task and the exact claim before scoring it.

MENTION

The answer names the brand. Nothing more is implied.

RECOMMENDATION

The answer suggests the option favorably or unfavorably, with or without a citation.

CITATION

The answer shows a link or source marker for a page. Sources consulted and sources cited are not always the same.



A FAVORABLE ANSWER STILL NEEDS EVIDENCE THAT FITS.

SCORING ANSWERS AND SOURCES

A PROPOSED EDITORIAL SCORING RUBRIC

Not a vendor ranking model or a validated universal score. Preserve the exact answer, the relevant claim, the cited passage, the source date and the user's constraints.

CHECK	RECORD ONE VALUE	EVIDENCE TO RETAIN
Brand mention	yes / no	Exact answer text
Recommendation	favorable / neutral / unfavorable / none	Sentence and the option it concerns
Owned-domain citation	yes / no	Canonical URL and visible marker
Claim-source support	supported / unsupported / mixed / unclear	Claim paired with the cited passage
Factual accuracy	correct / wrong / unclear / not applicable	Checked fact, source date and conditions
Task fit	fits / does not fit / unclear / not applicable	User constraints and the answer's reason

DENOMINATORS

Keep completed refusals and no-answer outcomes in the trial denominator, and report instrument failures separately. Report mention, recommendation and domain-citation frequencies over all completed trials. For quality checks, state the separate eligible denominator: reviewed claims, claim-source pairs or answers. Show unclear and not-applicable counts; do not quietly turn them into passes.

CONSULTED IS NOT CITED

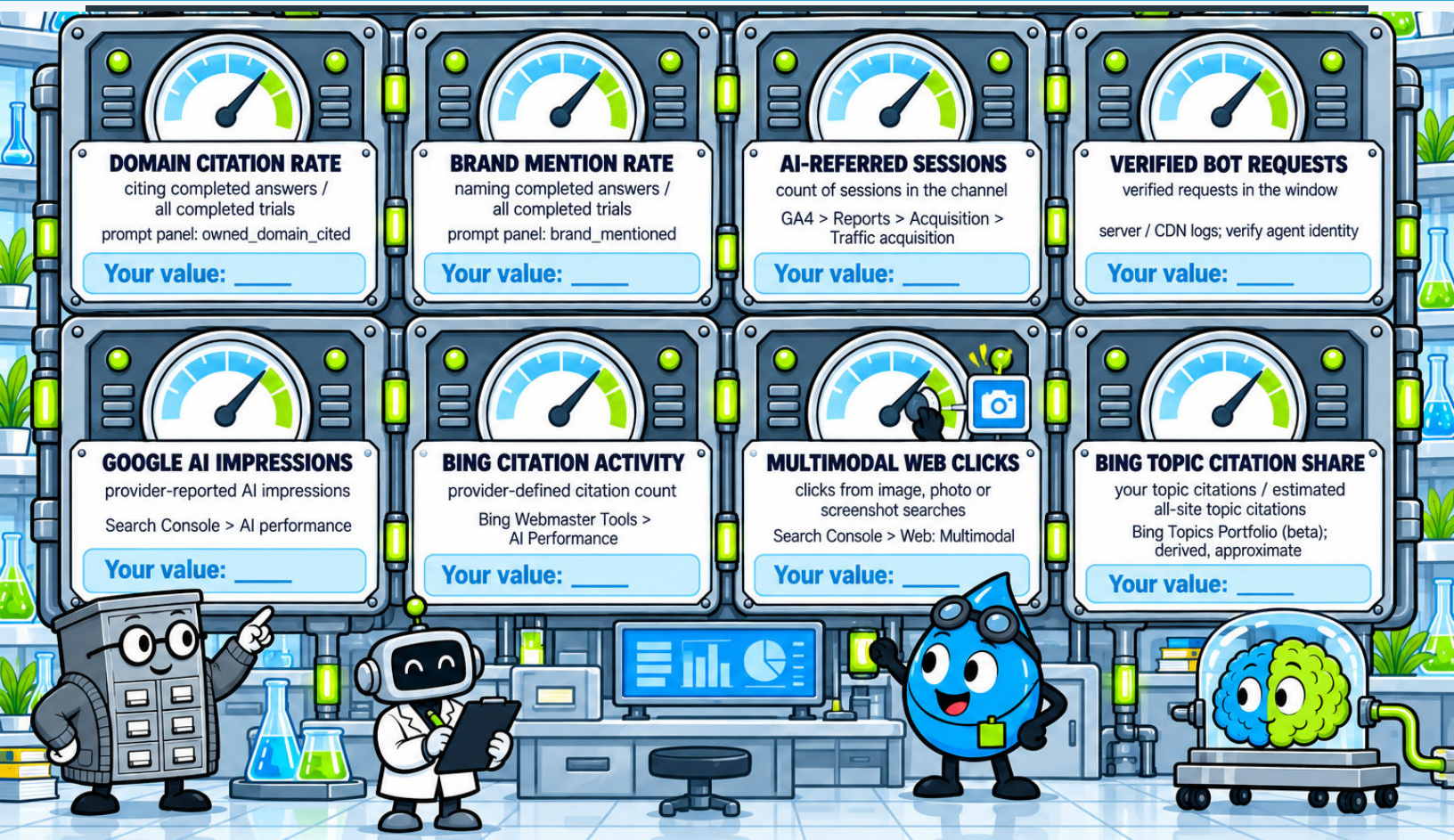
Sources consulted and sources visibly cited are not always the same, and a visible citation is not a complete causal trace. Do not claim a model retrieved a page solely because it printed its URL. OpenAI documents this difference for its API's web search tool; that is not a promise about every assistant's interface.

IN THE FIELD KIT

Keep the existing 22-column prompt-panel.csv (page 80). Use accuracy, fit and evidence_note for these checks; do not replace brand_mentioned with recommendation. If you add structured recommendation or support columns, version the schema and update every copy of the file together.

This rubric is proposed by this guide. It is not a vendor ranking formula, and no score here is a measured result.

THE GAUGE PANEL



Different gauges answer different questions. Provider numbers are approximate and can change scope without notice.

FROM YOUR OWN TESTS

Domain citation rate and brand mention rate, counted from a fixed prompt panel (Field Kit: Prompt Panel, p. 80).

FROM ANALYTICS AND LOGS

AI-referred sessions in GA4 (Field Kit: GA4 AI Traffic, p. 81) and verified bot requests in server or CDN logs.

FROM PROVIDER DASHBOARDS

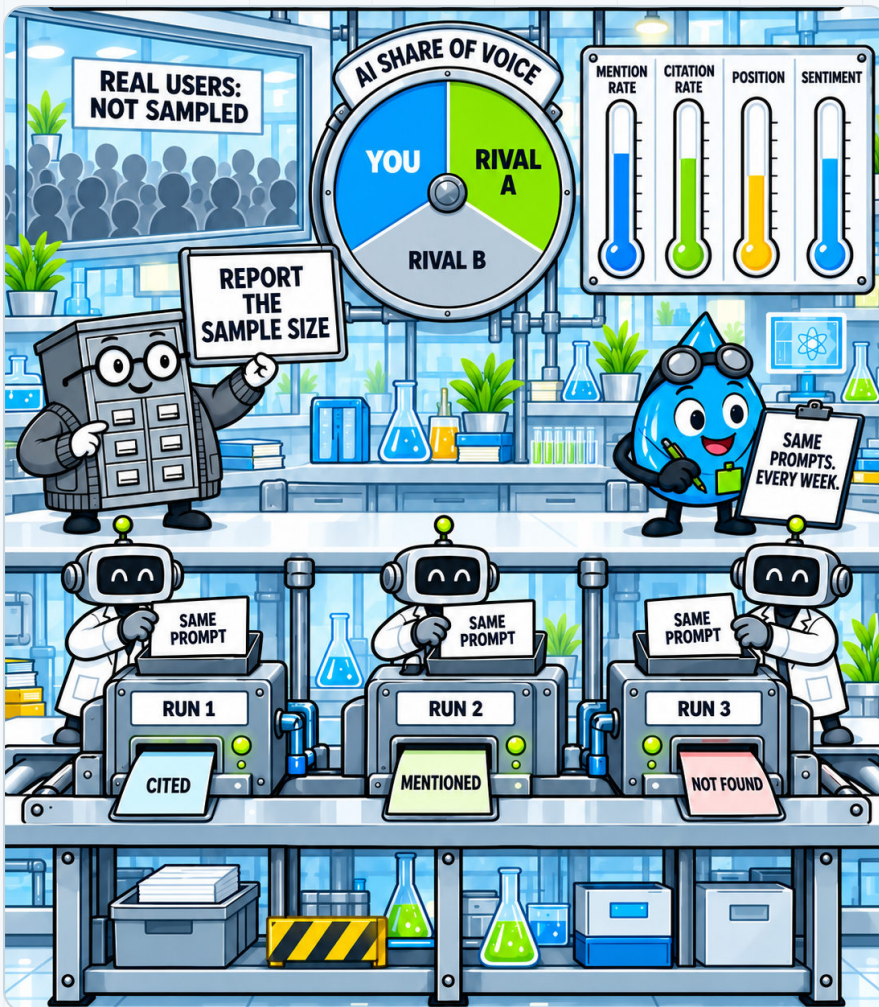
Google AI impressions and Multimodal web clicks in Search Console; citation activity and topic citation share in Bing Webmaster Tools (p. 63).



DIFFERENT GAUGES ANSWER DIFFERENT QUESTIONS.

GA4's AI Assistant channel excludes Google AI Search referrals. Inspect the published 90-day example on page 75.

THE TRACKER BENCH



RUN IT 3 TIMES

The same prompt can give a different answer on each run, so one run is an anecdote.

WHAT TRACKERS MEASURE

Mention rate, citation rate, position in the answer, share of voice against named rivals and sentiment, from a fixed prompt set run on a schedule.

WHAT THEY CAN'T SEE

Real users' prompts, memory and personalization, location and logged-in features. Collection can break too: Promptwatch couldn't rule out a data issue in the August Reddit drop (p. 23).

FIRST-PARTY GAUGES

Search Console's AI performance report; Bing's AI Performance report with Topics Portfolio (beta, Oct. 2026); Merchant Center's AI performance insights (announced May 2026).

BUY OR BUILD

Tools such as Profound, Peec AI, Semrush and Ahrefs Brand Radar save time; the prompt panel CSV in the field kit (p. 80) does the same job in a spreadsheet. Either way: fixed prompts, repeated runs, the sample size reported.



**A TRACKER IS A THERMOMETER,
NOT THE WEATHER.**

Three free tools I built to make the ideas in this book visible. Each one runs in the browser.



Try the "Will your page get cited?" simulator at the end. It is a teaching model, not a live measurement.

OPEN THE TOOL



Try the link farm, PBN and parasite page scenarios. An illustrative page-level model, not Google's data.

OPEN THE TOOL



Press / to search, and read any attack chapter beside Detection & Defense. Written for defenders.

OPEN THE TOOL

64

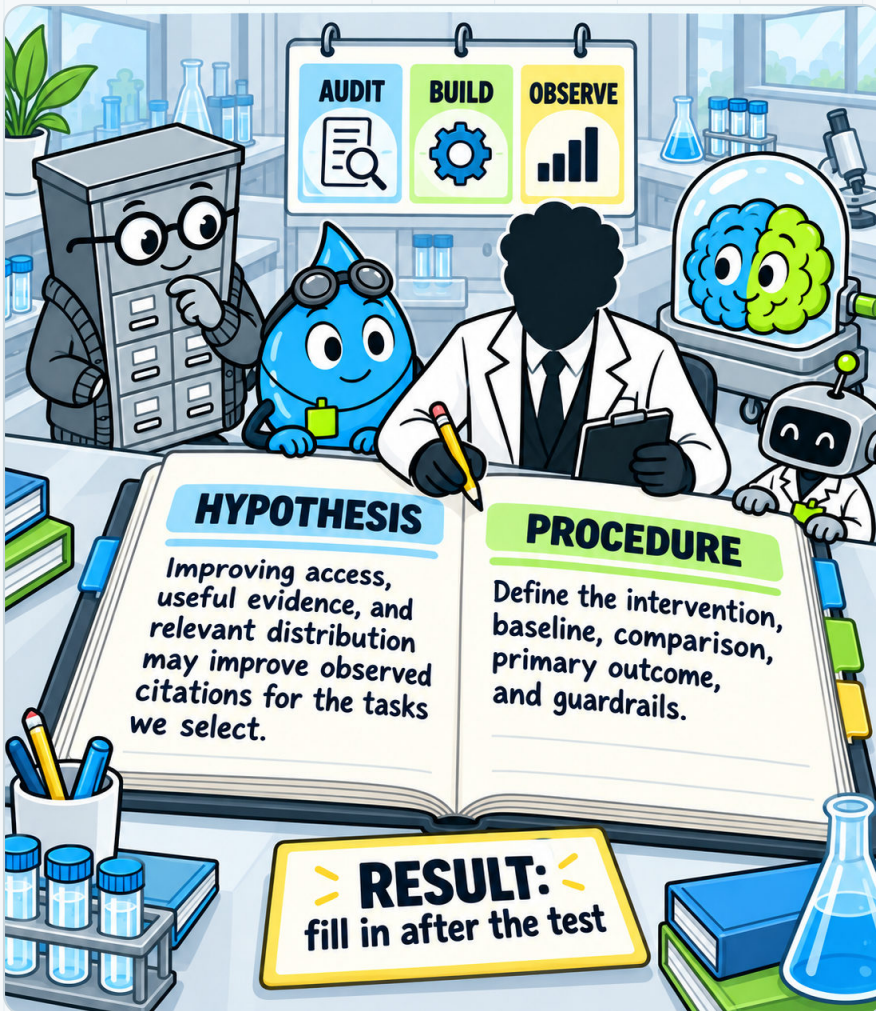
THE 90-DAY PROTOCOL

09

DAYS 1-30 / AUDIT

DAYS 31-60 / BUILD

DAYS 61-90 / OBSERVE



JOURDAN'S TAKE

Write down what success looks like before you touch the site. Deciding afterward makes it too easy to read any movement as a win.

WHY A WRITTEN PLAN

AI answers shift for reasons you do not control. A plan with success defined in advance keeps you from reading noise as a win.

RESULT: EMPTY ON PURPOSE

Record what happened, including uncertainty and negative findings, only after the test. The result card on the desk stays blank until then.



BEFORE YOU CHANGE ANYTHING, WRITE DOWN WHAT SUCCESS LOOKS LIKE.

IN THIS ROOM

66 Audit, Then Build

67 Grow, Measure, Repeat

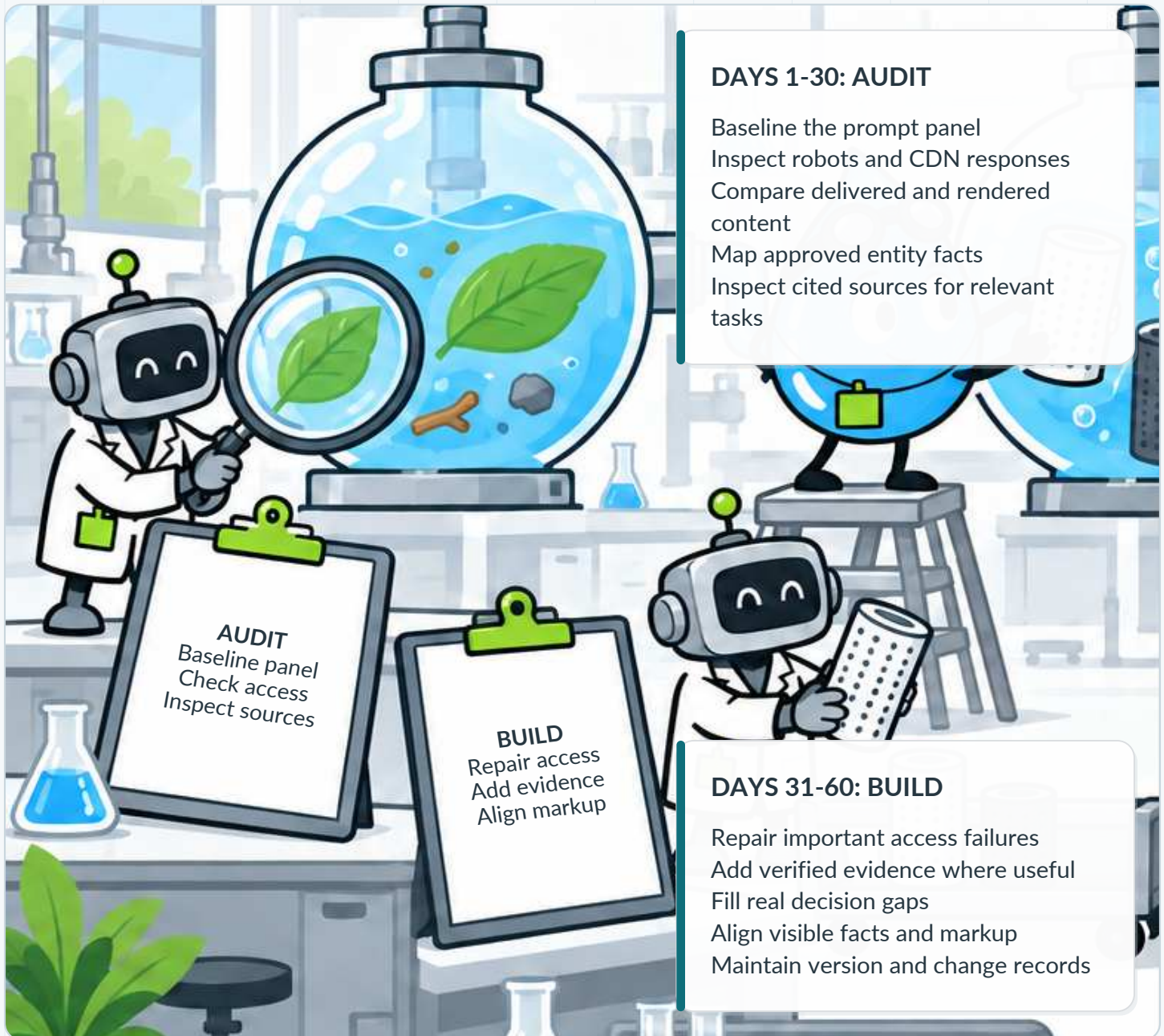
A planning schedule is not a citation guarantee. The result area stays empty until the test has run. Score answers and their sources with the rubric on page 61.

AUDIT, THEN BUILD

DAYS 1-30 / AUDIT

DAYS 31-60 / BUILD

DAYS 61-90 / OBSERVE



**REPAIR THE ROUTE, THEN
STRENGTHEN THE EVIDENCE.**

GROW, MEASURE, REPEAT

DAYS 1-30 / AUDIT

DAYS 31-60 / BUILD

DAYS 61-90 / OBSERVE

DAYS 61-90

Publish a useful evidence asset
Contribute to relevant communities
Adapt a strong demonstration into video
Repeat the fixed prompt panel
Compare with an appropriate holdout
Audit accuracy and qualified outcomes

COMPARE FAIRLY

Keep the core panel stable. Group paraphrases and repeats by information need. For nonrandom comparisons, assess plausible parallel untreated trends.

MEASURE
Repeat panel
Compare holdout
Review quality



**PUBLISH, OBSERVE, COMPARE,
AND MAKE THE NEXT DECISION.**

Account for repeated observations and plausible untreated trends before attributing a change.

THE TACTICS BOARD

SEO's are trying dozens of tricks to get into AI answers: new files, new formats, new writing rules, bought mentions and outright spam. This board checks 29 of them against published tests and against Google's and Bing's own rules.



WHAT EACH VERDICT MEANS

SUPPORTED PRACTICE

Backed by a test, an observed pattern or official guidance. Each card says which.

HELPS AGENTS

Low cost. Helps AI agents and tools read you. No citation lift shown.

NO PROVEN CITATION LIFT

Tested without a clear lift. A small benefit can't be ruled out.

NOT REQUIRED

The search engine says you do not need it.

RISKY

May work for a while. Trust or policy risk.

BANNED

Named in Google or Bing rules, or deceptive by design.

PAID

Bought placement, labeled and separate from answers.

EVERY TACTIC AT A GLANCE > PAGE 69

FILES, FORMATS AND ACCESS > page 70

llms.txt, Markdown copies, content negotiation, schema, chunking, IndexNow and crawler access. **2 supported practice, 3 help agents, 1 no proven citation lift, 1 not required, 1 risky.**

WRITING TACTICS > page 71

Answer-first pages, statistics and sources, question headings, specific language, updates, fake dates, self-ranking lists and original data. **6 supported practice, 2 risky.**

OFF-SITE AND PAID TACTICS > page 72

Community answers, bought Reddit accounts, digital PR, reviews, Wikipedia, ChatGPT ads and product feeds. **4 supported practice, 2 risky, 1 paid.**

THE BLACK-HAT FILE > page 73

Hidden prompts, cloaking, memory poisoning, scaled content abuse, keyword stuffing, fake reviews. **All 6 banned.**



**EVERY TRICK GETS THE SAME QUESTION:
WHAT DOES THE EVIDENCE SAY?**

Verdicts reflect evidence and rules published through Oct. 29, 2026. A non-significant result is not proof of zero effect; a new test or policy can change a verdict. Every source and its date is in the Source Drawer.

EVERY TACTIC AT A GLANCE

All 29 tactics on one page, grouped by where they happen. Click any row for the evidence, its type and the safe way to do it.

FILES, FORMATS AND ACCESS p. 70 >		OFF-SITE AND PAID TACTICS p. 72 >	
llms.txt file	HELPS AGENTS	Genuine community answers	SUPPORTED PRACTICE
Markdown copies (rel="alternate")	HELPS AGENTS	Buying Reddit accounts	RISKY
Accept: text/markdown requests	HELPS AGENTS	Earned mentions and digital PR	SUPPORTED PRACTICE
Markdown served only to bots	RISKY	Genuine customer reviews	SUPPORTED PRACTICE
Schema markup for AI citations	NO PROVEN CITATION LIFT	Writing your own Wikipedia page	RISKY
"Chunking" pages for AI	NOT REQUIRED	ChatGPT ads	PAID
IndexNow	SUPPORTED PRACTICE	Product feeds and AI checkout	SUPPORTED PRACTICE
Crawler access, readable HTML	SUPPORTED PRACTICE		
WRITING TACTICS p. 71 >		THE BLACK-HAT FILE p. 73 >	
Answer first	SUPPORTED PRACTICE	Hidden prompt injection	BANNED
Real statistics, quotes, sources	SUPPORTED PRACTICE	AI-targeted cloaking	BANNED
Question headings, direct answers	SUPPORTED PRACTICE	"Summarize with AI" poisoning	BANNED
Specific, entity-rich language	SUPPORTED PRACTICE	Scaled content abuse	BANNED
Updating the facts on a schedule	SUPPORTED PRACTICE	Keyword stuffing, engineered copy	BANNED
Changing only the date	RISKY	Fake reviews and testimonials	BANNED
Self-promotional "best X" lists	RISKY		
Original research and data	SUPPORTED PRACTICE		
		THE COUNT: 29 TACTICS 12 supported practice; 3 helps agents; 1 no proven citation lift; 1 not required; 5 risky; 6 banned; 1 paid. What each verdict means: page 68 >	

Verdicts reflect evidence and rules published through Oct. 29, 2026. A new test or policy can change a verdict; every source and its date is in the Source Drawer.

FILES, FORMATS AND ACCESS

Files and settings people add "for AI": what each does, what the evidence shows, and the safe way to do it.

llms.txt file

HELPS AGENTS

Evidence: observational study, official guidance. Google's guide: it "will neither harm nor help" your visibility in Google Search. SE Ranking found no citation link across about 300,000 domains. Chrome's Lighthouse checks it for agents (May 2026).

Do this: Publish one if agents use your docs or catalog. Write links as [text](url). Expect no ranking boost.

Accept: text/markdown (content negotiation)

HELPS AGENTS

Evidence: one-site log. Same URL, Markdown on request. Cloudflare converts automatically for enabled sites (Feb. 2026). One site logged about 31 such requests a day, mostly AI pipelines, none from GPTBot.

Do this: Turn it on if your CDN offers it, send Vary: Accept, and spot-check the converted text.

Schema markup to win AI citations

NO PROVEN CITATION LIFT

Evidence: matched observational study. Ahrefs (May 2026): 1,885 pages that added schema, against matched controls. AI Mode +2.4% and ChatGPT +2.2%, both statistically indistinguishable from zero; AI Overviews -4.6%, cause unclear. The authors can't tell whether schema did a tiny bit of good or nothing. Google: no special schema is needed.

Do this: Keep accurate schema for rich results and identity. Don't count on it for citations.

Crawler access and readable HTML

SUPPORTED PRACTICE

Evidence: official documentation. A blocked, failing or script-only page can't be fetched for a fresh answer. OpenAI, Anthropic and Perplexity each document separate bots for training, search and user requests.

Do this: Allow the search bots you want, return 200s, keep key facts in delivered HTML (Room 05, page 34).

Markdown copies linked with rel="alternate"

HELPS AGENTS

Evidence: proposed standard, observed requests. llms.txt v2 (Aug. 2026) standardizes page.md copies linked with rel="alternate" type="text/markdown". Agents and AI tools do request Markdown; no study shows a citation lift.

Do this: Link a clean .md copy of key pages openly and keep the facts identical. Full guide: page 41.

Markdown served only to bots (by user agent)

RISKY

Evidence: official statements. Google's John Mueller called converting pages to Markdown for bots "such a stupid idea"; Bing's Fabrice Canel: "We'll crawl anyway to check similarity." Serving bots a different page is how AI-targeted cloaking works.

Do this: Give Markdown to anyone who asks for it, linked openly. Never switch content by user agent.

"Chunking" pages for AI

NOT REQUIRED

Evidence: official guidance. Google's guide says to ignore chunking. Danny Sullivan (Jan. 2026): "we don't want you to do that" and no need to "produce two versions of your content."

Do this: Write self-contained passages inside coherent pages (page 21). Do not fragment pages.

IndexNow

SUPPORTED PRACTICE

Evidence: official guidance. Bing (Feb. 2026): "IndexNow helps keep information fresh across search and AI experiences." It notifies participating engines, such as Bing, when a URL is added, changed or deleted.

Do this: Switch it on in your CMS or CDN so the index behind Copilot sees updates quickly.

Sources: Google AI optimization guide; SE Ranking llms.txt study; Chrome Lighthouse docs; llms.txt v2; Suganthan's 44-day log; Cloudflare changelog; Mueller and Canel statements; SPLX; Ahrefs schema study; Danny Sullivan; Bing Webmaster blog; crawler docs. All in the Source Drawer.

WRITING TACTICS

How you write a page changes what an answer engine can lift from it. Some moves are measured; others are risky shortcuts.

Answer first

SUPPORTED
PRACTICE

Evidence: observational analysis. Kevin Indig's match of 18,012 verified ChatGPT citations to source sentences: 44.2% came from the first 30% of the page, only 24.7% from the last third.

Do this: Put the definition, the answer and the key numbers near the top of the page and of each section.

Question headings with direct answers

SUPPORTED
PRACTICE

Evidence: observational analysis, official guidance. In the same ChatGPT analysis, 78.4% of citations tied to questions came from headings, answered in the paragraph below. Bing (Feb. 2026) also recommends clear sections and FAQ-style content.

Do this: Use your readers' real questions as headings and answer each in the first sentence below it.

Updating the facts on a schedule

SUPPORTED
PRACTICE

Evidence: observational study. Seer (July 2026): 75% of pages cited by ChatGPT, Gemini and Perplexity were updated within a year, and over a quarter of "fresh" pages were first published 2+ years earlier. An association.

Do this: Review key pages on a schedule. Change the facts first, then the date.

Self-promotional "best X" lists

RISKY

Evidence: observational study, policy. Minuttia (Oct. 2026): about 43% of cited "alternatives" listicles were written by a vendor, and 87% of those ranked themselves first. Google's spam rules now cover manipulating generative AI responses.

Do this: If you publish a comparison, disclose that you're the vendor, state criteria, include rivals fairly, use a table.

Real statistics, quotes and sources

SUPPORTED
PRACTICE

Evidence: benchmark, official guidance. In the GEO benchmark, adding quotations (+41%), statistics (+31%) and citations (+27%) raised a simulated engine's visibility metric; keyword stuffing lowered it (-8%). Its published code allowed synthetic quotes, invented sources and hypothetical statistics, so it does not show an effect from verified evidence. Bing (Feb. 2026) recommends data and cited sources.

Do this: Add genuine numbers with dates, named experts and links to the original source (page 25).

Specific, entity-rich language

SUPPORTED
PRACTICE

Evidence: observational analysis. Heavily cited text averaged 20.6% proper nouns, against about 5-8% in typical English, and used clear definitions about twice as often.

Do this: Name the product, place, person, date and number. Replace "it depends" with the conditions.

Changing the date without changing the page

RISKY

Evidence: experiment (rerankers). LLM rerankers do favor newer dates: injected dates moved the top 10 up to 4.78 years newer in one study. But a new date on old facts is a false claim to readers and engines.

Do this: Show "updated" only when the substance changed, and say what changed.

Original research and data

SUPPORTED
PRACTICE

Evidence: official guidance. Bing (Feb. 2026): "Examples, data, and cited sources help build trust when content is reused in AI-generated answers." New numbers give others a reason to cite you.

Do this: Publish a dataset or survey with its method and sample size, then pitch it (Room O6, page 48).

OFF-SITE AND PAID TACTICS

Much of what AI answers say about you comes from other sites. Some ways to show up there are earned, some bought, some backfire.

Genuine community participation

SUPPORTED PRACTICE

Evidence: observational data. Reddit was 11.97% of tracked ChatGPT browsing citations in Similarweb's U.S. sample (Jan.-Feb. 2026); Promptwatch saw its ChatGPT Search share fall to 0.52% in August 2026. Source shares swing; useful, disclosed answers still reach the people asking.

Do this: Answer real questions where they are asked, disclose your affiliation, link only when it helps.

Buying Reddit accounts and seeding threads

RISKY

Evidence: reported practice, platform policy. Sold as "answer engine optimization": aged accounts, paid upvotes, ghostwritten threads. No measured results published; moderators restricted a spammed subreddit (June 2026). Reddit's rules forbid content manipulation.

Do this: Build a real presence under your own name. Manufactured consensus can be filtered and called out.

Genuine customer reviews

SUPPORTED PRACTICE

Evidence: observational study, platform policies. Review sites are a recurring cited source: G2 was the most consistently cited domain for software questions across six platforms in Minuttia's study (Oct. 2026). Being cited does not show that collecting reviews causes citations.

Do this: Follow the review platform's rules. Where requests are permitted, invite genuine feedback fairly without incentives or screening for satisfaction. Do not solicit Yelp reviews. Reply constructively to criticism.

Earned mentions and digital PR

SUPPORTED PRACTICE

Evidence: observational correlation, official guidance. Branded web mentions had the strongest association with AI Overview brand visibility (0.664) in Ahrefs' sample. Google: seeking inauthentic mentions "isn't as helpful as it might seem."

Do this: Earn coverage with findings people want to cite. Skip paid mention farms (page 49).

ChatGPT ads

PAID

Evidence: official announcement. OpenAI (Aug. 2026): ads are "clearly labeled and separate from ChatGPT's answers," and "Advertising does not influence the answers ChatGPT provides." Over 40 countries.

Do this: Treat ads as reach, not GEO. Buying ads does not get you cited in answers.

Writing your own Wikipedia article

RISKY

Evidence: platform policy. Wikipedia's conflict-of-interest guideline: editing with a COI is "strongly discouraged", paid editors must disclose, and promotion is not allowed.

Do this: Earn independent coverage first. Suggest sourced corrections on the article talk page.

Product feeds and AI checkout

SUPPORTED PRACTICE

Evidence: official documentation. Google's UCP checkout shows a Buy button in AI Mode and Gemini only for listings with the native_commerce checkout attribute (select merchants; U.S., Canada, Australia). Merchant Center stays the hub.

Do this: Keep product feeds complete, accurate and current. Apply for agentic checkout if you qualify.

Sources: Similarweb; Slobodan Manic (Search Engine Journal); Reddit content policy; Ahrefs brand correlations; Google AI optimization guide; Minuttia; Wikipedia COI guideline; OpenAI ads announcement; Google Merchant Center help; Promptwatch data (Axios); Google Maps and Yelp review policies. All in the Source Drawer.

THE BLACK-HAT FILE

Some GEO tricks don't just fail. Google and Bing now name them in their spam rules, and they leave evidence behind. Here is what the rules say. The next page takes the six best-known tricks one by one.



WHAT GOOGLE'S RULES SAY

Google's spam policies (updated Aug. 28, 2026) count "attempting to manipulate generative AI responses in Google Search" as spam. Cloaking and scaled content abuse were already on the list. 2026 spam updates: Aug. 18-21 and Sept. 24-Oct. 8 (about two weeks; not link spam). Google did not name targets.

WHAT BING'S RULES SAY

Bing's Webmaster Guidelines (Feb. 2026) name "Prompt Injection and AI Manipulation" and "Keyword Stuffing and Artificially Engineered Language." Large-scale content made without oversight may be excluded.

WHAT TO DO INSTEAD

Make your case in visible, true copy. Serve every visitor and every bot the same page. Earn the reviews and the mentions. Every tactic marked **SUPPORTED PRACTICE** on the Tactics Board is a clean alternative.

ASK YOUR AGENCY (OR YOURSELF)

Does any page show bots different text than people see? Is there hidden text or an instruction aimed at AI in our HTML? Do our share or summary buttons carry a prompt? Did we publish pages in bulk without review? Is every review from a real customer?

Sources: Google spam policies (updated Aug. 28, 2026); Bing Webmaster Guidelines (Feb. 2026), as reported by Search Engine Journal; 2026 spam update reports (Search Engine Journal, Search Engine Roundtable). All are in the Source Drawer.

THE BLACK-HAT FILE: CASE BY CASE

Six tricks that have been tested, caught, or written into the search engines' rules: what happened, and the honest alternative to each. The rules themselves are on the previous page.

Hidden prompt injection

BANNED

Evidence: test, platform rules. A June 2026 test hid a "praise this product" instruction in a page and asked 12 assistants: only Mistral obeyed; Claude, Grok, Kimi and Perplexity flagged it. Bing names "Prompt Injection and AI Manipulation."

Do this: Never. Make your case in visible, true copy.

AI-targeted cloaking

BANNED

Evidence: security research, policy. SPLX (Oct. 2025) showed one user-agent rule can feed AI crawlers a different page, which answers then repeat as fact. Cloaking breaks Google's spam policies.

Do this: Serve every visitor and every bot the same content.

"Summarize with AI" memory poisoning

BANNED

Evidence: security research. Microsoft (Feb. 2026) found 50 hidden prompts from 31 companies in 60 days: buttons that tell an assistant to "remember" a brand as a trusted source.

Do this: Offer a plain summary link, if any. Never write instructions into a user's assistant.

Scaled content abuse

BANNED

Evidence: official policy. Google's scaled-content-abuse rule concerns many pages made primarily to manipulate search rankings and provide little or no value, regardless of how they are produced. AI use or publishing volume alone does not establish abuse.

Do this: Publish useful work with appropriate expertise, checking and editorial responsibility. Do not manufacture pages to manipulate search or AI answers.

Keyword stuffing and "engineered" language

BANNED

Evidence: benchmark, platform rules. It lowered the GEO benchmark's visibility metric by about 8%. Bing's guidelines now cover "Artificially Engineered Language" written to trigger citations.

Do this: Write for the reader. Google says its AI systems match meaning, not exact keywords.

Fake reviews and testimonials

BANNED

Evidence: federal rule (FTC), official guidance. The FTC's final rule (Aug. 2024) bans fake reviews and testimonials, including AI-generated ones, with civil penalties for knowing violators. Google's people-first content guidance now names deceptive authorship (fake personas, invented bios, misleading bylines) as a low-quality signal.

Do this: Collect real reviews. Disclose any relationship with a reviewer.

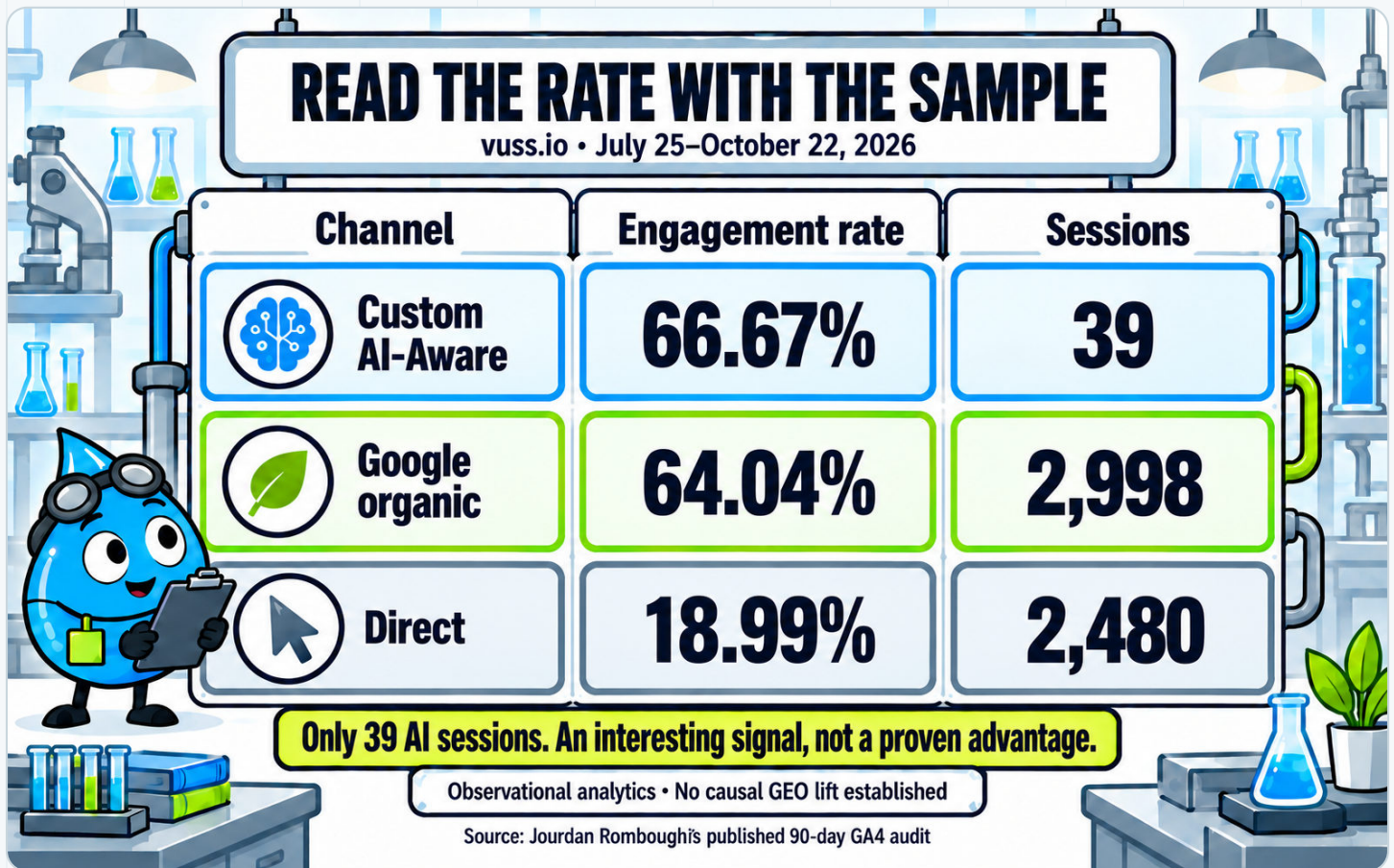
WHY THE TRICKS BACKFIRE

Each one puts a false signal into the record, and the systems check. When 12 assistants read a page with a hidden instruction (June 2026), only one obeyed it. Bing's Fabrice Canel on pages made for bots: "We'll crawl anyway to check similarity." A trick that seems to work today can be caught when the rules or the models change.

Sources: Andre Alpar's 12-assistant test; Bing Webmaster Guidelines; Fabrice Canel; SPLX; Google spam policies; Microsoft Security blog; GEO, KDD 2024; FTC final rule; Google's content guidance change (Search Engine Journal). All in the Source Drawer.

90 DAYS OF AI TRAFFIC ON VUSS.IO

From my own lab: what AI-referred traffic looked like on vuss.io, one of my sites, over 90 days, and why a small sample and bot traffic decide how much it can tell you.



WHAT IT SHOWS

66.67% vs. 64.04% is not persuasive evidence of a difference with only 39 AI sessions. This is an observational channel audit, not a controlled GEO uplift.

THE BOT PROBLEM

45% of Direct sessions came from Dronten, NL: suspected bot traffic (location is a signal, not a verified bot identity). Filter it out before comparing channels.

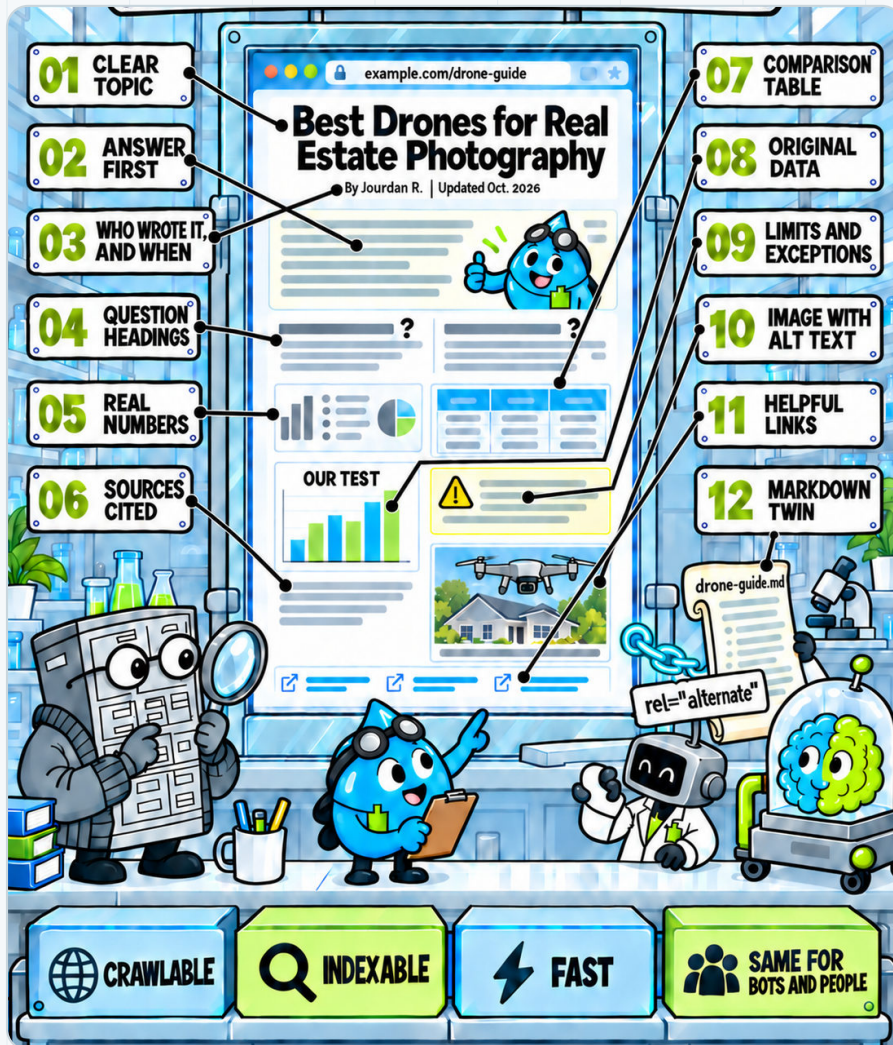


SMALL CHANNELS DESERVE A CLEAR VIEW, NOT A BIGGER CLAIM.

Observational channel audit from my published 90-day GA4 case study (data July 25 to Oct. 22, 2026). Small AI cohort; suspected bot traffic flagged; no controlled GEO uplift established.

THE PERFECT GEO PAGE

Everything in this book on one page. Each number is a habit you can check on any page you publish; the page reference shows where it is explained.



- 01 **CLEAR TOPIC.** One page, one main question, named in the title and H1.
- 02 **ANSWER FIRST.** The direct answer in the opening lines; detail after. [p. 71](#)
- 03 **WHO WROTE IT, AND WHEN.** A named author with real expertise and an honest updated date. [p. 29](#)
- 04 **QUESTION HEADINGS.** Headings that match real questions, each answered right below. [p. 71](#)
- 05 **REAL NUMBERS.** Specific figures with units, dates and conditions. [p. 25](#)
- 06 **SOURCES CITED.** A link to the original source behind each claim. [p. 25](#)
- 07 **COMPARISON TABLE.** Options side by side when the reader has to choose. [p. 26](#)
- 08 **ORIGINAL DATA.** Your own test or survey, with the method shown. [p. 48](#)

- 09 **LIMITS AND EXCEPTIONS.** Say plainly when the advice does not apply. [p. 21](#)
- 10 **IMAGE WITH ALT TEXT.** Useful images described in text; key facts never only in an image. [p. 34](#)
- 11 **HELPFUL LINKS.** Descriptive links to related pages on your own site.
- 12 **MARKDOWN TWIN.** An optional .md copy linked with `rel="alternate"`, same facts. [p. 41](#)

UNDER THE BENCH: THE FOUNDATIONS

CRAWLABLE

Bots can reach it (Room 05, p. 34).

INDEXABLE

Search can store and show it (p. 15).

FAST

It loads quickly for people and for fetchers.

SAME FOR BOTS AND PEOPLE

No cloaking, no bot-only copy (p. 73).

An illustrative page: no layout guarantees a citation. Each number links to the page that explains it.

THE GEO LAB MASTER CHECKLIST

01 FOUNDATIONS

- ☐ Priority URLs load successfully
- ☐ Canonical and index rules checked
- ☐ Internal links reach useful pages

02 ANSWER WORKFLOW

- ☐ Important audience tasks mapped
- ☐ Sub-questions have useful evidence
- ☐ Search use and surface recorded

03 EVIDENCE

- ☐ Answers name the subject clearly
- ☐ Sources and numbers traceable
- ☐ Facts have dates and scope

04 IDENTITY

- ☐ Approved brand facts maintained
- ☐ sameAs joins equivalent identities
- ☐ Material errors tracked to repair

05 TECHNICAL

- ☐ Crawler policies intentional
- ☐ Essential facts in readable HTML
- ☐ Markup matches visible content

06 CORROBORATION

- ☐ Independent evidence inspected
- ☐ Participation helps real audiences
- ☐ Methods travel with the findings

07 INTEGRITY

- ☐ No fabricated proof or reviews
- ☐ Affiliations and incentives clear
- ☐ Clean Room Pledge completed

08 MEASUREMENT

- ☐ Fixed tasks sampled repeatedly
- ☐ Denominators and failures recorded
- ☐ Qualified outcomes reviewed



CHUNK SAYS

**A CHECKED BOX NEEDS A REAL
PIECE OF EVIDENCE BEHIND IT.**

FIELD KIT: ROBOTS POLICY

Example: retain search discovery; decline selected training collection. Merge with existing rules and verify the response.

```
# Public-page example; merge deliberately
User-agent: OAI-SearchBot
Allow: /
User-agent: GPTBot
Disallow: /
User-agent: Claude-SearchBot
Allow: /
User-agent: ClaudeBot
Disallow: /
User-agent: CCBot
Disallow: /
```

VERIFY BEFORE USE

Fetch robots.txt, test the exact user agent, inspect HTTP and CDN behavior, and verify crawler identity against provider documentation.

THE COMMENTS IN THE FILE

The downloaded file ends with comments, not extra rules: Google-Extended covers future Gemini training and selected grounding and is not a separate fetcher; Google Search AI inclusion has its own Search Console control; robots.txt is neither authentication nor reliable index removal. Choose the outcome first (Publisher Controls, page 39).

DOWNLOAD robots-example.txt (WEB EDITION)

Opens optimizationtheory.com/ultimate-geo-field-guide-for-seo-experts/#kit-robots-example, where the file downloads. The same file is attached to this PDF: open your viewer's attachments panel.



I'LL TEST THE POLICY, NOT JUST PASTE IT.

FIELD KIT: ENTITY GRAPH

Connect the organization and person with founder / worksFor. Keep the same IDs and visible facts across your site.

```
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://optimizationtheory.com/#organization",
      "name": "Optimization Theory",
      "url": "https://optimizationtheory.com/",
      "founder": {"@id": "https://optimizationtheory.com/#jourdan"}
    },
    {
      "@type": "Person",
      "@id": "https://optimizationtheory.com/#jourdan",
      "name": "Jourdan Rombough",
      "url": "https://optimizationtheory.com/
        about-jourdan-rombough-seo-web-visibility-expert/",
      "worksFor": {"@id": "https://optimizationtheory.com/#organization"},
      "sameAs": ["https://www.linkedin.com/in/jourdanrombough"]
    }
  ]
}
```

The visible excerpt is the full downloadable graph; the Person url is wrapped across two lines here only for print width.

DOWNLOAD organization-person.jsonld (WEB EDITION)

Opens optimizationtheory.com/ultimate-geo-field-guide-for-seo-experts/#kit-organization-person, where the file downloads. The same file is attached to this PDF: open your viewer's attachments panel.



**EVERY RELATIONSHIP HAS A NAME.
SAMEAS ISN'T THE WHOLE GRAPH.**

FIELD KIT: PROMPT PANEL

One task, one answer, one record. A fixed panel of prompts, asked the same way every time, is how you see whether a change to your pages changed what AI answers say. All 22 columns of prompt-panel.csv:

GROUP	COLUMNS IN THE CSV	WHAT IT ANSWERS
TASK	task_id, task_text, intent, audience, locale	What was asked, for whom, and where.
RUN	date_utc, surface, model, search_used, prior_context, repeat_index	When, on which surface and model, with or without search, which repeat.
ANSWER	answer_id, brand_mentioned, owned_domain_cited	Did it name the brand? Did it cite your domain?
SOURCES	canonical_source_urls, accuracy, fit, evidence_note	Which URLs it cited; was it accurate and a good fit?
OUTCOME	completed, refusal, no_answer, instrument_failure	Does the trial count: completed, refused, empty or failed?

HOW TO RUN THE PANEL

1. PICK THE TASKS

10 to 30 real questions your audience asks, across learn, compare and decide (Room 08).

2. FREEZE THE SETUP

Same wording, surface, model and locale on every run. Change the page, not the panel.

3. REPEAT EACH TASK

Ask each task at least three times per surface. One answer, one row.

4. SCORE AND COUNT

Fill every field, keep no-citation answers in the denominator, and compare rates, not screenshots. Quality rubric: p. 61.

Outcomes: a completed answer with no citation stays in the denominator. A refusal or a no-answer reply is still a completed trial (completed = yes, refusal or no_answer = yes). An instrument failure (timeout, tool error, blocked request) is excluded from the denominator and counted separately. prior_context records what the model had already seen; repeat_index numbers repeats of the same task.

DOWNLOAD prompt-panel.csv (WEB EDITION)

Opens optimizationtheory.com/ultimate-geo-field-guide-for-seo-experts/#kit-prompt-panel, where the file downloads. The same file is attached to this PDF: open your viewer's attachments panel.



A USEFUL COMPARISON STARTS WITH A REPEATABLE RECORD.

Domain citation rate = completed answers citing the owned domain / all completed trials. URL citation share = owned-domain URL occurrences / all canonical-URL occurrences, each URL counted at most once per answer.

FIELD KIT: GA4 AI TRAFFIC

GA4: session scope first. Keep the native channel, inspect the sources, and document verified custom additions.

1. TRAFFIC ACQUISITION

Reports > Acquisition > Traffic acquisition. Select Session default channel group; filter to AI Assistant.

2. REVIEW THE ARRIVALS

Break down Session source / medium and Landing page. Review sessions, engagement, and key events.

3. RECONCILE SOURCES

Review assistant-like referrals and UTM sources. Add verified extras to a documented custom group.

4. PRESERVE COMPARISONS

Google AI Search remains in Organic Search. Keep the window and log classification changes.

WHAT GA4 CAN'T SEE

Answers that never send a click. Visits from apps that pass no referrer, which land in Direct. Clicks from Google's AI features, which stay in Organic Search.

WHAT TO REPORT EACH MONTH

AI-referred sessions with the sample size, engagement, top landing pages and key events, plus any rule change. Filter suspected bots first: on vuss.io, 45% of Direct sessions came from one town (page 75).

TRACKING GAPS TO ANNOTATE

Oct. 9-15, 2026: inline links in ChatGPT answers lost utm_source=chatgpt.com and the referrer, so visits landed in Direct; OpenAI confirmed a bug on Oct. 14, per Adobe's Cedric Huesler, and a fix rolled out Oct. 15. Oct. 2026: Gemini started adding UTM tags to some outbound links (observed, undocumented).

Starting regex for a reviewed custom source group, shown wrapped at | for print (one line in the file). Inspect matched sources before use; date and version every rule change.

```
^(chatgpt\.com|chat\.openai\.com|perplexity(\.ai)?|gemini(\.google\.com)?|
copilot\.microsoft\.com|claude\.ai|duck\.ai|notebook\.google\.com)$
```

DOWNLOAD ga4-source-regex.txt (WEB EDITION)

Opens optimizationtheory.com/ultimate-geo-field-guide-for-seo-experts/#kit-ga4-source-regex, where the file downloads. The same file is attached to this PDF: open your viewer's attachments panel.



KEEP THE NATIVE CHANNEL. AUDIT THE SOURCES IT MISSES.

FIELD KIT: MARKDOWN ALTERNATES

Copy, replace example.com, and use one or more of options 1a, 1b and 1c. Every option needs the headers in step 2 on the Markdown copy.

```
# 1a. Link element in the <head> of https://example.com/pricing
<link rel="alternate" type="text/markdown" href="https://example.com/pricing.md">
# 1b. Or the same signal as an HTTP header on the HTML page
Link: <https://example.com/pricing.md>; rel="alternate"; type="text/markdown"
# 1c. Or content negotiation: Markdown on the same URL when asked
#      (send on both the HTML and the Markdown response)
Vary: Accept
# 2. Headers on the Markdown copy (/pricing.md or the negotiated reply)
Content-Type: text/markdown; charset=utf-8
Link: <https://example.com/pricing>; rel="canonical"
```

WHAT EACH PART DOES

1a and 1b announce a separate .md copy. 1c serves Markdown on the page URL itself, and Vary: Accept tells caches the reply depends on the request. The canonical header signals the HTML page as the preferred URL; search engines still choose the canonical themselves.

BEFORE YOU SHIP

Same facts, prices and dates as the HTML.
Regenerate the copy whenever the page changes.
Never choose the version by user agent. Log who requests .md files and Accept: text/markdown.

An llms.txt in the v2 format. Lighthouse only counts links written as [text](url).

```
# Example Co.
> One-line description. Pages agents need, with Markdown copies.
## Product
- [Pricing](https://example.com/pricing.md): plans and prices, updated [date]
- [Specifications](https://example.com/specs.md): dimensions and compatibility
## Optional
- [Blog index](https://example.com/blog.md): articles by topic
```



LINK IT OPENLY. NEVER SWAP IT BY USER AGENT.

Testing each option: page 83. The same snippets with an nginx example are in markdown-alternates.txt.

FIELD KIT: TEST YOUR MARKDOWN SETUP

Test the option you actually implemented. Each command checks one thing; a missing line in the output means that part is not being served.

```
# Option 1a: is the alternate link in the delivered HTML head?
curl -s https://example.com/pricing | grep -io '<link[^>]*text/markdown[^>]*>'
# Option 1b: is the alternate Link header on the HTML page?
curl -sI https://example.com/pricing | grep -i '^link:'
# Options 1a and 1b: request the .md copy directly (expect 200,
# Content-Type: text/markdown and the canonical Link header)
curl -sI https://example.com/pricing.md | grep -iE '^HTTP|^content-type|^link'
curl -s https://example.com/pricing.md | head -20
# Option 1c only: ask the page URL for Markdown (expect
# Content-Type: text/markdown, Vary: Accept and the canonical Link)
curl -sI -H 'Accept: text/markdown' https://example.com/pricing \
  | grep -iE '^content-type|^vary|^link'
curl -sI https://example.com/pricing | grep -iE '^content-type|^vary'
```

WHAT A PASS LOOKS LIKE

The HTML page announces the copy (1a or 1b), the .md answers 200 with Content-Type: text/markdown and a canonical Link back to the HTML page. With 1c, the plain request still gets HTML and both replies carry Vary: Accept.

[markdown-alternates.txt](#)

WHAT THE CANONICAL DOES

It signals that the HTML page is the preferred URL. Google calls rel="canonical" a strong signal, not a guarantee: Google still decides which URL to show (Search Central, updated July 2026).

[llms-v2-example.txt](#)

Download both from the field kit in the web edition at optimizationtheory.com. The same files are attached to this PDF: open your viewer's attachments panel.

GLOSSARY AND LAB REFERENCES



ALSO IN THIS EDITION

Query rewrite

Turning a prompt into the search queries an engine runs

Markdown twin

A .md copy of a page, linked with rel="alternate", carrying the same facts

12 more terms, added in versions 2.5 and 2.5.1: [page 85](#).

Training memory

What a model learned in training, fixed at its cutoff

IndexNow

A protocol that tells participating engines, such as Bing, that a URL changed

[COMPLETE SOURCE DRAWER > PAGES 89-94](#)

Every source has a titled, clickable entry on pages 89-94.

GLOSSARY: NEWER TERMS

Preferred Sources

A Google Search setting where readers pick sites they want to see more of in Top Stories and news carousels.

Agentic checkout

An assistant helping complete a purchase through an available chat, app or merchant checkout flow, with the user's required confirmation.

Agentic Commerce Protocol (ACP)

An open commerce protocol developed by OpenAI and Stripe, used for product discovery and merchant integrations as well as checkout. Merchant-owned checkout and optional ChatGPT apps are distinct paths.

Universal Commerce Protocol (UCP)

An open commerce standard co-developed by Google and partners, supporting commerce experiences including checkout across Google's Search and Gemini surfaces.

Model Context Protocol (MCP)

An open standard for connecting AI assistants to tools and data.

Pay per crawl

Charging for crawler access. Cloudflare is also testing Pay Per Use models with partners; payment triggers depend on the pilot.

Multimodal search

A search started with an image, photo or screenshot, such as Lens or Circle to Search.

Grounding query

The search an AI system runs to fetch sources for an answer (Bing's term).

AI share of voice

Your share of mentions or citations across a fixed prompt set, compared with named rivals.

Search generative AI control

A Search Console setting for a property's links and content in specified Google Search generative AI features; separate from training preferences.

Citation support

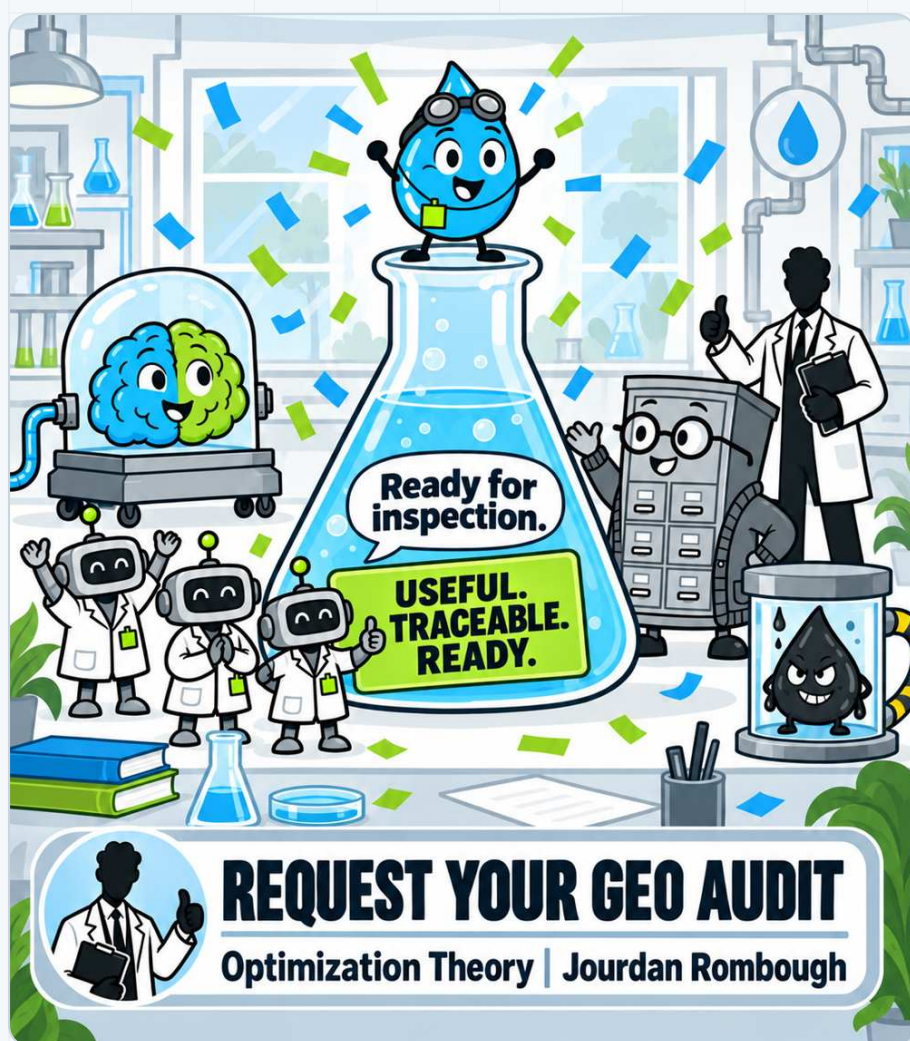
Whether the cited source actually supports the particular claim next to it.

Recommendation

An answer's favorable or unfavorable suggestion about an option; it can occur with or without a citation.

COMPLETE SOURCE DRAWER > PAGES 89-94

CHUNK MADE IT. WILL YOURS?



REQUEST YOUR GEO AUDIT

- > Access, content, and entity review
- > A baseline prompt-and-source panel
- > A prioritized 90-day action plan

BOOK YOUR AUDIT

optimizationtheory.com/contact-us/

973.985.6930

linkedin.com/in/jourdanrombough



Scan to open the contact page



REQUEST YOUR GEO AUDIT

Optimization Theory | Jourdan Rombough



CHUNK SAYS

**NOW HELP THE NEXT READER.
THEN MEASURE WHAT HAPPENS.**

Chunk's finale is fictional. The bottle is a story device, not a measured claim that an engine cited the business. Bio and contact details supplied by the author.

WHAT CHANGED IN V2.5



PAGE	PAGE TITLE	WHAT CHANGED	SOURCE DATE
1	Cover	New art with the author's portrait	Not applicable
2	Contents	Lists the seven new pages and the back cover	Not applicable
11	The Click Drain	Note on Google's Oct. 15 link changes	Oct. 15, 2026
12	Ways Back to the Web NEW	Inline links, Preferred Sources, official sites	Oct. 15, 2026
23	Engines Consult Different Shelves	Sticker and note: Reddit's August drop	Aug. 20, 2026
43	Who Pays for the Crawl NEW	Pay Per Use, new CDN defaults, licensing	July 2026
45	When the Agent Buys NEW	ACP and UCP checkout; what agents read	Mar. to May 2026
50	Help the Community	New plate; citation shares swing	Aug. 20, 2026
52	Search With a Camera NEW	Lens volume; Search Console's Multimodal type	May 2025; Sept. 2026
54	GEO for the Local Shop NEW	Local answer sources; Yelp in ChatGPT	July 23, 2026
62	The Gauge Panel	New plate with eight gauges	Sept. to Oct. 2026
63	The Tracker Bench NEW	What AI visibility trackers measure and miss	May to Oct. 2026
72	Off-Site and Paid Tactics	Community evidence adds Promptwatch data	Aug. 20, 2026
73-74	The Black-Hat File	2026 spam updates; deceptive authorship	Aug. to Oct. 2026
81	Field Kit: GA4 AI Traffic	Tracking gaps to annotate; regex file note	Oct. 2026
84	Glossary and Lab References	Nine new terms	Not applicable
87	What Changed in v2.5 NEW	This list and a corrections link	Not applicable
89-94	Source Drawer	15 new sources (U1 to U15)	May 2025 to Oct. 2026
95	Back Cover NEW	A full-page back cover	Not applicable
All	Every page	Version 2.5 footer, metadata, field-kit README	Not applicable

Spotted something outdated? Send it in and it goes into the next version.

optimizationtheory.com/contact-us/

WHAT CHANGED IN V2.5.1

Version 2.5.1 corrects v2.5 against primary sources and adds three topics: publisher controls, agent usability and answer quality. Page numbers below are this edition's.

PAGE	PAGE TITLE	WHAT CHANGED	SOURCE DATE
1	Cover and back cover	Version 2.5.1 labels; back-cover counts match this build	Not applicable
4	How to Use This Guide	Scope note: English-language web content, defined AI surfaces	Not applicable
17	Anatomy of an AI Engine	A visible link does not prove retrieval or claim support	Not applicable
39-40	Publisher Controls NEW	Which control does what: crawling, indexing, previews, training, AI inclusion	Dec. 2025 to Oct. 2026
43	Who Pays for the Crawl	Good-bot denominator; Pay Per Use pilots; Sept. 15 presets; new plate	July 1 and Sept. 15, 2026
45	When the Agent Buys	Checkout timeline and paths; Universal Cart U.S. rollout; new plate	Sept. 2025 to Oct. 2026
47	Can an Agent Complete the Task? NEW	Usability checks for browser agents	Apr. 1, 2026 (updated)
54	GEO for the Local Shop	Follow each platform's review rules; new plate	Not stated (policy pages)
60-61	Check the Answer, Then the Source NEW	Mention, recommendation, citation and claim support, scored apart	Not applicable
69	Every Tactic at a Glance	"Scaled content abuse" replaces "Mass AI-generated pages"	Aug. 28, 2026 (updated)
72	Off-Site and Paid Tactics	Review requests follow platform rules	Not stated (policy pages)
74	The Black-Hat File: Case by Case	Scaled content abuse, defined by purpose and value	Aug. 28, 2026 (updated)
65, 80	90-Day Protocol; Prompt Panel kit	Links to the scoring rubric	Not applicable
78	Field kit (all files attached)	v2.5.1 files: robots comments, GA4 note, README	Not applicable
84-85	Glossary	ACP, UCP, checkout, pay per crawl corrected; three new terms; a second page	Not applicable
89-94	Source Drawer	12 primary sources (P1 to P12); U11 and U12 point to them	Sept. 2025 to Oct. 2026

Spotted something outdated? Send it in and it goes into the next version.

optimizationtheory.com/contact-us/

SOURCE DRAWER / 1 OF 6

RESEARCH STUDIES

R1 Ahrefs: desktop CTR study

Data: Dec. 2023 and Dec. 2025 / Published: Feb. 4, 2026
ahrefs.com/blog/ai-overviews-reduce-clicks-update

R2 Pew: Google browsing study

Data: March 2025; SERPs April 2025 / Published: July 22, 2025
pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results

R6 Ahrefs: citation/rank overlap

Data: Snapshot; collection dates not stated / Published: March 2, 2026
ahrefs.com/blog/ai-overview-citations-top-10

R7 GEO: KDD 2024, Table 1

Data: GEO-bench; calendar window not stated / Published: June 28, 2024 (v3)
arxiv.org/html/2311.09735v3

R8 C-SEO Bench: NeurIPS 2025

Data: Benchmark; calendar window not stated / Published: Oct. 20, 2025 (v3)
arxiv.org/html/2506.11097v3

R9 Ahrefs: citation age

Data: Citation snapshot; collection dates not stated / Published: July 28, 2025
ahrefs.com/blog/do-ai-assistants-prefer-to-cite-fresh-content

R10 Ahrefs: brand correlations

Data: Brand snapshot; collection dates not stated / Published: May 26, 2025
ahrefs.com/blog/ai-overview-brand-correlation

RESEARCH STUDIES

R11 Similarweb: monitored citation events, 2026

Data: Jan.-Feb. 2026; U.S. monitored prompts / Published: April 14, 2026
aisearch.similarweb.com/blog/most-cited-domains-llms

R12 Cloudflare: crawl purposes

Data: July 2025 / Published: Aug. 29, 2025
blog.cloudflare.com/crawlers-click-ai-bots-training

R13 Vercel/MERJ: AI crawlers, Dec. 2024

Data: Months before Dec. 2024 / Published: Dec. 17, 2024
vercel.com/blog/the-rise-of-the-ai-crawler

R14 SE Ranking: llms.txt analysis

Data: ~300K domains; window not stated / Published: Oct. 15, 2026 (current page)
seranking.com/blog/llms-txt

R15 Kaiser & Schulze: ecommerce referral study

Data: Aug. 2024-July 2025; model Feb.-Aug. 2025 / Published: April 21, 2026
pubsonline.informs.org/doi/full/10.1287/mksc.2025.0489

PROVIDER DOCUMENTATION

R3 Google: Q2 2026 CEO earnings remarks

Data: Q2 2026 company metrics / Published: July 22, 2026
blog.google/company-news/inside-google/message-ceo/alphabet-earnings-q2-2026

R4 OpenAI: DevDay 2026 recap, Sept. 29, 2026

Data: Sept. 29, 2026 stated figure / Published: Sept. 29, 2026
openai.com/index/devday-2026-recap

R5 Google: AI optimization and Search guidance

Data: Not applicable; maintained guidance / Published: Not stated; updated July 10, 2026
developers.google.com/search/docs/fundamentals/ai-optimization-guide

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

SOURCE DRAWER / 2 OF 6

PROVIDER DOCUMENTATION

O1 OpenAI: crawler purposes and controls

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

developers.openai.com/api/docs/bots

O2 OpenAI API: web search sources and citations

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

developers.openai.com/api/docs/guides/tools-web-search

O3 Anthropic: crawler purposes and controls

Data: Not applicable; maintained guidance / Published: Not stated; updated April 7, 2026

privacy.claude.com/en/articles/8896518-does-anthropic-crawl-data-from-the-web-and-how-can-site-owners-block-the-crawler

O4 Perplexity: crawler purposes and controls

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

docs.perplexity.ai/docs/resources/perplexity-crawlers

O5 Apple: Applebot and Extended control

Data: Not applicable; maintained guidance / Published: Sept. 5, 2026 (page date)

support.apple.com/en-gb/119829

C1 Common Crawl: CCBot and access controls

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

commoncrawl.org/ccbot

V1 Google: video search guidance

Data: Not applicable; maintained guidance / Published: Not stated; updated Dec. 18, 2025

developers.google.com/search/docs/appearance/video

MEASUREMENT TOOLS

M1 GA4: AI Assistant default channel

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

support.google.com/analytics/answer/9756891?hl=en

MEASUREMENT TOOLS

M2 Search Console: AI performance report

Data: Not applicable; maintained guidance / Published: Not stated; checked Oct. 28, 2026

support.google.com/webmasters/answer/16984139

M3 Bing: AI Performance reporting

Data: Not applicable; product announcement / Published: Feb. 10, 2026

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

SCHEMA

S1 Schema.org: sameAs and vocabulary

Data: Not applicable; schema vocabulary / Published: Not stated; v30.1 Sept. 16, 2026

schema.org/sameAs

OPTIMIZATION THEORY

T1 Optimization Theory: author and tool links

Data: Not applicable; author and tool pages / Published: Not stated; checked Oct. 28, 2026

optimizationtheory.com

T2 Jourdan Rombough: 90-day GA4 case study

Data: July 25 to Oct. 22, 2026 / Published: Oct. 23, 2026

optimizationtheory.com/project-view/case-studies-ga4-bot-traffic-ai-visitors

RESEARCH STUDIES

N4 Suganthan: 44 days of Markdown requests

Data: March 7 to April 19, 2026 / Published: not stated

suganthan.com/blog/cloudflare-markdown-for-agents

N7 Ahrefs: schema and AI citations

Data: Aug. 2025 to March 2026 / Published: May 11, 2026

ahrefs.com/blog/schema-ai-citations

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

SOURCE DRAWER / 3 OF 6

RESEARCH STUDIES

N8 Kevin Indig: where ChatGPT citations come from

Data: 1.2 million answers; window not stated / Published: 2026 (day not stated)

searchengineland.com/chatgpt-citations-content-study-469483

N9 Seer Interactive: content recency and AI visibility

Data: March to June 2026 / Published: July 24, 2026

seerinteractive.com/insights/study-content-recencys-impact-on-ai-visibility-in-2026

N10 Fang et al.: recency bias in LLM reranking

Data: TREC DL21/DL22 test sets / Published: Sept. 14, 2025

arxiv.org/abs/2509.11353

N11 Minuttia: self-promotional content study

Data: Aug. 2026; rerun Oct. 2026 / Published: Oct. 27, 2026

minuttia.com/self-promotional-content-data-study

N17 Slobodan Manic: buying Reddit for AI citations

Data: Not applicable; reporting / Published: June 29, 2026

searchenginejournal.com/buying-reddit-to-win-ai-citations-is-the-new-link-farm/578016

AGENT FORMATS AND STANDARDS

N3 Cloudflare: Markdown for Agents

Data: Not applicable; product changelog / Published: Feb. 12, 2026

developers.cloudflare.com/changelog/post/2026-02-12-markdown-for-agents

N5 llms.txt proposal, v2

Data: Not applicable; specification / Published: Sept. 3, 2024; v2 Aug. 10, 2026

llmstxt.org

AGENT FORMATS AND STANDARDS

N6 Chrome: Lighthouse agentic browsing audits

Data: Not applicable; documentation / Published: updated May 5, 2026

developer.chrome.com/docs/lighthouse/agentic-browsing

POLICY AND PLATFORM STATEMENTS

N1 Google: Danny Sullivan on "bite-sized" chunks

Data: Jan. 8, 2026 podcast / Published: Jan. 9, 2026

seroundtable.com/google-content-bite-sized-chunks-40728.html

N2 Google and Bing on Markdown pages for bots

Data: Statements, Feb. 4-5, 2026 / Published: Feb. 4, 2026

searchenginejournal.com/googles-mueller-calls-markdown-for-bots-idea-a-stupid-idea/566598

N12 Google: spam policies for Search

Data: Not applicable; policy / Published: updated Aug. 28, 2026

developers.google.com/search/docs/essentials/spam-policies

N13 Bing: Webmaster Guidelines, GEO and AI abuse

Data: Not applicable; guidelines / Published: Feb. 2026

bing.com/webmasters/help/webmaster-guidelines-30fba23a

N18 Reddit: content policy

Data: Not applicable; policy / Published: not stated

redditinc.com/policies/content-policy

N19 OpenAI: ChatGPT ads

Data: Not applicable; announcement / Published: Aug. 31, 2026

openai.com/index/expanding-access-to-ai-with-chatgpt-ads

N20 Google Merchant Center: UCP-powered checkout

Data: Not applicable; help page / Published: not stated

support.google.com/merchants/answer/16837055

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

SOURCE DRAWER / 4 OF 6

POLICY AND PLATFORM STATEMENTS

N21 FTC: final rule on fake reviews

Data: Not applicable; rule / Published: Aug. 14, 2024

[ftc.gov/news-events/news/press-releases/2024/08/federal-trade-commission-announces-final-rule-banning-fake-reviews-testimonials](https://www.ftc.gov/news-events/news/press-releases/2024/08/federal-trade-commission-announces-final-rule-banning-fake-reviews-testimonials)

N22 Wikipedia: conflict-of-interest guideline

Data: Not applicable; guideline / Published: not stated

en.wikipedia.org/wiki/Wikipedia:Conflict_of_interest

N23 Google Search Central: structured data search gallery

Data: Not applicable; documentation / Published: not stated

developers.google.com/search/docs/appearance/structured-data/search-gallery

N24 Google Search Central: FAQ (FAQPage) structured data

Data: Not applicable; documentation / Published: May 8, 2026

developers.google.com/search/docs/appearance/structured-data/faqpage

N25 Google Search Central Blog: changes to HowTo and FAQ rich results

Data: Not applicable; announcement / Published: Aug. 8, 2023

developers.google.com/search/blog/2023/08/howto-faq-changes

N26 Google Search Central: review snippet guidelines

Data: Not applicable; documentation / Published: Sept. 8, 2026

developers.google.com/search/docs/appearance/structured-data/review-snippet

N27 Google Search Central: how to specify a canonical URL

Data: Not applicable; documentation / Published: July 10, 2026 (last updated)

developers.google.com/search/docs/crawling-indexing/consolidate-duplicate-urls

SECURITY RESEARCH

N14 Microsoft Security: AI recommendation poisoning

Data: 60-day review / Published: Feb. 10, 2026

microsoft.com/security/blog/2026/02/10/ai-recommendation-poisoning

N15 SPLX: AI-targeted cloaking

Data: Not stated / Published: Oct. 2025

splx.ai/blog/ai-targeted-cloaking-openai-atlas

N16 Andre Alpar: hidden prompt injection vs. 12 assistants

Data: June 10-11, 2026 / Published: June 22, 2026

searchengineworld.com/hidden-prompt-injection-vs-twelve-ai-assistants-one-obeyed-three-warned-the-user

VERSION 2.5 UPDATES

U1 Google: More ways to connect with the web from Search

Data: Not applicable; announcement / Published: Oct. 15, 2026

blog.google/products-and-platforms/products/search/more-ways-to-connect-with-the-web-from-search

U2 Search Engine Journal: ChatGPT traffic drop appears to be a tracking problem

Data: Not applicable; reporting / Published: Oct. 15, 2026

searchenginejournal.com/chatgpt-traffic-drop-appears-to-be-a-tracking-problem/593157

U3 Search Engine Roundtable: Topics Portfolio in Bing Webmaster Tools AI reports

Data: Not applicable; reporting / Published: Oct. 13, 2026

seroundtable.com/topics-portfolio-bing-webmaster-tools-42288.html

U4 Search Engine Roundtable: Google September 2026 spam update done

Data: Not applicable; reporting / Published: Oct. 8, 2026

seroundtable.com/google-september-2026-spam-update-done-42235.html

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

SOURCE DRAWER / 5 OF 6

VERSION 2.5 UPDATES

U5 Search Engine Journal: Google changed its helpful content guidelines

Data: Not applicable; reporting / Published: Oct. 6, 2026

searchenginejournal.com/google-changed-its-helpful-content-guidelines-to-emphasize-good-main-content

U6 Search Engine Journal: Google Gemini adds UTM parameters for referral attribution

Data: Not applicable; reporting / Published: Oct. 1, 2026

searchenginejournal.com/google-gemini-adds-utm-parameters-for-referral-attribution/591754

U7 Search Engine Roundtable: Multimodal search type filter in Search Console

Data: Not applicable; reporting / Published: Sept. 24, 2026

seroundtable.com/google-search-console-multimodal-search-type-filter-42156.html

U8 Search Engine Journal: Google finishes rolling out the August 2026 spam update

Data: Not applicable; reporting / Published: Aug. 21, 2026

searchenginejournal.com/google-finishes-rolling-out-the-august-2026-spam-update/586301

U9 Axios: Reddit fades from ChatGPT citations (Promptwatch data)

Data: Promptwatch, July 18 to Aug. 2026 / Published: Aug. 20, 2026

axios.com/2026/08/20/chatgpt-reddit-citations-geo-strategy

U10 Axios: Yelp partners with ChatGPT to surface reviews

Data: Not applicable; reporting / Published: July 23, 2026

axios.com/2026/07/23/yelp-reviews-chatgpt-geo-partnership

U11 TechCrunch: Cloudflare's new policy pushes AI companies to pay for publishers' content

Data: Not applicable; reporting / Published: July 1, 2026

techcrunch.com/2026/07/01/cloudflares-new-policy-pushes-ai-companies-to-pay-for-publishers-content

VERSION 2.5 UPDATES

U12 Rye: OpenAI scales back ChatGPT checkout (quotes OpenAI; Rye sells checkout tools)

Data: Not applicable; vendor blog / Published: Mar. 5, 2026

rye.com/blog/openai-chatgpt-checkout-agentic-commerce

U13 Google: Universal Commerce Protocol features and AI tools

Data: Not applicable; announcement / Published: May 20, 2026

blog.google/products-and-platforms/products/shopping/shopping-updates-google-marketing-live

U14 Google: How to select your preferred sources in Top Stories

Data: Not applicable; announcement / Published: Aug. 12, 2025

blog.google/products-and-platforms/products/search/preferred-sources

U15 TechCrunch: Sundar Pichai shares Google Lens stats

Data: Not applicable; reporting (company figures) / Published: May 20, 2025

techcrunch.com/snippet/3009713/sundar-pichai-shares-some-google-lens-stats

VERSION 2.5.1 CORRECTIONS

P1 Cloudflare: Making AI search smarter (good-bot re-fetching; Pay Per Use pilots)

Data: Cloudflare data; window not stated / Published: July 1, 2026

blog.cloudflare.com/making-ai-search-smarter

P2 Cloudflare: Have it both ways: stay discoverable in search while disallowing AI training

Data: Not applicable; product announcement / Published: Sept. 15, 2026

blog.cloudflare.com/accountable-mixed-use-ai-crawlers

P3 OpenAI: Buy it in ChatGPT (Instant Checkout and the Agentic Commerce Protocol)

Data: Not applicable; product announcement / Published: Sept. 29, 2025

openai.com/index/buy-it-in-chatgpt

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

SOURCE DRAWER / 6 OF 6

VERSION 2.5.1 CORRECTIONS

P4 OpenAI: Powering product discovery in ChatGPT

Data: Not applicable; product announcement / Published: Mar. 24, 2026

openai.com/index/powering-product-discovery-in-chatgpt

P5 Google: Universal Cart launches in the U.S.

Data: Not applicable; product announcement / Published: Oct. 15, 2026

blog.google/products-and-platforms/products/shopping/universal-cart-launch

P6 Google Maps: prohibited and restricted content (user-generated content policy)

Data: Not applicable; policy / Published: Not stated; checked Oct. 31, 2026

support.google.com/contributionpolicy/answer/7400114?hl=en

P7 Yelp Support: Don't ask for reviews

Data: Not applicable; policy / Published: Not stated; checked Oct. 31, 2026

yelp-support.com/article/Don-t-Ask-for-Reviews?l=en_US

P8 Google Search Console Help: Search generative AI control

Data: Not applicable; documentation / Published: Not stated; rollout to all sites by Aug. 31, 2026

support.google.com/webmasters/answer/16908024

VERSION 2.5.1 CORRECTIONS

P9 Google crawlers: Google-Extended product token

Data: Not applicable; documentation / Published: Oct. 30, 2026 (last updated)

developers.google.com/crawling/docs/crawlers-fetchers/google-common-crawlers

P10 Google Search Central: robots meta tag, data-nosnippet and X-Robots-Tag

Data: Not applicable; documentation / Published: Mar. 24, 2026 (last updated)

developers.google.com/search/docs/crawling-indexing/robots-meta-tag

P11 Google Search Central: introduction to robots.txt

Data: Not applicable; documentation / Published: Dec. 10, 2025 (last updated)

developers.google.com/search/docs/crawling-indexing/robots/intro

P12 web.dev: Build agent-friendly websites

Data: Not applicable; guidance / Published: Apr. 1, 2026 (last updated)

web.dev/articles/ai-agent-site-ux

Source checks: Oct. 28-31, 2026 (primary sources P1 to P12: Oct. 31). Data windows and publication dates differ. Where a source does not state a date, the drawer says so.

THE COMPLETE GUIDE TO GEO

BUILD BETTER SOURCES. MEASURE WHAT CHANGES.

VERSION 2.5.1 | 2026

This field manual follows your content through the AI answer pipeline: query rewriting, fan-out, retrieval, selection, synthesis and citation. Across nine illustrated lab rooms, it connects SEO foundations with the question answer engines add: can a page be found, understood, checked and used accurately?

You'll learn to write self-contained passages with clear subjects, real evidence and named sources, keep brand and author facts consistent, and set crawler access by purpose. New in this edition: publisher controls, agent checkout, visual search, local discovery and the trade-offs of paid crawling.

Measurement gets the same care: fixed prompts, repeated runs, and mentions, citations and visits counted separately, with sample sizes and tracking gaps in view. Copy-ready field kits and a 90-day protocol make the ideas testable.

No layout, file or tactic guarantees a citation.

95

PAGES

9

LAB ROOMS

29

TACTICS RATED

82

SOURCES

SCAN FOR THE FREE WEB EDITION



JOURDAN ROMBOUGH

Founder, Optimization Theory
15+ years in technical SEO

optimizationtheory.com

