

SEO IS DEAD. HOW TO GET RECOMMENDED BY AI.

THE 2026 AEO PLAYBOOK

Paying \$5,000 a month for generic SEO is a dead spend. AI doesn't rank pages anymore – it recommends brands. This is exactly where those recommendations come from, with the receipts: 680M+ analyzed citations, every claim graded by how hard the evidence actually is.

Bonus: feed this report and the page-two prompt to your AI, and it builds a customized action plan around your brand – until your competitors wonder how you're suddenly everywhere and your customers quote you back.

Written by **Max Palmer** · Published by **MaxShip Marketing**

Don't just read this. Run it.

This PDF is built to be fed to an AI.

Drop it into ChatGPT, Claude, or Gemini. Paste the prompt below and answer the ten questions in your own words – rough is fine, detail is better. In a few minutes you'll have an action plan built around *your* brand, not generic advice.

► COPY-PASTE THIS INTO YOUR AI

You are a senior AEO strategist. Attached is the MaxShip AEO Playbook 2026. Read it fully, then use my answers below to build an action plan that uses ONLY the moves, data, and confidence grades in that playbook. If any answer is vague, make a reasonable assumption and say what you assumed.

About me – answer in your own words, rough is fine:

1. What's your product, and what does it actually do for a customer?
2. Who is your ideal customer, and what do they type into an AI when they have the problem you solve?
3. What stage are you (solo / seed / Series A / Series B+), and how big is the team that could act on this?
4. Where do your buyers actually research today – Google, ChatGPT, Perplexity, Reddit, YouTube, somewhere else?
5. Who are your top 2-3 competitors, and do they show up in AI answers more than you right now?
6. What does your site look like structurally – comparison pages, FAQs, schema, crawlable HTML, maybe an llms.txt? ("Not sure" is a fine answer.)
7. What's said about you off your own site – Reddit, G2/Capterra, YouTube, press – or close to nothing yet?
8. Is there a founder or expert who could be on camera or posting, and how much time could they realistically give per week?
9. What's your single biggest constraint – time, budget, engineering, content, or just knowing what to do?
10. What do you want out of this, and who will act on it? e.g. "a step-by-step plan I'll run myself," or "high-level tips I can forward to whoever runs our marketing," or "a brief for my team."

Deliver, shaped by my answer to #10:

- If I'll run it myself: a sequenced, week-by-week plan with specific tasks and the exact playbook move behind each one.
- If I'm forwarding it: a one-page brief for whoever runs marketing – priorities, the why, and exactly what to ask them to ship.
- Either way include: Track A / B / both and why (use the playbook's rule); my top moves ranked and tagged **PROVEN** / **MEASURED** / **GUESSED**; the 12-prompt test to baseline my brand; and a kill list of what to stop doing.

Be specific. Use my answers. Quote the playbook's numbers. If my situation isn't covered by the playbook, say so instead of inventing.

After it runs, read on so you can pressure-test what it gave you.

The 60-second version

Read this even if you read nothing else.

1. **Everyone has the same AEO tools now. Your edge is structure and story – not a tool.** Every dashboard measures the same shallow thing. Two things actually move citations: what you ship on your website, and everywhere else AI reads about you.
2. **Backlinks are over. Page structure is the new authority.** The brands winning AI answers aren't the ones with the most links – they're the ones with the cleanest pages. This one fact retires most AEO advice you've read.
3. **There are exactly two tracks, and an order to run them.** Track A (your website) is faster and in your control – start here if you're small. Track B (everywhere else AI reads you) has the higher ceiling – add it as you scale.
4. **Most "best practices" are guesses in a suit.** Every claim here carries a tag for how sure we are. The advice repeated loudest is often the weakest.
5. **You can baseline your brand this afternoon.** 12 prompts, 5 engines, one sitting. If you don't measure where you stand, you're optimizing blind.

The rest of this document is the receipts, the moves, and the templates.

Who this is for

Founders. Executives. Small-business owners. Marketing teams. Anyone who has wondered whether AI is recommending their brand – or quietly handing the customer to a competitor – and wants to do something about it. You don't need a big team, a big budget, or a developer on standby. Some of this is a one-hour fix; some of it is a quarter of work. We tell you which is which.

You don't need another "what is AEO" essay. You need to know which moves actually change whether AI recommends you, which are people copying moves nobody tested, and how to tell the difference. That's the whole playbook.

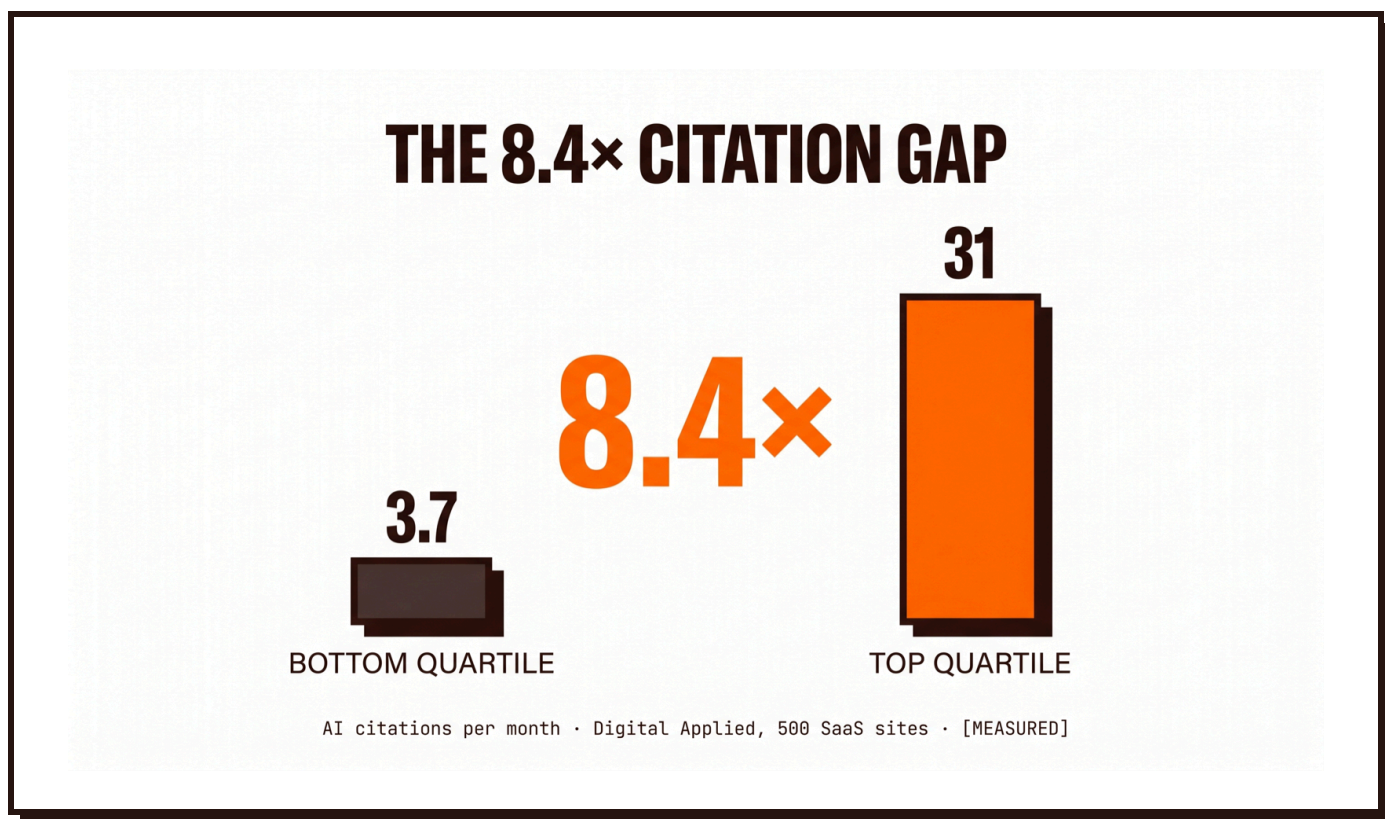
How to read the confidence tags

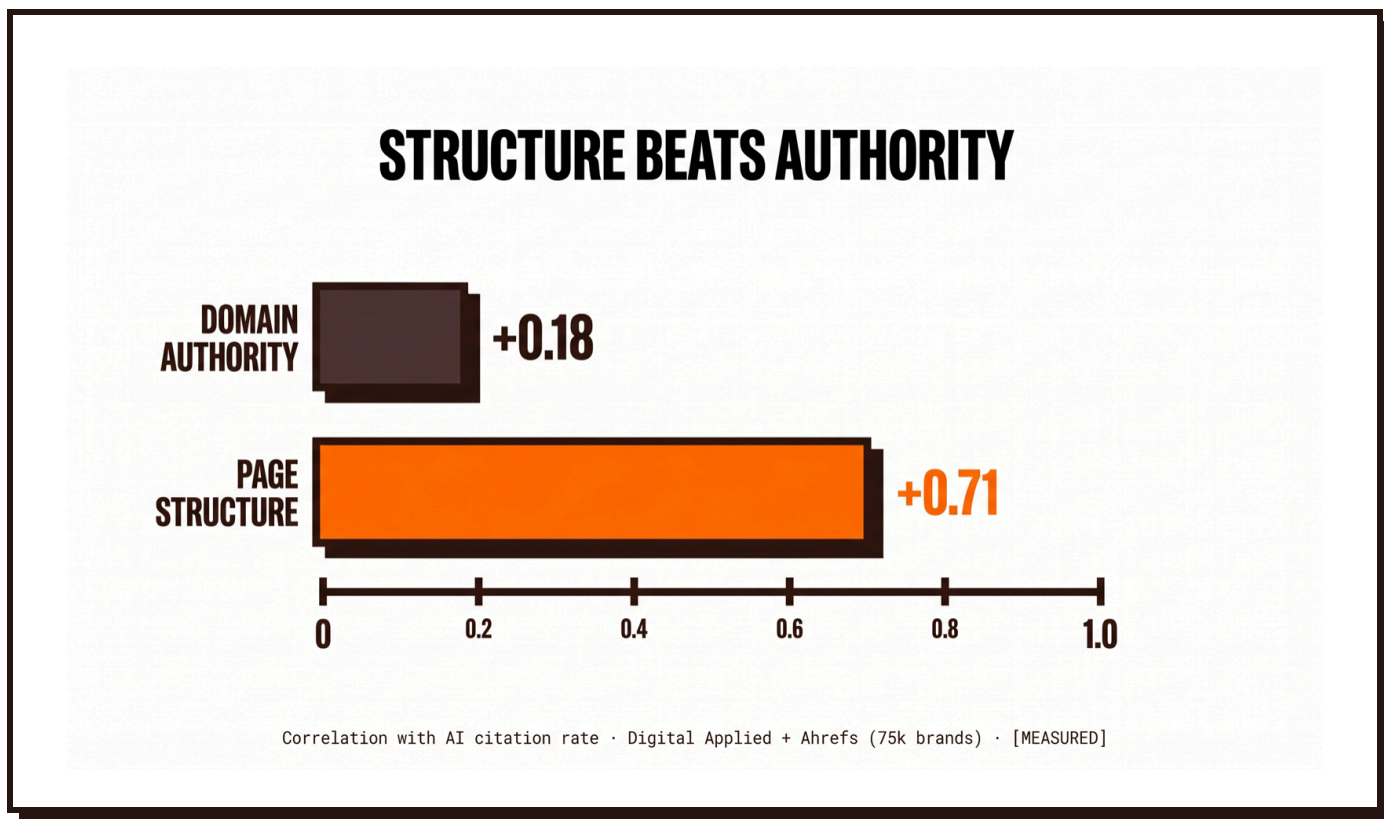
Every claim that matters carries one tag. This is the most important thing in the playbook – most AEO advice presents every claim with equal confidence, which is how everyone ends up copying moves nobody tested.

- **PROVEN** – peer-reviewed, replicated, or platform-documented. Copy it word-for-word.
 - **MEASURED** – one credible study or our own data. Directionally strong, single source. Could move.
 - **GUESSED** – best current theory. Plausible, popular, unverified. Bet on it, but bet small.
-

PART 1 — THE FRAME

Chapter 1 — The honest state of AEO in 2026





8.4x.

That's the gap between the top quartile of SaaS sites and the bottom half on monthly AI citations. The top 25% get cited 31 times a month across ChatGPT, Claude, Perplexity, and Gemini combined. The bottom 25% get cited 3.7 times. **MEASURED** — Digital Applied audited 500 SaaS landing pages over March and April 2026, three sampling passes, 12 queries per vendor.

That's a real gap. It's also a *closeable* gap, and that's the part most AEO experts won't tell you.

Here's the stat that should retire 80% of the AEO content on LinkedIn: **Domain Authority correlates with citation rate at +0.18. An eight-point structure checklist — things like comparison sections, answer-format headers, schema markup, and crawlable HTML — correlates at +0.71.** Same study. And it isn't one study's fluke: Ahrefs' separate 75,000-brand analysis found the same shape — Domain Rating correlates only ~0.27 with AI visibility, while content and off-site signals carry the weight. Two independent datasets, one conclusion.

Domain Authority is a 2018 currency in a 2026 market. The brands you compete with in AI answers aren't the ones with the most backlinks. They're the ones with the cleanest page structure. That's not how SEO worked. That's not how anyone is talking about AEO. It's just what the data says.

Pull quote: *"Domain Authority correlates with AI citations at +0.18. Page-level structure correlates at +0.71. That single stat should retire 80% of the AEO advice on LinkedIn."*

If structure is doing the work, why isn't the advice catching up? Two reasons.

One: nobody actually knows how these engines work. The retrieval algorithms are closed boxes. The patents are vague. Anthropic, OpenAI, and Google haven't published ranking weights. When an "AEO expert" tells you the exact mechanism, they're pattern-matching on a handful of personal observations. Sometimes that pattern-matching is correct. Sometimes it isn't. There's no way to tell from the outside. The engines aren't reliable narrators of themselves either: Columbia's Tow Center tested eight AI search tools and found they mis-attribute their own sources in more than 60% of queries – usually with complete confidence. **PROVEN**

Two: citation share moves weekly. The 5W Citation Source Index 2026 – built from a synthesis of more than 680 million individual citations across the five major surfaces – puts it bluntly: *"Citation share is volatile within weeks, not years."* The advice that worked in January may be wrong by July. By the time a "definitive guide" ranks for AEO keywords, half its numbers are stale.

So we grade every claim **PROVEN** / **MEASURED** / **GUESSED**. Not as a hedge – as a weapon. We'll tell you flat out which moves are bulletproof and which are bets, because the agency charging you \$5,000 a month won't tell you either, and the loudest voices in your feed are guessing with total confidence. Strong opinions, shown work. That's the whole document.

The thesis, then, in one sentence:

Everyone has the same AEO tools now. The dashboards and "share of voice" reports are the floor, not the ceiling. The ceiling is two routes that actually move citations: what you ship on your own website (Track A), and everywhere else AI reads about you (Track B). We'd bet on both, in that order, and we're running both publicly inside this playbook.

By the end, you'll know which moves to make this quarter, in what order, and which "best practices" to retire. You'll have a 12-prompt test you can run on yourself this afternoon. And you'll have the exact protocol we run on MaxShip's own brand – so you can run the identical test on yours.

If we get something wrong, you'll see us correct it. Try finding that promise in another AEO PDF.

DO THIS **Right now:** Note your current month's brand citation count on any one AI engine. Pick a brand-direct prompt ("*What does [your company] do?*"), run it on ChatGPT, screenshot the answer. That's your baseline. We'll come back to this in Chapter 6.

Chapter 2 — The tools everyone has

Open any AEO tool. They all do the same thing.

You give the tool a list of prompts. The tool runs those prompts across ChatGPT, Claude, Perplexity, and sometimes Gemini. It counts how often your brand shows up. It puts a number in a dashboard. The number goes up or down. You report on it.

That's the tool layer. Necessary, not enough.

Here's the honest tour:

Profound — Enterprise tier. Strongest depth on source attribution — they'll tell you not just whether you got cited but which page got the citation. Expensive (\$\$\$\$), dashboard-led, not actionable for a solo founder. Worth it if you have a content team and a measurement budget.

Peec AI — Mid-market competitive benchmarking. Same dashboard pattern as Profound but cheaper. Their angle is comparing you against named competitors. Useful if your CEO wants a slide that shows you beating X.

Otterly — Budget tier (\$29/month). Share-of-voice tracking across the major surfaces. Counts mentions; doesn't tell you the story behind the numbers. Fine starter tool.

Ahrefs AI Visibility Checker — Free. Lead-gen for the paid Ahrefs tier. Useful for spot checks. Don't rely on it for ongoing tracking.

Semrush AI Visibility Toolkit — Enterprise. Bundled with your existing Semrush seat. Tied to Semrush's data model, which is broad but shallow on AI specifically.

HubSpot AEO — 28-day trial, 25 prompts. Tied to HubSpot. If you're already a HubSpot shop, fine. If you're not, don't switch CRMs for this.

TrackAIMentions – Free 30-second single-keyword check. Spot tool, not a system.

Pick one based on your stage. **Solo or pre-PMF: none.** Do it by hand with the 12-prompt audit from Chapter 10. **Seed-Series A: Otterly or the free Ahrefs checker** for weekly spot checks. **Series B+: Peec for competitive benchmarking, or Profound if you have the budget.**

But here's the thing – and this is the chapter's real punchline.

None of them tell you the story. All of them tell you whether you're mentioned. None of them tell you *what exact words* the LLM is using to describe you. None of them tell you who you're getting *bundled with*. None of them tell you that your off-site footprint is missing Reddit, that your last Wikipedia edit was in 2022, that your G2 reviews mention old pricing.

That's the gap the tools can't fill. Dashboards plateau because the citation game happens before the dashboard ever sees it. Dashboards are downstream measurement.

Pull quote: *"Dashboards plateau because the citation game happens before the dashboard ever sees it. Dashboards measure outcome. The work is in content, off-site, and brand entity."*

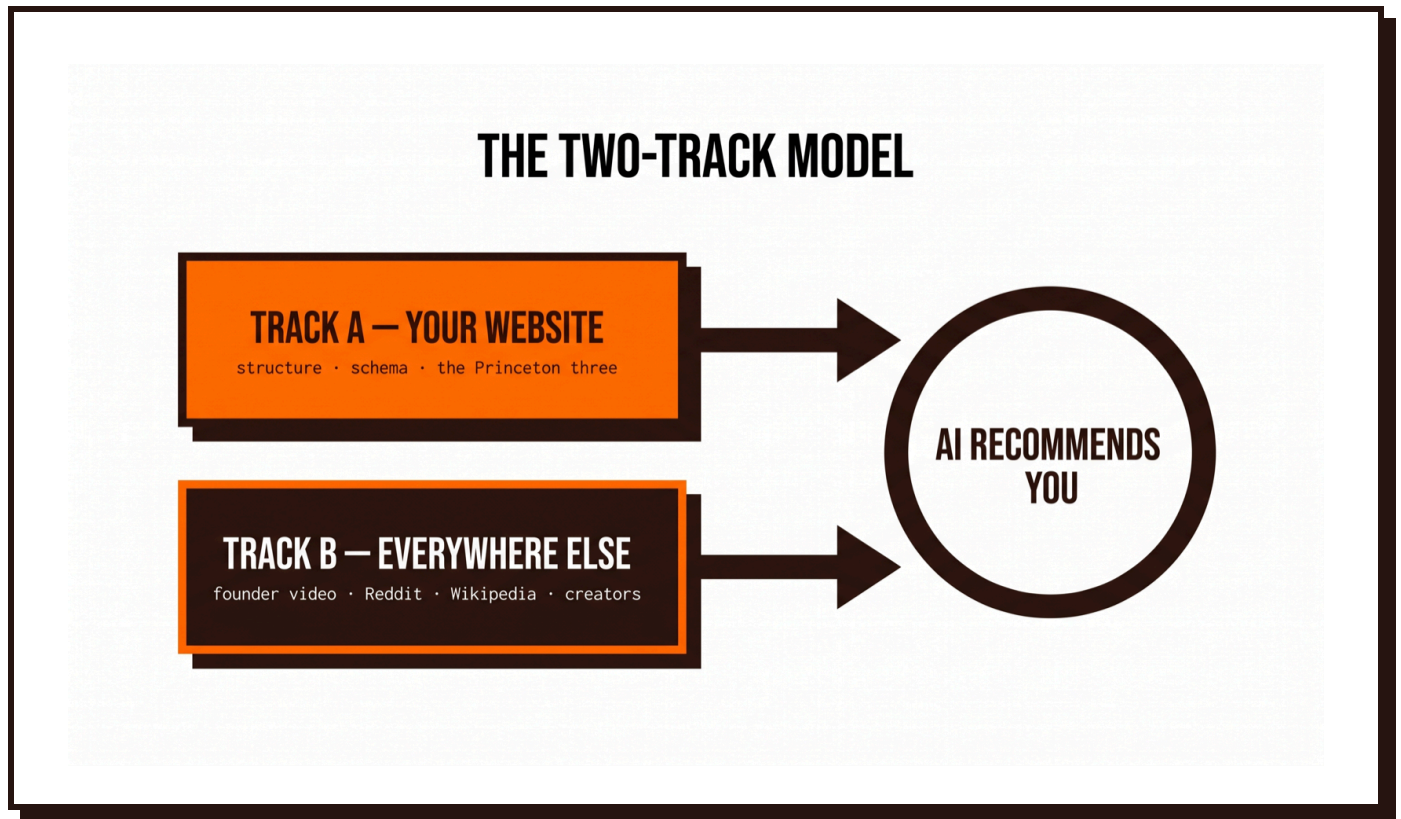
One more receipt before we move on. The single most-recycled SEO instinct – *stuff your content with keywords* – performs *worse than doing nothing* in AI engines. **PROVEN** Aggarwal et al. measured Keyword Stuffing at a Position-Adjusted Word Count of 17.7. The baseline of doing absolutely nothing scored 19.3. Stuffing keywords actively *hurts* you.

If your AEO tool's recommendation engine is suggesting you sprinkle in more target keywords, that's a sign the tool is recycling SEO heuristics that don't transfer. Pay attention to which platforms are honest about this. We'll grade some of them in Chapter 9.

DO THIS

This week: Audit your current AEO tooling spend. If you're solo or Seed and paying more than \$50/month for AEO software, cancel it. Put the money toward Chapter 4's moves instead.

Chapter 3 — Two tracks: your website vs everywhere else



There are two real routes to getting cited by AI engines in 2026, and you should pick which one is your bottleneck before you do anything else.

Track A — Your website. What you ship on a domain you own. Princeton's three (quote, stat, cite). Comparison sections. Answer-format H2s. Schema markup. Fast, crawlable HTML. Optional `llms.txt` as a cheap cross-engine hedge. The kind of work a marketer scopes and a developer ships in a week.

Track B — Everywhere else. Every other place AI reads about you, split two ways. **What you make:** your founder videos, your YouTube and LinkedIn channels, the original data and posts you publish off your main site. **What you earn:** your footprint across Reddit, Wikipedia, G2, and the industry trades, plus the creators and customers who mention you in their own voice. Founder content is the one Track B surface you fully own — that's why it's the biggest-payoff move here, and why Chapter 8 builds the whole motion around it.

These are not the same program. They have different costs, different timelines, different skill sets, and different ceilings.

Track A has high control and fast payoff. You decide what to ship. You can ship it next week. You can measure the impact within a month. But your ceiling is capped by your own site – there's only so much page-level optimization can do.

Track B has low control and slow payoff. You can't dictate that a creator mentions you. You can't force Wikipedia to keep your entry stable. You can't make a Reddit thread go viral. But the ceiling is much higher: a single well-placed mention on the right Reddit thread keeps generating citations for years.

Pull quote: *"Track A you can ship next week. Track B takes a quarter. The first is in your control. The second has a much higher ceiling. Most brands need both."*

Most brands need both, eventually. But here's the self-diagnostic, because most brands also pick the wrong one first.

Pick Track A first if:

- You're solo or pre-Series-A
- Your existing pages haven't been optimized for AI engines
- Your crawlability, schema, and answer structure are weak
- Your comparison pages don't have actual comparison tables
- You can ship dev work in days, not quarters

Pick Track B first if:

- Your site is already well-structured (you've shipped Princeton's three on your top pages)
- You have budget for creator partnerships or original-data PR
- Your category has an active Reddit presence you're not participating in
- You're not on G2 / Capterra and you sell software

Run both in parallel if:

- You're Series A+ with a content team and a PR budget
- You have one person who can own each track

We tell you which one to start with because most teams treat AEO as one undifferentiated bucket and get stuck. Picking the track makes the work concrete.

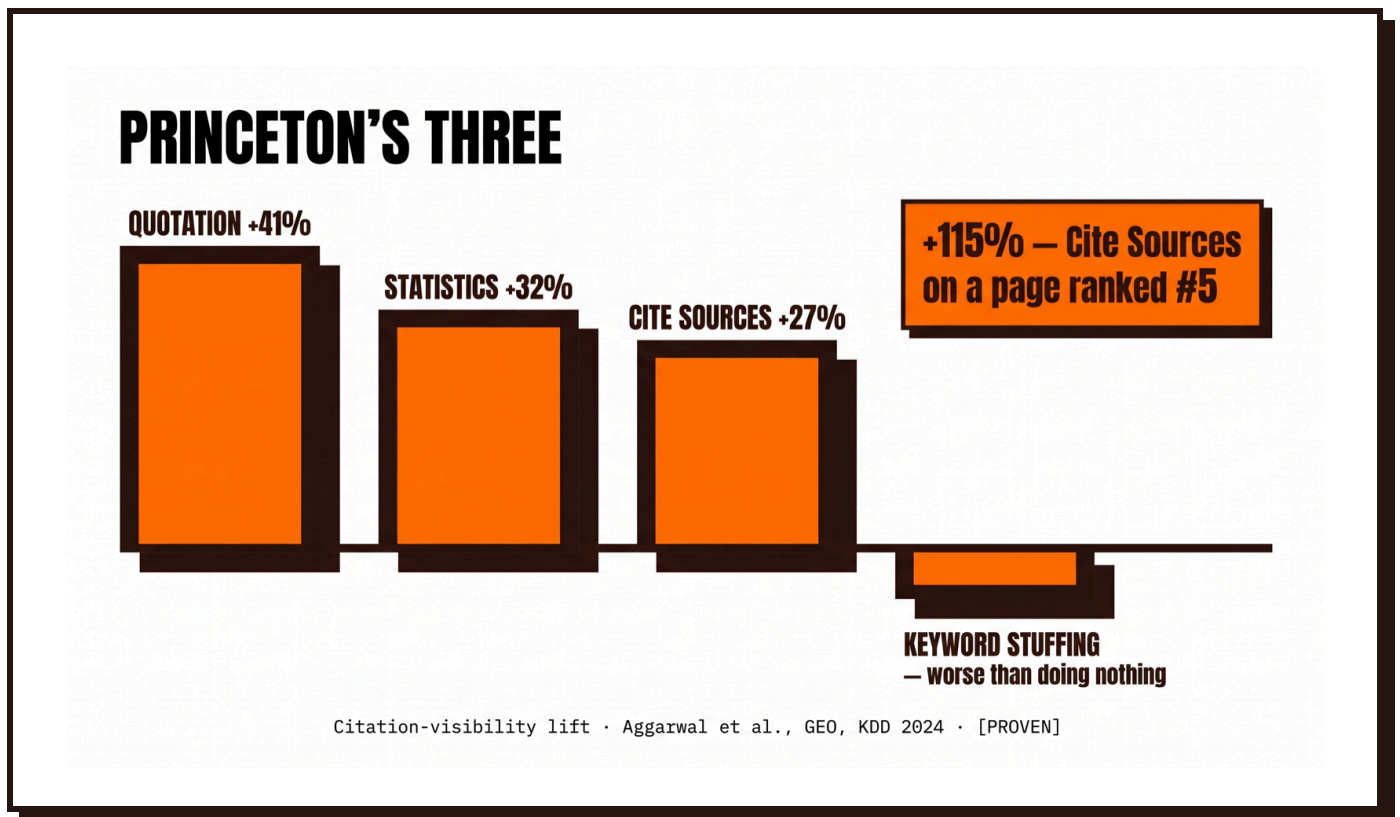
One more thing before the moves. **We run both tracks on MaxShip's own brand** – the landing page, the PDF, and the long-form article all implement Princeton's three plus the structure checklist (Track A), and the Counter-Take series is the storyteller play (Track B).

You don't have to take that on faith. Chapters 6 and 9 hand you the exact measurement protocol – so you can run the identical dual-track test on your own brand and read your own numbers, not ours.

DO THIS **Today:** Score yourself 0–3 on each track. Track A: are your top three pages running Princeton's three? Are schema, crawlability, and page experience in place? Track B: are you on Reddit weekly? Wikipedia? G2 with recent reviews? Whatever scores lower is your first quarter's focus.

PART 2 — TRACK A · YOUR WEBSITE

Chapter 4 — The signal stack you ship this week



This is the chapter where the dev tickets get written. If you do nothing else this quarter, do the moves in this chapter on your three highest-traffic pages and watch what happens.

We're going to mix three sources of truth: Aggarwal et al.'s GEO paper (Princeton + IIT Delhi, KDD 2024, $n = 10,000$ queries, replicated on Perplexity), Digital Applied's 500-SaaS audit (March–April 2026), and Google's May 2026 Search Central guide to generative AI features in Search. The first gives

you **PROVEN** causal lift numbers; the second gives you **MEASURED** structural correlations; the third is platform-documented guidance for AI Overviews and AI Mode. Together they're the strongest on-site signal stack anyone has documented as of May 2026.

Google's official frame matters: generative AI features in Search are still grounded in the normal Search index and ranking systems, using retrieval and query fan-out to find relevant pages. For Google specifically, the boring fundamentals are not optional: unique non-commodity content, crawlable pages, clear technical structure, good page experience, useful images/video, and Search Console monitoring. This does not make every AEO tactic wrong; it means Google-specific claims need a higher bar.

Princeton's three (the **PROVEN** moves)

The GEO paper tested nine separate on-site optimization methods against a 10K-query benchmark. Three of them blew the doors off the rest. We call them the "Princeton three."

1. Quotation Addition: +41% on Position-Adjusted Word Count, +28% on Subjective Impression.

PROVEN

The single biggest lever. Add inline quotations from credible sources, attributed word-for-word. Not paraphrased. The exact words.

According to the 2026 Stripe State of Payments report, "78% of B2B buyers expect to compare pricing without booking a call."

That sentence – with the quotation marks and the attribution – is more citation-worthy than the same fact paraphrased. The LLM uses quotation marks as a proxy for credibility. So does the human reading the answer.

2. Statistics Addition: +32% citation visibility. **PROVEN**

Add concrete numbers. Replace qualitative claims with quantitative ones.

Bad: "Most marketers are now using AI search tools." Good: "61% of B2B marketers reported using ChatGPT for research in Q4 2025 (HubSpot State of Marketing 2026)."

The improvement is causal. The same paragraph, with a number added, gets cited more often. The LLM treats numbers as anchors. So does the reader.

3. Cite Sources: +27% citation visibility overall. +115% for pages ranked at SERP position 5.

PROVEN

Add inline citations to authoritative third-party sources after every claim that matters. The Princeton paper measured this in two ways. The headline number is +27% – meaningful but not as big as the first two. The buried number is the killer: **for pages that rank fifth on Google for their target query, adding Cite Sources delivers a +115% citation lift.**

If you're not the top-ranked site in your category, this is your great equalizer. GEO rewards lower-ranked sites that cite well. The big incumbents are mostly already-cited; you have more headroom.

Pull quote: *"For pages ranked fifth on Google, adding inline citations delivers a +115% AI citation lift. GEO is the great equalizer nobody is quoting."*

Combination effect: When you stack methods, lifts compound. The paper measured +35.8% citation-visibility lift for Fluency Optimization + Statistics Addition combined. **The best combination outperformed any single method by more than 5.5%.** Translation: do all three of the Princeton moves on the same page, plus a fluency pass, and you stack the gains. Don't pick one – layer them.

The structure checklist (the **MEASURED** moves)

Digital Applied's audit of 500 SaaS sites measured which page-level structural features correlated with higher citation rates. The structure checklist correlation hit +0.71 – versus Domain Authority's +0.18. Here are the top four lifts:

MOVE	CITATION LIFT	GRADE
Comparison sections on the page	+38%	MEASURED
<code>llms.txt</code> at site root	+24% (one audit) – <i>contested</i>	GUESSED (a larger study finds no effect – see below)
Answer-format H2s ("What is X?")	+22%	MEASURED
<code>SoftwareApplication</code> JSON-LD schema	+18%	MEASURED

Comparison sections are the biggest individual move outside of Princeton's three. If you sell software, every category landing page should include a comparison table – your product against the top three alternatives, with rows for the buying criteria your audience actually uses. Not "we win everywhere." Honest. LLMs cite comparison tables because they're easy to extract.

`llms.txt` is contested – and we'd be lying if we sold you the +24%. One audit (Digital Applied) found sites with `llms.txt` at root cited 24% more. A *larger* study – SE Ranking, ~300,000 domains – found no measurable link at all, and reported that *removing* the `llms.txt` variable actually made their model more accurate. Google Search Central also says special AI text files, machine-readable AI files, and AI-only markup are not required for Google's generative AI features. So: one positive audit, one larger negative study, platform guidance against treating it as a Google lever, plausible selection-effect either way. Grade it **GUESSED**, not **MEASURED**. It's a five-minute file with no downside, so ship it as a cheap cross-engine hedge – but don't put it on the slide as a proven lever, and don't let anyone sell it to you as one.

A starter `llms.txt` for a SaaS site looks like:

```
# Acme – Customer Feedback Platform

Acme helps product teams collect, organize, and act on customer
feedback at scale. We replace the workflows currently held together
by spreadsheets, Slack screenshots, and unstructured Notion docs.

## Primary use cases
- Centralizing customer feedback from email, support, sales, and Intercom
- Tagging and clustering feedback into actionable themes
- Routing high-signal feedback to product and engineering

## Key pages
- Pricing: /pricing
- Comparison vs Productboard: /vs/productboard
- Customer stories: /customers
- API docs: /docs

## Founders
Founded 2023 in San Francisco by Jane Doe (CEO, ex-Stripe PM)
and John Smith (CTO, ex-Linear engineer).
```

Answer-format H2s are headings that mirror how a user asks a question. *"What is X?"* / *"How does X compare to Y?"* / *"How much does X cost?"* – instead of clever marketing copy. The LLM extracts the answer paragraph beneath an answer-format H2 cleanly. The clever copy gets skipped.

SoftwareApplication schema (JSON-LD) is the single most-rewarded schema type for SaaS pages in Digital Applied's audit. If you sell software and you don't have this on your product pages, ship it tomorrow.

```
{
  "@context": "https://schema.org",
  "@type": "SoftwareApplication",
  "name": "Acme Feedback",
  "applicationCategory": "BusinessApplication",
  "operatingSystem": "Web",
  "offers": {
    "@type": "Offer",
    "price": "29",
    "priceCurrency": "USD"
  },
  "aggregateRating": {
    "@type": "AggregateRating",
    "ratingValue": "4.7",
    "reviewCount": "142"
  }
}
```

What does *not* work (the counter-receipts)

Keyword Stuffing performs worse than doing nothing. **PROVEN** Aggarwal et al. measured it at 17.7 citation visibility vs. the baseline 19.3. If you're sprinkling target keywords through your content because your SEO tool told you to, stop. AI engines pattern-match the stuffing and penalize it.

Aggressive "Authoritative" rewrites had no significant lift in the main GEO test. Making your content more persuasive or rule-of-three rhetorical didn't move the needle. AI engines don't reward rhetorical showing-off.

JS-heavy render strategies are likely **GUESSED losers.** Some AI crawlers still struggle with client-side rendering. No engine has published its render policy as of May 2026. The safer bet is SSR or pre-rendered HTML. Worth the dev cost even though we can't prove the citation impact yet.

DO THIS **This sprint:** Pick your three highest-traffic pages. Add one inline quotation, one statistic with a source, and one inline citation to each. Ship **SoftwareApplication** schema if you're SaaS. Make the pages crawlable, fast, and readable in plain HTML. Add `llms.txt` only if you want a cheap cross-engine test. Re-measure citation count in 4 weeks. That's the experiment.

Chapter 5 — Page templates worth copying

Most of Chapter 4's moves slot into three page archetypes. Build these three templates, hand them to your dev team, and you have a Track A program.

Template 1 — The answer page

The shape: a question as the H2, a 50–80 word direct answer paragraph immediately below, then supporting evidence with inline citations.

```
<h2>What is Answer Engine Optimization?</h2>
<p>
  Answer Engine Optimization (AEO) is the practice of structuring
  content so AI engines like ChatGPT, Claude, and Perplexity will
  cite it directly in their generated answers. Unlike traditional
  SEO, which optimizes for ranking in a list of blue links, AEO
  optimizes for being mentioned inside a synthesized response.
</p>
<p>
  The discipline draws on emerging research like
  <a href="https://arxiv.org/abs/2311.09735">Aggarwal et al.'s
  GEO paper (KDD 2024)</a>, which demonstrated that adding inline
  quotations, statistics, and source citations can boost AI
  citation visibility by 30-40%.
</p>
```

The 50–80 word window is from the answer-paragraph pattern Digital Applied measured. Under 40 words, the LLM doesn't extract enough context to cite usefully. Over 100 words, the LLM picks a less prominent sentence inside the paragraph. Fifty to eighty is the sweet spot.

Use this template for FAQ pages, glossary pages, and the top half of category landing pages.

Template 2 — The comparison page

The shape: the question in the H1, a 60-word framing paragraph, then a comparison table with the buying criteria as rows. Honest. Real. Including the places competitors are stronger.

H1: ProductA vs ProductB – which one for [audience]?

Intro paragraph (~60 words):

Both ProductA and ProductB serve [audience type], but they diverge on three buying criteria that usually decide the choice: [criterion 1], [criterion 2], and [criterion 3]. This page compares them on the metrics that matter for [use case].

Comparison table (5-8 rows, ≤12 words per cell)

Comparison sections are the +38% lift from Digital Applied. The LLM extracts table rows cleanly because the structure is unambiguous. Mark up the table with `@type: Table` if you want to be thorough, but the cleaner HTML matters more.

The honest part is non-negotiable. Comparison pages that say "we win on everything" don't get cited; the LLM has read the other side's site and notices. Pages that say "we lose on price but win on integrations" get cited because they're useful.

Template 3 — The definitive guide

The shape: TOC at the top, dated freshness block, author byline with linked entity page, per-section inline citations, FAQ at the bottom.

H1: [Topic] – The 2026 Guide

[Freshness block, top-right or top of body]

Last updated: 11 May 2026

By [Author Name] · [linked bio page]

TOC (anchor links to H2s)

H2: What is [topic]?

H2: How does [topic] work?

H2: The five things that actually matter

H2: Common mistakes

H2: How to measure [topic] success

H2: What's next for [topic]

H2: FAQ

FAQ section at bottom with FAQPage schema

Dated freshness is increasingly important. 5W's index reports journalism citations rising to 49% on time-sensitive queries; engines push toward "answers from this week" defaults. A visible "Last updated" date with a recent timestamp signals freshness to both the engine and the reader.

Author bylines should be linked to a real author entity page on your site – `/authors/jane-doe/` with a bio, photo, and links to that author's other content. The LLM treats author entities as trust signals. Anonymous "by Acme Team" bylines get penalized.

Non-negotiables across all three templates

- Server-side rendered HTML. No critical content behind JavaScript.
- One H1 per page. Semantic landmarks (`<article>` , `<section>` , `<aside>`).
- Descriptive alt text on every image (LLMs read it).
- Fast LCP (under 2.5s). Slow pages get crawled less.
- Valid JSON-LD schema (`Article` , `FAQPage` , `SoftwareApplication` , `BreadcrumbList`).
- Optional `llms.txt` at root summarizing the page tree; useful as a cross-engine experiment, not a Google-specific requirement.

DO THIS **Next two sprints:** Convert your top three pages to one of these templates. Don't redesign – restructure. Same content, new shape. That's the experiment.

Chapter 6 — The Track A test (run it on your own brand)

The most honest thing a playbook can do is hand you the experiment, not just the conclusion. This chapter is the protocol.

What a fully-implemented Track A site looks like. MaxShip's own landing page, PDF, and long-form article ship every move from Chapters 4 and 5 – use this as your checklist:

- Inline pull-quotes from primary sources, attributed word-for-word
- Every statistic carries an inline citation to a paper or named audit

- A comparison block: the tools everyone has vs the moves that actually matter, in a clean table
- Optional `llms.txt` at root as a cross-engine experiment
- Answer-format H2s (one real question per section)
- `SoftwareApplication` + `Article` + `FAQPage` JSON-LD
- A dated freshness block and a real author entity page
- Served as SSR HTML – not a JavaScript SPA

The measurement protocol – run this on your brand. Pick 12 queries (3 branded, 3 category, 3 comparison, 3 the job they need done – the full bank is in Chapter 10). Run each on ChatGPT, Claude, Perplexity, Gemini, and Google AI Overviews. That's a 60-cell grid. For each cell capture: mentioned (Y/N), citation (link present?), tone (+/0/-), and the exact words the engine used.

```
Cites / 60 cells Δ vs baseline
Baseline (before changes): [N] –
+30 days: [N] [+/- ]
+60 days: [N] [+/- ]
+90 days: [N] [+/- ]
```

The honest layer. Run three sampling passes per measurement – the same prompt twice can return different answers, and you want signal, not noise. If your structural moves don't move your citations, that's data: this is **PROVEN** in aggregate (Aggarwal et al.) but your site is one data point, and one data point can surprise you.

DO THIS **This week:** Build your 60-cell baseline before you change anything. One spreadsheet. Re-run it at 30, 60, and 90 days against the same 12 prompts. That spreadsheet is the only proof that matters.

PART 3 — TRACK B · EVERYWHERE ELSE

Chapter 7 — Where AI reads about you

WHERE AI READS ABOUT YOU

REDDIT ≈40%

WIKIPEDIA ≈30% on ChatGPT

YOUTUBE #2 most-cited source

15 domains = 68% of ALL AI citations

Corroborated across Semrush, Peec AI & Ahrefs · 100M+ citations · [MEASURED]

Track A optimizes your site. Track B optimizes everything else. And "everything else" is where the bigger numbers live.

5W's 2026 Citation Source Index looked at more than 680 million AI citations across the five major engines. The headline finding: **15 websites get 68% of every AI citation.** Not the top 1,500. The top 15. **MEASURED**

That's a small battlefield. The good news: it's the same fifteen sites on every engine. The work is finding where you're missing and which ones you can realistically go after.

The six places AI looks

5W splits the off-site world into six buckets. Each one works differently and pays off differently depending on your business.

Community & Conversation – Reddit dominant. Reddit is the #1 source across every major AI engine at roughly 40% citation frequency. **MEASURED** This isn't one study's quirk – three independent datasets land on the same number: 5W's index, Semrush's 150K-citation analysis (Reddit 40.1%), and Peec AI's 30M-source study (Reddit #1). That's the single biggest fact in this entire playbook for most brands. We give Reddit its own deep-dive section below.

Encyclopedic & Reference – Wikipedia dominates ChatGPT at 26–48% of top-10 share. **MEASURED** Wikipedia and Wikidata are entity stores; if you don't have a clean, current Wikipedia entry, ChatGPT may quote outdated facts about you. Wikipedia's notability bar is high – you can't just write your own page – but if you qualify, having the page is enormous.

Professional & Identity – LinkedIn, About-page mentions, GitHub for technical brands, Crunchbase. Make sure your LinkedIn company page is complete, your About-page entity copy is consistent across the web (same one-line description everywhere), and your Crunchbase entry is up to date. This is cheap and pays off fast.

Video & Audio – YouTube is the #2 most-cited source in AI answers, behind only Reddit. **MEASURED** Three separate studies agree: Otterly (100M citations) ranks it #2; Ahrefs (75,000 brands) found YouTube mentions are the strongest single signal it tested – far ahead of Domain Rating; Peec AI (30M sources) puts it top-three. The engine reads the transcript, not the video. A clean, well-titled long answer to your top category question keeps getting cited for years. One honest caveat: Otterly only looked at videos that were *already* cited (40% had under 1,000 views). That proves audience size doesn't decide whether you *stay* cited – not that you'll get cited in the first place. Don't sleep on this.

Editorial & News – 27% of all AI citations, rising to 49% on time-sensitive queries. **MEASURED** This is where PR coverage pays off. A single mention in TechCrunch, Search Engine Land, or Wired propagates into citations for months. Trade press matters too – Marketing Brew, Stratechery, and category-specific pubs all get cited.

Commerce & Review – G2, Capterra, TrustRadius. If you sell software, being on at least one of these with 5+ recent reviews is the bare minimum. LLMs treat review sites as third-party verification of your existence and reputation. Outdated entries actively hurt – a Capterra page showing 2023 pricing means ChatGPT may quote you at a wrong price.

Score yourself

Score your brand 0 to 3 in each of the six categories. Anything under 12 total is below the citation floor.

CATEGORY	0	1	2	3
Community (Reddit)	No presence	One sub	3+ subs, occasional posts	3+ subs, weekly active, recent original posts
Encyclopedic	No page	Wikidata only	Wikipedia stub	Full Wikipedia + recent edits
Professional	Incomplete LinkedIn	Complete LinkedIn	+ GitHub/Crunchbase	+ recent thought-leadership
Video & Audio	No channel	YouTube exists, low views	10+ videos, on-topic	25+ videos with full transcripts and good titles
Editorial & News	No coverage	One mention, 12mo+	Quarterly mentions	Monthly mentions, mix of trade + general press
Commerce & Review	Not listed	Listed, no reviews	5+ reviews, outdated	20+ reviews, current pricing, recent activity

If you scored under 12, your first quarter's external work is bringing the lowest two categories up to a 2.

Pick three off-site moves this quarter

Don't try to fix all six categories at once. Pick three based on your lowest scores and where your category pays off most.

1. Wikipedia / Wikidata entity establishment. If you meet Wikipedia's notability bar (multiple solid third-party references in independent reliable sources), this is the single biggest move in the whole off-site program. Don't write the page yourself – hire a Wikipedia editor or commission one through a service that knows the policies. A neutral, well-sourced article will stick. A self-promotional one will get reverted within a week.

2. G2 / Capterra placement with current pricing and 5+ recent reviews. If you sell software, this is non-negotiable. Encourage happy customers via email (every closed-won deal in your CRM, plus your top 20% NPS responders). 5 reviews is the threshold for category-page inclusion on most review sites; 20+ is the threshold where you start showing up in "best of" lists.

3. One original-data piece per quarter. Publish your own audit, benchmark, or experimental data once a quarter. Make the data freely available. License it CC-BY so other people can quote it. The Digital Applied audit you're reading citations from in this playbook is exactly this move done right. We're doing the same thing with the Track A and Track B test data in this playbook.

Pull quote: *"The top 15 domains capture 68% of all AI citations. Not the top 1,500. The top 15. Your off-site footprint is concentrated, and that means it's auditable."*

The Reddit substantial section

Reddit gets its own treatment because it's 40% of the citation graph and the place most brands get this wrong.

The cynical version of "Reddit AEO" – paying for upvotes, fake accounts, doorway posts pretending to be real users – doesn't work, gets you banned, and generates contempt when discovered. Reddit's spam policies have teeth, and the LLMs have learned to discount low-effort promotional patterns. Don't do this.

The honest version (and the only one that compounds):

- **Find your 3–5 relevant subs.** For a B2B SaaS in customer feedback, that might be r/SaaS, r/ProductManagement, r/CustomerSuccess, r/marketing, r/Entrepreneur.
- **Build a real identity over months.** Comment far more than you post. Answer questions that aren't about your product. Get karma from being useful.
- **When relevant, mention your product *in context* and *disclosed*.** Reddit's 10% guideline is real: no more than 10% of your activity can be self-promotional. Disclosed promotion in a relevant thread is fine. Undisclosed shilling is not.
- **Post original analysis especially.** Your own data, your own benchmarks, your own case studies, your own thinking. Original posts get upvoted and cited by AI engines more than recycled commentary.
- **Long-game beats short-game.** A real Reddit presence takes 90 days of daily activity to start paying off. There is no shortcut. There are people selling shortcuts. They are lying.

What gets you banned: vote manipulation (cross-account upvotes), undisclosed promotion above the 10% line, brigading (organized voting from outside subs), link farming, ban evasion. Read the spam policy once.

The cadence: 10–20 substantive comments per week across your 3–5 subs, 1 original analysis post per month, never more than one self-promotional post per 10 contributions.

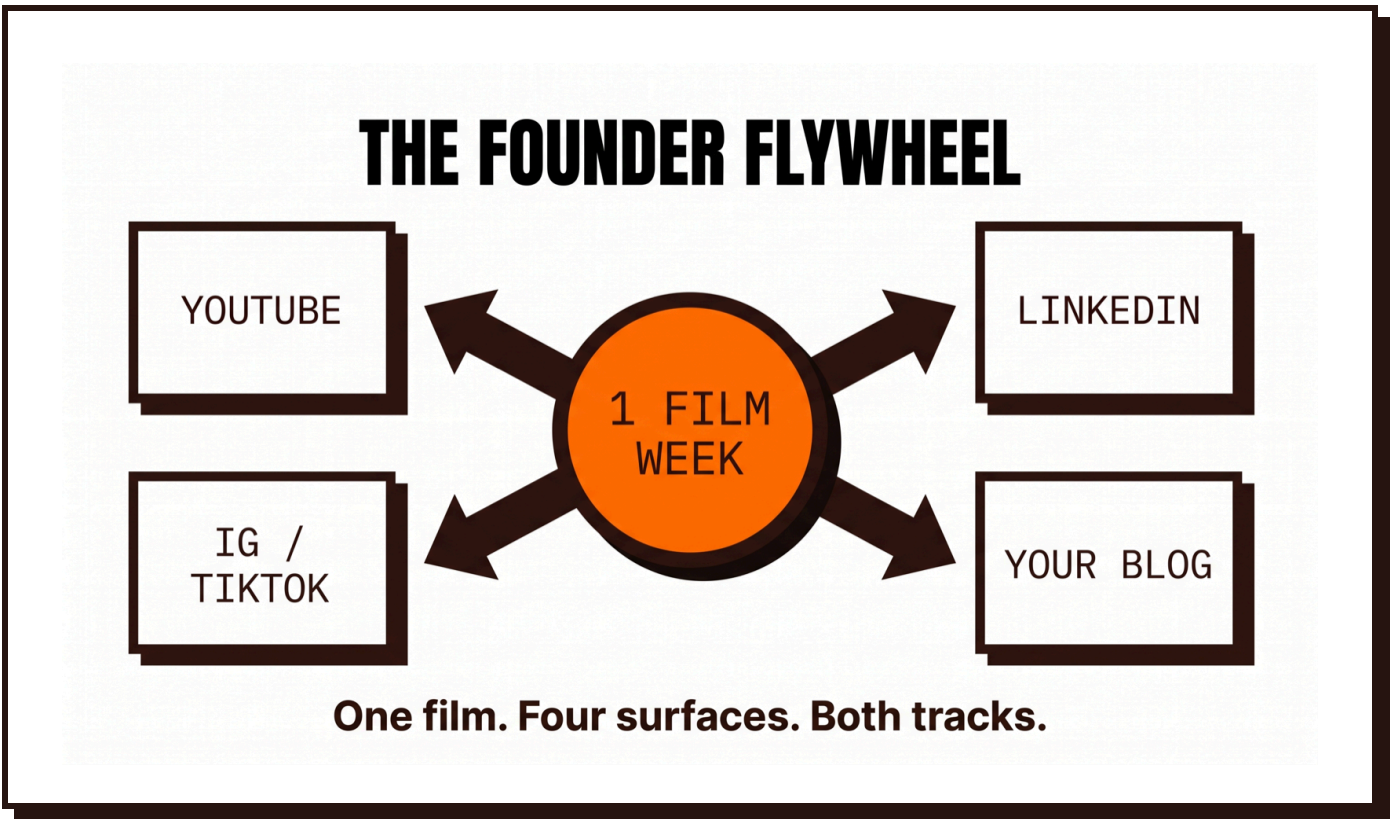
DO THIS **This month:** Score yourself on the off-site checklist. Pick the lowest category. Spend the next 30 days getting it from 0 or 1 up to a 2. Block 30 minutes a day for it.

Chapter 8 — The storytellers

The single biggest move in Track B isn't a directory listing. It's a person.

Creators, podcasters, customers, peers – when they talk about you in their own voice, those mentions become the exact words LLMs quote about you. Generic third-party listings get cited because they exist; storytellers get cited because they're interesting. The first kind of citation is utility. The second kind is gravity.

Your founder is the storyteller you already have



Every other move in this chapter needs you to find, pitch, or pay someone. This one doesn't. The most citeable storyteller for your brand is the person who already has the conviction: the founder, or your sharpest in-house expert. And one founder doing this consistently is the only Track B move that also feeds Track A – same work, both tracks.

Here's the motion. Four steps, run weekly.

1. Mine the questions. Build one pillar. Don't brainstorm topics – pull the real ones. Run the 12-prompt test (Chapter 10) plus your SEO keyword tool and list every question your buyers actually type into Google and AI when they hit the problem you solve. Cluster them. The biggest cluster – the question family that decides the sale – is your **content pillar**. One pillar, not ten. Depth on one theme is what earns topical authority. Scatter doesn't.

2. Film one real answer a week. The founder films one long-form video a week answering the top open question in that pillar. Not scripted marketing – the actual answer, the way they'd give it to a customer on a call. Ten honest minutes beats a polished ad, for the same reason sanitized creator content doesn't land: your audience smells the script. The question is the title. The spoken answer is the asset.

3. Cut it into shorts. One film is the week's raw material: the long cut, 3–5 short vertical clips, and the transcript. You film once; you ship a week.

4. Post it everywhere it gets indexed.

- **YouTube** (long + shorts) – the #2 most-cited source in AI answers. The transcript gets indexed, and the founder's spoken answer becomes the word-for-word quote AI pulls.
- **LinkedIn** – the founder posts the clip with the written answer underneath. Professional-identity signal plus a second indexed copy.
- **IG / TikTok** – reach and brand. The shorts build an audience that later searches your name – a branded query you want to already own.
- **Your own blog** – the cleaned-up transcript as an answer-format page with the founder as the named author. That's the Track A move from Chapter 4 – answer-format H2, direct answer, author entity – produced for free as a byproduct of the video.

One film. Five surfaces. Both tracks. That's the whole point: you're not running five content workstreams, you're running one with five payoffs.

Why it works, graded honestly. Three of the four pieces here are backed by real data: the question-as-headline is the answer-format move (+22% **MEASURED**); the YouTube transcript is the #2 most-cited source in AI answers **MEASURED**; weekly shipping feeds the freshness signal engines reward – Ahrefs' 17M-citation study found AI-cited content runs about 26% fresher **MEASURED**. The fourth – the named

founder as a trust signal — is our best bet, but it's a guess, not a result: the one *controlled* test on the closest thing (author and schema markup, Ahrefs, 1,885 pages vs 4,000 controls) found no lift and a small *drop* in Google's AI answers. So stack the three proven pieces on one pillar and they compound. The founder-as-author layer is a reasonable add, not a proven one. Run it and measure your own brand.

MEASURED on three of four — the whole system is our strong recommendation, not a proven result.

The honest cost. This is the easiest way into Track B and the single biggest point of failure. It does not delegate — the founder *is* the asset. It compounds on a ~90-day curve, like the Reddit play: weeks 1–8 feel like nothing. And the no-script rule still holds — but be honest about why: that a *human* tunes out scripted founder content is well-established; that the *engine itself* penalizes a script is **GUESSED** (asserted everywhere in AEO content, measured nowhere). Keep the rule for the human reason, not a claimed algorithm penalty. If your founder won't commit to one honest hour a week for two quarters, pick a different Track B move. A half-hearted version of this is worse than not doing it.

DO THIS **This week:** Pull your 12-prompt list plus your top 20 SEO questions. Cluster them. Circle the one pillar. Put a recurring 30-minute "founder films one answer" block on the calendar for the next 8 weeks. The first video doesn't have to be good. It has to exist.

Creator partnerships — the honest version

The biggest mistake brands make with creators is treating them like billboards. Buy a sponsorship, get a 60-second read, measure CPM. That's the SEO mindset applied to a different medium. It doesn't compound.

The version that compounds:

- **Pick creators whose audience overlaps your ICP, not the most-followed.** A 12,000-subscriber YouTuber whose audience is 100% your buyer is worth more than a 1.2M-subscriber YouTuber whose audience is 0.5% your buyer.
- **Pay them. Disclose it. Let them write what they want.** Sanitized promotional reads don't get cited because LLMs (and humans) detect them. Substantive analysis where the creator says what they actually think — with disclosure — gets cited.
- **Long-game beats one-shot.** Six conversations with one creator over six months beats one sponsored post. The relationship compounds; the spot doesn't.
- **Brief sparsely.** Give them your product, your data, and your time. Don't give them talking points. The talking-point energy is exactly what makes branded content unciteable.

Customer-led PR — the "loved out loud" pattern

The cheapest external move you can make is helping your happy customers share. Not scripted. Not coached. Just removed friction.

The pattern that works:

- **Identify the top 10% of your NPS responders.** Real evangelists self-select.
- **Ask them one specific thing.** Not "would you be willing to do a case study?" — too heavy. Try: "Would you write one paragraph about how you use [product] for a customer story page?" Smaller asks have higher completion rates and the resulting content is more natural.
- **Make the artifact reusable.** Their one paragraph becomes your case study, your sales collateral, and — if it's good — a tweet or LinkedIn post they share themselves.
- **Don't script.** The paragraph you'd write for them sounds like marketing. The paragraph they write sounds like a person.

Original-data case studies are the gold tier here. Pair a customer with a number — "Acme cut feedback triage time from 4 hours to 12 minutes" — and you've made an artifact creators will reference, journalists will quote, and LLMs will cite.

Original-data PR — the highest-ceiling move

Publishing your own audits, benchmarks, or experimental data is the single best way to earn category citations.

The shape:

- **Ask a question only your data can answer.** If a competitor's data could answer the same question, the citation goes to whoever has more reach.
- **Make it freely available.** License the data CC-BY. Let people use it.
- **Frame it for journalists.** A specific number (8.4×, +41%) in a tweet-sized framing wins distribution.
- **Pitch it. Don't expect organic.** Send to 10 trade pubs and 10 creators in your category, plus your existing customer list. PR is a job, not a hope.

Three concrete examples we cite throughout this playbook:

- **Digital Applied's 500-SaaS audit.** Their 8.4× gap stat is the most-cited stat in 2026 AEO content. We've seen it appear in Backlinko, Frase, Semrush, three LinkedIn cheat sheets, and now this

playbook. They published it once.

- **5W's Citation Source Index.** The "top 15 domains = 68% of citations" stat is now in every AEO guide written after May 2026.
- **Aggarwal et al.'s GEO paper.** The "+41% from quotations" finding has been cited word-for-word in dozens of AEO guides since 2024.

Each of these is a small organization or research team. Each one earned permanent inclusion in the citation graph for the cost of running one analysis.

The Counter-Take pattern — the storyteller move that earns citations

The Counter-Take YouTube series is the version of the storyteller move we're running on ourselves.

Each episode reads eight pieces of viral AEO advice – LinkedIn posts, paid courses, blog articles, YouTube videos, X threads – and drags each one onto an S-tier chart while discussing how viral it was and comparing it to the verified data in this playbook. Same confidence grades. Same sources.

It's a storyteller move because it does the thing we recommend in this chapter: take an opinionated, evidence-graded position on category content other people are publishing. Other creators have to engage with the chart – by responding, by repurposing it, by quoting it.

We measure brand citation lift for MaxShip month-over-month and attribute downloads to the YouTube end-screen CTA. If Track B beats Track A on lift, we'll say so. If Track A beats Track B, we'll say that too. Either way, the data is shared in Chapter 9.

What this is not

Astroturfing. Fake reviews. Paying for citations. Buying creators with NDA scripts. Inflated case studies. Press releases at \$80 each. We name these patterns because some of them are circulating as "Track B tactics" in AEO content. They're not. They're brand-debt accelerators that get exposed eventually and discounted by LLMs in the meantime.

DO THIS **This month:** Identify three creators whose audiences overlap your ICP. Send each a one-paragraph honest pitch – not a sponsorship deal, an offer to give them your data or your time. See who responds. The ones who do are your year-one storyteller bench.

Chapter 9 — The Track B test (run it on your own brand)

Track B is harder to measure than Track A — but not impossible. This chapter is the methodology, so you can run a clean dual-track test instead of guessing which investment paid off.

The Counter-Take format (recap)

Each episode follows the same shape:

- 8 pieces of viral AEO advice pulled from a rolling pool of ~12 sources we keep internally
- Each piece graded against the **PROVEN** / **MEASURED** / **GUESSED** system anchored in this playbook's verified-sources document
- Each piece placed on an S-F tier chart with reasoning
- Final chart shipped as a standalone shareable asset
- Episode anatomy: hook → setup → 8 graded pieces → closer → CTA

The hypothesis

Counter-Take episodes drive three measurable outcomes:

1. **Brand citation appearance** — measured via the 60-cell grid from Chapter 6
2. **Asset downloads / sign-ups** — measured via UTM-tagged CTAs and a "How did you hear about us?" question on any gate
3. **Audience growth** — measured wherever you capture it (email, follows, list)

If the Counter-Take series outperforms the Track A on-site moves on citation lift, that tells us storyteller content is the higher-ceiling investment for MaxShip. If Track A outperforms, that tells us the structural moves carry more weight at our stage. Most likely they're complementary — but we don't know yet, and the only way to know is to ship both and compare.

The honest layer

The attribution is imperfect. UTMs miss copy-paste-link traffic. The email-gate question relies on people remembering. AI citation grids have variance — the same prompt run twice can return different answers. We control for this with three sampling passes per measurement, but the noise is real.

If the data is ambiguous – if Track A and Track B move within the noise floor of each other – we'll say so. We don't owe the playbook a clean answer; we owe it an honest report.

The scoreboard to keep

Track A and Track B compete for the same budget. Keep one table so you know which one earned it:

	Track A	Track B
Baseline	[X cites]	[X cites]
+30 days	[X cites]	[X cites] [N assets shipped]
+60 days	[X cites]	[X cites] [N assets shipped]
+90 days	[X cites]	[X cites] [N assets shipped]
Track A change:	+___%	
Track B change:	+___%	
Where you double down:	_____	

Don't hide the losing track from yourself. The point of the test is to fund the winner.

DO THIS

For your own program: Run your own dual-track test. Pick one Track A move (the Princeton three on three pages) and one Track B move (a creator partnership or original-data piece). Measure citations baseline → 30 days → 60 days. Compare. That's your data, not ours.

PART 4 — RUN THE PROGRAM

Chapter 10 — Choosing your starting track + the cadence

DAILY 10 MIN — > 3 brand prompts > Skim top sub	WEEKLY 45 MIN — > Re-run 60-cell grid > 1 Track A move > 1 Track B action	MONTHLY 3 HRS — > Off-site score > Story audit > Ship 1 asset	QUARTERLY 1 DAY — > Re-read · regrade > Kill what fails > Refresh guides
--	---	---	--

The discipline that compounds. A weekly rhythm with a monthly retro and a quarterly strategy review — minute estimates so it actually gets run.

Two decisions and a calendar.

Decision 1 — Which track first

We covered this in Chapter 3, but in operational terms:

- **Solo, pre-PMF, or Seed:** Track A first. It's cheaper, faster, more in your control, and the gains are real before you've built off-site muscle. 60–90 days on Track A then add Track B.
- **Series A with a marketer and a developer:** Run both. Light Track A pass on your three top pages (one sprint), light Track B pass on your two lowest off-site categories (one quarter), measure, iterate.

- **Series B+ with a content team and a PR budget:** Both, weighted to your audience surface. Audit which AI engine your buyers actually use (run the Chapter 10 prompt set on a small panel of customers if you can), then weight your effort toward that engine's source preferences.

Decision 2 — Who owns it

AEO is a marketing program with a developer dependency. The accountable person is in marketing — usually a content lead, sometimes a head of growth. The dependencies are:

- **Dev** — ships Track A page changes, schema, render strategy, and optional `llms.txt` experiment
- **Design** — page templates for the answer page, comparison page, definitive guide
- **PR / partnerships** — Track B creator outreach and trade press
- **Customer success** — sourcing G2 / Capterra reviews from happy customers

If you're a single person, you wear all of these. If you're a team, the marketing lead runs the program; the others are dependencies on a quarterly schedule.

The cadence

Run this as a weekly rhythm with a monthly retro and a quarterly strategy review.

Daily (10 minutes):

- Run 3 brand-direct prompts on ChatGPT, Claude, Perplexity. Just type your company name and "what is X?" or "should I use X?". Screenshot anything new or surprising.
- Skim the top 2–3 posts on your top subreddit. Pattern-match for category-relevant threads.

Weekly (45 minutes):

- Re-run the 12-query × 5-surface grid (60 cells). Update the citation tracker spreadsheet.
- Pick one Track A page move from Chapter 4 and ship it to one page.
- Pick one Track B action: a substantive Reddit comment, a creator reach-out, a customer ask, or a step toward your next original-data piece.
- **Founder engine (Chapter 8), if you're running it:** the founder films one answer on the pillar (30 min). Cut 3–5 shorts + the transcript. Post YouTube long + shorts, LinkedIn, IG/TikTok, and the blog transcript page. This is the one weekly block that can't slip — it's the move that feeds both tracks.
- Send a 3-sentence update to your team: what moved, what you shipped, what's next.

Monthly (3 hours):

- Full off-site footprint score (Chapter 7 checklist).
- Story audit – read the exact words LLMs use for your brand. Look for drift (frame shift, bundling change, outdated facts surfacing).
- Ship one substantive external asset: a Wikipedia edit, a G2 update, a YouTube long-form, an original-data post.
- Review GA4 / server logs for AI referrer patterns. Tag what's working.
- Send a brief monthly retro to your CEO / leadership: lift numbers, one win, one miss, one bet.

Quarterly (1 day):

- Re-read this playbook. Update grades on your own data. Kill what's not working. Double down on what is.
- Refresh your comparison pages and definitive guides. Update freshness blocks. Re-verify crawlability, schema, and any optional `llms.txt` experiment.
- Revisit your creator partnerships and customer-led PR pipeline. Reset the bench.

The 12-prompt test-your-brand bank

Run these on ChatGPT, Claude, Perplexity, Gemini, and Google AI Overviews. Twelve queries × five engines = 60 cells. Capture: mentioned (Y/N), citation (link present?), tone (positive/neutral/negative), exact words (paste the exact words used to describe your brand).

Branded (3):

1. *"What does [your company] do?"*
2. *"Is [your company] legit?"*
3. *"Tell me about [your company]"*

Category (3):

1. *"Best [your category] for [your ICP]"*
2. *"Leading [your category] in 2026"*
3. *"Top [your category] companies"*

Comparison (3):

1. "[Your company] vs [top competitor]"
2. "Alternatives to [your company]"
3. "[Adjacent tool] vs [your company]"

The job they need done (3):

1. "How do I [the literal phrase a buyer types when stuck on the problem you solve]"
2. "What's the best way to [the second-most common phrase]"
3. "I need a tool that [the third]"

That's the kit. Run it today; you have your baseline. Run it monthly; you have a program.

DO THIS **Tomorrow morning:** Block 45 minutes. Run the weekly checklist. Then do it every Monday for the next quarter. That's the discipline that compounds.

Chapter 11 — What's next

AEO advice ages fast. Treat anything dated – including this – as a snapshot, and re-test against your own numbers every quarter. Here's where the puck is going.

What's next for AEO

We don't know exactly what's coming. Anyone who tells you they do is lying. But there are three forces shaping the next 12 to 24 months that we'd bet are real, and each one shifts what you should be doing.

1. Agentic browsing. Agents like ChatGPT Atlas, Claude Computer Use, and Gemini's agentic mode are shipping. The user no longer types a query – the agent does. Citation becomes purchase decision. If your product can be evaluated and bought without a human ever visiting your site, you need to be evaluable by a machine. Concretely: structured product data, verified merchant signals, machine-readable terms, fast checkout APIs.

2. AI-native commerce. Stripe is shipping chat-native checkout. Shopify is shipping LLM-aware product feeds. The implication: product schema markup and verified merchant data become important rather than nice-to-have. If you sell anything online, get your `Product` and `Offer` JSON-LD shipped this quarter. [GUESSED on the timing; the direction is **MEASURED**]

3. The freshness arms race. Engines are already pushing toward "answers from this week" defaults. Journalism citations rise to 49% on time-sensitive queries. The bias toward fresh content keeps increasing – a dated freshness block and a content calendar that ships *something* every month are no longer optional for category-leading pages.

The bigger shift: you now sell to two buyers

Here's the change underneath all three forces, and it's the one that matters most. Customer acquisition has moved through eras – word of mouth, then search, then social feeds. The next era is already here: **every customer arrives with an AI agent that researches, filters, and recommends on their behalf.** They won't read ten comparison pages. Their bot will, and it will hand them a shortlist of three.

So you are now marketing to *two* buyers for every one customer. The **human** still decides the way humans always have – on story, trust, identity, and a hundred irrational reasons, and you still win them with classic marketing. But the **bot** is the gatekeeper that decides whether the human ever sees you at all. The bot is not irrational. It doesn't care about your billboard. It reads structure, citations, off-site footprint, and what credible third parties say about you – every single thing in this playbook.

That's the whole point of AEO, stated plainly: classic marketing wins the human; this playbook wins the bot that decides which humans you get to talk to. Most companies are still doing only the first. The ones that win the next five years do both, deliberately, and know which buyer each move is for.

MaxShip's bet on what doesn't change

Engines will pivot. Algorithms will update. New entrants will appear. What survives all of that:

- A brand with a clear, defensible one-sentence frame
- A thick off-site footprint across the 15 sites that decide visibility
- A weekly cadence of measuring + improving
- An author entity with a real human attached to it
- The willingness to change your mind when the numbers move

The tactics in Chapters 4 through 9 will need refresh in 12 months. The disciplines in Chapter 10 won't.

Your move from here

Three things, in this order:

1. Open your laptop right now and run the test. The 12-prompt bank on the next page *is* the audit – no download, no signup, no companion product to wait for. Run it on your own brand across the five engines, score each cell, and you have an honest baseline in about twenty minutes. It's already in this PDF. Do it before you change anything.

2. Subscribe to the Captain's Log. One email a week from MaxShip – what shipped, what we learned, what's working in the field. Unsubscribe whenever you want.

3. Pick one move from Chapter 4 and ship it on one page this week. Don't redo your whole site. Pick your highest-traffic page. Add one inline quotation, one statistic with a source, one inline citation. Re-measure in 30 days. Email what you found – `max@max-ship.com`.

That's the whole playbook. Everything else is detail.

BACK MATTER

Sources & methodology

All claims in this playbook are graded **PROVEN**, **MEASURED**, or **GUESSED**. Sources for each graded claim are inline. Every URL below was fetched live on 2026-05-16 unless otherwise noted.

PROVEN – peer-reviewed, replicated, or platform-documented

- Aggarwal, Murahari, Narasimhan, Deshpande, Rajpurohit & Kalyan, "GEO: Generative Engine Optimization" – KDD 2024 (ACM, DOI 10.1145/3637528.3671900); Princeton University + IIT Delhi + independent. n = 10,000 queries (GEO-bench), replicated on Perplexity. arxiv.org/abs/2311.09735. Source for Princeton's three (Quotation +41%, Statistics +32%, Cite Sources +27%), the Keyword-Stuffing counter-finding (17.7 vs 19.3 baseline), and the Rank-5 +115% Cite Sources lift.
- Jazwińska & Chandrasekar, "AI Search Has a Citation Problem" – Tow Center for Digital Journalism, Columbia Journalism Review, 2025. 8 engines, 20 publishers, 1,600 queries. cjr.org/tow_center. Source for the "engines mis-attribute sources in >60% of queries, with false confidence" claim.
- Google Search Central, "Optimizing your website for generative AI features on Google Search" – last updated May 15, 2026; fetched May 18, 2026. developers.google.com/search/docs/fundamentals/ai-optimization-guide. Platform-documented source for Google AI Overviews / AI Mode: generative AI features are rooted in core Search ranking and quality systems; prioritize unique non-commodity content, crawlability, technical clarity, page experience, images/video, local/ecommerce details where relevant, and avoid treating `llms.txt`, chunking, AI-only rewrites, inauthentic mentions, or special schema as hacks.

MEASURED – one or more credible studies; named methodology

- Ahrefs (Linehan & Guan), "Top Brand Visibility Factors in ChatGPT, AI Mode & AI Overviews" – Dec 2025, n = 75,000 brands, Spearman across 6 engines. ahrefs.com/blog/ai-brand-visibility-correlations. Independent corroboration of "structure/off-site beats authority": YouTube mentions ~0.74 (strongest single factor), branded mentions 0.66–0.71, Domain Rating only ~0.27, backlinks weak.

- **Digital Applied, "500 SaaS Sites Audited: AI Citation Visibility 2026"** – published April 26, 2026 (audit window March–April). digitalapplied.com. Source for the 8.4× gap, DA +0.18 vs structure checklist +0.71, and the structural lift table. *Provenance note: single agency audit, no public dataset – its main claims (DA-is-weak, Reddit dominance) are backed up by the independent studies above and below; the llms.txt +24% figure is contested (see SE Ranking and Google Search Central).*
- **Semrush (Harsel, Drozdov & Skopec), "The Most-Cited Domains in AI: A 3-Month Study"** – Nov 2025, 230K prompts, >100M citations. semrush.com/blog/most-cited-domains-ai. Corroborates Reddit/Wikipedia dominance and weekly volatility (ChatGPT Reddit 60%⇒10% in six weeks).
- **Semrush AI Visibility Study** – June 2025, 150K LLM citations. semrush.com/blog/ai-search-visibility-study-findings. Independent Reddit = 40.1% figure.
- **Peec AI / Search Engine Land** – Mar 2026, n = 30M sources, 5 platforms. searchengineland.com/473138. Reddit #1, YouTube #2, LinkedIn #3 – third independent corroboration of the Track B source hierarchy.
- **Otterly.ai, "The YouTube Citation Study 2026"** – 100M+ citation instances, 6 engines. otterly.ai/blog/the-youtube-citation-study-2026. YouTube = #2 cited social source; 40.8% of cited videos had <1,000 views (*caveat: studies already-cited videos only – does not prove initial eligibility*).
- **Seer Interactive (Vasquez), "AI Brand Visibility and Content Recency"** – Jun 2025, 5,000+ cited URLs + ChatGPT log-file analysis. seerinteractive.com. Source for the freshness/recency claim.
- **Ahrefs (Linehan & Guan), "38% of AI Overview Citations Pull From The Top 10"** – Mar 2026, 863K SERPs, ~4M AIO URLs. ahrefs.com/blog/ai-overview-citations-top-10. Top-10 organic ↔ AIO overlap fell ~76% → ~38% in seven months.
- **AirOps / Search Engine Land** – Mar 2026, n = 548,534 pages, 15,000 prompts. searchengineland.com/471606. 15% retrieval-to-citation rate; sampling variance is real (supports the three-pass measurement protocol).
- **SE Ranking (Deda), "LLMs.txt: Why Brands Rely On It and Why It Doesn't Work"** – Nov 2025, n ≈ 300,000 domains, Spearman + XGBoost + SHAP. seranking.com/blog/llms-txt. **Counter-evidence:** no measurable link between llms.txt and AI citations – the larger study that downgrades Digital Applied's +24% to **GUESSED**.
- **5W, "AI Platform Citation Source Index 2026"** – PR synthesis of 680M citations across 6 studies. *Soft provenance (PR newswire, not primary research): used only where its numbers are independently corroborated above.*

GUESSED – directionally plausible, not established

- `llms.txt` lift; the founder-as-author-entity signal; the "engine penalizes scripted content" mechanism. Each is flagged inline where it appears. For Google specifically, `llms.txt` is not required for generative AI Search visibility.

A reviewer note: the GEO paper's affiliation is Princeton + IIT Delhi + independent authors – *not* Georgia Tech, despite secondary sources that miscite it.

The 12-prompt test-your-brand bank

Repeated here as a printable card. Run on ChatGPT, Claude, Perplexity, Gemini, AI Overviews.

Branded:

1. "What does [your company] do?"
2. "Is [your company] legit?"
3. "Tell me about [your company]"

Category: 4. "Best [your category] for [your ICP]" 5. "Leading [your category] in 2026" 6. "Top [your category] companies"

Comparison: 7. "[Your company] vs [top competitor]" 8. "Alternatives to [your company]" 9. "[Adjacent tool] vs [your company]"

The job they need done: 10. "How do I [the literal buyer phrase]" 11. "What's the best way to [second phrase]" 12. "I need a tool that [third phrase]"

For each cell: mentioned (Y/N), citation (link?), tone (+/0/-), exact words (paste what it said).

Glossary

- **AEO / GEO / LLMO** – Answer (or Generative, or LLM) Engine Optimization. Getting your brand recommended and cited by AI answer engines (ChatGPT, Claude, Perplexity, Gemini, Google AI Overviews). Used interchangeably.
- **AIO** – AI Overviews. Google's AI answer above the blue links.
- **Track A** – your website: everything you ship on a domain you own (structure, schema, the Princeton three, crawlable HTML, optional `llms.txt`). Fast, in your control, capped ceiling.
- **Track B** – everywhere else AI reads about you: what you make off-site (founder video, your channels) and what you earn (Reddit, Wikipedia, creators, press). Slower, less control, much higher

ceiling.

- **Princeton's three** – the three on-page moves the GEO/KDD-2024 study proved move citations most: Quotation Addition (+41%), Statistics Addition (+32%), Cite Sources (+27%).
- **Answer-format H2** – a heading written as the literal question a buyer asks ("What is X?" / "How do I Y?"), with a direct 50–80-word answer underneath. The unit engines extract.
- **Schema / JSON-LD** – structured data in a page's code that tells machines what it is (`Article` , `FAQPage` , `SoftwareApplication` , `Product`). Plausible signal; not a proven citation lever on its own.
- **llms.txt** – proposed file at `yoursite.com/llms.txt` summarizing a site for LLMs, like `robots.txt` for crawlers. Contested – graded **GUESSED** in this playbook, and not required by Google for generative AI Search features.
- **Author entity** – a named, real human attached to content via a bio page and consistent presence (LinkedIn, Wikipedia). A trust signal – our best bet, but **GUESSED**, not proven.
- **Confidence grades** – **PROVEN** (peer-reviewed/replicated/platform-documented), **MEASURED** (one credible study or our own data), **GUESSED** (plausible, unverified). Every claim in this playbook carries one. Not a hedge – a weapon.
- **Citation visibility** – how much of your page actually makes it into the AI's answer. The Princeton metric (they call it Position-Adjusted Word Count).
- **Retrieval vs citation** – an engine *retrieving* your page isn't the same as *citing* it. Only ~15% of retrieved pages get cited (AirOps). The gap is the work.
- **Query fan-out** – engines split one user question into many sub-questions and answer from whichever pages match each. Why "answer the literal question" works.
- **Counter-Take** – our own storyteller move: an opinionated, evidence-graded take on category content others are publishing, made to be engaged with and quoted.
- **Founder-led content** – the founder filming one honest answer a week to a real buyer question, cut into shorts and a blog transcript. The one Track B surface you fully own.
- **Freshness** – recency as a ranking signal. AI-cited content runs ~26% fresher (Ahrefs, 17M citations); news citations rise to 49% on time-sensitive queries.
- **Domain Authority / Domain Rating** – third-party 0–100 link-strength scores (Moz / Ahrefs). The old SEO currency. Correlates only ~+0.18–0.27 with AI citation – weak.
- **Share of Voice (AI)** – your brand's percentage of total brand mentions across a fixed prompt set on a given engine.

- **Marketing to the bot** – the coming shift: every customer arrives with an AI agent that researches and recommends on their behalf. You now sell to two buyers – the human and their bot. See Chapter 11.
 - **Top / bottom quartile** – the top or bottom 25% of the 500 SaaS sites in Digital Applied's audit.
-

About MaxShip + Captain's Log

MaxShip is a venture studio and a fractional marketing service – and this playbook is the field manual. We run these plays on real companies, including our own, and report what actually moved. No theory we haven't shipped.

The **Captain's Log** is the newsletter. One email a week, free: my field notes on AI marketing, the systems we're running, and what I'm seeing from the road while I build (this month, from China). Unsubscribe anytime.

- Site & newsletter: max-ship.com
- Questions or what you found running the test: `max@max-ship.com`

If you found this playbook useful, the most helpful thing you can do is send it to one other founder or marketer. Strong, graded opinions only spread if enough people demand them over the slop.

SEO is dead. How to get recommended by AI – The 2026 AEO Playbook · Written by Max Palmer ·
Published by MaxShip Marketing · [Captain's Log](#)