

A WORKING BOOK

Answering Machines

How to be the answer when your customer stops searching and starts asking. A plain-English guide to AEO, GEO and AI visibility — built on what has actually been measured.

Published by Acclaira · August 2026

Every figure in this book carries its source and a confidence rating.

Nothing here is invented. Where the evidence is thin, it says so.

CONTENTS

What is in this book

Twenty-three chapters, ordered by how much each one will change your results — not by category. If you read only the first nine, you will have the majority of the value.

FRONT

- How to read this book (and how to trust it)

PART ONE • THE MACHINE

- 1 What actually happens in the eight seconds after someone asks
- 2 The five engines, and who reads what
- 3 Passages, not pages: the shift that explains everything else
- 4 The new economics: fewer clicks, better clicks

PART TWO • THE BIG LEVERS, IN ORDER

- 5 Lever 1 — Be talked about somewhere other than your own website
- 6 Lever 2 — Be readable by a machine that will not wait
- 7 Lever 3 — Be in the index the engine actually queries
- 8 Lever 4 — Answer first, elaborate second
- 9 Lever 5 — Evidence beats adjectives
- 10 Lever 6 — Own a number nobody else has
- 11 Lever 7 — Be one clear entity, everywhere
- 12 Lever 8 — Live where the engines already trust

PART THREE • THE CRAFT

- 13 The page types AI actually cites
- 14 How to write a paragraph a machine can quote
- 15 Freshness is a habit, not a date field
- 16 Local businesses and service businesses
- 17 Playing each engine differently

PART FOUR • MEASUREMENT AND CONTROL

- 18 Measuring without fooling yourself
- 19 When the AI is wrong about you
- 20 What does not work — and why it keeps being sold
- 21 How to read an AEO statistic

PART FIVE • DOING IT

- 22 The first ninety days
- 23 What is coming next

APPENDICES

- A Glossary — every term, in one line
- B The evidence table — every number, sourced and rated
- C The one-page checklist

How to read this book (and how to trust it)

This field is two years old and full of confident numbers, most of which are marketing. Before anything else, here is the rule I applied to every figure in this book.

AEO is being sold hard. Because it is new, almost nobody checks the sources, and a statistic invented in a blog post in March is quoted as fact by forty other blog posts by June. I have read a great deal of that material to write this, and I want to be straight with you about what survived.

Every number in this book carries a tag. When you see one, it means:

TAG	WHAT IT MEANS	HOW TO TREAT IT
TIER A	Peer-reviewed research, a platform's own documentation, or a study with a published method and a large sample that I could inspect.	Act on it.
TIER B	A named vendor's study. Real data, real sample, but the vendor sells a product that the finding happens to support.	Use the direction, distrust the decimal point.
TIER C	Widely repeated, no traceable method. I have included a few of these <i>only</i> to warn you off them.	Do not quote it. Ever.

Appendix B lists every figure in the book with its source and tag, so you can check me. If you ever intend to put one of these numbers in front of a client, use that table, not your memory of this chapter.

A WARNING ABOUT ROUND NUMBERS

Three claims circulate constantly in this industry that are simply false: that AI answers are "more trusted than Google search results", that a fixed number of brand mentions guarantees inclusion, and that any tool has access to a model's internal ranking. Google states plainly in its own documentation that **no third-party tool has access to its internal ranking or AI systems.** TIER A If a vendor tells you otherwise, that is the end of the meeting.

How the chapters are built

Each chapter follows the same shape, so you can skim it or study it:

- **An analogy** — something from the physical world that behaves the same way.
- **The mechanism** — what is actually happening inside the machine.
- **What to do** — concrete, in order.
- **How to check** — how you know it worked, without buying anything.
- **If you only do one thing** — the single highest-value action in that chapter.

Three words you will see constantly

The industry has not settled on a name, which tells you how young it is. In practice they mean the same job:

- **AEO** — Answer Engine Optimisation. Being the answer.
- **GEO** — Generative Engine Optimisation. Same thing, named after the 2024 academic paper that coined it.
- **AI visibility** / **AI search** — the plain-English versions, and the ones I would use with a client.

Google's own position, published in 2026, is that all of this "is still SEO" and uses the same ranking systems. TIER A That is broadly true and worth holding on to. It is also incomplete, and Chapter 3 explains exactly where it stops being true.

PART ONE

The Machine

Four chapters on what is physically happening when someone asks an AI a question. Understand this and roughly eighty per cent of the advice in the rest of the book stops being a rule you memorise and starts being something you could have worked out yourself.

What actually happens in the eight seconds after someone asks

People assume the AI "knows" the answer from its training. For anything that matters commercially, it almost never does. It goes and looks — and the looking is where you win or lose.

THE ANALOGY

Imagine you send a very fast, very literal research assistant into a library. You ask: *"Who should I use to clear customs at Port Klang?"*

The assistant does not read your question and recite from memory. They do four things. They **break your question into smaller questions** — who operates there, what does clearance cost, who is licensed, who do people rate. They **run to the shelves** and pull down whatever the catalogue points them at. They **photocopy a few paragraphs** — not whole books, paragraphs. Then they come back and **write you one page** in their own words, with little footnotes saying where each bit came from.

Your job, as a business, is not to be a good book. It is to be a paragraph that gets photocopied.

The mechanism, step by step

Nearly every AI answer that cites sources is produced by a pattern called **retrieval-augmented generation**, or RAG. Google confirms in its own developer documentation that its AI features work this way — they ground responses in indexed search results rather than in the model's memory. TIER A Here is the sequence.

1 Decide whether to search at all

A classifier inside the product decides if this question needs live information. "What is 12 times 8" does not. "Best freight forwarder in Klang" does.

2 Fan the question out

Your one question is rewritten into a set of narrower ones. Google calls this *query fan-out*. Reported estimates put it in the range of eight to twelve sub-queries for a typical AI Mode question. TIER B You are not competing for one query any more. You are competing for a dozen you never saw.

3 Retrieve candidate passages

Each sub-query runs against an index. What comes back is not a list of ten websites — it is a pile of *passages*, pulled from many pages, scored for how well they match that specific sub-question.

4 Rank and filter

The passages are re-scored for quality, authority and agreement with each other. Perplexity, which has published more about its stack than most, runs candidates through several reranking layers and will discard the whole set and re-query rather than answer from weak sources. **TIER B**

5 Synthesise and attribute

The model writes one short answer constrained to the evidence it was handed, attaching citations to individual claims. In the Pew Research study of real Google usage, the median AI summary was

67 words

and 88% of them cited three or more sources. **TIER A**

You are not trying to rank. You are trying to be one of roughly three paragraphs that survive a filter you cannot see, for a question that was never typed.

Why this reframes everything

Three consequences fall straight out of the mechanism, and they drive most of this book.

First: the unit of competition is the passage, not the page. A 400-word answer on a small site can beat a 5,000-word guide from a large one, if that specific paragraph is a sharper match for a specific sub-question. This is genuinely new, and it is the best news small businesses have had from search in fifteen years. Chapter 3 is entirely about this.

Second: you are being judged on agreement, not assertion. The engine pulls passages from many places at once. If your site says you are the leading customs broker in Selangor and nothing else on the internet says anything about you, you are a single unsupported claim. If four other sources describe what you do in roughly the same terms, you become a fact. This is why Chapter 5 — off-site mentions — is the biggest lever in the book, ahead of anything you can do to your own website.

Third: retrieval happens under time pressure. These systems fetch at enormous scale with tight timeouts. They do not wait for slow pages and they do not run your JavaScript. Chapter 6 covers this and it is the single most common way a good business is invisible for a purely technical reason.

HOW TO CHECK — FIVE MINUTES, NO TOOLS

Open ChatGPT, Gemini and Perplexity side by side. Ask each the exact question your best customer would ask before they knew your name — not your company name, the *problem*. Note three things: are you mentioned at all; who *is* mentioned; and which websites are cited underneath. That citation list is your real competitor set. It is usually not the list you expected.

IF YOU ONLY DO ONE THING

Write down the ten questions a customer asks before they know you exist, and run all ten through two AI tools this week. Everything else in this book is a response to what you find.

The five engines, and who reads what

"AI search" is not one thing. Five systems dominate, they read different indexes, and being invisible in one tells you nothing about the others. This chapter is the map.

THE ANALOGY

Five different newspapers cover your town. They all print stories about local business, but one only reads the council register, one only reads the chamber of commerce newsletter, one sends its own reporter, and two share a wire service. Getting into the council register does nothing for the paper that reads the newsletter. You need to know which door each one uses.

The map

ENGINE	WHERE IT GETS PAGES	ITS CRAWLER	WHAT THIS MEANS FOR YOU
ChatGPT Search + browsing	Primarily Bing's index, supplemented by OpenAI's own crawling	OAI-SearchBot (search) GPTBot (training) ChatGPT-User (live fetch)	If you are not in Bing , you are largely invisible here regardless of your Google rankings. Bing Webmaster Tools is not optional.
Google AI Overviews The summary box	Google's main index and ranking systems	Googlebot	Standard Google SEO is the entry ticket. Google says explicitly there are no additional requirements. TIER A
Google AI Mode The conversational tab	Same index, but heavy query fan-out and passage-level retrieval	Googlebot	Same ticket, different game. Ranking #1 for the head term does not guarantee you serve any of the sub-queries.
Perplexity	Its own index, rebuilt in-house on Vespa in 2025, plus licensed third-party crawls	PerplexityBot	An independent front door. You can be strong here and weak elsewhere, and vice versa.
Claude & Copilot	Claude uses a search partner; Copilot uses Bing	ClaudeBot , Claude- SearchBot	Copilot inherits your Bing position. Claude is worth spot-checking because it skews toward documentation and primary sources.

Sources: Google Search Central AI features documentation; Vercel/MERJ crawler telemetry; Perplexity engineering statements; OpenAI crawler documentation. See Appendix B.

The two-word distinction that matters most

Notice that OpenAI runs several bots, and they do different jobs. This trips up almost everyone, including people who should know better, so it is worth being precise:

- **Training crawlers** — `GPTBot`, `Google-Extended`, `Applebot-Extended`. These gather text that may be used to train future models. Blocking them is a defensible business decision and it does *not* remove you from AI search results.
- **Search crawlers** — `OAI-SearchBot`, `PerplexityBot`, `Claude-SearchBot`. These gather pages to *cite in answers*. Blocking these removes you from those answers.
- **User-triggered fetchers** — `ChatGPT-User`. Fires when a person or an agent asks the tool to go and look at a specific page, live.

THE MOST EXPENSIVE MISTAKE IN THIS CHAPTER

In 2023 and 2024 a great many businesses pasted a "block the AI bots" snippet into `robots.txt` to protect their content from training. A large share of them blocked the *search* bots at the same time, and then spent 2026 wondering why they never appear in ChatGPT. Go and read your `robots.txt` today. It takes ninety seconds and it is the highest-return ninety seconds in this book.

Where they overlap, and where they do not

Two practical points follow from the map.

The first is that **Bing is now load-bearing infrastructure** in a way it has not been for a decade. Not because anyone searches on Bing, but because ChatGPT and Copilot retrieve through it. Verifying your site in Bing Webmaster Tools, submitting a sitemap, and pushing updates through IndexNow costs an hour and unlocks two of the largest answer engines. **TIER B**

The second is that **results diverge sharply between engines**, and you should expect that rather than treat it as an error. The same prompt run on ChatGPT, Gemini and Perplexity will often return three substantially different shortlists. Chapter 17 covers what to do about it; for now, simply stop assuming one check is a measurement.

HOW TO CHECK

- Visit `yourdomain.com/robots.txt` and read every `Disallow` line. Confirm no search-class bot is blocked.
- Search Bing for `site:yourdomain.com`. If the count is far below your Google equivalent, that is your ChatGPT problem in one number.
- Ask your developer, or check your server logs, for hits from `OAI-SearchBot` and `PerplexityBot` in the last 30 days. Zero hits is a finding.

IF YOU ONLY DO ONE THING

Verify your site in Bing Webmaster Tools and submit your sitemap. It is free, it takes under an hour, and for most businesses it is the difference between existing and not existing inside ChatGPT.

Passages, not pages: the shift that explains everything else

Traditional search ranked documents. Answer engines retrieve fragments. Almost every practical difference between SEO and AEO comes from this one change.

THE ANALOGY

A librarian used to hand you a book and say "the answer is in here somewhere." That was Google in 2015: ten books, ranked.

The new librarian hands you a single index card with four sentences on it, each copied from a different book, each with the book's name written underneath. Your book being excellent overall is now almost irrelevant. What matters is whether any single passage inside it was worth copying onto the card.

Long, meandering chapters that build to a point are the enemy. The passage has to make sense *lifted out of the book entirely*, because that is exactly what happens to it.

What "self-contained" really means

When a retrieval system scores your content, it does so on a fragment — typically a section under a heading, or a few hundred words. That fragment arrives at the model with no surrounding context. If your paragraph says *"This makes it the most cost-effective of the three options"*, the machine has no idea what "this" is, what the three options were, or what you are talking about. It is unusable, and it will not be used.

X Not retrievable. "As we saw above, this approach tends to be significantly cheaper for most importers, particularly given the considerations discussed in the previous section."

✓ Retrievable. "Consolidating LCL shipments through a single forwarder typically costs importers 15–30% less than booking each shipment separately, because the forwarder buys container space in bulk."

The second sentence survives being torn out of the page. That is the whole test.

The honest nuance — and it matters

You will be told by a lot of agencies that you must "chunk your content for AI." Google's official documentation says the opposite in plain words: **"There's no requirement to break your content into tiny pieces"**, and that its systems understand pages covering several topics and can extract the relevant part themselves. TIER A

Both things are true, and here is how they fit together, because this is the kind of nuance that separates someone who understands the field from someone repeating it:

- **Chunking is not a requirement.** You will not be excluded for writing long, flowing prose. Google will find the relevant part. Do not let anyone sell you a "chunking service."
- **Clarity is still a tiebreaker.** When two passages answer a sub-query equally well, the one that reads cleanly as a standalone statement is more useful to a model that has to quote it in 67 words. You are not unlocking a gate; you are improving your odds at the margin.
- **The real win is for humans anyway.** Clear headings and answer-first paragraphs make your page better for readers. That is a benefit you keep no matter what the engines do next.

So: write for people, in complete thoughts, with descriptive headings. That is the entire technique. It is not a new discipline — it is good editing, which has always been rare.

Why this favours the small business

Under document ranking, a large site with fifteen years of backlinks beat you on almost every term, structurally. Under passage retrieval, each sub-question is scored somewhat independently, and a sharp, specific, first-hand paragraph from a small operator can be the best available match for a narrow sub-question that the big site answers only vaguely.

The original GEO research found exactly this pattern experimentally: sites ranked lower in the results benefited *substantially more* from these techniques than sites already at the top. In one condition, adding source citations lifted visibility for a fifth-ranked site by 115%. TIER A The authors described the effect as having the potential to democratise the space. That is an unusually strong claim for an academic paper and it has held up reasonably well.

This does not mean authority is dead. Domain-level trust still influences which passages get surfaced, and Chapter 5 is about exactly that. It means the gap is crossable in a way it was not before.

HOW TO CHECK

Take your most important page. Copy any single paragraph out of the middle of it into a blank document, and read it cold. Does it state a complete, useful fact on its own? Now do it with five more paragraphs. Most business websites fail this on four out of six. That is your editing list.

IF YOU ONLY DO ONE THING

Rewrite the first paragraph under every heading on your top five pages so that it answers the heading directly, in one or two sentences, before any context or build-up.

The new economics: fewer clicks, better clicks

If you measure this work in traffic, you will conclude it is failing. The traffic genuinely is falling. The reason to do this anyway is that what remains has changed character entirely.

THE ANALOGY

You run a shop on a busy street. The council builds a covered walkway that lets people see your window display without coming inside. Footfall drops by half. You panic.

Then you look at the till. The people who *do* come in now have already seen the display, already decided, and already know roughly what they want. Your conversion rate has doubled and your average sale is up.

You have not lost customers. You have lost browsers, and been handed buyers. But only if your window display was any good — and the walkway shows your competitor's window too.

The bad news, measured properly

The most credible measurement here is Pew Research Center's study of real browsing behaviour: 900 US adults who agreed to have their actual URL history collected, covering 68,879 Google searches in March 2025, of which 12,593 produced an AI summary. That is observed behaviour, not a survey. TIER A

WHAT PEW FOUND	FIGURE
Searches where a user clicked a traditional result, <i>without</i> an AI summary	15%
Searches where a user clicked a traditional result, <i>with</i> an AI summary	8%
Searches where the user clicked a source cited <i>inside</i> the AI summary	1%
Share of all Google searches that produced an AI summary (March 2025)	18%
AI summaries citing three or more sources	88%

Read the third row again. **One per cent.** Being cited in an AI Overview is, in click terms, worth very little on its own. Anyone selling you AEO on the promise of citation traffic is selling you the wrong thing.

The trend since has continued in the same direction. Ahrefs' measurement of click-through-rate loss for the top organic result moved from roughly 34.5% in April 2025 to roughly 58% by February 2026. TIER B SparkToro's clickstream work puts the US zero-click share at 68% in early 2026, up from around 60% in 2024. TIER B Directionally these agree with Pew, which is the main reason I trust them.

The good news, also measured

Two things are true at once. Volume is down. Value per visit is up, substantially.

Semrush's clickstream analysis found AI-search visitors worth roughly **4.4 times** the average traditional organic visitor, and around three times as likely to convert. **TIER B** The mechanism is obvious once stated: someone arriving from an AI answer has already had the basic question resolved. They are not researching what a customs broker does. They are clicking through to check whether *you* are the one. They arrive later in the journey and closer to a decision.

Meanwhile the absolute volume is still small and you should not be told otherwise. ChatGPT's referral share of global website traffic was measured at around **0.32%** in May 2026 **TIER B** — growing fast, from a very low base. Anyone projecting that AI referrals will replace your organic traffic this year is guessing.

The honest position: AI search will not save your traffic numbers. It decides whether you are on the shortlist. Those are different businesses, and the second one is worth more.

What you should actually be optimising for

This changes the goal, and therefore the scoreboard. Three targets, in order of how much they matter:

- 1 Inclusion in the recommendation**
Are you named when someone asks for options in your category? This is the whole game. It produces no traffic and no analytics event, and it is still the most valuable thing on this list.
- 2 Accuracy of what is said about you**
Being named with the wrong services, wrong location or wrong price is worse than not being named. Chapter 19.
- 3 Referral quality, not referral quantity**
Track conversion rate and value of AI-sourced sessions separately. If you judge them by volume you will kill a channel that is outperforming per visit.

For a B2B or considered-purchase business, there is a further finding worth knowing. G2's 2026 buyer research reports that around half of B2B software buyers now begin research with an AI chatbot, that 69% ended up choosing a different vendor than they had originally planned, and that a third bought from a vendor they had never heard of before. **TIER B** Treat the decimals with suspicion; treat the direction seriously. The shortlist is being formed somewhere you cannot see, before anyone contacts you.

HOW TO CHECK

In your analytics, build a segment for referrals from `chatgpt.com`, `perplexity.ai`, `gemini.google.com` and `copilot.microsoft.com`. Compare conversion rate and average value against organic — not sessions against organic. If it is a tiny, high-converting trickle, that is the expected, healthy result. Chapter 18 explains why the session count will still understate reality.

IF YOU ONLY DO ONE THING

Change the question you ask about this work. Not "how much traffic did it bring" but "when a stranger asks the AI for someone like us, are we on the list?" Everything else follows from asking the right one.

PART TWO

The Big Levers, In Order

Eight chapters, deliberately ranked. Chapter 5 will move your results more than Chapter 12, and both will move them more than anything in Part Three. If your budget or your attention runs out halfway, run out at the bottom of this list rather than the top.

Be talked about somewhere other than your own website

This is the biggest lever, and it is not on your website. The strongest measured correlate of AI visibility is how often your brand is mentioned across the rest of the internet — ahead of backlinks, ahead of domain authority, ahead of everything you control directly.

THE ANALOGY

You are hiring, and two candidates hand you a CV. Both CVs claim excellence. One candidate also has six people in your industry who, unprompted, describe their work in similar terms.

You do not believe the CV. You believe the corroboration. Not because the CV is lying, but because a claim that only exists in one place — the place with an interest in making it — is not evidence of anything.

An answer engine assembles a picture of you from many sources at once. Your website is the CV. Everything else is the reference check. Most businesses spend ninety per cent of their marketing effort polishing the CV.

The evidence

Ahrefs studied 75,000 brands, measuring which signals correlate with being mentioned in Google's AI Overviews. They used Spearman rank correlation and published the full table. This is the single most useful dataset in the field. **TIER A**

SIGNAL	CORRELATION	READ IT AS
Branded web mentions	0.664	How often your name appears anywhere online
Branded anchor text	0.527	Links that use your actual name as the link text
Branded search volume	0.392	How many people search for you by name
Domain Rating	0.326	Classic authority score
Referring domains	0.295	How many sites link to you
Branded traffic	0.274	Visits from people who searched your name
Total backlinks	0.218	The thing SEO spent twenty years on
URL rating	0.180	Page-level authority
Number of pages on site	0.170	Publishing volume

Ahrefs, "An Analysis of AI Overview Brand Visibility Factors", 75,000 brands, Spearman correlation. A companion Ahrefs report found YouTube mentions – brand name appearing in a video title, description or transcript – correlating at approximately 0.737, the highest single figure they published. **TIER B**

The headline: **brand mentions correlate roughly three times more strongly with AI visibility than backlinks do.** For twenty years, an unlinked mention of your company was considered worthless by SEO orthodoxy — a wasted opportunity, something to chase down and convert into a link. That orthodoxy is now backwards.

READ THE CAVEAT, BECAUSE IT IS LOAD-BEARING

Ahrefs themselves say it: **correlation is not causation**, and they describe these as moderate-to-weak relationships. Big, well-known brands get mentioned a lot *and* appear in AI answers a lot, and both may be caused by simply being a big, well-known brand. Do not treat 0.664 as a dial you can turn.

What makes the finding actionable anyway is that it agrees with the mechanism from Chapter 1. Retrieval systems assemble answers from multiple sources and reward agreement between them. A model with no corroborating evidence about you has nothing to be confident about. The correlation and the mechanism point the same way, which is the strongest form of evidence available in a field this young.

What to do about it

The instruction "get mentioned more" is useless without a method. Here is the method, in rough order of return for a small or mid-sized business.

1 Claim and complete every directory and profile that exists in your category

Industry associations, chambers, trade bodies, licensing registers, local business directories, LinkedIn company page, Crunchbase, Google Business Profile, Apple Business Connect. These are boring, free, permanent, and they are the cheapest corroboration you will ever buy. Do all of them before you do anything clever.

2 Get onto the review and comparison platforms your buyers use

For software: G2, Capterra, TrustRadius. For services: Google reviews, industry-specific directories. Review platforms appeared in roughly 34.5% of commercial AI Overviews in one vendor analysis **TIER B** — and structurally it makes sense, because they are the only third-party source with organised, current, category-level data on vendors.

3 Be a source, not an advertiser

Journalist-request platforms, trade press, podcast guesting, industry panels, association newsletters. The objective is a named mention in a piece of content someone else published. You do not need a link. You need your name, in context, next to your category.

4 Show up on YouTube — even badly

Given the strength of the YouTube signal, and given that AI engines can read titles, descriptions and transcripts, a modest channel of genuinely useful videos is disproportionately valuable. This is not about production quality or subscriber counts. It is about your brand name existing inside a transcript on a platform the engines trust.

Participate where your industry actually talks

Reddit, specialist forums, LinkedIn. Chapter 12 covers this properly, including how to do it without getting banned and without embarrassing yourself.

DO NOT BUY MENTIONS

Google's own guidance names this explicitly as a myth to avoid: seeking **inauthentic mentions** across the web does not work, and their systems are built to filter manipulation. **TIER A** Paid mention networks, spun press releases pushed to two hundred syndication sites, and AI-generated "citations" on sites nobody reads are the 2026 equivalent of buying links in 2011. They failed then. They are failing faster now, because agreement between low-quality sources is exactly the pattern these systems are tuned to discount.

HOW TO CHECK

- Search Google for your brand name in quotes, minus your own domain: "Your Company" - site:yourdomain.com . Count the genuinely independent results on the first three pages. That number is your starting position.
- Do the same for your two strongest competitors. The gap is your work list.
- Ask an AI directly: "What do you know about [company name]?" The length, accuracy and confidence of the answer is a rough proxy for how much corroboration exists.

IF YOU ONLY DO ONE THING

Spend one working day claiming and fully completing every free profile, directory and register that applies to your business, with identical wording for your name, address and description on each. It is dull, it is free, and it is the highest-return day in this book.

Be readable by a machine that will not wait

The most common reason a good business is invisible to AI has nothing to do with content. Their site is built so that the words only appear after JavaScript runs — and the major AI crawlers do not run JavaScript.

THE ANALOGY

You send a courier to collect a document. They have ninety seconds and a hundred other stops. They will take whatever is sitting on the reception desk.

Your document is in a safe. The safe opens automatically, but it takes four seconds and requires the building's electricity. Google's courier has a key and is willing to wait. The AI couriers are not. They look at the desk, see nothing, and leave.

Your document is perfect. It has never been collected.

The evidence, and it is unusually clear

Vercel and MERJ instrumented real crawler traffic across their network and published the telemetry. This is not a survey or an inference — it is a log of what the bots actually did. TIER A

CRAWLER	MONTHLY FETCHES OBSERVED	EXECUTES JAVASCRIPT?
Googlebot	4.5 billion	Yes
GPTBot (OpenAI)	569 million	No
ClaudeBot (Anthropic)	370 million	No
AppleBot	314 million	Yes
PerplexityBot	24.4 million	No

Vercel / MERJ, "The rise of the AI crawler". The study found zero evidence of JavaScript execution by GPTBot across its sample. The bots do sometimes *download* JavaScript files — around 11.5% of ChatGPT's fetches and 23.8% of Claude's — but downloading is not executing.

This is not an oversight anyone is racing to fix. Rendering JavaScript at the scale of hundreds of millions of fetches a month is expensive in compute and slow. Skipping it keeps retrieval fast and cheap. Treat it as a design decision that will hold for the foreseeable future, not a temporary gap.

If your product descriptions, pricing, FAQs or service pages only exist after JavaScript runs, they do not exist for ChatGPT, Claude or Perplexity. Not "rank poorly". Do not exist.

Who this hits

You are at risk if your site is a single-page application, or built on a JavaScript framework in client-rendered mode, or if key content loads through tabs, accordions, "load more" buttons, or third-party widgets. Common examples:

- A React or Vue site rendered entirely in the browser.
- Pricing tables pulled from an API after page load.
- Reviews or testimonials injected by an external widget.
- FAQ answers hidden inside accordions that only populate on click.
- Product specifications loaded into a tab that is empty until selected.

You are probably fine if your site is WordPress with a conventional theme, plain HTML, Shopify's standard storefront, or a static-site generator. "Probably" is doing work in that sentence — check anyway, because plugins and widgets reintroduce the problem quietly.

HOW TO CHECK — THE DEFINITIVE TEST, FREE, TWO MINUTES

In your browser, open your most important page and press **Ctrl+U** (Cmd+Option+U on Mac). This shows the raw HTML exactly as a non-rendering crawler receives it — *not* the inspector, which shows the page after JavaScript has run. That distinction is the whole test.

Now press Ctrl+F in that view and search for a distinctive sentence from your page body. If you find it, you are fine. If the raw HTML is mostly `<div id="root"></div>` and a pile of script tags, your content is invisible to three of the five engines.

Repeat for: your homepage, your main service page, your pricing page, and your contact page.

What to do if you fail the test

In order of preference:

- 1 Server-side render or pre-render**

Every serious framework supports it — Next.js, Nuxt, SvelteKit, Astro. The content is generated into the HTML before it is sent. This is the correct fix and it also makes the site faster for humans.
- 2 Move critical facts into the server-rendered HTML even if the rest stays dynamic**

You do not need to rebuild everything. Prices, service descriptions, locations, opening hours and FAQ answers are the content that gets retrieved. Get those into the initial HTML and you have captured most of the value.

3

Un-hide accordion and tab content

Content inside a collapsed accordion is usually in the HTML and simply hidden with CSS — which is fine. Content that is *fetches* when you click is not. Know which kind you have.

4

Publish a plain HTML version of your most important content

A pragmatic stopgap if a rebuild is months away. Not elegant, entirely effective.

THE SECOND FINDING FROM THAT STUDY, WHICH NOBODY TALKS ABOUT

Around **34.8%** of ChatGPT's crawler fetches and **34.2%** of Claude's landed on 404 pages, against 8.2% for Googlebot. **TIER A** These crawlers are working from stale or guessed URL lists and burning a third of their effort on pages that do not exist. Two implications: keep your sitemap accurate and current, and always redirect old URLs properly instead of letting them 404. You are cleaning up someone else's mess, but it is your visibility.

IF YOU ONLY DO ONE THING

Press Ctrl+U on your four most important pages today and search the raw source for your own sentences. If they are not there, nothing else in this book will help until they are.

Be in the index the engine actually queries

Readable is not the same as reachable. This chapter is the plumbing: robots files, sitemaps, Bing, and the security layer that may be blocking the bots without anyone telling you.

THE ANALOGY

Your shop is beautiful, well stocked and open. It is also not in the phone book, the front gate has a padlock the delivery driver cannot open, and your building manager has quietly instructed security to turn away anyone in an unfamiliar uniform.

None of these are shop problems. All of them are why nobody comes.

The four gates, in the order they fail

GATE ONE — ROBOTS.TXT

Covered in Chapter 2, repeated here because it is the most common single failure. Open `yourdomain.com/robots.txt`. You are looking for anything that blocks `OAI-SearchBot`, `PerplexityBot`, `Claude-SearchBot`, or a blanket `User-agent: *` with a broad `Disallow`.

A reasonable default for a business that wants AI visibility but does not want to donate to model training:

✓ **Allow the search bots** — `OAI-SearchBot`, `PerplexityBot`, `Claude-SearchBot`, `Googlebot`, `Bingbot`. These are how you get cited.

✗ **Block the training bots if you wish** — `GPTBot`, `Google-Extended`, `Applebot-Extended`, `CCBot`. This is a genuine choice with no direct visibility cost.

You will see claims that blocking any given bot costs a precise percentage of citations. Those numbers are invented. The honest statement is: blocking a search bot removes you from that engine; blocking a training bot does not.

GATE TWO — YOUR FIREWALL AND CDN

This is the invisible one. Cloudflare, AWS WAF, Sucuri and similar services block traffic they consider non-human, and AI crawlers frequently qualify. Your `robots.txt` can be perfect while your edge quietly returns 403 to every bot that matters. Nothing in your analytics will tell you.

The landscape here is also shifting under you commercially. Cloudflare moved through 2026 from a pay-per-crawl model toward paying based on whether content was actually used in an answer, and announced that from **15 September 2026** its default settings would block mixed-use crawlers from pages hosting ads unless the site owner changes the setting. **TIER B** The specifics will keep moving. The lesson will not: *somebody other than you may be deciding whether AI can read your website, and the default may not be what you want.*

GATE THREE – BING

Because ChatGPT and Copilot retrieve substantially through Bing's index, a site that is well-indexed in Google and poorly indexed in Bing is invisible to a large share of AI users. This is the single most under-done piece of technical work in the field, because Bing has been irrelevant for so long that nobody checks it.

- Verify the site in **Bing Webmaster Tools**. You can import directly from Google Search Console, which takes minutes.
- Submit your sitemap.
- Turn on **IndexNow** — an open protocol that pushes new and changed URLs to Bing instantly rather than waiting for a crawl. Most CMS platforms have a plugin. It is free.
- Confirm **Bingbot** is not blocked at the firewall. It very often is.

One honest caveat: IndexNow tells Bing a URL changed. It does not guarantee the page is indexed or surfaced. It removes a delay; it does not manufacture relevance.

GATE FOUR – SITEMAP AND URL HYGIENE

Given that roughly a third of AI-crawler fetches hit 404s, an accurate sitemap is worth more now than it was two years ago. Practical rules:

- Your sitemap should list only live, indexable, canonical URLs. No redirects, no 404s, no **noindex** pages.
- When you move a page, 301 it. Permanently. Do not delete it.
- Keep **lastmod** dates truthful. Falsifying them across the whole site to fake freshness is a well-known and easily detected pattern.

HOW TO CHECK – THE THIRTY-MINUTE AUDIT

1. Read `robots.txt` line by line.
2. Compare `site:yourdomain.com` counts on Google and Bing. A large gap is a finding.
3. Ask whoever manages your hosting to grep the server logs for `OAI-SearchBot`, `PerplexityBot` and `ClaudeBot` over the last 30 days. Zero visits means a gate is shut.
4. Open your sitemap and click ten URLs at random. Count how many are not a live 200 page.
5. Check your CDN or WAF bot-management settings for anything set to "block AI bots" or "block unknown bots".

IF YOU ONLY DO ONE THING

Get someone to look at your raw server logs for AI crawler hits. It is the only way to know for certain that the machine is being let in, and it answers in one query what a month of guessing will not.

Answer first, elaborate second

Most business writing builds towards its point. Retrieval systems take the first useful thing they find. Inverting that habit is the cheapest content change available and it improves the page for humans too.

THE ANALOGY

A journalist writes the story so that if the editor cuts it anywhere, what remains still makes sense. Headline, then the fact, then the detail, then the background. They call it the inverted pyramid and it exists because column space used to get cut without warning.

An answer engine cuts without warning. It takes one paragraph and discards the other forty. Write like the cut is coming, because it is.

The pattern

For every question your content addresses, the structure is the same three moves:

- 1 **A heading that is the question, or names the thing precisely**
 "How much does customs clearance cost in Port Klang?" beats "Our Approach to Pricing". The heading defines the topical boundary of the passage underneath it.
- 2 **A direct answer in the first one or two sentences**
 Complete, specific, and comprehensible with nothing above it. Include the numbers. Include the qualifier. This is the sentence that gets quoted.
- 3 **Then the context, caveats, examples and nuance**
 This is where you earn trust with the human who clicked through. It is not wasted. It is simply not the part that gets retrieved.

X Typical business page. "Pricing. At DNE we understand that every shipment is unique. Our team works closely with each client to understand their requirements and deliver a tailored solution. Contact us for a quotation."

Contains zero retrievable facts. Answers no sub-query. Will never be cited by anything.

✓ Same page, rewritten. "How much does customs clearance cost at Port Klang? Standard import clearance for a single 20ft container at Port Klang typically runs RM450–RM850 in agent fees, excluding duties, taxes and port charges. Cost varies with commodity classification, whether a permit is required, and whether the shipment is FCL or LCL. Below is a breakdown of each component..."

Answers a real sub-query. Contains a number, a range, and the conditions. Survives being lifted out.

Things that quietly make a page unretrievable

- **Pronouns and back-references across paragraphs.** "This", "the above", "as mentioned". Each one breaks a passage that would otherwise stand alone.
- **Answers split across two sections.** The question in one place and the number in another means neither passage is complete.
- **Vague headings.** "Our Services", "Why Choose Us", "Solutions". They define no topical boundary, so the passage underneath belongs to no sub-query.
- **Facts trapped in images.** A price list rendered as a JPEG is a blank space to a text retriever. Put the numbers in HTML text, use the image as decoration.
- **Burying the specific under the general.** Three paragraphs of philosophy before the first concrete claim means the retriever scores your philosophy.

The tension with Google's advice, resolved

Google says you do not need to write in a special way for generative AI, and that its systems understand synonyms, semantic meaning and multi-topic pages. **TIER A** That is true and you should believe it.

What it means practically is that this chapter is not a compliance exercise with a gate at the end. There is no penalty for prose. What there *is*, is a competition between passages for each sub-query, and clarity wins ties. You are not obeying a rule. You are making it easier for a machine to choose you over an equally good competitor — and simultaneously making the page better for the human who arrives.

That is the correct mental model for almost every content recommendation in this book. Nothing here is a gate. Everything here is a tiebreaker, and in a field where thousands of pages are competing for three citation slots, tiebreakers decide outcomes.

HOW TO CHECK

Print your top service page. Take a highlighter and mark every sentence that states a specific, checkable fact — a number, a timeframe, a named process, a condition. If the page has fewer than five highlighted sentences, it is unretrievable regardless of how well it reads.

IF YOU ONLY DO ONE THING

Turn every vague heading on your site into a question a customer would actually ask, and make the sentence directly beneath it answer that question completely.

Evidence beats adjectives

This is the most replicated finding in the field, and the only one that comes from a controlled experiment rather than a correlation. Content that cites, quotes and quantifies gets pulled into answers substantially more often than content that asserts.

THE ANALOGY

Two suppliers pitch you. The first says they are reliable, experienced and committed to excellence. The second says they handled 4,200 containers last year, that their average clearance took 1.8 days, and quotes their largest customer by name saying so.

You did not verify either. You believe the second one anyway, because specificity is expensive to fake and vagueness is free. A language model runs the same instinct, for the same reason: it was trained on text written by people who apply it constantly.

The experiment

The 2024 paper that named this field — *GEO: Generative Engine Optimization*, presented at KDD 2024 — is the closest thing to a controlled trial we have. The authors built GEO-BENCH, a benchmark of 10,000 queries drawn from nine sources including MS MARCO, Natural Questions, ORCAS and real Perplexity queries, then rewrote source content nine different ways and measured which versions got pulled into generated answers more often. TIER A

REWRITING METHOD	VISIBILITY CHANGE	WHAT IT MEANS
Add quotations	+40.5%	Quote a named, credible person
Add statistics	+32.3%	Replace qualitative claims with numbers
Improve fluency	+28.7%	Make the writing flow properly
Cite sources	+27.6%	Reference credible external sources
Authoritative tone	varies	Confident, declarative writing — best in debate and history topics
Easy-to-understand	modest gain	Simplify language, keep the information
Technical terms	modest / mixed	Add domain-specific vocabulary
Unique words	modest / mixed	Rarer vocabulary
Keyword stuffing	−9.1%	The classic SEO move. It actively hurts.

Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan & Deshpande, "GEO: Generative Engine Optimization", KDD 2024. Figures shown are on the Position-Adjusted Word Count metric; the paper also reports a Subjective Impression metric where Statistics (+29.8%) and Quotations (+27.8%) again lead. Headline finding: up to 40% visibility improvement.

Three things in that table deserve to be read twice.

Keyword stuffing produced a negative result. The single most instinctive SEO reflex — say the target phrase more often — made content *less* likely to be used. If you take one thing from this chapter into a conversation with an agency, take that.

Fluency scored as highly as citation. Simply making the writing good was worth roughly as much as adding sources. This is not usually how optimisation works, and it is a reassuring signal about where the systems are pointed.

The effect was strongest for the underdog. Sites ranked lower in results gained far more than sites at the top. Adding source citations lifted a fifth-ranked site's visibility by 115.1% in the paper's measurement. If you are not the biggest name in your category, this is your chapter.

TWO HONEST LIMITS ON THIS STUDY

It was published in 2024 and tested against the generative engines of that moment; the systems have changed. And its visibility metrics are proxies, constructed by the authors, not revenue. Treat the *ranking of methods* as durable and the *exact percentages* as a snapshot. The ranking has been broadly corroborated by every vendor study since, which is why I rate it Tier A while still telling you not to quote "+40.5%" as though it were a law of physics.

What this looks like on a real page

✗ "We have extensive experience handling hazardous goods and a strong track record with the authorities."

✓ "In 2025 we filed 3,928 import declarations covering 10,400 containers at Port Klang, for 56 customer groups. Roughly two-thirds of those containers were 20ft."

One sentence became four verifiable statistics. None of them is a claim about how good the company is — which is precisely why they carry weight. Every figure here comes from DNE's operational records; see the method note in Chapter 10.

The four moves, in order of return

- **Quote someone real and name them.** Your operations manager, a customer, a regulator's published statement, an industry body. A name and a role attached to a sentence changes its status.
- **Turn every adjective into a number.** "High volume" is nothing. "10,400 containers across 3,928 declarations in 2025" is a fact. Go through your site and hunt adjectives.
- **Cite the source of every external claim.** Link to the regulation, the study, the official figure. It costs a sentence.
- **Then edit for fluency.** Read it aloud. Fix what makes you stumble. This scored as highly as the clever techniques.

HOW TO CHECK

Take any page and count two things: the number of specific figures, and the number of external citations. Most business pages score zero and zero. Set a floor — say three figures and two citations per substantial page — and work through your site against it.

IF YOU ONLY DO ONE THING

Go through your five most important pages and replace every unsupported superlative with a number you can defend. If you cannot defend a number, delete the claim rather than softening it.

Own a number nobody else has

Everything in the last chapter can be copied by a competitor within a week. This one cannot. Publishing data that exists nowhere else is the only durable position in AI visibility.

THE ANALOGY

Ten shops on the same street sell the same product. Nine of them write blog posts explaining what the product is. Anyone can write that post; the tenth shop writes it too, and so does an AI in four seconds.

The tenth shop also publishes, every quarter, the actual average price paid on that street, because it is the only one that tracks it. When anyone — a customer, a journalist, a machine — wants that figure, there is exactly one place to get it. They must name the source, because there is no alternative source.

That is not content marketing. That is a monopoly on a fact.

Why this works mechanically

Return to Chapter 1. The engine retrieves passages and prefers ones that answer a sub-query with specific, corroborated information. Now consider what happens with a genuinely proprietary figure:

- It cannot be sourced from a competitor, because they do not have it.
- It cannot be synthesised from the model's training, because it did not exist at training time.
- If the engine wants to use it, it has to cite you — or omit the point entirely.

Every other tactic in this book makes you *more likely* to be chosen among substitutes. This one removes the substitutes. It is the difference between being a good option and being the only option.

Commodity content competes. Proprietary data compels.

You already have the data

Most businesses believe they have no data worth publishing. Almost all of them are wrong. You are sitting on operational records that nobody outside your company has ever seen, and which answer questions people genuinely ask. Some patterns:

Logistics / freight business	Average clearance times by port and month. Detention and demurrage rates by shipping line. Seasonal volume patterns. Typical cost breakdown for a standard shipment.
Professional services firm	How long a typical engagement takes. The three problems that account for most enquiries. Anonymised outcome benchmarks across your client base.
Local trade or contractor	Actual local price ranges for common jobs. Which faults you see most often in this region. Seasonal demand patterns.
Agency or consultancy	Aggregated, anonymised client results. Before-and-after benchmarks. Time-to-result distributions.
Retailer or e-commerce	What actually sells together. Return-rate patterns by category. Regional demand differences.

How to publish it so it actually gets used

- 1 State the method, briefly and honestly**

"Based on 3,928 import declarations, covering 10,400 containers, filed at Port Klang between January and December 2025." One sentence. It is what makes the number credible rather than a marketing claim, and it is what a careful writer needs before they will cite you.
- 2 Put the headline number in plain HTML text, in the first paragraph**

Not in a chart image. Not in a downloadable PDF behind a form. A retriever cannot read either. Chart it as well, by all means — but the number must exist as text.
- 3 Give it a stable, permanent URL and a clear title**

"Port Klang Customs Clearance Benchmark 2026". Something a person could cite in a sentence. Update the same URL each period rather than creating a new one — you compound authority instead of splitting it.
- 4 Repeat it on a schedule**

Quarterly or annually. A one-off is a curiosity; a series becomes the reference. The second edition is worth more than the first, and the fourth is worth more than the first three combined.
- 5 Tell people it exists**

Trade press, LinkedIn, your association, the journalists covering your sector. Data with no distribution is a diary. This is also how Lever One gets fed — original data is the most reliable way to earn genuine, unpaid mentions.

THE THREE WAYS THIS GOES WRONG

- **Publishing something nobody asked about.** Interesting to you is not the test. Start from a question you are actually asked, then find the data that answers it.
- **Locking it behind a form.** Gated data is invisible data. If lead capture matters, publish the findings openly and gate the full dataset or the detailed methodology.
- **Breaching confidence.** Aggregate, anonymise, and check your contracts before publishing anything derived from client work. This is the one part of this chapter with legal exposure.

You will find vendor claims that original research achieves citation rates of 38–65% against 6–15% for standard blog posts. **TIER C** I cannot trace those figures to a published method and I would not repeat them. The mechanical argument above stands on its own without them, which is exactly why I have made it that way.

HOW TO CHECK

Ask yourself one question about any page you are about to publish: *could a competent competitor write this page tomorrow?* If yes, it is commodity content — publish it if it serves customers, but do not expect it to build a position. If no, and specifically if the reason is that you hold data they do not, you are building something that lasts.

IF YOU ONLY DO ONE THING

Find one question you are asked constantly, pull the answer out of your own operational records, and publish it with the method stated. One number, honestly derived, is worth more than a year of blog posts.

Be one clear entity, everywhere

Machines do not recognise companies. They recognise entities — resolved, consistent, cross-referenced identities. If your business appears under four slightly different names and three descriptions, you are not one entity. You are noise.

THE ANALOGY

You are trying to reconcile a customer list. One record says "DNE Total Sdn Bhd", another "DNE Total", another "D.N.E. Logistics", another "DNE Group (Klang)". A human glances at it and says obviously these are the same company.

A machine does not know that. It sees four small, unimpressive companies instead of one substantial one — and the reputation, mentions and evidence that should have accumulated to a single identity are scattered across four ghosts.

Entity work is deduplicating yourself in the eyes of every system that will ever describe you.

What an entity is, in practice

Search and AI systems maintain structured representations of real-world things: a company, a person, a product, a place. Each has attributes — what it is, where it is, what it does, who runs it — and links to other entities. Google's Knowledge Graph is the best-known example, and it draws heavily on Wikidata, which in turn draws on Wikipedia.

When a model answers "who is X", the quality of its answer depends almost entirely on whether X resolves cleanly to one well-described entity. When it does not, the model does what models do with ambiguity: it produces the most statistically plausible answer, confidently, and gets things wrong. That is the subject of Chapter 19, and this chapter is the prevention.

The work, in order

1 Decide your canonical name and never deviate

One legal name, one trading name, one way of writing each. Pick the form and use it identically everywhere: website footer, LinkedIn, directories, invoices, email signatures, press. Same for address and phone. This is genuinely the highest-value step and it costs nothing but discipline.

2 Write one description and reuse it verbatim

Two sentences: what you do, for whom, where. Paste the same two sentences into every profile. Variation reads as inconsistency to a machine that is checking whether sources agree.

3 Add Organization schema with sameAs

One block of JSON-LD on your homepage stating your name, logo, address, and a `sameAs` list pointing at every profile you own — LinkedIn, Wikidata if you have an entry, Crunchbase, your Google Business Profile, your registry listing. This is the explicit machine-readable statement that all these identities are one thing.

4 Claim your Google Business Profile and complete it fully

Google names Business Profiles explicitly in its AI optimisation guidance as the way to make business information available to its AI features. **TIER A** Fill every field. Categories, hours, service areas, description, photos, services list.

5 Get your people named alongside the company

Named authors with real credentials, a proper about page, LinkedIn profiles that state the company name identically. People are entities too, and a company with recognisable humans attached resolves more confidently than a faceless one.

6 Wikidata and Wikipedia, if — and only if — you genuinely qualify

A Wikidata item is achievable for many established businesses and feeds the Knowledge Graph directly. Wikipedia has strict notability requirements and attempting to game it will backfire publicly. Do not commission a "Wikipedia page service". If you are notable, the article will be accepted; if you are not, no amount of money changes that.

Where schema fits — the honest version

You will be sold schema markup as an AI visibility unlock. The evidence does not support that framing, and I would rather you hear it here than pay for it.

Ahrefs ran the closest thing to a controlled test: 1,885 pages that added JSON-LD schema between August 2025 and March 2026, against 4,000 matched pages that did not, measured 30 days either side.

TIER A The results:

PLATFORM	CHANGE IN CITATIONS AFTER ADDING SCHEMA
Google AI Overviews	−4.6%
Google AI Mode	+2.4%
ChatGPT	+2.2%

All three are within noise. Ahrefs' conclusion: adding schema produced no major uplift on any platform. Separately, 53% of AI-cited pages do carry schema — but that is correlation. Sites that implement schema tend to be sites that invest in everything else too.

Google's own documentation agrees: structured data **is not required** for generative AI search and there is **no special schema.org markup** for it. **TIER A** They also say continuing to use it is a good idea for classic rich results.

So the correct position is narrow and worth stating precisely. **Use schema for what it is good at** — declaring your identity unambiguously (Organization, `sameAs`), and earning rich results in classic search (Product, Review, FAQ, LocalBusiness, Article). **Do not use it as an AI visibility tactic**, do not pay anyone to "optimise your schema for AI", and do not expect citations to move because you added it.

HOW TO CHECK

- Search your brand name and open the first ten results. Note every variation of your name and description. Consistency is the metric.
- Ask three different AI tools "what is [your company]?" Compare the answers to each other and to reality. Divergence between them means your entity is unresolved.
- Run your homepage through Google's Rich Results Test to confirm your Organization markup parses.
- Search Wikidata for your company name. If nothing is there and you are an established business, that is an opportunity.

IF YOU ONLY DO ONE THING

Write your canonical name, address, phone and two-sentence description in a single document, then spend an afternoon making every profile you own match it exactly, character for character.

Live where the engines already trust

A small number of platforms account for a very large share of everything AI engines cite. You will not change that list. You can decide whether you are present on it.

THE ANALOGY

A new journalist arrives in your city and needs a story about your industry by tomorrow. They do not walk the streets. They call the three people everyone calls, read the two trade publications everyone reads, and check the forum where practitioners argue.

Being excellent and unknown to those five sources means you are not in the story. Not because anyone judged you — because you were not where they looked.

The concentration is real, the numbers are not settled

Multiple independent trackers agree on the shape: citations cluster heavily into a small set of domains, and the same names appear in every study — **Reddit, Wikipedia, YouTube, LinkedIn**, major review platforms, and a handful of large publishers.

They disagree substantially on the magnitudes. One 2026 aggregate index puts Reddit at around 40% of multi-engine citation frequency; another study reports YouTube overtaking Reddit in the same period; a third reports Wikipedia at 47.9% of ChatGPT citations. **TIER B** These cannot all be right, and the variation tells you something more useful than any of the individual figures:

WHAT THE DISAGREEMENT ACTUALLY MEANS

Citation sources are **volatile**. They shift with model updates, with content-licensing deals, with product changes, and with the mix of queries a given study happened to run. Any number you are quoted about "the most-cited domain" is a photograph of a moving object.

Plan for the durable pattern — *a few large platforms dominate, and they are the ones where people discuss things publicly* — and never build a strategy that depends on one platform's share holding.

What to actually do on each

REDDIT AND SPECIALIST FORUMS

The most-cited and the most dangerous. Reddit communities detect and punish marketing with unusual speed, and a ban is permanent and public. The only approach that works:

- Participate as a named human with an occupation, not as a company account.
- Answer questions in your area of genuine expertise, thoroughly, with no link.

- Disclose your affiliation plainly when it is relevant. "I run a forwarding agent in Klang, so take this with that in mind" earns credibility rather than costing it.
- Never post a promotional thread. Never use multiple accounts. Never pay someone to mention you.

This is slow and it is not delegable to an agency. It is also the highest-trust surface on the internet right now, precisely because it is hard to fake.

YOUTUBE

Given the strength of the correlation reported for YouTube mentions, and given that engines read titles, descriptions and transcripts, this is the most under-exploited channel for most B2B and service businesses. Note what is being measured: *your brand name appearing in video text*. That includes videos other people make. Practical moves:

- Answer your ten most common customer questions as ten plain videos. A phone and good light is enough.
- Write real descriptions — several hundred words that restate the answer in text, because the description is retrievable content.
- Upload accurate transcripts rather than relying on auto-captions.
- Get yourself onto other people's channels: interviews, panels, industry podcasts that publish to YouTube.

REVIEW AND COMPARISON PLATFORMS

For B2B software, G2 and Capterra. For services, Google reviews plus whatever vertical directory your buyers use. The mechanism is straightforward: these are the only sources with structured, current, category-organised third-party data about vendors, which makes them ideal retrieval targets for "best X for Y" questions.

- Claim the profile and complete every field, especially category placement.
- Build a steady, honest flow of reviews rather than a burst. Bursts look manufactured to platforms and to readers.
- Respond to reviews, including bad ones. Responses are text on a trusted domain that mentions your brand.

WIKIPEDIA AND WIKIDATA

Covered in Chapter 11. Wikidata is achievable and worth doing. Wikipedia is not something you buy.

LINKEDIN

A consistently cited domain and one of the few where you fully control the content. A complete company page, employees with matching profiles, and regular substantive posts from named people all contribute mentions on a high-trust domain.

HOW TO CHECK

Run your five most important buyer questions through Perplexity — it shows sources most clearly — and write down every domain cited across all five. That list *is* your target platform list, specific to your industry, and it beats any generic ranking including the ones in this chapter. Do it again in three months and watch it move.

IF YOU ONLY DO ONE THING

Pick the one platform where your buyers genuinely discuss their problems, and spend thirty minutes a week there being useful under your own name. One year of that beats any campaign you can buy.

PART THREE

The Craft

Part Two was about direction. This part is about execution — which pages to build, how to write them, how to keep them alive, and how the job changes if you are a local business or if you are chasing one engine in particular.

The page types AI actually cites

Some page formats get retrieved far more than others, and the reason is structural rather than mysterious. If you only ever build six kinds of page, build these.

THE ANALOGY

A researcher on a deadline reaches for the reference book, the comparison table and the FAQ sheet. They do not reach for the company brochure, however handsome it is.

Not because the brochure is bad. Because it does not contain the shape of information they need to extract: a direct comparison, a specific number, a clear definition.

The six formats, and why each one works

FORMAT	WHY IT GETS RETRIEVED	EXAMPLE
Comparison & "alternatives"	Matches the highest-intent question people ask AI: "X versus Y", "alternatives to X". Tables are dense with extractable facts.	"Air freight vs sea freight from Shanghai to Port Klang: cost, transit time and when each makes sense"
"Best / top" listicles	Directly answers the shortlist question. AI answering "best X" retrieves lists of X. Also the format most saturated with low-quality content, so quality shows.	"Seven questions to ask a customs broker before you appoint one"
Original data & benchmarks	Chapter 10. Unsubstitutable, so it must be cited.	"Port Klang import volumes, 2025: a review of 3,928 declarations"
Structured Q&A / FAQ	Question-and-answer pairs are, almost literally, pre-formatted retrieval units.	"Importing into Malaysia: 24 questions we are asked every week"
Definitions & glossaries	"What is X" is one of the most common query shapes. A clean definition in the first sentence is trivially quotable.	"What is a Delivery Order, and who issues it?"
Step-by-step process guides	Numbered steps are self-contained by construction, and "how do I..." is a huge share of AI questions.	"How to clear an LCL import at Port Klang, step by step"

Vendor studies consistently rank comparison content, listicles, FAQs and data-driven articles at the top of citation rates, though the specific multipliers they publish (2x, 3.2x, "95% citation rate") are untraceable and should not be repeated. TIER C The ranking is corroborated across sources; the numbers are not.

The one page most businesses are missing

Of the six, the format with the widest gap between value and adoption is the **honest comparison page** — including comparisons where you do not always win.

Most companies will not write "when you should use a different kind of provider", because it feels like giving away business. But that page answers the exact question a buyer asks an AI at the moment of highest intent, it is the kind of content that gets cited because it reads as informative rather than promotional, and the buyer it sends elsewhere was never going to buy from you anyway. Meanwhile the buyer it keeps arrives already convinced, because you told them the truth first.

The page where you admit what you are not best at is usually the page that earns the most trust — from readers and from the systems trained on how humans write when they are being straight.

What not to build

- **Generic explainers of things already explained a thousand times.** "What is logistics?" You will never win it, and there is no reason you should.
- **Thin AI-generated volume.** Publishing two hundred near-identical pages was a bad strategy in 2020 and is a worse one now, because the systems are explicitly built to filter commodity content.
- **Pure brochure pages with no facts.** Keep them for humans who are already convinced. Do not expect them to be retrieved.
- **Gated PDFs.** Invisible. If it matters, publish it as a page.

HOW TO CHECK

List every page on your site and tag it against the six formats. Most businesses find they have twenty brochure pages and zero of the six. That inventory is your content plan, and it is usually a three-month plan rather than a three-year one.

IF YOU ONLY DO ONE THING

Write one genuinely honest comparison page covering the choice your buyers actually agonise over — including the cases where the answer is not you.

How to write a paragraph a machine can quote

Chapter 8 covered the structure. This is the sentence-level craft — eight editing habits that make the difference, all of which also make the writing better for people.

THE ANALOGY

Think of each paragraph as a business card. It will be handed to someone who has never met you, in a room you are not in, with no explanation attached. Everything they need has to be on the card.

The eight habits

1 Name the subject in every paragraph

Not "it", "this", "the company", "the process" — the actual name. It reads slightly repetitive to you because you have the context. To a passage lifted out of the page, it is the only thing that makes the passage mean anything.

2 Put the number in the sentence, not in the chart

"Volumes grew (see figure 3)" is unretrievable. "Port Klang import containers rose 2.3% between the August–December windows of 2024 and 2025" is a fact that travels.

3 Front-load the qualifier

"For FCL imports under 20 tonnes, clearance typically takes two working days" is more useful than the same claim with the condition buried at the end. It also stops the engine quoting you out of context and misrepresenting you.

4 Write the definition sentence explicitly

Somewhere on any page about a concept, there should be a sentence of the form "A *[thing]* is a *[category]* that *[distinguishing feature]*." Plain, almost dictionary-like. It will be quoted more than anything else you write.

5 One idea per paragraph, one topic per heading

Paragraphs carrying three ideas match no sub-query cleanly. Split them.

6 Use real tables for real comparisons

An HTML table is highly structured, unambiguous, dense with facts, and readable without rendering. It is one of the few formatting choices that is genuinely mechanically advantageous.

7 Cut the throat-clearing

"In today's fast-paced business environment..." Delete every such opening. It delays the answer, adds no information, and is the single clearest marker of content written to fill space.

Read it aloud

The GEO experiment found plain fluency worth roughly as much as adding citations. Reading aloud is how you find the sentences that do not work. It remains the best editing tool anyone has invented.

A worked rewrite

X Before. "In today's competitive logistics landscape, businesses need a partner they can rely on. Our team has been serving clients across the region for many years, and we pride ourselves on our attention to detail. Whatever your needs, we have the expertise to help. This is why so many companies choose us."

✓ After. "**What size of importer does DNE handle at Port Klang?** DNE filed 3,928 import declarations covering 10,400 containers at Port Klang in 2025, for 56 customer groups — an average of 2.6 containers per declaration. Roughly two-thirds of those containers were 20ft and one-third 40ft, which is the profile of regular industrial importers rather than one-off shippers. Shipments requiring a permit from a controlling agency are handled by the importer's own permit holder; DNE does not advise on or obtain permits."

Fifty-four words became eighty. The first version contains no facts and answers no question. The second answers a real question, carries four numbers, names its evidence base, draws the inference a buyer actually cares about, and states a scope boundary. Every sentence in it survives being lifted out.

On using AI to write the content

You will be tempted, and used carefully it is fine. The rule that keeps you safe is simple: **AI is good at structure and bad at substance.**

- **Reasonable:** asking it to restructure your draft answer-first, to draft headings from your list of customer questions, to turn your messy notes into a clean table, to find the passive sentences.
- **Not reasonable:** asking it to write the content. It has no access to your operational data, your first-hand experience or your judgement — which, per Chapters 9 and 10, is the entire source of your advantage. What it produces is by definition the average of what already exists, which is precisely the commodity content these systems are built to skip.

Google's guidance points the same way: it asks for content with a unique point of view, first-hand reviews and expert perspective, and specifically warns against commodity content. TIER A

HOW TO CHECK

Paste one of your pages into an AI tool with the prompt: *"Based only on this text, what specific factual claims could you cite, and what question would each one answer?"* If it struggles, so will a retriever. This is a free, surprisingly brutal editing test.

IF YOU ONLY DO ONE THING

Delete every opening sentence on your site that says nothing, and replace it with the answer.

Freshness is a habit, not a date field

AI answers skew towards recent content, which tempts people into faking dates. That does not work. What works is a small, boring maintenance routine.

THE ANALOGY

A restaurant reprints its menu with today's date but serves the same tired dishes. Nobody is fooled twice. The restaurant next door changes three dishes a season and quietly becomes the one people recommend.

What the evidence supports

There is a real recency effect, and it is more modest than the industry claims. Ahrefs' analysis of around 17 million citations across seven AI platforms found cited pages averaged **1,064 days old**, against **1,432 days** for traditional organic results — a meaningful skew towards fresher content, but note that both averages are measured in *years*. TIER B

You will also see claims that content under 30 days old earns "3.2× more citations", or that a refresh programme moved a citation rate from 12% to 47%. TIER C I could not trace a method behind either. Ignore them.

The honest summary: **recency helps at the margin; it does not override relevance or evidence**. A three-year-old page with a unique number will beat a one-week-old page of generic advice.

What actually counts as a refresh

Changing the date is not a refresh. It is worse than nothing, because if it becomes a site-wide pattern it is trivially detectable and it destroys the credibility of your dates generally. A refresh means at least one of:

- New data, new figures, or updated prices.
- A regulation, process or requirement that has changed.
- A new section answering a question you have started being asked.
- Removal of something that is no longer true.

When you do make a substantive change, then *do* update both the visible "last updated" date on the page and the `dateModified` field in your Article schema, and keep them consistent with each other.

THE BULK-RESTAMPING TRAP

Several publishing tools and SEO plugins will helpfully set `dateModified` to the build date across your entire site every time you deploy. Your whole archive then claims to have been updated this morning. This is the single most common way a well-intentioned site ends up with dishonest dates, and nobody notices because no human reads schema. Check what your build actually stamps.

A maintenance routine that fits a real business

1 Identify your twenty pages that matter

The ones tied to revenue. Ignore the rest.

2 Review five of them each quarter

Twenty pages, four quarters, five a quarter. Each one gets looked at annually. That is a sustainable rhythm, not a project.

3 Ask three questions per page

Is anything on it now false? Is there a number that should be this year's? Is there a question we now get asked that this page does not answer?

4 Change something real, then update the date

In that order. Always in that order.

5 Push the change to Bing via IndexNow

So the update reaches ChatGPT's retrieval layer without waiting for a crawl.

IF YOU ONLY DO ONE THING

Put a recurring quarterly appointment in the calendar to review five pages. The routine matters far more than the individual edits.

Local businesses and service businesses

If you serve a place rather than the internet, the levers reorder. Your profile data and your reviews matter more than your website, and the data flows through channels most people have never heard of.

THE ANALOGY

Someone new in town asks where to get a car serviced. They do not read garage websites. They ask three people, glance at a map app, and look at star ratings.

An AI does the same thing, and the three people it asks are your profile data, your reviews, and whatever aggregator feeds its location layer.

What changes

For a local business, the ranking of levers shifts:

1. **Profile completeness and consistency** across every map, directory and aggregator — this becomes lever one.
2. **Reviews** — volume, recency, and your responses. This is your corroboration layer.
3. **Location-and-service specificity on your own site** — actual place names, actual service names, in text.
4. Everything else from Part Two, in the same order.

The aggregator point most people miss

Different AI products source location data differently, and not all of them go through Google. Apple Maps, Yelp, Foursquare and various national aggregators feed different assistants. You will see a specific claim that ChatGPT draws "over 70%" of its local business data from Foursquare and similar aggregators

TIER C — I cannot verify that figure and would not repeat it.

The underlying point stands on its own logic, though: **a perfect Google Business Profile is necessary and not sufficient**. Claim your listing on Apple Business Connect, Bing Places, and whichever aggregators are significant in your market, and make the data identical on all of them.

ON "98.8% OF LOCAL BUSINESSES ARE INVISIBLE TO AI"

This figure circulates widely with no method attached. **TIER C** It is the kind of number that exists to sell a service. Do not use it, and be wary of anyone who does — it is a reliable indicator of how carefully the rest of their pitch was assembled.

The practical list

- **Google Business Profile:** every field completed, correct primary category, service areas defined, services itemised, hours accurate including holidays, photos current.
- **Identical NAP everywhere:** name, address, phone, character for character, on every listing. Chapter 11's discipline, applied locally.
- **Reviews as a routine:** a steady, honest stream. Ask every satisfied customer, by a repeatable process. Reply to all of them — replies are indexable text containing your name on a trusted domain.
- **Name the places in your text:** "Port Klang", "Shah Alam", "Klang Valley" — written out, in sentences, not just in a footer. If you serve multiple areas, give each a real page with genuinely different content, not a template with the town name swapped.
- **LocalBusiness schema:** address, geo, hours, service area. This is the case where schema is genuinely worth the effort, because it is declaring identity and location rather than trying to buy a citation.
- **Answer local questions on your site:** "Do you deliver to Bukit Raja?" "What are your Saturday hours?" These are exactly the sub-queries a local AI answer fans out into.

HOW TO CHECK

Ask ChatGPT, Gemini and Perplexity: *"Recommend a [your service] in [your town]"*, then *"What are [your business]'s opening hours?"* and *"Where is [your business] located?"* Wrong answers to the second and third are more urgent than absence from the first — a confident wrong phone number costs you real calls.

IF YOU ONLY DO ONE THING

Build a repeatable habit of asking every happy customer for a review. It is the highest-leverage local action, it compounds, and it cannot be bought convincingly.

Playing each engine differently

The same prompt gives different answers on different engines. Rather than chase all of them, understand what each rewards and pick the ones your buyers use.

ENGINE	TENDS TO REWARD	YOUR LEVER
ChatGPT	Being present in Bing's index; established, well-referenced sources; clear comparison and list content	Bing Webmaster Tools + IndexNow. Then Chapters 5 and 13.
Google AI Overviews	Conventional Google ranking strength; heavy weighting to sites already performing organically	Classic SEO. Nothing exotic. Google says so itself.
Google AI Mode	Depth of coverage across many sub-questions rather than one strong page	Topic clusters: one strong pillar page plus specific pages answering each sub-question properly.
Perplexity	Fresh, specific, well-sourced pages; more willing to cite smaller sites	Original data and recency. Often the easiest engine for a small business to break into.
Claude	Documentation, primary sources, careful and complete explanations	Thorough, honest, well-structured reference content.
Copilot	Whatever Bing ranks	Inherits your Bing work. No separate effort required.

How to choose where to spend

Do not spread across six. Pick using two questions:

- **Where are your buyers?** Consumer and small-business audiences skew heavily to ChatGPT and to Google's AI features because they are the default. Technical and enterprise buyers use Perplexity and Claude more than their market share suggests. Ask five customers what they use; it is a better answer than any market-share chart.
- **Where are you closest to breaking through?** If you are already cited occasionally on Perplexity and never on ChatGPT, the marginal effort on Perplexity is far cheaper. Push where you have traction.

For most businesses reading this the honest answer is: **get ChatGPT and Google right, spot-check the others quarterly.** That is not a hedge, it is where the users are.

DO NOT READ ONE ANSWER AS A MEASUREMENT

These systems are non-deterministic. The same prompt, same engine, five minutes apart, can produce different sources. Personalisation, location and session history all move results, and asking from your own logged-in account is close to worthless as a measurement because the system knows who you are.

If you want a number rather than a vibe, you need the same prompt set, run repeatedly, from a clean session. Chapter 18.

IF YOU ONLY DO ONE THING

Ask five actual customers which AI tool they use. Optimise for that one. The market-share statistics are about the world; your customers are about your business.

PART FOUR

Measurement and Control

How to know whether any of this is working, how to fix it when the machine says something false about you, what to refuse to pay for, and how to tell a real statistic from a manufactured one. This part will save you more money than the rest of the book combined.

Measuring without fooling yourself

AI visibility is genuinely hard to measure, your analytics understate it structurally, and most of the dashboards being sold measure something adjacent to what you care about. Here is what is real.

THE ANALOGY

You want to know whether people are recommending your restaurant at dinner parties. You cannot attend the dinner parties. You can count bookings, ask new customers how they heard about you, and occasionally get invited to a party and listen.

That is the situation exactly. There is no dashboard for a conversation you are not in. There are three imperfect instruments, and using all three beats trusting any one.

Instrument one — your own prompt set (the most valuable, and free)

Build a fixed list of 20 to 40 questions your buyers actually ask, run them on a schedule, and record what happens. This is a research panel, not a ranking report, and treated that way it is the most honest signal available.

- 1 **Write the prompts as real people phrase them**
Not keywords. Full questions, in natural language, the way someone types into a chat box. "Who can handle DG cargo clearance at Port Klang?" not "DG cargo clearance Port Klang".
- 2 **Cover the five prompt types**
Informational ("what is..."), comparative ("X vs Y", "alternatives to..."), instructional ("how do I..."), brand ("what is [you]", "is [you] any good"), and transactional ("best [category] in [place]"). Tag each by funnel stage so you can tell an awareness gap from a shortlist gap.
- 3 **Run them from a clean session**
Logged out, no history, incognito. Otherwise the engine is answering for *you*, not for a stranger, and you will flatter yourself.
- 4 **Record four fields per prompt**
Were you mentioned (yes/no). Was the mention accurate (yes/no). Who else was mentioned. Which domains were cited. Nothing else. Four columns in a spreadsheet.
- 5 **Repeat monthly, same prompts, same way**
The absolute number is meaningless. The trend, and the competitor comparison, is the whole point.

Two derived numbers are worth tracking: your **presence rate** (the share of prompts where you appear at all) and your **share of voice** (your mentions as a proportion of all brand mentions across the set). Both are crude. Both move in the right direction when the work is working, which is all you need from a metric.

Instrument two — Search Console and server logs

Google launched **Search Generative AI performance reports** in Search Console on 3 June 2026, giving a dedicated view of impressions inside AI Overviews, AI Mode and generative features in Discover. TIER A Know its limits before you rely on it:

- Impressions, pages, countries, devices and dates only — **no clicks, no CTR, no query data** at launch.
- Data starts **18 May 2026**. There is no history before that.
- It rolled out to a subset of site owners first, beginning in the UK, expanding since.

Alongside it, your raw server logs answer a question no dashboard will: are `OAI-SearchBot`, `PerplexityBot` and `ClaudeBot` actually fetching your pages, which ones, and how often? This is the only direct evidence that the machines are reading you, and it is free.

Instrument three — referral traffic, read correctly

Segment sessions from `chatgpt.com`, `perplexity.ai`, `gemini.google.com` and `copilot.microsoft.com`. Then apply three corrections before you believe anything:

- **It undercounts, structurally.** Many AI referrals arrive with no referrer header and land in "direct". Your true AI traffic is higher than the number you see, by an amount you cannot determine.
- **Engine splits are unreliable.** Because of that stripping, the apparent breakdown between engines is partly an artefact of which ones happen to pass a referrer. Do not make decisions on it.
- **Judge it on quality, not volume.** Conversion rate and value per session, compared against organic. Volume will look disappointing for a long time and that is not the story.

WHAT TO REFUSE TO BELIEVE

Any tool claiming to show your "ranking" inside an AI model, your "AI algorithm score", or data drawn from a platform's internals. Google states directly that **no third-party tool has access to its internal ranking or AI systems.** TIER A Every legitimate visibility tool works the same way you would manually: it runs prompts, at scale, and counts what comes back. That is genuinely useful — it is sampling, not telemetry, and a vendor who blurs the difference is telling you what they are.

Should you buy a tool?

The category — Profound, Peec, Otterly, Semrush's AI toolkit, Ahrefs Brand Radar and others — automates instrument one across more engines and more prompts than you would run by hand, and adds competitor tracking. That is real value at scale.

My honest advice for a business under, say, fifty people: **run the spreadsheet by hand for three months first.** You will learn more from reading forty answers yourself than from any dashboard, you will discover that your assumed competitor set is wrong, and you will know exactly what you need before you pay for it. Then buy a tool if the manual version has become the bottleneck.

IF YOU ONLY DO ONE THING

Create a spreadsheet with 20 real customer questions and four columns. Fill it in this month. Fill it in again next month. That is a measurement programme.

When the AI is wrong about you

Being described inaccurately is more damaging than being absent, it happens constantly, and there is no complaints department. There is a method, though, and it works — slowly.

THE ANALOGY

A rumour is circulating that you closed your Klang office. You cannot phone the rumour. You can only be so visibly, consistently open — same address on every door, every letterhead, every listing — that repeating the rumour becomes obviously wrong to anyone who checks.

Correction is not a request. It is out-evidencing the error.

Why it happens

Three causes, and the fix differs for each:

- **Gaps.** The model has no good information about you and generates the statistically plausible thing instead. Confidently. This is the most common cause for small businesses and the easiest to fix.
- **Contradictions.** Different sources say different things — an old address on a directory, a former service on a partner's site, a discontinued price in a cached PDF. The model picks one, sometimes the wrong one.
- **Staleness.** The information was true. You changed. The internet has not caught up, and neither has the training data.

THERE IS NO CORRECTION FORM

OpenAI, Google and Anthropic do not operate a brand-correction portal. There is no ticket to raise, no dashboard, no escalation path. Anyone offering to "get your AI listing corrected" through a contact at a platform is describing something that does not exist.

The correction method

1 Write down the exact wrong claim, and where

Which engine, which prompt, what it said. Screenshot it. You need a baseline or you will never know whether it improved.

2 Find the source

Perplexity and Google AI Overviews show their citations directly — start there. For ChatGPT, search the web for the wrong claim itself; there is usually a page saying it. Very often it is your own site: an old page, a stale PDF, a legacy subdomain nobody remembers.

- 3 Fix the source if you control it**
Update, correct, or remove the page. Then redirect it if you removed it. Roughly half of brand hallucinations trace back to something the company itself published years ago and forgot.
- 4 Correct the source if you do not control it**
Directories, partner sites, old press. Email and ask. Most will fix it; it costs them nothing. This is unglamorous work and it is the work.
- 5 Publish the correct fact clearly and unmistakably**
A page on your own site that states the truth in plain, quotable language. "DNE Total operates from Klang, Selangor. We do not provide permit advisory services." Direct statements, easy to retrieve, hard to misread.
- 6 Push the corrected fact out to the corroboration layer**
Update every profile. Refresh your Wikidata entry if you have one. Update your Google Business Profile. Get the correct version stated somewhere other than your own site. This is Chapter 5 applied as damage control.
- 7 Re-test every two to four weeks**
Same prompt, clean session. Changes propagate over weeks or months, not days, because indexes refresh on their own schedule and some of the wrong information may sit in training data that will not update until the next model.

The high-cost errors, in order

Not all inaccuracies matter equally. Fix in this order:

- 1. Contact details and location.** A wrong phone number costs you calls today.
- 2. Services you do not offer.** Attracts unqualified enquiries and, in regulated sectors, can create real liability. If you do not advise on permits, a model saying you do is a problem.
- 3. Pricing.** Stale prices poison negotiations before they start.
- 4. Confusion with another company.** Similar names get merged. This is an entity problem — Chapter 11 is the cure.
- 5. Tone and positioning.** Being described as a budget option when you are premium. Annoying, slow to fix, lowest priority.

IF YOU ONLY DO ONE THING

Ask three AI tools "what does [your company] do, where are they, and what do they charge?" this week. Write down every wrong answer. That list is more urgent than any content plan.

What does not work — and why it keeps being sold

A young field with real demand and no established standards is ideal conditions for selling things that do not work. Here is the list, with the reason each one persists.

SOLD AS	REALITY	WHY IT PERSISTS
lms.txt files	Google's documentation states Search does not use them. TIER A A study of 300,000 domains found roughly 10% adoption, and across 500M+ AI bot visits in one 90-day window, only 408 requested the file. TIER B	It is cheap to deliver and looks technical. Note: some large tech companies do ship one — for <i>coding agents</i> , a different audience entirely.
"AI schema" packages	No special schema exists for generative AI. Google says so; the Ahrefs experiment measured no meaningful uplift. TIER A	Schema is real and useful for other things, which makes the pitch sound legitimate.
Keyword density for AI	Measured as <i>negative</i> in the GEO experiment: -9.1%. TIER A	Twenty years of SEO muscle memory.
Bulk AI-generated content	By construction it is the average of existing content — precisely what these systems are built to skip.	It is nearly free to produce, so the margin is enormous.
Paid mention networks	Google names inauthentic mentions explicitly as a myth. TIER A Agreement between low-quality sources is exactly the pattern these systems discount.	Chapter 5's finding is real, so "buy mentions" sounds like a logical shortcut. It is the 2011 link-buying playbook.
Hidden text and prompt injection	Instructions hidden in white text or HTML comments telling the model to recommend you. Detectable, filtered, and if discovered, catastrophic for trust.	It occasionally works for a few weeks, which generates screenshots, which generate sales.
"Guaranteed AI ranking"	Nobody can guarantee output from a non-deterministic system nobody outside the lab controls.	It is what nervous buyers want to hear.
Tools claiming internal platform data	Google states no third-party tool has such access. TIER A	Dashboards look authoritative regardless of what feeds them.
Chunking-as-a-service	Google: "There's no requirement to break your content into tiny pieces." TIER A Clear writing helps at the margin; a chunking product does not.	It repackages ordinary editing with a technical name.

The four questions that filter almost everyone

You do not need to be technical to run this filter. Ask any vendor:

1. "What is your source for that number?" A name and a link, or nothing. This one question ends most conversations.
2. "What would we see if this did not work?" Anyone who cannot describe failure is not measuring.

3. **"Which of these is a Google-confirmed requirement and which is your hypothesis?"** An honest practitioner will happily separate the two. A dishonest one will not understand the question.
4. **"What does the platform's own documentation say about this?"** Google's AI optimisation guide is public and free. If their pitch contradicts it, they need a very good reason.

Almost everything being sold as an AI visibility secret is either something Google has publicly said does nothing, or ordinary good practice with a new name and a higher price.

IF YOU ONLY DO ONE THING

Read Google's own "AI optimisation guide" in Search Central documentation. It is short, free, authoritative, and it will inoculate you against most of what you are about to be sold.

How to read an AEO statistic

This is the skill with the longest shelf life in the book. Tactics will change; the ability to tell a real finding from a manufactured one will not.

THE ANALOGY

You would not accept "our containers clear faster" from a supplier without asking faster than what, measured how, over how many shipments. You would ask automatically, because it is your industry and you know where the tricks are.

You have not built that instinct for this field yet. This chapter is the shortcut.

The seven tells of a manufactured statistic

1. **No sample size.** "Studies show pages with X get 3× more citations." Which pages? How many? Over what period? A real study leads with its sample because it is the strongest thing about it.
2. **A suspiciously clean multiple.** "2×", "3.2×", "10×". Real measurements are messy. Clean multiples are usually reverse-engineered from a conclusion someone already wanted.
3. **The seller benefits precisely.** A chunking vendor publishing that chunking yields 2–4× improvement. Not automatically false — automatically requiring the method.
4. **Circular sourcing.** Follow the citation. Blog A cites Blog B cites Blog C cites Blog A. This is extremely common in this field, and once you start following links you will find it constantly.
5. **A percentage with no baseline.** "Improved citations by 292%" — from what to what? From one citation to four is technically 300%.
6. **Precision that cannot exist.** "Each blocked bot costs 18–34% of potential citations on that engine." Nobody outside the platform could measure that. The range exists to look researched.
7. **Contradicts the platform's own documentation without acknowledging it.** If a vendor's claim conflicts with Google's published guidance and they do not mention the conflict, they either have not read it or hope you have not.

The three-question test

Before you repeat any number from this field — including one from this book — ask:

1 Who measured it, and what do they sell?

Not disqualifying. Vendors run the best studies in this field because they have the data. But it sets your prior.

2 Can I find the method in one click?

Sample size, time period, what was compared, control group. If it takes more than one click, it usually does not exist.

3 Does it agree with the mechanism?

The strongest test. Given how retrieval works — Chapter 1 — does this finding make sense? A claim that contradicts the mechanism needs extraordinary evidence. A claim that fits it needs less.

Worked examples, using numbers from this book

✓ **"Pew found 8% click-through with an AI summary versus 15% without."** Named non-commercial research organisation. Method published: 900 US adults, real browsing data, 68,879 queries, March 2025. Observed behaviour rather than a survey. Fits the mechanism — if the answer is on the page, fewer people click. *Trust it.*

✓ **"Branded web mentions correlate 0.664 with AI Overview visibility."** Named vendor, 75,000 brands, stated method, full table published, and the authors themselves warn against reading causation into it. Fits the mechanism. *Trust the direction; do not treat it as a dial.*

X **"Content chunking produces 2–4× improvement in citation rates."** No sample, no method, clean multiple, published by people selling chunking, and it contradicts Google's own documentation saying chunking is not required. *Discard.*

X **"98.8% of local businesses are invisible to AI."** No method, alarming, precise to one decimal place, and it exists to create urgency. *Discard.*

A live example, from building this book

The worked examples in Chapters 9, 10 and 14 use real figures from DNE Logistics' operational records. Pulling them produced a textbook illustration of this chapter, so it is worth showing you the working.

The first query returned Port Klang import volumes for 2024 and 2025:

YEAR	IMPORT DECLARATIONS	CONTAINERS
2024	1,718	4,597
2025	3,928	10,400
Apparent change	+128.6%	+126.2%

A 126% year-on-year jump. It is a wonderful number. It is also completely false, and it would have gone into this book unchallenged if the next query had not been run.

Breaking the same data down by month showed that **2024 contains no records before August**. The operational system was adopted mid-year. "2024" in that table is five months of data being compared against twelve. Volumes are, in reality, close to flat:

COMPARABLE WINDOW	DECLARATIONS	CONTAINERS
August–December 2024	1,718	4,597
August–December 2025	1,745	4,704
Real change	+1.6%	+2.3%

+126% became +2.3%. Same database, same query language, same afternoon. The difference was one additional question: *does my denominator cover the period I am claiming?*

WHY THIS IS IN THE BOOK

Nobody fabricated anything. A competent analyst ran a correct query against real data and got a wildly misleading answer, because a system-adoption date created a coverage gap that the totals concealed. This is how most bad statistics are actually born — not from dishonesty, but from a denominator nobody interrogated.

It is also why the Tier system exists. A vendor publishing "+126%" here would not be lying. They would simply not have run the second query, and you would have no way to tell from the outside. That is exactly why "what is your method?" is the only question that reliably works.

Apply it to this book

Everything here carries a tag and Appendix B lists the sources. Some of what I have rated Tier B will turn out to be wrong. The GEO paper's exact percentages are from 2024 systems that no longer exist, and I have said so in the chapter rather than in a footnote.

What I have tried to give you is not a list of facts to memorise but a mechanism to reason from, and the habit of asking where a number came from. The tactics in this book have a shelf life of perhaps eighteen months. The mechanism and the scepticism should last considerably longer.

IF YOU ONLY DO ONE THING

The next time anyone quotes you an AI-visibility statistic, ask one question: "where is that from?" The answer will tell you everything you need to know about the rest of the meeting.

PART FIVE

Doing It

A ninety-day plan built from the levers in order, and an honest assessment of what is coming next — including which parts of it you should act on now and which you should simply watch.

The first ninety days

Sequenced so that the cheapest, highest-certainty work happens first, and nothing depends on a budget you have not approved. Roughly two to four hours a week after the first fortnight.

Weeks 1–2 • Find out where you actually stand

Do not change anything yet. Measure.

- Write your 20-prompt list — real questions, five types, clean sessions. Run it on ChatGPT, Gemini and Perplexity. Record presence, accuracy, competitors, cited domains. (*Chapter 18*)
- Press Ctrl+U on your four most important pages. Search the raw source for your own sentences. (*Chapter 6*)
- Read your `robots.txt` line by line. (*Chapter 2*)
- Compare `site:` counts on Google and Bing. (*Chapter 7*)
- Ask three engines what your company does, where it is, and what it charges. Log every wrong answer. (*Chapter 19*)

Output: one page listing what is broken, what is wrong, and who is being recommended instead of you.

Weeks 3–4 • Unblock the plumbing

Everything here is technical, cheap, and pointless to postpone — no content work matters until the machine can read you.

- Fix any search-bot blocks in `robots.txt`.
- Check the CDN or firewall for bot-blocking rules and whitelist the search crawlers.
- Verify in Bing Webmaster Tools, submit the sitemap, enable IndexNow.
- If the Ctrl+U test failed: get the critical content server-rendered. This is the one item that may need a developer and a budget. It is also the one item without which nothing else works.
- Clean the sitemap. Redirect anything that 404s.
- Claim and fully complete the Google Business Profile.

Output: the engines can reach you, read you, and find you in the index they use.

Weeks 5–6 • Become one entity

- Write the canonical name, address, phone and two-sentence description. One document.
- Make every profile you own match it exactly. Every directory, every social account, every registry, every partner listing.

- Add Organization schema with a full `sameAs` list to the homepage.
- Claim every free directory, association listing and register in your category that you do not already hold.
- Fix the highest-cost inaccuracies from week 1 at their source.

Output: you resolve to one thing, described consistently, in more places than before.

Weeks 7–10 • Build the pages that get cited

Four pages. Not forty. Choose from Chapter 13's six formats, weighted to what your week-1 prompt run showed you were missing.

1. **One honest comparison page** covering the choice your buyers actually agonise over.
2. **One original-data page** — a number from your own records, with the method stated. (*Chapter 10*)
3. **One substantial FAQ page** built from the questions you are genuinely asked, answered directly.
4. **One process guide** walking through how the thing actually works, step by step.

Write all four to the Chapter 8 and 14 rules: question headings, answer-first, numbers in text, sources cited, no throat-clearing. Then rewrite the opening paragraph under every heading on your five existing key pages.

Output: four pages that can be retrieved, plus a materially improved existing site.

Weeks 11–12 • Start the off-site engine, and re-measure

- Pick one platform from Chapter 12 where your buyers genuinely are. Commit to thirty minutes a week, under your own name, permanently.
- Claim and complete your profiles on the review platforms that matter in your category.
- Start the review-request habit if you are a local or service business.
- Distribute the original-data page: trade press, LinkedIn, your association, anyone who covers your sector.
- **Re-run the exact same 20 prompts, the same way.** Compare against week 1.

WHAT TO EXPECT AT DAY 90

Probably: small movement on presence rate, visible improvement in accuracy, and one or two prompts where you now appear and did not before. Possibly: nothing measurable yet, particularly if your starting point was zero mentions and low brand awareness.

That is a normal result, not a failure. The technical fixes are instant but the corroboration layer — Chapter 5, the biggest lever — compounds over quarters, not weeks. Anyone promising a transformation in ninety days is selling something. What you will definitely have by day 90 is a site that can be read, an identity that resolves, four pages worth citing, and a measurement habit. That is the foundation, and most of your competitors will not have built it.

Then: the standing routine

- **Monthly:** re-run the prompt set. Twenty minutes.
- **Monthly:** one new page in a Chapter 13 format.
- **Weekly:** thirty minutes on your chosen off-site platform.
- **Quarterly:** refresh five pages properly. Re-check `robots.txt` and crawler logs. Re-run the brand-accuracy questions.
- **Annually:** republish your data study. It compounds.

IF YOU ONLY DO ONE THING

Do weeks 1 and 2. The measurement alone will tell you which of the remaining ten weeks actually apply to your business.

What is coming next

Four shifts are underway. Two you should prepare for now; two you should watch and not spend money on yet. Knowing which is which is the point of this chapter.

1. Agents that act, not just answer — prepare now, lightly

The current generation answers questions. The next books, buys and fills in forms. Three competing standards emerged through 2025 and 2026: **UCP** (Google and Shopify, for discovery and cart), **ACP** (OpenAI and Stripe, for in-chat checkout) and **AP2** (now governed by the FIDO Alliance, for proving who authorised a payment). They are designed to stack rather than compete outright. **TIER B** Google's own documentation now points site owners at agent-friendly site practices. **TIER A**

What to do now: if you sell products online, keep your product data clean, structured and accurate — that is the input every one of these protocols consumes, and it is work you should be doing anyway. If you sell services, make sure your enquiry form works without JavaScript gymnastics and that your key information is in plain HTML. Do not implement a protocol yet. Three competing standards means at least one will lose.

2. Someone else deciding whether AI can read you — watch closely

Cloudflare's shift through 2026 from pay-per-crawl toward paying based on whether content was actually used, and its move to block mixed-use crawlers by default on ad-bearing pages from September 2026, is the visible edge of a much larger negotiation between publishers and AI companies. **TIER B**

What to do now: know what your infrastructure provider's defaults are and check them after every platform change. The risk is not that you make a bad decision — it is that a decision gets made for you by a default setting and you find out six months later. Add "check bot-management defaults" to the quarterly routine.

3. Measurement getting real — prepare now

Search Console's generative AI reports launched in June 2026 with impressions only. Clicks, queries and richer breakdowns are stated as coming. **TIER A** When they arrive, the guesswork drops sharply and the vendor tools that currently sell inference will have to compete with free first-party data.

What to do now: make sure Search Console is properly verified for every property you own, including subdomains, so you have history when the data widens. Do not sign a long tool contract on the assumption that this gap stays open.

4. The commodity content collapse — act now

This is the shift with the clearest implication. As models get better at synthesising, the value of content that merely explains something already explained falls towards zero. The AI simply answers it. What stays valuable is what a model *cannot* generate: your operational data, your first-hand experience, your judgement about edge cases, your specific local knowledge.

What to do now: stop investing in explanatory content that competes with the model, and move that budget into the things only you have. This is Chapter 10, and it is the single most future-proof position in this book.

Every shift on this list points the same way: the durable asset is not your content, it is your evidence. Everything else is being commoditised in real time.

What will not change

Worth ending here, because it is where your effort is safest:

- **Machines will need to read you.** Whatever the format, being mechanically readable will always be prerequisite.
- **Corroboration will always beat assertion.** Any system that assembles an answer from multiple sources will reward agreement between them.
- **Specificity will always beat vagueness.** Numbers, names, conditions and sources are what makes a claim usable.
- **Being genuinely good at something, and demonstrably so, will keep working.** Unfashionable, and the only strategy that has survived every platform shift so far.

IF YOU ONLY DO ONE THING

Shift budget out of content that explains things and into content that proves things. That single reallocation is the correct response to all four shifts at once.

APPENDICES

Reference

The glossary, the full evidence table so you can check every number in this book, and a single-page checklist you can print.

Glossary

| Every term used in this book, in one line each.

TERM	MEANING
AEO	Answer Engine Optimisation — making your business the answer an AI gives.
GEO	Generative Engine Optimisation — the same job, named after the 2024 academic paper.
AI Overviews	Google's AI summary box at the top of search results.
AI Mode	Google's conversational search tab. Heavier fan-out, passage-level retrieval.
RAG	Retrieval-Augmented Generation — the model retrieves real documents before answering, rather than answering from memory.
Query fan-out	One user question expanded into many narrower sub-queries, run in parallel.
Passage / chunk	The fragment of a page that is actually retrieved and scored — not the whole page.
Citation	The source link an AI attaches to a claim in its answer.
Share of voice	Your brand's mentions as a proportion of all brand mentions across a fixed prompt set.
Presence rate	The share of your tracked prompts where you appear at all.
Prompt set	A fixed list of questions you re-run on a schedule to measure visibility.
Entity	A resolved, machine-readable identity for a real-world thing — your company, your people, your locations.
Knowledge Graph	Google's database of entities and the relationships between them.
Wikidata	Open structured database that feeds the Knowledge Graph. Easier to enter than Wikipedia.
Schema / JSON-LD	Machine-readable markup describing what a page is about. Useful for identity and rich results; not an AI unlock.
sameAs	The schema property listing your other official profiles, declaring they are all the same entity.
NAP	Name, Address, Phone — the trio that must be identical everywhere.
SSR / pre-rendering	Generating page content on the server so it exists in the HTML before any JavaScript runs.
SPA	Single-Page Application — a site rendered in the browser. The highest-risk architecture for AI visibility.
robots.txt	The file telling crawlers what they may fetch. Where most AI visibility is accidentally destroyed.
IndexNow	Open protocol pushing changed URLs to Bing instantly. Free, and matters because ChatGPT retrieves via Bing.
GPTBot / OAI-SearchBot	OpenAI's training crawler and search crawler respectively. Blocking the second removes you from ChatGPT answers; blocking the first does not.
Google-Extended	The user-agent controlling whether your content trains Gemini. Blocking it does not affect Search or AI Overviews.
Zero-click	A search that ends without the user visiting any website.
Hallucination	A confident, plausible, false statement — often about your business.

TERM	MEANING
UCP / ACP / AP2	Competing agentic-commerce standards for discovery, checkout and payment authorisation.

The evidence table

Every figure quoted in this book, with its source and rating. If you are going to put one of these in front of a client, take it from here.

CLAIM	SOURCE	TIER
Quotations +40.5%, Statistics +32.3%, Fluency +28.7%, Cite Sources +27.6%, Keyword Stuffing –9.1% on generative visibility; up to 40% overall; +115.1% for a 5th-ranked site adding citations; GEO-BENCH = 10,000 queries	Aggarwal et al., "GEO: Generative Engine Optimization", KDD 2024 (arXiv 2311.09735)	A
GPTBot, ClaudeBot and PerplexityBot do not execute JavaScript. Monthly fetches: Googlebot 4.5bn, GPTBot 569m, Claude 370m, AppleBot 314m, PerplexityBot 24.4m. ChatGPT 34.82% and Claude 34.16% of fetches hit 404s vs Googlebot 8.22%	Vercel / MERJ, "The rise of the AI crawler"	A
Click-through 8% with an AI summary vs 15% without; 1% clicked a cited source; 18% of searches produced a summary; 88% of summaries cited 3+ sources; median summary 67 words	Pew Research Center, July 2025. 900 US adults, real browsing data, 68,879 queries, March 2025	A
Branded web mentions 0.664, branded anchors 0.527, branded search volume 0.392, DR 0.326, referring domains 0.295, backlinks 0.218	Ahrefs, "An Analysis of AI Overview Brand Visibility Factors", 75,000 brands, Spearman	A
Adding schema: AI Overviews –4.6%, AI Mode +2.4%, ChatGPT +2.2% — no meaningful uplift. 53% of AI-cited pages carry schema (correlation)	Ahrefs, 1,885 treated pages vs 4,000 controls, Aug 2025 – Mar 2026	A
No special schema, no llms.txt, no chunking requirement, no AI-specific writing style, no inauthentic mentions; no third-party tool has internal access; Business Profiles feed AI features	Google Search Central, AI optimisation guide + AI features documentation	A
Search Console generative AI performance reports launched 3 June 2026; impressions/pages/countries/devices/dates only; data from 18 May 2026	Google Search Central blog	A
DNE Logistics, Port Klang imports 2025: 3,928 declarations, 10,400 containers, 56 customer groups, 6,372 × 20ft and 3,267 × 40ft. Comparable Aug–Dec window 2024→2025: declarations +1.6%, containers +2.3%. (2024 records begin in August — full-year 2024 is not comparable.)	DNE Logistics operational database (consolidated_jobs), filtered job_type='import' and pod_name='PORT KLANG (MALAYSIA)' on eta_date . Queried 8 August 2026 against the daily 02:30 MYT clone	A
YouTube mentions correlate ≈0.737 with AI visibility	Ahrefs companion report, 75,000 brands	B
AI Overviews cut top-result CTR ≈34.5% (Apr 2025) rising to ≈58% (Feb 2026)	Ahrefs	B
US zero-click share 60.45% (2024) → 68.01% (early 2026)	SparkToro clickstream	B
AI-search visitors ≈4.4× the value of an average organic visitor; ≈3× more likely to convert; AI search projected to overtake traditional search by 2028	Semrush clickstream analysis	B
ChatGPT referral share ≈0.32% of global website traffic, May 2026	SE Ranking / Similarweb	B
llms.txt adoption ≈10.13% of 300,000 domains; 408 targeted requests across 500M+ AI bot visits in a 90-day window	SE Ranking; independent crawler telemetry	B
Google AI Mode uses roughly 8–12 sub-queries per question	Multiple SEO analyses; not confirmed by Google	B

CLAIM	SOURCE	TIER
Cited pages average 1,064 days old vs 1,432 for organic results (≈17m citations, 7 platforms)	Ahrefs	B
Review platforms appear in ≈34.5% of commercial AI Overviews	Vendor analysis, method not fully published	B
≈50% of B2B software buyers start research with an AI chatbot; 69% switched intended vendor; 1/3 bought from a vendor they had not heard of	G2 buyer research, 2026	B
AI Overviews trigger on ≈48% of queries (Mar 2026); 36% informational vs 8% commercial vs 5% transactional	Semrush; Seer Interactive. Varies enormously by keyword panel	B
Perplexity rebuilt search in-house on Vespa (2025); hybrid BM25 + dense retrieval with multiple rerank layers	Perplexity engineering statements	B
Cloudflare shifted pay-per-crawl → pay-per-use (July 2026); default blocking of mixed-use crawlers on ad pages from 15 Sept 2026	Cloudflare announcements, trade press	B
Malaysia: 67% of 2,556 adults used an AI tool in the previous three months; 81% of those used AI-assisted search	Vodus AI Consumption Study, 2026	B
Reddit ≈40% of multi-engine citations / YouTube overtaking Reddit / Wikipedia 47.9% of ChatGPT citations	Three separate 2026 indices that contradict each other. Cited in Ch.12 <i>as evidence of volatility</i> , not as fact	B
"Chunking gives 2–4× more citations"	Untraceable. Quoted only to warn against it	C
"Original research achieves 38–65% citation rates"	Untraceable. Quoted only to warn against it	C
"Content under 30 days old gets 3.2× more citations"	Untraceable. Quoted only to warn against it	C
"Each blocked bot costs 18–34% of citations"	Untraceable. Quoted only to warn against it	C
"98.8% of local businesses are invisible to AI"	Untraceable. Quoted only to warn against it	C
"ChatGPT draws 70%+ of local data from Foursquare"	Untraceable. Quoted only to warn against it	C

Compiled August 2026. Tier definitions are in the front matter. Where a Tier B source is a vendor with a commercial interest in the finding, that is noted in the chapter as well as here.

The one-page checklist

Print this. Everything in the book, compressed to what you actually do.

CAN THE MACHINE READ YOU?

- Ctrl+U on your four key pages — your sentences appear in the raw HTML
- robots.txt allows OAI-SearchBot , PerplexityBot , Claude-SearchBot , Googlebot , Bingbot
- CDN / firewall is not blocking AI crawlers
- Server logs show real AI crawler visits in the last 30 days
- Sitemap is clean; old URLs 301 rather than 404

CAN IT FIND YOU?

- Verified in Bing Webmaster Tools, sitemap submitted, IndexNow on
- Google Search Console verified for every property
- Google Business Profile claimed and every field completed

DO YOU RESOLVE TO ONE ENTITY?

- One canonical name, address, phone, description — identical everywhere
- Organization schema with a full sameAs list
- Every free directory, register and association listing claimed
- Named people with real credentials attached to the business

IS THERE ANYTHING WORTH QUOTING?

- Headings are questions; the sentence beneath answers them directly
- Every paragraph names its subject and survives being lifted out
- Numbers in HTML text, not trapped in images or PDFs
- Adjectives replaced with figures; external claims cited
- At least one page built on data only you have
- One honest comparison page, one real FAQ, one process guide

DOES ANYONE ELSE SAY SO?

- Profiles complete on the review platforms your buyers use
- A steady, honest flow of reviews — and replies to all of them
- Thirty minutes a week on one platform where your buyers talk, under your own name
- Videos answering your top customer questions, with real descriptions and transcripts

ARE YOU MEASURING?

- A fixed 20-prompt set, run monthly, clean sessions, four columns
- Brand-accuracy questions re-run quarterly
- AI referrals segmented and judged on conversion, not volume
- Quarterly: refresh five pages, re-check `robots.txt` and crawler logs

ARE YOU REFUSING THE RIGHT THINGS?

- No llms.txt package, no "AI schema", no chunking service
- No bulk AI-generated content, no paid mention networks, no hidden text
- No tool claiming internal platform data, no guaranteed AI rankings
- No statistic repeated without its source

Answering Machines · Compiled August 2026. Every figure sourced in Appendix B. Tactics have a shelf life of roughly eighteen months; the mechanism in Chapter 1 and the scepticism in Chapter 21 should last considerably longer.