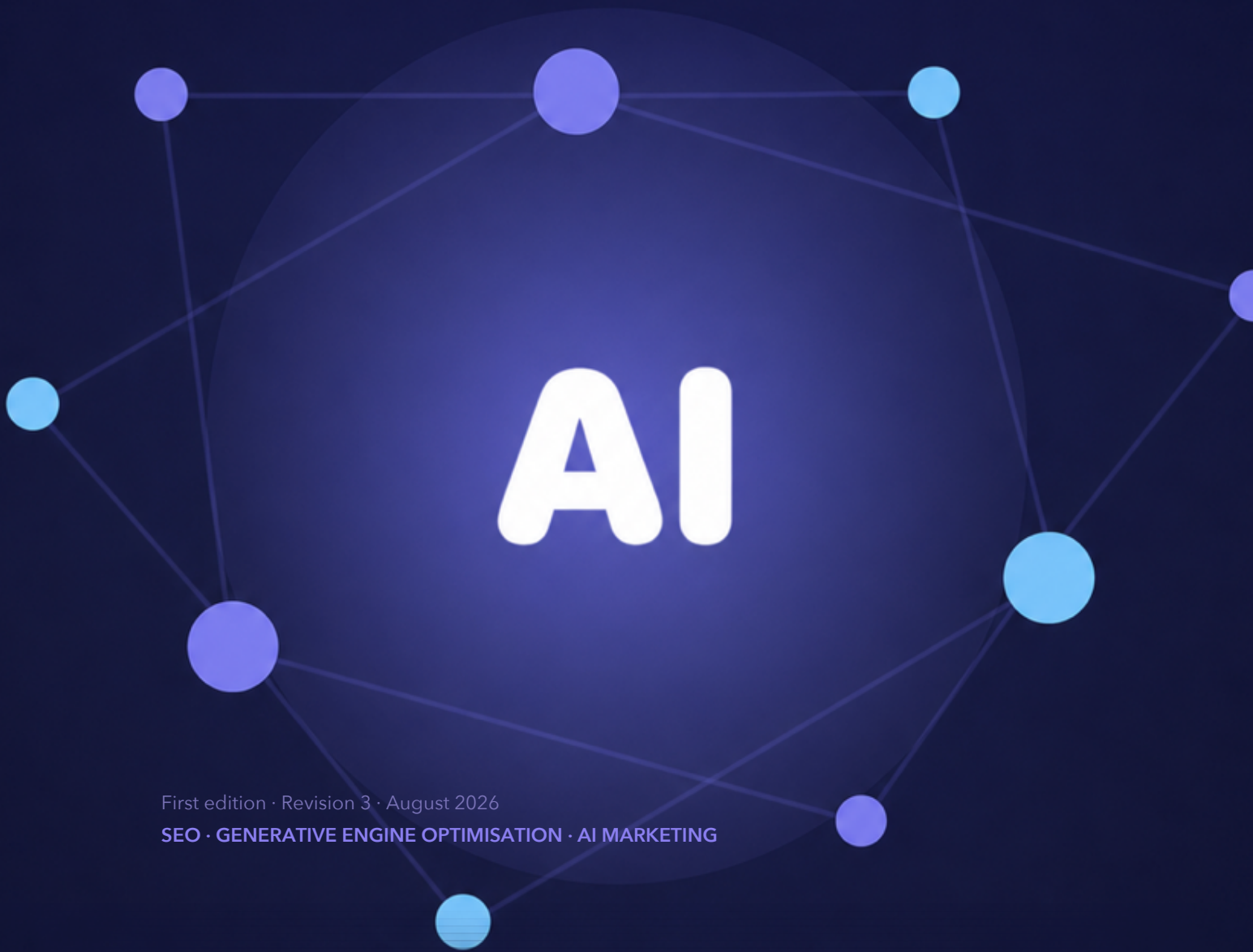
A COMPLETE, SELF-CONTAINED GUIDE

The AI Search Handbook

How search really works now — and how to be found in it

LUKE MCLAUGHLIN



First edition · Revision 3 · August 2026

SEO · GENERATIVE ENGINE OPTIMISATION · AI MARKETING

About this book

This handbook was compiled in August 2026 for a reader with no coding background who wants the complete skill set behind search visibility in the age of generative engines. It is designed to be self-contained: every concept it depends on is explained inside it, so you never need to leave the page to follow an argument.

It is deliberately unexcited. AI search has generated an enormous amount of confident advice, much of it untested. Where the evidence is strong, this book says so. Where a widely repeated tactic has no support — or has been actively contradicted by data — it says that too, and explains how to tell the difference yourself. That habit will outlast any individual tactic in these pages.

How to read it

Parts I and II build the mental model. Do not skip them: almost every bad piece of AI-SEO advice comes from misunderstanding that machinery. Parts III to V are the practical craft — technical foundations, content, authority. Part VI covers selling, and the visitor a citation earns you. Part VII is measurement, where most practitioners are weakest. Part VIII turns it into audits and a 90-day plan. Part IX is instrumentation: measuring citation properly, building the tracker yourself, and reading your server logs as evidence. Part X looks further out, and Part XI is reference — glossary, checklists, templates, worked answers and an index.

Few people read a handbook of this length front to back. Four routes through it:

If you have	Read
One afternoon	Chapters 1, 4, 8, 13 and 29 — the mental model, the pipeline, the access check, the writing craft, and how to measure any of it.
A week	Parts I to IV in order, then the audit in Chapter 32. That is the core of the discipline.
A specific problem	Start at Chapter 32. Each audit stage names the chapter that covers what it finds.
An interview	Chapters 1, 7, 28 and 38 — what changed, how to judge a claim, what can honestly be measured, and how to say so out loud.

The running example

A single company runs through the whole book. Kettleworth is a Leeds specialty coffee equipment retailer — twelve people, 2,200 products, an editorial section that used to be their main acquisition channel until an AI Overview started answering the questions it was built to win. Each chapter shows what they did, including the things that failed and the sprint they wasted. Follow that thread and you have one situation worked end to end.

Kettleworth is a composite, not a client. It is built from patterns that recur across real engagements, but the company does not exist and none of its figures are data. They are illustrative — chosen to show the shape of a problem and the size of a plausible result, not to be quoted. This book spends a chapter teaching you to ask what was actually measured, so it owes you the same distinction about itself: the boxes marked **By the numbers** are sourced and dated, and the case study is a teaching device.

It is deliberately consumer e-commerce, because that exercises the technical, content, commercial and international material at once. If you work in B2B the mechanics transfer unchanged but the platforms do not, and the chapters say so where it matters: Chapter 22 on platforms, and B2B translations in Chapters 26 and 29.

The furniture

Element	What it means
Key idea	The load-bearing concept in a section. If you remember one thing, this.
Myth vs reality	A widely repeated claim, checked against the evidence.
In practice	A worked example with real numbers or real copy.
Common mistake	The specific error most people make here, and the fix.
By the numbers	Data, with its source and its limitations stated.
Case study	The running Kettleworth example, chapter by chapter.
Chapter summary	Takeaways in compressed form, for revision.
Check yourself	Questions to test whether it actually landed.
Answers	Worked responses to the self-checks, in Chapter 60.

A note on shelf life

The mechanics in Parts I-IV change slowly; retrieval, chunking and entities will still matter in five years. The specific products, percentages and interface behaviours in Parts II and VII will date quickest. Every time-sensitive figure carries its date and source, so you can re-check it rather than trust it indefinitely.

Corrections and the current build live at **github.com/lukekevinmclaughlin-oss/SEO-Ai-Marketing**. If you find an error, that is where to report it, and where a corrected edition will appear.

Contents

Part I: Foundations 21

1. What Actually Changed . . . 22

The shift from ranking to retrieval . . . 22

What the data actually shows . . . 24

Two games, one board . . . 25

Why "AI SEO" has so much bad advice in it . . . 27

2. How Search Engines Actually Work . . . 30

Stage one: crawling . . . 30

Stage two: indexing . . . 31

Stage three: ranking . . . 32

The vocabulary you now have . . . 33

3. How Language Models Actually Work . . . 36

What the model is actually doing . . . 36

Tokens . . . 36

Embeddings: meaning as position . . . 37

The context window . . . 38

Training versus inference: the distinction that matters most . . . 38

Why models hallucinate . . . 39

4. How AI Search Systems Work . . . 42

Stage 1: The question arrives . . . 42

Stage 2: Query fan-out . . . 43

Stage 3: Retrieval . . . 44

Stage 4: Chunking — the passage is the unit . . . 44

Stage 5: Reranking and selection . . . 45

Stage 6: Synthesis and citation . . . 46

Part II: The Landscape 49

5. Google: AI Overviews and AI Mode . . . 50

AI Overviews versus AI Mode . . . 50

How an AI Overview gets built . . . 51
What happens to the organic results . . . 52
The controls you actually have . . . 52

6. The Assistants . . . 55

The five that matter . . . 55
When each engine decides to retrieve . . . 56
What you can actually capture from each . . . 58
The differences that change what you do . . . 58

7. Who Gets Cited, and How to Judge a Claim . . . 62

What the citation studies show . . . 62
A method for judging any claim . . . 63
Running your own tests . . . 65

Part III: Technical Foundations 67

8. Crawlability and the AI Crawlers . . . 68

The two jobs crawlers do . . . 68
The crawlers, precisely . . . 69
Reading `robots.txt` . . . 69
Beyond robots.txt: the blocks you did not know about . . . 71
Should you block AI crawlers at all? . . . 72

9. Rendering, JavaScript and What Bots Can Read . . . 74

What rendering means . . . 74
How to check what a bot sees . . . 75
The specific patterns that break things . . . 76
What to ask a developer for . . . 77

10. Architecture and Internal Linking . . . 79

What internal links actually do . . . 79
Topic clusters, and why fan-out makes them work harder . . . 80
Content modelling: deciding what your pages *are* . . . 81
URLs, headings and machine-readable context . . . 82
Depth and orphans . . . 83

11. Structured Data and Schema . . . 85

What it is . . . 85
What it actually does . . . 86
The types that carry weight . . . 87

12. llms.txt and the Emerging Standards . . . 90

The proposal . . . 90
What the data shows . . . 90
How to think about emerging standards . . . 92

Part IV: Content 95

13. Writing for Retrieval . . . 96

The test . . . 96
The answer-first pattern . . . 96
Specificity is now a technical requirement . . . 98
Formatting that helps . . . 99
Restructuring without breaking what works . . . 99

14. Entities and Topical Authority . . . 102

What an entity is . . . 102
How to become a recognised entity . . . 103
Topical authority, mechanically . . . 104
The consistency check . . . 105

15. E-E-A-T, Trust and Why It Now Matters More . . . 107

Why it matters more in AI search . . . 107
Experience — the winnable one . . . 108
Expertise . . . 108
Authoritativeness . . . 109
Trust . . . 110

16. Citable Assets . . . 112

Why original data wins . . . 112
Formats that earn citations . . . 113
Producing research without a budget . . . 113
Making an asset easy to cite . . . 115

17. Content Decay and the Refresh Discipline . . . 117

Why decay happens faster now . . . 117
The audit . . . 118
Consolidation . . . 119
The publishing ratio . . . 119

18. International and Multilingual AI Search . . . 121

Why non-English AI search is weaker . . . 121
Citation sources do not travel . . . 122
Translation versus transcreation . . . 123
The technical layer . . . 124
Measuring across markets . . . 125

19. Multimodal GEO . . . 127

The general rule . . . 127
Images . . . 128
Video . . . 129
Audio and podcasts . . . 129
Voice, and why it is different . . . 129
Knowledge panels and the knowledge graph . . . 130

Part V: Authority Beyond Your Site 132

20. Links, Mentions and What Changed . . . 133

What links still do . . . 133
Why mentions now carry weight . . . 133
Co-occurrence . . . 134
What link building is worth now . . . 135
Where to redirect effort . . . 135

21. Share of Model and Brand Presence . . . 138

Share of model, defined . . . 138
How brand strength feeds generative answers . . . 138
The brand search feedback loop . . . 140

22. The Platforms That Feed the Models . . . 142

Reddit . . . 142
Wikipedia . . . 143

YouTube . . . 143
LinkedIn . . . 144
Which platform for which business . . . 144

23. Digital PR for AI Visibility . . . 147

What changes . . . 147
What makes a story citable . . . 147
Working with journalists . . . 148
Measuring PR against AI visibility . . . 149

24. When the Answer Is Wrong About You . . . 151

Five causes, five different fixes . . . 151
Tracing the wrong answer . . . 152
The correction ladder . . . 152
When it is damaging: the crisis playbook . . . 154
What you cannot fix . . . 154

Part VI: Commerce and Place 157

25. E-commerce and Product Discovery . . . 158

Two pipelines, not one . . . 158
The machine-readable layer . . . 159
Reviews: the most valuable invisible asset . . . 161
Winning the comparison query . . . 162

26. After the Citation . . . 164

A different arrival state . . . 164
What the page needs to do . . . 165
The re-answering mistake . . . 166
Measuring it . . . 167

27. Local and Service-Area GEO . . . 169

Why local is different . . . 169
The three signal families . . . 170
How assistants answer local questions . . . 171
Multi-location businesses . . . 171

Part VII: Measurement and Proof 174

28. The Measurement Problem 175
What you lost 175
The four dimensions 175
Why proxies are legitimate 176
Talking about uncertainty 177

29. Building a Measurement System 180
Step 1: Build the prompt set 180
Step 2: Establish the baseline 182
Step 3: Compute the metrics 183
Step 4: The monthly routine 183
Step 4b: How many runs, and when a movement counts 183
Step 4c: Recording the configuration, not just the engine 184
Step 5: Supporting instrumentation 185
Do you need a paid tool? 185

30. Attribution and Proving Value 187
The attribution gap 187
Building the case 188
Leading with the right number 189
When you are asked for a number you do not have 189

31. Testing and Experiment Design 192
Why you cannot A/B test AI search 192
Designing a change you can attribute 193
The confounds that will ruin it 194
Sample size, when the population is a prompt set 195

Part VIII: Practice 197

32. The Audit Playbook 198
Stage 0: Establish the baseline (half a day) 198
Stage 1: Access (one hour, catches the most catastrophic problems) 198
Stage 2: Readability (one hour) 199
Stage 3: Structure (half a day) 199

Stage 4: Entity and trust (two hours) . . . 199
Stage 5: Content coverage (half a day) . . . 200
Stage 6: Off-site (two hours) . . . 200
Migrations and replatforming . . . 200
Turning findings into priorities . . . 201
From finding to verified . . . 201

33. The GEO Toolkit . . . 205

The free tier does most of the work . . . 205
What paid tools add . . . 206
Assembling a stack rather than a subscription pile . . . 207
Automation: what to and what not to . . . 207
What an API is, and whether you need one . . . 208

34. Using AI to Do the Work . . . 210

Where models genuinely help . . . 210
Where they quietly damage you . . . 211
Detection, disclosure and the quality bar . . . 212
A working process . . . 213
Prompting well enough . . . 213

35. The Prompt Library . . . 215

Research and planning . . . 215
Auditing and improving content . . . 216
Measurement . . . 218
Technical . . . 218
Strategy . . . 219
What good output looks like . . . 219

36. Strategy and the First 90 Days . . . 222

The sequencing principle . . . 222
Days 1-30: Baseline and access . . . 222
Days 31-60: Structure and foundation . . . 223
Days 61-90: Authority and proof . . . 223
Deciding what not to do . . . 224
Adapting to your situation . . . 224

37. The GEO Practitioner's Week . . . 226

The cadence . . . 226
Allocating the time . . . 227
How the year unfolds . . . 228
Solo versus in a team . . . 228
Skills, in the order worth building them . . . 229

38. Working with Stakeholders . . . 231

The four stakeholders you will meet . . . 231
The objections, answered . . . 232
Briefing an agency or freelancer . . . 233
Setting expectations that survive . . . 234
Positioning yourself . . . 234

Part IX: Instrumentation 236

39. Measuring Citations in Practice . . . 237

Three different things people call a citation . . . 237
Recording an answer so it can be re-examined . . . 238
Finding the passage that was lifted . . . 239
The three numbers worth reporting . . . 240
What a defensible claim sounds like . . . 241

40. Building a Citation Tracker . . . 243

Why build rather than buy . . . 243
What you are building . . . 244
Step 1: Set up the sheets . . . 244
Step 2: Add the script . . . 245
Step 3: Store the key safely . . . 247
Step 4: Run it, then schedule it . . . 247
Step 5: Turn rows into the three numbers . . . 248
What this will not do . . . 249

41. Watching the Crawlers . . . 251

Getting hold of the logs . . . 251
Reading a log line . . . 252
Verifying a crawler is genuine . . . 253
Connecting crawl to citation . . . 254

When you cannot get logs at all 254

Part X: Forensics 256

42. Evidence Status: Actual, Inferred, Unknown 257

The failure this prevents 257

The three labels 258

Applying it to what you already do 258

Saying it out loud 259

43. Citation Decision Forensics 262

Why the earlier exercise is not enough 262

The thirteen-part audit 263

Scoring a source, and scoring the one that lost 264

Turning it into work 266

44. Claim-Level Validation 268

The gap between adjacency and support 268

What outruns the evidence 269

The hostile audit 270

When it is your page being misused 271

45. Source Lineage and Independence 273

The collapse 273

Tracing a chain 274

What to do with the finding 275

46. Counterfactual Testing 277

The question these tests answer 277

The six reruns 278

Reading the pattern 279

47. The Correction Runbook 282

Twelve steps 282

The three requests 284

The incident log 285

48. The Evidence Archive 288

Why this decays faster than you expect . . . 288
What to preserve, per answer . . . 289
A structure that survives handover . . . 289
How much of this to do . . . 290

Part XI: Playbooks 293

49. Copy-Paste Implementation . . . 294
The robots.txt . . . 294
The four schema blocks . . . 295
Where it goes, by platform . . . 297
The one command . . . 298

50. The Page Teardown . . . 301
The page before . . . 301
The page after . . . 302
The writer's rubric . . . 303

51. GEO at Scale . . . 306
Where the method breaks . . . 306
Finding the work . . . 306
Bulk transformation with a human gate . . . 307
Structured data by template and by feed . . . 307
Decay at scale . . . 308

52. Industry Playbooks . . . 310
B2B software . . . 310
Local and service-area . . . 311
Health, finance and legal . . . 311

53. The Testing Lab . . . 314
The retrieval simulator . . . 314
The citation-likelihood check . . . 315
Fan-out with gap mapping . . . 316
The sceptical reviewer . . . 317

54. The Capstone . . . 319
Choosing the subject . . . 319

The ten deliverables . . . 319
Judging it . . . 320

Part XII: Horizons 323

55. Agentic Search . . . 324
What an agent actually is . . . 324
What breaks . . . 325
What is genuinely new . . . 325
How an agent chooses between options . . . 326
What it does to content . . . 327
Recognising agent traffic . . . 327
The economics nobody has settled . . . 328
An honest assessment of the uncertainty . . . 329

56. Ethics and Second-Order Effects . . . 331
Concentration risk . . . 331
The broken value exchange . . . 332
Built-in bias . . . 332
The feedback loop . . . 333
The case for, stated fairly . . . 333
When you are asked to do something you think is wrong . . . 334
Where the regulation is heading . . . 334
What one practitioner can actually change . . . 335
Holding a position . . . 335

Part XIII: Reference 338

57. Glossary . . . 339
A-C . . . 339
D-I . . . 340
J-R . . . 341
S-Z . . . 342

58. Checklists . . . 344
The ten-minute emergency check . . . 344
Crawler access . . . 344

Rendering . . . 344
Page structure . . . 345
Entity and trust . . . 345
Structured data . . . 345
Measurement . . . 346
Content refresh . . . 346

59. Templates . . . 347

Prompt set tracker . . . 347
Monthly measurement note . . . 350
Quarterly stakeholder report . . . 351
Developer request . . . 351
Content brief . . . 352

60. Answer Key . . . 353

Part I — Foundations . . . 353
Part II — The Landscape . . . 355
Part III — Technical Foundations . . . 356
Part IV — Content . . . 359
Part V — Authority Beyond Your Site . . . 362
Part VI — Commerce and Conversion . . . 364
Part VII — Measurement . . . 366
Part VIII — Practice . . . 368
Part IX — Instrumentation . . . 372
Part X — Forensics . . . 374
Part XI — Playbooks . . . 378
Part XII — Horizons . . . 381

61. Sources and Further Reading . . . 383

Primary evidence used . . . 383
How to keep this current . . . 384
A closing note . . . 384

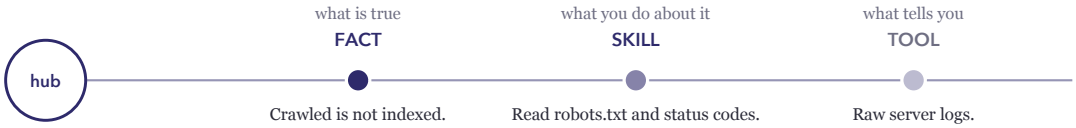
62. Index . . . 386

The book on three webs

Everything that follows hangs on three frames. Each has six strands, and each strand is read from the hub outward: the **fact** that is true, the **skill** it demands, and the **tool** that answers it. The three rings carry that structure — they are not decoration.

Classic SEO is the qualifying round. **GEO** is the layer that turns qualifying into being named. **The overlap** is where one piece of work is paid by both, and it is the frame to run first.

Come back to these whenever a chapter starts to feel like detail. Every one of the 62 chapters is an expansion of a single strand.



Every strand of every web reads this way, hub outward. The example above is strand 1 of the first web.

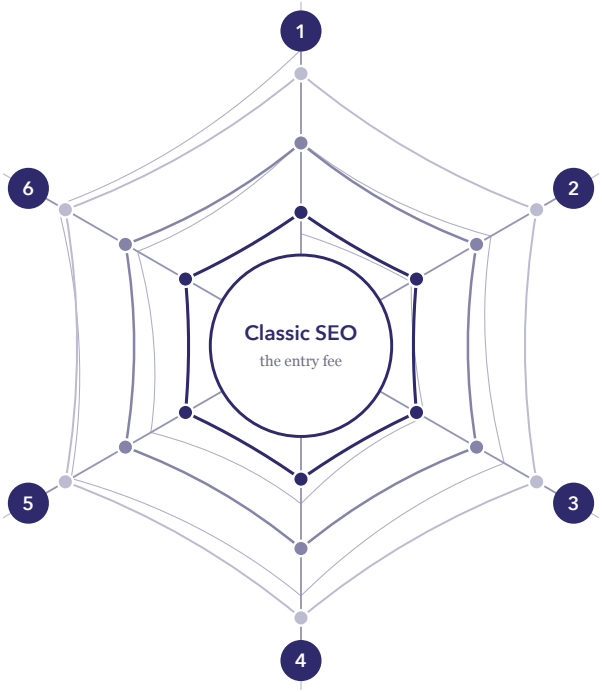
KEY IDEA

Where the three webs meet

Five gates stand between your content and a citation: **crawled**, **indexed**, **retrieved**, **selected**, **cited**. The first web carries you through the first three. The second converts the last three.

Retrieved is the hinge, because it is the only gate both games are playing for. That is why the third web is the one to run first — and why this book is not called *Replace SEO With GEO*.

The first web — classic SEO. The qualifying round. It no longer pays on its own, and nothing downstream happens without it.



1 CRAWL
- Crawled is not indexed, and indexed is not ranked.
- Read robots.txt and status codes before anything else.
- Raw server logs

2 INDEX
- "Crawled — currently not indexed" means it read the page and declined.
- Fix canonicals; consolidate the thin pages.
- Search Console › Pages

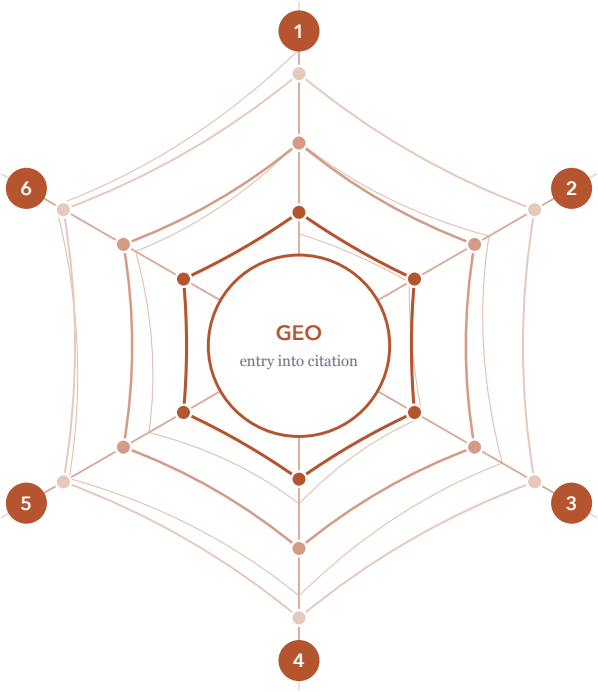
3 RANK
- There is no single algorithm. Weighting is query-dependent.
- Match the intent, not the keyword.
- Search Console › Performance

4 ARCHITECTURE
- An orphan page is invisible however good it is.
- Cut click depth; build real topic clusters.
- A site crawler, above ~200 pages

5 LINKS
- Quality dominates quantity. Links are discovery and endorsement at once.
- Earn contextual mentions; stop buying volume.
- An SEO platform, for competitors

6 RENDERING
- Googlebot renders JavaScript, but on a delay.
- Ask for the content in the initial HTML.
- View Page Source

The second web — GEO. The layer that turns candidacy into being named. All of it assumes the first web already works.



1 THE PIPELINE
- One question becomes 8–12 sub-queries before anything is retrieved.
- Cover the sub-questions, not the phrase.
- A model, to draft the fan-out

2 THE PASSAGE
- The passage replaced the page as the unit of competition.
- Answer first; make each section survive being lifted out alone.
- The lift-out test

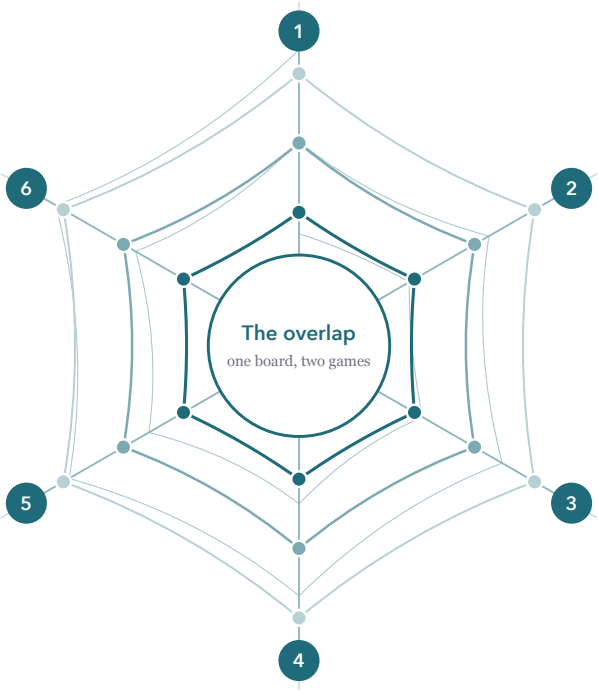
3 ENTITIES
- Consistency across independent sources is the whole mechanism.
- Align every profile; put named people behind the work.
- Wikidata, schema sameAs

4 CITABLE ASSETS
- Original data competes with nothing, because it is the source.
- Survey your own audience. Never gate the finding.
- A spreadsheet

5 PLATFORMS
- Reddit is 24% of Perplexity's citations and 0.1% of Gemini's.
- Participate as a person, disclosed, roughly thirty to one.
- The assistants, logged out

6 MEASUREMENT
- There is no rank, no impression count, and often no click.
- Fix 40–60 prompts and never change the set.
- Share of model, citation rate, concentration

The third web — the overlap. Where one piece of work is paid twice. Every strand here is load-bearing for both games.



1 INDEX MEMBERSHIP
- The generative layer retrieves its candidates from the classical index.
- Win the qualifying round; you just no longer get paid for winning it.
- Search Console › Indexing

2 CRAWLER ACCESS
- CDN bot rules block AI crawlers silently — the leading cause of unexplained invisibility.
- Refuse training, permit retrieval, then check it held.
- Raw server logs

3 RENDERING
- Most AI crawlers run no JavaScript at all.
- Put it in the initial HTML response.
- View Page Source

4 STRUCTURE
- Topic clusters map directly onto query fan-out.
- Phrase every heading as the question it answers.
- Rich Results Test

5 TRUST
- Self-contradiction across sources costs you in both games.
- E-E-A-T — experience is the winnable one.
- A quarterly decay audit

6 EVIDENCE
- Only topical relevance and context position replicate.
- Run the five-question filter before acting on any claim.
- A matched, untouched control group

Quick reference

The book compressed onto one page. Each line names the chapter that explains it.

1 **Can the AI crawlers reach you?**
robots.txt — and ask about CDN bot rules · Ch 8

2 **Can they read the page?**
View Page Source. Price and reviews in the HTML? · Ch 9

3 **What do assistants say about you?**
Five questions, logged out, four engines · Ch 3

4 **What is your actual exposure?**
Share of impressions on triggering queries · Ch 5

5 **Does each section stand alone?**
Lift it out. Does it still answer something? · Ch 13

6 **Is the answer in the first sentence?**
Heading asks; first line answers · Ch 13

7 **Are your facts consistent everywhere?**
Site, LinkedIn, directories, registry · Ch 14

8 **Which third parties do engines cite?**
Fix your accuracy on those pages first · Ch 21

9 **Do you have a fixed prompt set?**
40-60 questions, never changed once set · Ch 29

10 **Can you prove your work caused it?**
A matched control group, left untouched · Ch 31

KEY IDEA

The test that governs everything

If a machine lifted this section out and showed it to a stranger with nothing else from the page — would it make complete sense, and would it answer something? Everything else in this book is downstream of that.

Three numbers worth carrying

	What it is	The caveat that matters
~50%	of US Google queries show an AI Overview (Google's own figure, Feb 2026)	Trackers range 25-60%; health runs near 88%. Calculate your own.
68%	of US searches ended without a click in early 2026 (SparkToro)	Up from 60% in 2024. The click was always the exception.
~5%	the ceiling any single domain reaches in an engine's citations	A long tail, not winner-takes-all. There is room.

PART I

Foundations

The machinery underneath. Almost every bad piece of AI-SEO advice comes from misunderstanding one of these four chapters.

CHAPTER 1

What Actually Changed

Why "the click is dying" is the wrong headline, and what the real shift is

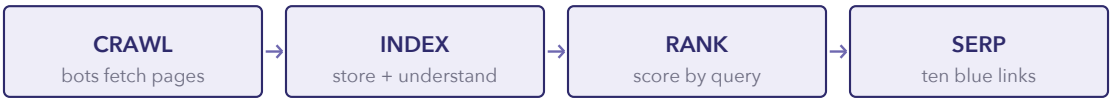
WHAT YOU'LL LEARN

- The single structural change that explains almost everything else: retrieval replacing ranking as the thing you compete in
- What the click-through and zero-click data genuinely shows, and where the widely quoted numbers are softer than they look
- Why **being cited** is now a distinct objective from **being ranked**, requiring different work and different measurement
- How to hold both models in your head at once, because the old one has not gone away

For twenty-five years, search worked one way. You typed some words. A system matched those words against a store of pages. It handed you a list, and you chose. Your entire job as a marketer was to get onto that list, as high as possible.

That model has not disappeared. It still runs underneath everything in this book. But a second layer now sits on top of it, and that layer does something categorically different: instead of handing you a list to choose from, it reads the list on your behalf and writes you an answer.

The consequences of that one change ripple through every practice in this field. It changes what you write, how you structure it, what counts as authority, and — most awkwardly for anyone with a job to defend — what you can measure.



The user chooses which result to click. Your page is the destination.

Figure 1.1 – The classical pipeline. Four stages, one output: a list of destinations you compete to appear on.

The shift from ranking to retrieval

In classical search, the unit of competition is the **page**. A page has a rank. That rank is a position on a list, and everyone can see it. You could open a browser, type a query, and count down until you found yourself. The whole discipline of SEO grew around that single

observable number.

In generative search, the unit of competition is the **passage**. A passage is a chunk of a page — usually a few hundred words — that a system has decided is relevant to one specific facet of a question. Passages do not have ranks. They are either retrieved into the answer-writing process or they are not, and you usually cannot see which.

Dimension	Classical search	Generative search
Unit of competition	The page	The passage
What the user receives	A list of links	A written answer
Your objective	Rank in the top three	Be retrieved, then cited
Primary feedback signal	Clicks, impressions, position	Citations, mentions, share of answers
Visibility of your standing	Directly observable	Inferred from sampling
Winner structure	Roughly winner-takes-most	Long tail; several sources per answer

> KEY IDEA
>
> The unit of competition changed from the **page** to the **passage**. Almost everything else in this book — how you structure content, what technical work matters, how you measure — is downstream of that single sentence.

Why this makes small sites competitive again

Here is the consequence that surprises people most. In classical search, a large, authoritative publisher with thousands of links beats a small consultancy almost every time on a competitive query. Authority is applied at the level of the whole domain and the whole page.

In passage-level retrieval, the competition is narrower. The system is not asking "which page is best about CRMs?" It is asking "which passage best answers *what does a CRM cost for a five-person agency?*" A 600-word post from a small consultancy that answers exactly that question, in a clean self-contained paragraph, can be retrieved ahead of a 4,000-word guide from a major publisher in which the same fact is buried in the middle of a section about something else.

This is not a guarantee, and authority still matters a great deal — we will come back to that repeatedly. But it is a real structural opening, and it is the reason a small operator can compete in AI search in a way that had become almost impossible in classical search.

MYTH vs REALITY

"Long-form content always wins"

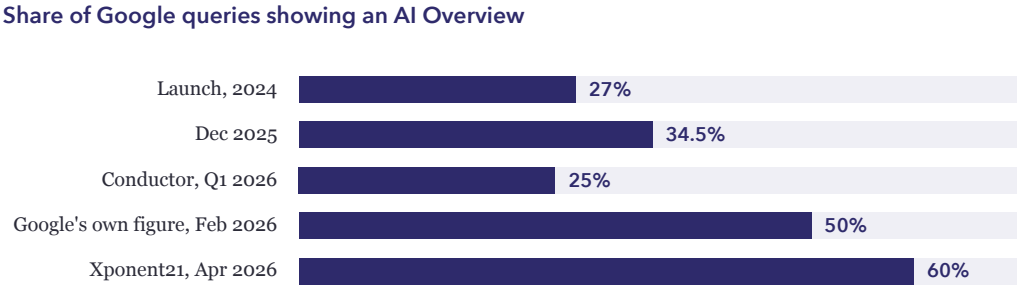
Length was never the causal factor. Long content correlated with thoroughness, and thoroughness correlated with rankings, but the length itself did nothing.

In passage-level retrieval, length can actively work against you. A diffuse 3,000-word page may contain no single self-contained passage that cleanly answers any specific sub-question. A tight 800-word page built from four clean sections may contain four.

The fix is structure, not brevity. A long page made of well-formed, self-contained sections beats both the rambling long page and the thin short one.

What the data actually shows

You will see a lot of alarming statistics about AI search. Most of them point in the same direction, and that direction is real. But the precision implied by the numbers is usually false, and you should understand why before you quote any of them to a boss.

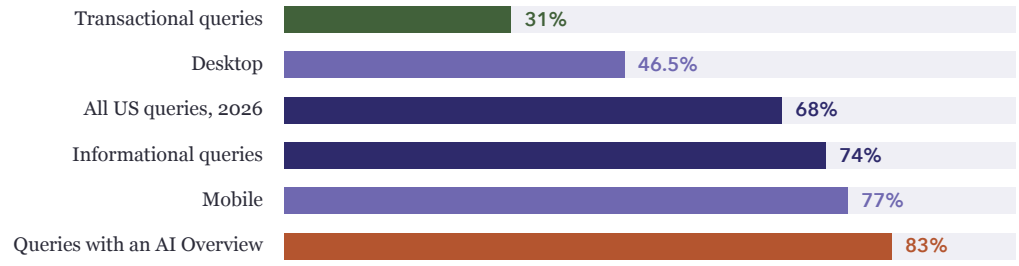


Trackers disagree because query mix differs. Google's own number is the safest to quote.

Figure 1.2 – The growth of AI Overviews on Google. Trackers disagree on the level; none disagree on the trend.

The spread between trackers is enormous — Conductor measured 25% in Q1 2026, Xponent21 60% in April, and Google itself says roughly 50%. That gap is not incompetence. It is a methodological problem: the figure depends entirely on **which queries you sample**. A tracker monitoring commercial keywords for e-commerce clients will see a very different rate from one monitoring health queries, where trigger rates run near 88%. There is no neutral query set, so there is no single true number — which is why Chapter 5 asks you to calculate your own exposure rather than adopt anyone else's headline.

Zero-click rate: searches that end without a click



The click was always the exception. Your exposure depends entirely on your query mix.

Figure 1.3 – Zero-click rates. Note that the majority of searches ended without a click long before AI arrived.

BY THE NUMBERS

Figures worth knowing, each with its caveat:

- **~50%** of US Google queries show an AI Overview, per Google's own statement in February 2026 — up from ~34.5% in December 2025. *Caveat: independent trackers disagree sharply, from Conductor's 25% to Xponent21's 60%. The spread is a sampling artefact, not incompetence.*
- **68%** of US Google searches ended without a click across the first four months of 2026, up from 60.45% in 2024 (SparkToro). On queries that trigger an AI Overview the figure is **~83%**. *Caveat: "zero-click" includes searches that were never going to produce a click, such as checking the weather.*
- Zero-click varies enormously by intent: **74%** on informational queries against **31%** on transactional. By device, **77%** on mobile against **46.5%** on desktop.
- **30-50%** average reduction in organic click-through rate for the top three positions on AI-Overview queries, per Semrush and Ahrefs analyses. *Caveat: averages hide enormous variance; reported declines range from 15% to 89% by query type.*
- Sites cited *inside* an AI Overview report roughly a **35%** lift in organic clicks and a **15%** lift in branded search. *Caveat: this is correlational. Sites good enough to be cited may have been growing anyway.*

The honest summary is this: **clicks per search are falling, the fall is steepest on informational queries, and being cited partially offsets it.** That is the defensible version. Anything more precise than that is overclaiming.

COMMON MISTAKE

The most common error in this field is quoting a single dramatic statistic — "AI killed 65% of organic traffic" — without its query-mix caveat. When a stakeholder later sees a different number from a different tracker, your credibility takes the damage, not the tracker's. Always state the range and the source. It makes you look more rigorous, not less certain.

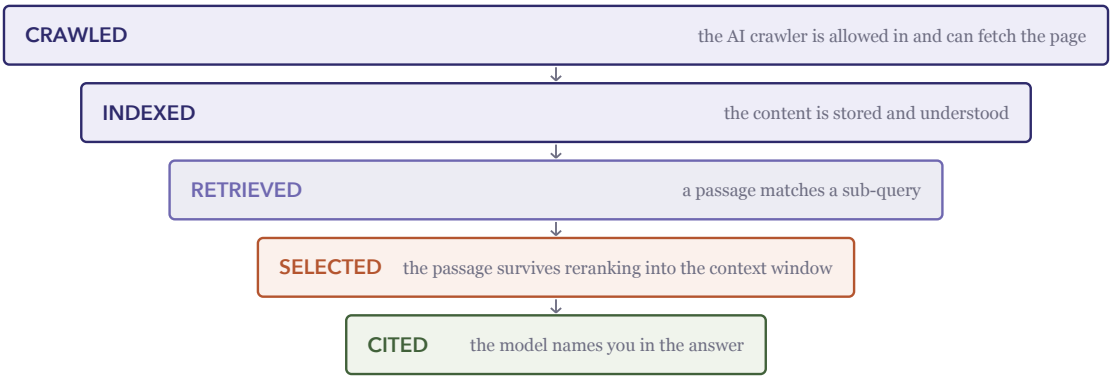
Two games, one board

The mistake that follows from all this is to conclude that classical SEO is over and should be abandoned. It is not, for a reason that becomes obvious once you understand the machinery in Chapter 4: **the generative layer gets its candidate passages from the classical index.**

Google's AI Overviews retrieve from Google's search index. ChatGPT's search mode retrieves from a web index. Perplexity retrieves from an index. If your page cannot be crawled, is not indexed, and does not rank for anything, it is not in the candidate pool from which the answer is assembled. Classical SEO is now the entry fee rather than the prize.

> *Classical SEO stopped being the whole game and became the qualifying round. You still have to win it. You just no longer get paid for winning it.*

Five gates. Failing any one makes you invisible.



Most SEO effort targets gates 2 and 3. Gates 4 and 5 are where AI search is won.

Figure 1.4 – The five gates between your content and a citation. Traditional SEO gets you through gates one to three.

This is why the framing of this book is not "replace SEO with GEO". It is: keep doing the foundational work, because it now buys you entry rather than results, and then add the layer of work that converts entry into citation.

IN PRACTICE

What this changes on Monday morning

Concretely, for a page you already have ranking on page one:

1. **Do not** delete it, shorten it drastically, or stuff it with statistics because a blog post said statistics improve GEO.
2. **Do** open it and ask of each section: *if a machine lifted this section out and showed it alone, would it make sense and answer something?*
3. **Do** rewrite the sections that fail that test so the answer comes first and the context comes second.
4. **Do** check that the AI crawlers are allowed to fetch it (Chapter 8) — a surprising number of sites block them by accident.

That sequence costs an afternoon and addresses gates three, four and five without touching anything that is currently working for you.

Why "AI SEO" has so much bad advice in it

It is understanding why this field is unusually full of confident nonsense, because it will help you filter everything — including this book.

The foundational academic work on Generative Engine Optimisation came from a 2023 paper by researchers at Princeton, Georgia Tech, the Allen Institute and IIT Delhi. It reported that certain content changes — adding statistics, adding citations, adding quotations — could lift a source's visibility in generated answers by up to 40%. That number spread everywhere.

A critical survey published in July 2026, reviewing 45 studies across the field, found that the original result was valid **within its experimental setting** but conditional on a source already being present in a fixed context window. In plain terms: the study showed that *if a passage has already been retrieved*, certain edits make it more likely to be quoted. It did not show that those edits help you get retrieved in the first place, and it did not measure whether any of it produced durable traffic.

The same survey found that only two levers showed reproducible effects across studies — **topical relevance** and **position within the retrieved context**, and concluded that no reviewed technique demonstrated a stable, longitudinal, cross-platform causal effect on organic discoverability.

MYTH vs REALITY

"Adding statistics to your content boosts AI visibility by 40%"

What was shown: among passages already retrieved into a model's context, those containing statistics and citations were quoted more often, in one experimental setup.

What was not shown: that adding statistics helps a passage get retrieved; that the effect holds across different engines; that it persists once competitors do the same thing; or that it produces traffic.

What to do instead: add statistics when they genuinely make the passage a better answer. That is defensible on its own merits and happens to align with the evidence. Do not bolt numbers onto content as a ranking tactic.

This is the pattern to watch for throughout your career in this field: a real but narrow finding, stripped of its conditions, sold as a universal tactic. Chapter 7 gives you a method for evaluating claims like this yourself.

CASE STUDY · KETTLEWORTH

Meet Kettleworth: a Leeds-based specialty coffee equipment retailer. Twelve people, about 2,200 products, and an editorial section of buying guides that had been their main acquisition channel for six years.

In January their organic traffic fell 34% year on year. Rankings had barely moved — they still held position two or three for their biggest terms. What had changed was that those terms now returned an AI Overview, and the summary answered the question.

Their first instinct was that they had been penalised. They had not. They had been summarised. This book follows what they did about it.

CHAPTER SUMMARY

- The structural change is that **retrieval replaced ranking** as the competition you are in; the passage replaced the page as the unit
- Classical SEO is now the **entry fee**: it gets you into the candidate pool from which answers are assembled
- Click-through is falling, but every headline figure is sample-dependent — quote ranges, not points
- Passage-level retrieval structurally re-opens competition for small sites, because it evaluates sections rather than domains
- The academic evidence for most GEO tactics is far weaker than the marketing around it; only topical relevance and context position replicate reliably

CHECK YOURSELF

1. Explain, without using the word “algorithm”, why a page ranking fourteenth could be cited in an AI answer when the page ranking second is not.
2. A stakeholder says “AI Overviews have destroyed 65% of our clicks.” What two questions would you ask before accepting that number?
3. Why is it a mistake to conclude from this chapter that you should stop doing technical SEO?
4. What was the actual, narrow finding of the original GEO paper, and what did it not show?

CHAPTER 2

How Search Engines Actually Work

The three-stage machine that still decides whether you exist

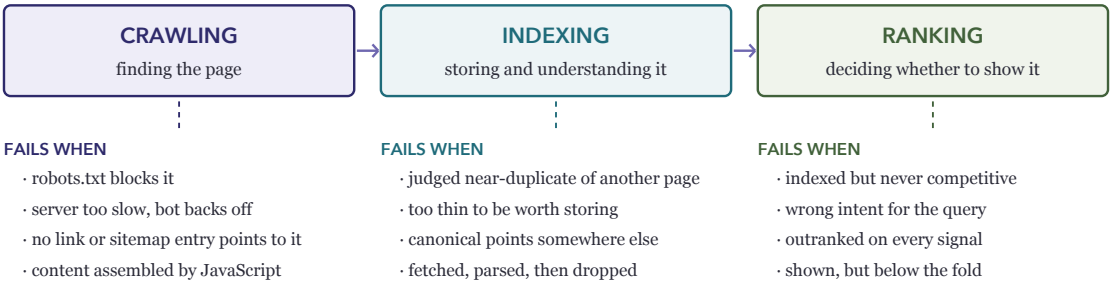
WHAT YOU'LL LEARN

- The three stages every search engine runs — crawling, indexing, ranking, and what can go wrong at each
- What a crawl budget is, and why it matters more on large sites than small ones
- How ranking signals actually combine, and why "the algorithm" is a misleading singular
- Enough of the vocabulary that the technical chapters later in this book read easily

You cannot reason about AI search without understanding classical search, because the generative layer sits directly on top of it and inherits all of its constraints. This chapter is the foundation. If you have no technical background, this is the chapter that makes the rest of the book readable.

A search engine does three things, in order. It finds pages. It stores and organises what it found. It decides what to show for a given query. Everything else is detail hanging off those three.

Three stages, and where each one actually breaks



The stages are sequential, so a failure at stage one makes stages two and three unreachable.
This is why crawl problems outrank content problems: nothing downstream of a block can be fixed by writing better.

Figure 2.1 – The three stages, and the specific failure that stops a page at each one.

Stage one: crawling

A crawler — also called a bot or a spider — is a program that fetches web pages, reads them, extracts the links, and adds those links to a queue of things to fetch next. Googlebot is the best known. There are now dozens that matter, including several operated by AI companies, which we cover in detail in Chapter 8.

Crawling is deceptively simple in concept and full of failure modes in practice:

- **The crawler must be allowed in.** A file at the root of your site called `robots.txt` tells crawlers which paths they may fetch. It is a request, not an enforcement mechanism, but all major crawlers honour it. Misconfiguring this file is one of the most common catastrophic SEO errors.
- **The server must respond.** If your server is slow, or returns errors under load, the crawler backs off and fetches less.
- **The crawler must be able to find the page.** A page with no internal links pointing to it and no entry in a sitemap is effectively invisible, however good it is.
- **The crawler must be able to read the content.** If your content is assembled in the browser by JavaScript, the crawler has to run that JavaScript to see it — an expensive extra step that not every crawler performs. This matters enormously for AI crawlers, most of which do *not* run JavaScript. Chapter 9 covers this.

Crawl budget

Search engines allocate a finite amount of fetching to each site — informally, a *crawl budget*. It is a function of how important the engine thinks your site is and how much load your server can take.

For a 200-page site, crawl budget is irrelevant; everything gets fetched often. For a 500,000-page e-commerce site, it is a central concern: the crawler can spend its whole allocation on faceted-navigation URLs — the filter combinations behind "colour: black, size: medium, sort by price", which generate tens of thousands of near-identical pages. Your actual products then get crawled rarely, and updates take weeks to register.

IN PRACTICE

Spotting a crawl budget problem

In Google Search Console, open **Settings › Crawl stats**. Look at the breakdown by response and by file type. Warning signs:

- A large share of crawl requests returning 404 or 301 — the crawler is burning its allocation on pages that no longer exist
- Crawl requests heavily weighted towards URLs with query parameters (`?colour=red&size=4`) rather than real pages
- Average response time trending upwards over months

The fix is usually some combination of: blocking parameter URLs in `robots.txt`, fixing internal links that point to redirects, and removing dead pages from sitemaps.

Stage two: indexing

Once fetched, a page is processed and stored. This is more than filing it away. The engine parses the HTML, extracts the main content from the boilerplate, works out what the page is about, identifies the entities mentioned, checks whether it duplicates something already stored, and decides whether it is worth keeping at all.

That last point is the one people miss. **Being crawled does not mean being indexed.** Search engines discard an enormous amount of what they fetch — near-duplicates, thin pages, pages that appear to serve no user need. In Search Console this appears as "Crawled — currently not indexed", which is the engine telling you, politely, that it fetched your page and decided against it.

> KEY IDEA
>
> Crawled, indexed and ranking are three different states, and a page can fail at any of them. Diagnosing which one has failed is the first move in almost every technical SEO investigation, and the same logic applies to AI visibility.

Canonicalisation

When several URLs show substantially the same content, the engine picks one to keep, and that choice is called *canonicalisation*. You can express a preference with a `link rel="canonical"` tag, but it is a hint, not an instruction; the engine can and does override it.

This matters because duplicate content splits your signals. If the same article is reachable at four URLs, any links and engagement scatter across all four rather than concentrating on one.

Stage three: ranking

Here is where the mythology accumulates. There is no single "algorithm". There are many systems, running in sequence and in parallel, each contributing a signal, continuously retrained and adjusted. What follows is the honest simplification.

Broadly, ranking systems assess three families of thing:

Relevance — does this page address the query?
Historically this was largely lexical: does the page contain the words in the query? Modern systems are semantic — they work with meaning rather than string matching, which is why exact-match keyword density stopped being a useful lever more than a decade ago. Chapter 3 explains the mechanism behind this.

Quality — is this page trustworthy and useful?
This bundles a great many signals: the quantity and quality of links from other sites, the reputation of the domain, the apparent expertise behind the content, whether the page seems designed for people or for machines. Google describes the human-judgement

framework behind this as E-E-A-T, covered in Chapter 15.

Usability — can a person actually use this page?

Loading speed, layout stability, mobile rendering, and intrusive interstitials — pop-ups and overlays that cover the content a visitor arrived to read. These are real but comparatively weak signals; they act more as tie-breakers than as primary drivers. A slow page with the best answer will usually still beat a fast page with a worse one.

> **COMMON MISTAKE**
>
> A persistent error is treating ranking factors as a checklist with fixed weights — "links are 30%, content is 40%". No such fixed weighting exists. The systems are query-dependent: what determines rankings for *"chest pain causes"* (heavily weighted towards medical authority) is very different from what determines rankings for *"pizza near me"* (heavily weighted towards proximity and freshness). Always ask "for what kind of query?"

What links actually do

A link from another site is two things at once: a path for a crawler to discover you, and a signal that someone thought you were worth pointing at. The second is why links became the currency of SEO.

The useful mental model is that a link passes a portion of the linking page's own standing. A link from a page that is itself widely referenced is worth far more than a link from a page nobody references. This is why link *quality* dominates link *quantity*, and why bought links — which come from pages with no genuine standing — do so little.

Links still matter in AI search, but their role shifts in a way Chapter 20 covers in full: **the unlinked brand mention becomes nearly as valuable as the link**. A language model reading text does not need a hyperlink to learn that your brand is associated with a topic.

The vocabulary you now have

You will meet these terms constantly. Here they are in one place.

Term	What it means
Crawler / bot / spider	A program that fetches pages automatically
`robots.txt`	Root file requesting which paths crawlers may fetch
Index	The engine's processed store of pages
Canonical	The URL chosen as the master copy among duplicates

Term	What it means
SERP	Search engine results page
Rendering	Running a page's JavaScript to see the final content
Sitemap	An XML file listing URLs you want crawled
Anchor text	The visible, clickable words in a link
Nofollow	A link attribute asking engines not to pass standing
Core update	A periodic broad revision of ranking systems

Every term this book uses is defined in the glossary, Chapter 57. If a word in a later chapter is unfamiliar, look there before assuming you have missed something earlier.

CASE STUDY · KETTLEWORTH

Kettleworth's first technical finding had nothing to do with AI. Their crawl stats showed 71% of Googlebot's requests going to URLs like `/grinders?colour=black&burr=conical&sort=price` — the filter combinations on their category pages, generating tens of thousands of near-identical URLs.

Their actual product pages were being crawled roughly once every six weeks. Price changes took a month and a half to register anywhere.

Blocking the filter parameters in `robots.txt` and removing them from internal links cut crawl waste to 18%. Product pages went to a four-day crawl cycle. Nothing about this was new or AI-specific — it was simply overdue, and everything later in this book depended on it.

CHAPTER SUMMARY

- Search engines crawl, index and rank — three distinct stages, each with its own failure modes
- Being crawled does not mean being indexed; being indexed does not mean ranking
- There is no single algorithm and no fixed weighting of factors; systems are query-dependent
- Links function both as discovery paths and as endorsement signals, and quality dominates quantity
- All of this still runs underneath AI search, which is why it remains the entry fee

CHECK YOURSELF

1. A page is not appearing in search at all. What three states would you check, in what order?
2. Why does crawl budget matter for a large e-commerce site but not for a small brochure site?
3. Why is “keyword density” no longer a meaningful lever?
4. Explain why duplicate content is a problem even when no page is penalised for it.

CHAPTER 3

How Language Models Actually Work

Enough of the mechanism to reason about it, with no maths and no code

WHAT YOU'LL LEARN

- What a language model is doing when it produces text, in mechanical rather than mystical terms
- Tokens, embeddings and context windows — the three concepts that explain most model behaviour
- Why models hallucinate, and why that is a structural property rather than a bug to be fixed
- The critical distinction between what a model *learned* and what it *retrieves*, which underpins all of AI SEO

You do not need to understand transformers mathematically to work in this field. You do need to understand four things, because nearly every practical question — why models cite what they cite, why freshness matters, why they get your brand wrong — resolves to one of them.

What the model is actually doing

A large language model is a system that predicts what text should come next. That is the whole of it. Given a sequence of text, it produces a probability distribution over what the next fragment might be, picks one, appends it, and repeats.

It learned to do this by being shown an enormous quantity of text and repeatedly adjusting billions of internal numerical parameters to get better at the prediction. Nothing in that process stores documents. The model does not contain a copy of the web. It contains a compressed statistical representation of the patterns in what it read.

KEY IDEA

A language model does not look things up. It generates text that is statistically plausible given its training. This is why it can be fluent and wrong at the same time, and it is precisely the problem that retrieval, covered in Chapter 4, exists to solve.

Tokens

Models do not read words. They read **tokens** — fragments of text, usually between one character and one word. “Marketing” might be a single token; “generative” might split into “gener” and “ative”; an unusual brand name might split into four or five.

Two practical consequences:

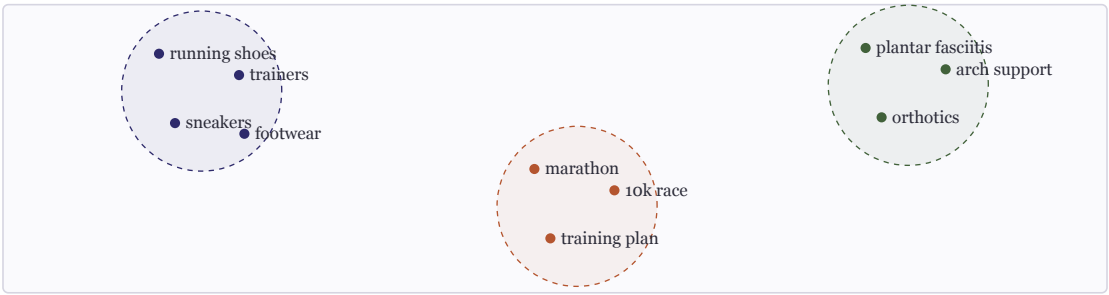
- **Unusual brand names are handled worse than common words.** A name that fragments into several tokens is harder for a model to associate reliably with a concept. This is one reason distinctive-but-pronounceable brand names do better in AI search than deliberately misspelled ones.
- **Everything has a token cost.** Context windows, pricing and processing limits are all measured in tokens, not words. A rough conversion for English: 1,000 tokens is about 750 words.

Embeddings: meaning as position

This is the single most useful concept in the chapter, and it explains why modern search stopped caring about exact keywords.

An **embedding** is a list of numbers representing a piece of text’s meaning. Text with similar meaning gets similar numbers. You can think of it as a position in a space with many dimensions, where distance corresponds to semantic similarity.

Meaning as position: how models group related ideas



Terms that mean similar things sit near each other. This is why exact-match keywords stopped mattering.

Figure 3.1 – Semantically related terms cluster together. “Trainers” and “sneakers” sit near each other without sharing a single character.

The consequence for search is profound. A system using embeddings can tell that “how much do trainers cost” and “running shoe prices” are asking the same thing, despite sharing no words. It can also tell that a passage about plantar fasciitis is topically adjacent to running shoes without either mentioning the other.

This is why:

- Exact-match keyword optimisation stopped working — the system is matching meaning, not strings
- Writing naturally about a topic outperforms engineering keyword variants

- **Topical coverage** matters: covering the neighbouring concepts around your subject makes your content a better semantic match for a wider range of questions
- Synonyms and paraphrases are not tricks you need to insert; the system already understands them

MYTH vs REALITY

"You need to include every keyword variant"

Under lexical matching this was true. Under embedding-based retrieval it is close to useless, and the resulting stilted copy ("running shoes trainers sneakers footwear for running") actively harms readability, which harms the human signals that do matter.

Write for the concept, not the string. If you have covered the topic properly, the variants take care of themselves.

The context window

The **context window** is the amount of text a model can consider at once — its working memory for a single conversation. Everything the model is reasoning about must fit inside it: your question, any retrieved documents, its own previous responses.

This is where AI SEO becomes concrete. When an AI search system answers a question, it retrieves passages and places them into the context window alongside the question. The model then writes an answer using what is in that window.

So the competition is: **does your passage make it into the context window, and once there, is it used?** Those are two different battles. The 2026 critical survey found that most GEO research studies only the second.

Context windows are large now — hundreds of thousands of tokens, but they are not infinite, and crucially, position within the window affects how much influence a passage has. Material at the beginning and end of a long context tends to be weighted more heavily than material buried in the middle. This is one of the two effects that replicated reliably across studies.

Training versus inference: the distinction that matters most

This is the concept that resolves the largest number of practical confusions, so it is worth being precise.

Training happened in the past. The model was shown a vast corpus, and its parameters were adjusted. This is where it learned language, facts, reasoning patterns and — relevantly for you — general associations about brands and topics. Training data has a cutoff date. Anything after it is not in the model's parameters.

Inference is happening now. You ask a question; the model generates an answer. If the system has retrieval attached, it can fetch current documents at inference time and put them

into the context window.

Why models cite at all: retrieval-augmented generation



Figure 3.2 – Retrieval-augmented generation. The model supplies the language; the retriever supplies the facts.

KEY IDEA

Your content can influence a model in two completely different ways: by being in its **training data** (slow, indirect, uncontrollable, affects general brand associations) or by being **retrieved at inference** (fast, direct, and the thing you can actually optimise for).

Almost all practical AI SEO targets the second. When someone tells you to "optimise for the training data", they are describing something with a multi-year feedback loop that you cannot measure.

Why models hallucinate

A model generates plausible continuations. When it lacks the information to be accurate, it does not know that it lacks it — there is no internal "I don't know" signal in the base mechanism. It produces the most plausible-sounding text, which may be confidently, fluently false.

This is structural, not a bug awaiting a patch. It follows directly from what the system is: a plausibility engine, not a database.

Three practical implications for you:

1. **Retrieval reduces hallucination but does not eliminate it.** Grounding the model in fetched documents helps enormously, which is why AI search systems cite sources. But the model can still misread, over-generalise or blend two sources.
2. **Models will state things about your brand that are wrong.** Confidently. Chapter 29 covers monitoring and correcting this.
3. **Clear, unambiguous, well-structured source content reduces the error rate.** If your pricing page states a figure plainly in a sentence, it will be reported correctly far more often than if the figure only exists inside a comparison table image.

IN PRACTICE

Testing what a model believes about you

Open ChatGPT, Claude, Gemini and Perplexity. Ask each the same five questions:

1. "What does *[your company]* do?"
2. "How much does *[your product]* cost?"
3. "Who are *[your company]*'s main competitors?"
4. "What are the downsides of *[your product]*?"
5. "Is *[your company]* a good choice for *[your core customer type]*?"

Record the answers verbatim, and note whether each was retrieved (with citations) or generated from training. Two things will usually surprise you: how much is wrong, and how much of the wrong material traces back to a single outdated third-party page.

This exercise takes fifteen minutes and is the most useful first diagnostic in this entire book.

CASE STUDY · KETTLEWORTH

Before changing anything, Kettleworth ran the fifteen-minute diagnostic. The results were worse than expected.

Asked what Kettleworth sold, two of four assistants described them as a coffee *shop* — a café — rather than an equipment retailer. Asked about the K3 grinder's price, three gave £149; the correct figure was £189. Asked who they competed with, one named two companies that had ceased trading.

None of this was mysterious once they understood the mechanism. The café confusion came from an eight-year-old local directory listing. The price came from a marketplace listing nobody had maintained since 2024. The models were not wrong so much as faithfully reporting stale sources.

CHAPTER SUMMARY

- A language model predicts plausible next text; it does not store or look up documents
- Tokens are the unit of processing; embeddings represent meaning as position, which is why semantic beats lexical
- The context window is the model's working memory, and getting your passage into it is the actual competition
- Training (past, slow, uncontrollable) and inference-time retrieval (now, fast, optimisable) are different routes to influence
- Hallucination is structural; retrieval and clear source content reduce it but cannot remove it

CHECK YOURSELF

1. Why can a model be completely fluent and completely wrong at the same time?
2. Explain to a non-technical colleague why stuffing keyword variants no longer works, using the idea of embeddings.
3. What is the practical difference between influencing training data and influencing retrieval? Which can you measure?
4. Your pricing is being reported incorrectly by ChatGPT. Name two plausible causes rooted in this chapter.

CHAPTER 4

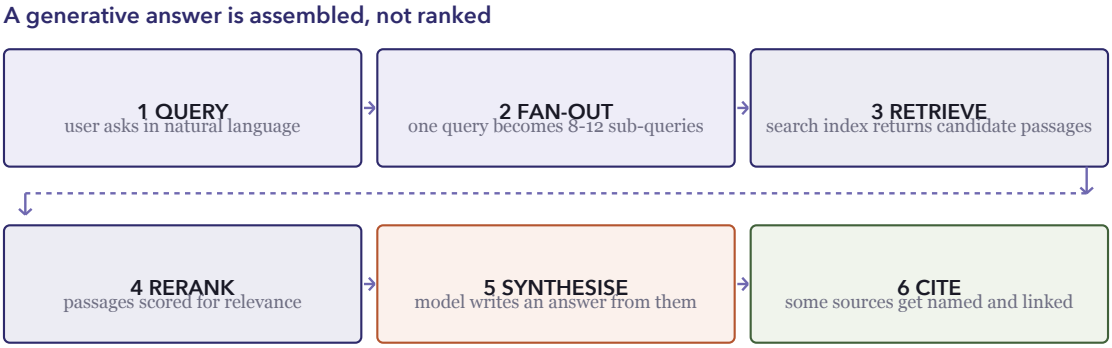
How AI Search Systems Work

Fan-out, chunking and reranking — the pipeline you are actually competing in

WHAT YOU'LL LEARN

- The full pipeline from question to cited answer, stage by stage
- Query fan-out: how one question becomes eight to twelve, and why that multiplies your opportunities
- Chunking: why retrieval operates on passages and what the optimal passage looks like
- Reranking and selection — the least understood and most decisive stage

This is the most important chapter in the book. Everything practical in Parts III to V is an attempt to influence one of the stages described here. If you understand this pipeline, you can derive most tactics yourself — and, more usefully, you can tell which tactics are nonsense.



You are competing to be a retrieved passage, not a ranked page.

Figure 4.1 – From question to citation. Six stages, each of which can exclude you.

Stage 1: The question arrives

Someone types or speaks a question. Note immediately how different this input is from a classical search query. People type *"best CRM"* into Google. They type *"I run a five-person creative agency and we're drowning in follow-ups, what CRM should we use and roughly what will it cost"* into ChatGPT.

Longer, more conversational, more context, and often containing several distinct questions at once. This has a direct consequence: the keyword research methods you may have learned do not capture these inputs, because nobody was ever typing them into a keyword tool. Chapter 28 covers how to build a prompt set instead.

Stage 2: Query fan-out

The system does not search for the question as typed. It decomposes it into multiple sub-queries and runs them in parallel. Google's AI Mode uses a Gemini model for this decomposition, generating roughly **8-12 sub-queries** for a standard question, and potentially hundreds for a deep research request.

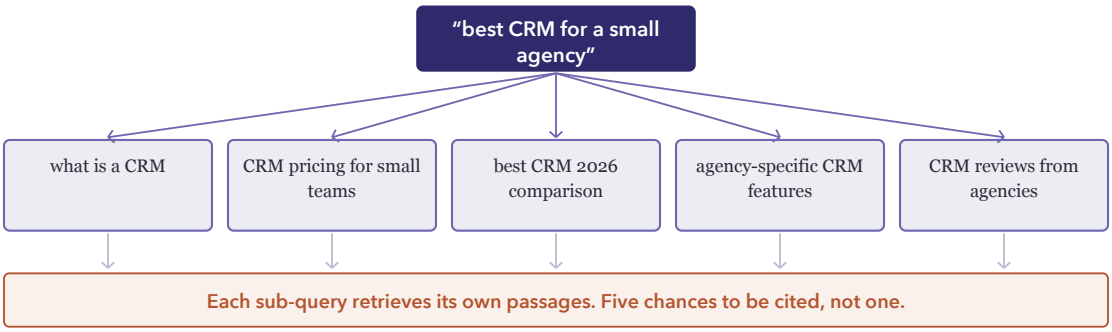


Figure 4.2 – Query fan-out. One question becomes many searches, each retrieving its own candidate passages.

This is good news, and it is under-appreciated. You do not need to be the single best answer to the whole question. You need to be the best answer to **one facet** of it. A page that comprehensively covers agency-specific CRM features may be retrieved for that sub-query even if it says nothing about pricing.

> **KEY IDEA**
>
> Fan-out converts one competitive slot into eight to twelve. Content strategy in AI search is therefore about **covering the facets of a question thoroughly** rather than trying to rank a single page for a single head term.

IN PRACTICE

Reverse-engineering the fan-out

You cannot see the actual sub-queries, but you can approximate them well:

1. Take your target question: *"best project management tool for agencies"*
2. Ask a model directly: *"If you were answering this question, what specific sub-questions would you need to research first? List 12."*
3. Compare that list against your content. Which sub-questions do you answer in a clean, self-contained passage? Which do you not answer at all?

The gaps are your content plan. This is a more reliable brief-building method than keyword volume, because it maps to how the system actually retrieves.

What this is, and is not. The list is a *hypothetical research decomposition* — what one model says it would ask. It is not a record of the sub-queries any assistant actually issued when answering a real user, and no product exposes those. Label the output INFERRED and use it to plan content, never as evidence of what happened. Chapter 42 explains why the distinction is worth this much fuss.

NOTE

Three of the six stages in this pipeline are things you can never observe directly: the actual sub-queries, the reranking, and which candidates were considered and dropped. That does not make the model useless — it makes it a map rather than a log. Part X turns the map into an audit method that says out loud which cells are observed and which are reconstructed.

Stage 3: Retrieval

Each sub-query hits an index and returns candidate documents. This is where classical SEO does its work: the index being searched is, for Google's AI features, Google's own index. If you are not in it, or rank nowhere for anything related, you are not a candidate.

But documents are not what gets used. They get **chunked**.

Stage 4: Chunking – the passage is the unit

Retrieved documents are split into semantic passages, typically in the region of 200-500 tokens each. The system then works with those passages, not with whole pages.

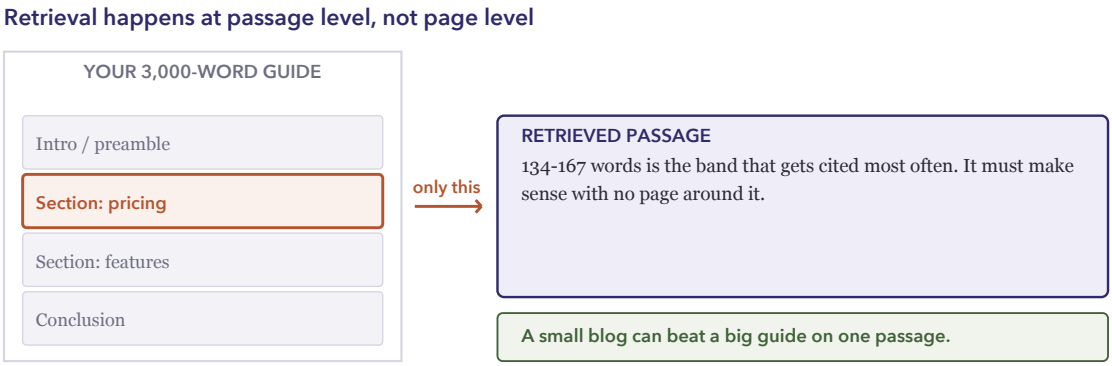


Figure 4.3 – A page is not retrieved. A passage from it is.

Research by Wellows analysing 15,847 AI Overview results found that passages in the **134-167 word** band achieved the highest citation rates. Treat that as a useful heuristic rather than a magic number, but the underlying principle is solid and worth internalising:

A retrievable passage is one that makes complete sense with no surrounding page.

That is the test. Read any section of your content in isolation, with the heading, and ask whether a stranger would understand it and get an answer. If it begins "As we saw above…" or "This is why the third factor matters…", it fails. It cannot be lifted.

In Google's AI Mode, the system may also pull several chunks either side of a matched one to supply context — which means a strong passage can drag its neighbours in with it. Well-structured content compounds.

> COMMON MISTAKE
>
> The common overcorrection here is to chop everything into 150-word fragments with a heading on each. This produces content that is unpleasant to read and often incoherent as a whole, which damages the human engagement signals that still feed the classical layer.
>
> The goal is **self-contained sections of natural length**, not uniform 150-word blocks. If a topic genuinely needs 400 words, give it 400 words and make sure it opens with the answer.

Stage 5: Reranking and selection

Retrieved passages are scored again, more expensively, for how well they actually answer the sub-query. The best survivors are placed into the context window. This is the decisive stage and the one you have least visibility into.

What is known to influence it:

- **Topical relevance** — the strongest and most reliably replicated factor across studies
- **Position in the assembled context** — the second reliably replicated factor

- **Semantic directness** — passages that answer the question rather than discussing it
- **Internal coherence** — self-contained passages score better than fragments requiring outside context
- **Source authority** — carried over from the classical layer, still a real input
- **Freshness** — heavily weighted for anything time-sensitive

MYTH vs REALITY

"There is a GEO checklist that guarantees citation"

The July 2026 critical survey of 45 GEO studies concluded that **no reviewed technique showed a stable, longitudinal, cross-platform causal effect** on organic discoverability. It also flagged three problems that should shape how you read any tactic:

- **Low reproducibility** — generic heuristics transfer poorly between engines and topics
- **Competitive erosion** — gains shrink as more people apply the same tactic
- **Quality degradation** — rewriting content specifically for citation can make it *worse* at being retrieved

The two levers that did replicate are topical relevance and context position. Notice that both are about genuinely being the right answer, not about formatting tricks.

Stage 6: Synthesis and citation

The model writes an answer from what is in its context window, and attributes some of it. Not every passage used gets cited — this is important and frequently missed. A model may absorb information from your page, use it to shape its answer, and cite a different source, or none.

The critical survey formalises this with a **visibility vector** of four distinct dimensions:

Dimension	What it means	Can you measure it?
Discoverability	Your content can be found and retrieved at all	Partially — via crawl logs
Citation	You are explicitly named and linked	Yes — via tracking tools
Absorption	Your information shapes the answer without attribution	Barely — the hardest problem in the field
Economic outcome	It produces business value	Only indirectly

Absorption is the uncomfortable one. Your work can be influencing answers, driving demand, and shaping how a category is described, while producing no citation and no click.

Chapter 30 deals with how to make a defensible business case in the face of this.

IN PRACTICE

Tracing a single answer end to end

Pick a question your customers ask. Then:

1. Ask it in Google AI Mode, ChatGPT, Perplexity and Claude. Save all four answers.
2. For each, list the cited sources. Note which appear in more than one.
3. For each cited source, open it and find *the specific passage* the answer drew on. Read that passage in isolation.
4. Compare those passages with the equivalent section on your own site.

You will find the cited passages are consistently more direct, more self-contained and more specific than your equivalent. That gap is your work list, derived from evidence rather than from a blog post.

CASE STUDY · KETTLEWORTH

Kettleworth took their most commercially important question — *"what grinder should I buy for espresso at home?"*, and asked a model for the twelve sub-questions it would research.

The list included: burr type differences, grind consistency and why it matters, stepped versus stepless adjustment, retention, noise, price bands, and whether entry-level grinders are adequate for espresso at all.

They had one page. It was 2,800 words, titled "The Ultimate Guide to Coffee Grinders", and it addressed nine of those twelve, but in flowing prose with headings like "Getting Started" and "A Note on Burrs". No section answered any sub-question in a form that could be lifted.

Nine answers, zero retrievable passages.

CHAPTER SUMMARY

- The pipeline is: question › fan-out › retrieval › chunking › reranking › synthesis and citation
- Fan-out turns one question into 8-12 sub-queries, multiplying the ways you can be relevant
- Retrieval operates on **passages**, not pages; a passage must stand alone to be liftable
- Reranking is the decisive stage; topical relevance and context position are the two factors that replicate
- Citation and absorption are different — your content can shape an answer with no attribution at all

CHECK YOURSELF

1. Walk through all six stages for the question “is oat milk actually better for you than dairy”. Where could a good page fail?
2. Why does query fan-out favour thorough topical coverage over a single optimised page?
3. What is the practical test for whether a passage is retrievable?
4. Define absorption, and explain why it makes reporting on AI search genuinely difficult.

PART II

The Landscape

The engines answering your customers' questions, how each one differs, and how to tell a real finding from a confident guess

CHAPTER 5

Google: AI Overviews and AI Mode

The surface where most of the volume still is

WHAT YOU'LL LEARN

- The difference between AI Overviews and AI Mode, and why it matters commercially
- How the AI Overview is assembled, and what determines which sources appear in it
- What happens to the classic organic results underneath, and what they are still worth
- Which Google-specific controls you have, including the ones that do not do what people think

Google still handles the overwhelming majority of search volume, so whatever else changes, this is the surface that moves your numbers most. It now has two distinct AI behaviours, and conflating them causes a lot of confused analysis.

AI Overviews versus AI Mode

AI Overviews are the generated summaries that appear at the top of an otherwise normal results page. The user did not ask for them. They typed a normal query and got an answer block above the links. Below it, the familiar organic results continue.

AI Mode is a separate, opt-in conversational surface. The user is deliberately in a chat-style interface, asking longer questions and following up. There are no traditional organic results in the usual sense — it behaves much more like ChatGPT.

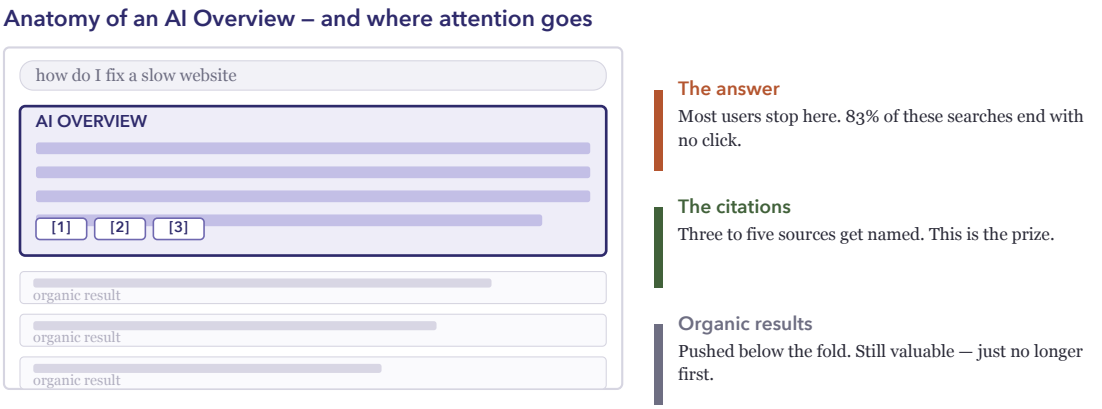


Figure 5.1 – Anatomy of an AI Overview page, and where user attention goes.

The commercial difference is significant:

	AI Overviews	AI Mode
How users arrive	Default, unrequested	Deliberately chosen
Query style	Short, keyword-like	Long, conversational
Organic results present	Yes, below the fold	Not in the traditional form
Volume	Very high	Growing, much smaller
Fan-out intensity	Moderate	Heavy — including Deep Search
Your measurable signal	Search Console still reports impressions	Almost nothing reported

The practical implication: **AI Overviews are where the volume damage and the citation opportunity both are today.** AI Mode is where the behaviour is heading, and it is worth understanding, but it is not yet where your traffic is being decided.

How an AI Overview gets built

The mechanism is the pipeline from Chapter 4, running against Google's index. The query is fanned out, passages are retrieved from indexed pages, reranked, and synthesised into the summary. Three to five sources are typically surfaced as citation chips.

Two Google-specific points are worth knowing.

First, **the sources cited in an AI Overview correlate with, but are not identical to, the top organic results.** Studies consistently find substantial overlap — often around half — with the rest drawn from pages ranking further down, sometimes far down. This is the passage-level effect from Chapter 1 made visible.

Second, **AI Overviews trigger unevenly by query type** — and the spread is much wider than most people assume. Reported 2026 trigger rates by sector: around **88%** for healthcare queries, **83%** for education, and **37%** for entertainment. They appear heavily on informational and comparative queries, and much less on transactional, navigational and local ones.

This is why the aggregate number is close to useless for planning. A healthcare publisher and a ticketing site face completely different situations, and both are inside the same "~50% of queries" headline. If your business lives on "plumber near me", your exposure is low. If it lives on "how do I fix a dripping tap", your exposure is very high.

IN PRACTICE

Mapping your own exposure

Do this before you plan any AI-search work. It takes an afternoon and prevents you from solving a problem you do not have.

1. Export your top 200 queries from Search Console by impressions.
2. Classify each by intent: informational, commercial-investigation, transactional, navigational, local.
3. Sample 30 of them manually in an incognito browser and record whether an AI Overview appears.
4. Calculate the share of your *impressions* — not your query count — that sit on AI-Overview-triggering queries.

If that figure is 15%, you have a monitoring job. If it is 70%, you have a strategic problem that needs a plan this quarter. Most people have never calculated it and are arguing about tactics without knowing which situation they are in.

What happens to the organic results

They are still there, still clicked, and still worth having. What has changed is the volume reaching them and the shape of the traffic.

The clicks that do come through are disproportionately from people whose question was *not* satisfied by the summary — meaning they have a more specific or more serious need. This is the mechanism behind the widely quoted finding that post-AI-Overview clicks convert at much higher rates. Fewer visitors, better qualified.

MYTH vs REALITY

"AI Overviews mean organic rankings no longer matter"

Ranking is now a strong predictor of being cited, because the AI Overview retrieves from the same index and correlates substantially with top results.

Abandoning the work that gets you ranked is therefore the fastest way to *also* lose your citations. Ranking is not the goal any more, but it remains the mechanism.

The controls you actually have

This is where a lot of misinformation circulates, so be precise.

- `nosnippet`, `max-snippet` **and** `data-nosnippet` — these do limit what Google can show from your page, and Google has confirmed they apply to AI Overviews. Using them will reduce or prevent your content appearing in the summary. It will also strip your normal snippets, which usually costs more than it saves.

- `Google-Extended` — this robots.txt token controls whether your content is used for Gemini model training. It does **not** control AI Overviews. Many sites have added it believing they were opting out of AI Overviews; they were not.
- `noindex` — removes you from the index entirely, and therefore from everything. A blunt instrument.

COMMON MISTAKE

There is no clean "appear in search but not in AI Overviews" switch. Google has been explicit that AI Overviews are part of Search. The only mechanisms are snippet controls that also damage your ordinary listings. Anyone selling you a tidy opt-out is mistaken.

CASE STUDY · KETTLEWORTH

Before planning anything, Kettleworth calculated their actual exposure rather than assuming it.

They exported their top 200 queries by impressions, classified them by intent, and sampled 30 manually. The result: 61% of their impressions sat on queries that triggered an AI Overview — overwhelmingly the informational ones ("how to", "what is", "best … for") that their editorial guides had been built to win.

Their transactional queries — brand and product names — triggered almost nothing.

That single number reframed the whole project. They did not have a general traffic problem. They had a specific problem with the top of their funnel, and their bottom of funnel was intact.

CHAPTER SUMMARY

- AI Overviews are unrequested summaries on a normal SERP; AI Mode is a separate conversational surface
- Citations correlate with top rankings but pull from further down too — passage quality creates genuine openings
- Trigger rates vary enormously by query intent; calculate your own exposure before planning anything
- Organic clicks fall in volume but rise in quality
- `Google-Extended` controls Gemini training, **not** AI Overviews — a very common and costly misunderstanding

CHECK YOURSELF

1. What is the difference between AI Overviews and AI Mode, and which affects your traffic more today?
2. How would you calculate your business's actual exposure to AI Overviews?
3. A colleague adds `Google-Extended: Disallow` to opt out of AI Overviews. What have they actually done?
4. Why do post-AI-Overview clicks convert better?

CHAPTER 6

The Assistants

ChatGPT, Perplexity, Claude, Gemini and Copilot — five engines with genuinely different behaviour

WHAT YOU'LL LEARN

- How each major assistant retrieves, and where they differ in ways that change your tactics
- Why optimising for "AI" as a single thing is a category error
- Which engines matter for which kinds of business
- How to prioritise when you cannot chase all of them

The single biggest tactical mistake in this field is treating "AI search" as one destination. These systems have different indexes, different retrieval behaviour, different source preferences and radically different citation patterns. A tactic that works on Perplexity may do nothing on Gemini.

The five that matter

ChatGPT

The largest by consumer usage by a wide margin. It answers from training data when it can and retrieves when the question needs current information, which means **the same question can produce a retrieved, cited answer one day and a generated, uncited one the next**, depending on how it judges the need.

Its search index draws on a partnership with Microsoft Bing plus its own crawling. Two distinct crawlers matter: `GPTBot` (training) and `OAI-SearchBot` (search retrieval). Blocking the first does not block the second, and confusing them is a common and expensive error.

Citation behaviour: heavily weighted towards Wikipedia and Reddit, each around 12-13% of US citations by some measures, with Reddit appearing in over 5% of responses.

Perplexity

Much smaller in absolute usage but disproportionately important because it is retrieval-first by design. Perplexity always searches and always cites — it was built as an answer engine rather than a chatbot with search bolted on.

This makes it the **best testing ground you have**. If you want to understand what a retrieval system thinks about your topic, Perplexity shows you its working more

transparently than anything else.

Citation behaviour: extremely social-heavy. Around 24% of its citations came from Reddit alone in January 2026, and roughly 31% from social media overall.

Google Gemini

Google's assistant, sharing infrastructure with AI Mode. Notably, its citation profile is completely unlike the others: Reddit appears in around **0.1%** of Gemini responses, against 24% for Perplexity. It leans far more on established web sources, and in AI Mode, Fandom, Wikipedia and YouTube lead.

Claude

Anthropic's assistant, strong in professional and technical contexts, with meaningful enterprise adoption. Its web search cites sources. Crawlers: `ClaudeBot` for training, `Claude-SearchBot` for search.

Microsoft Copilot

Integrated into Windows, Edge and the Microsoft 365 suite. Its significance is distribution rather than enthusiasm — it reaches enormous numbers of corporate desktops by default. For B2B, this matters more than its consumer mindshare suggests.

Five engines, five different behaviours

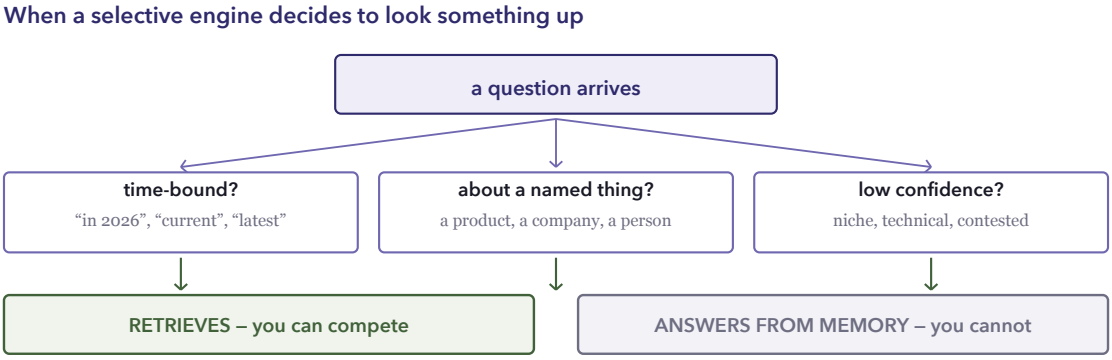
ENGINE	RETRIEVES	LEANS ON	USE IT FOR
ChatGPT	when it judges it needs to	Wikipedia, Reddit, mainstream	volume reality check
Perplexity	always, and always cites	Reddit and social, heavily	diagnosis – it shows its working
Gemini	always	established web, YouTube	the Google-ecosystem proxy
Claude	on demand	documentation, primary sources	B2B and technical signal
Copilot	always, Bing-backed	Bing index, corporate sources	B2B distribution check

Reddit is 24% of Perplexity's citations and 0.1% of Gemini's. There is no single "AI search".

Figure 6.1 – The five engines side by side. Optimising for "AI" as one thing is a category error.

When each engine decides to retrieve

The single most useful thing to understand about these products is *when* they go and look something up, because it determines whether your content can influence the answer at all.



This is why definitional explainer content is the weakest category in AI search.

Figure 6.2 – What tips a selective engine into looking something up.

Always-retrieve engines

Perplexity and Gemini search on essentially every question. There is no path to an answer that does not pass through retrieval, which means your content is always eligible, and that a question about your category will always surface *someone's* pages, whether or not they are yours.

Selective-retrieve engines

ChatGPT and Claude decide. They weigh whether the question needs current information, whether the model's own knowledge is likely to suffice, and how specific the request is. This produces a behaviour that confuses people constantly: **the same question can yield a retrieved, cited answer one day and a generated, uncited one the next.**

Two practical consequences:

- **When measuring, record whether each answer was retrieved or generated.** A month where your share of model falls may simply be a month where the engine retrieved less, which is not something you did.
- **Questions that signal recency retrieve more.** "Best CRM in 2026", "current pricing", "has X changed" all push an engine towards looking. Including a year or a freshness cue in your tracked prompts will raise retrieval rates — which is useful for diagnosis, and a distortion if your real customers do not phrase questions that way.

What tips a selective engine into retrieving

Broadly: anything time-bound, anything about a specific named product or company, anything where the model's confidence is low, and anything the user has explicitly asked it to look up. Broad conceptual questions — "what is a CRM" — are the least likely to trigger retrieval, which is precisely why generic explainer content performs so poorly in AI search. The model already knows that, and does not need you.

KEY IDEA

Explainer content aimed at definitional questions is the worst-performing category in AI search, because those are the questions engines answer without retrieving anything. Your opportunity is in the specific, the current and the comparative — the questions a model cannot answer from memory.

What you can actually capture from each

The comparison above is about how the engines behave. This is about what you can *record* when one of them answers — which decides what your measurement and forensics are able to say later.

For each product you sample, establish these seven things once, and write the answers down:

1. Can you tell whether web search occurred at all?
2. Are the search queries it issued visible?
3. Does it expose every source retrieved, or only the ones cited?
4. Do inline citations map to specific claims, or are sources pooled at the end?
5. Is there a sources panel, and does it survive being copied?
6. How do you export the answer with its citations intact?
7. What is simply unavailable — and therefore `UNKNOWN` in your records?

The last one is the point of the exercise. A product that shows a sources panel and nothing else supports citation counting and does not support source-selection analysis, and knowing that in advance stops you writing a forensic conclusion the evidence cannot carry.

COMMON MISTAKE

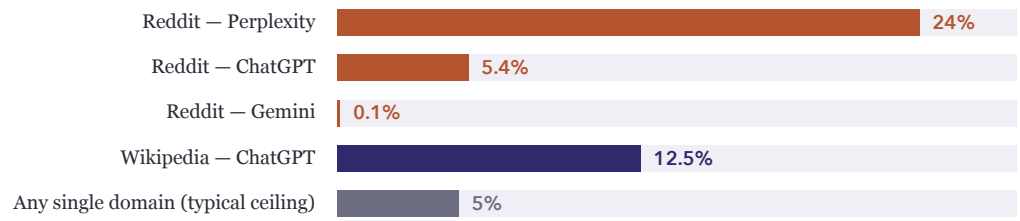
Product interfaces change faster than print. Treat your answers to those seven questions as a note with a date on it, re-check them quarterly, and never assume that because a product exposed its search steps in March it still does in September. Chapter 42 is the discipline; this is where it starts.

NOTE

Pooled citations deserve particular suspicion. When an answer lists five sources at the bottom rather than attaching them to sentences, you cannot tell which source supports which claim — so claim-level validation (Chapter 44) becomes inference rather than checking. Record which style each product used, because it changes what your audit is worth.

The differences that change what you do

Where AI assistants get their sources (share of citations)



Platforms differ enormously. Optimise for the engines your buyers actually use.

Figure 6.3 – Citation sources vary enormously by engine. There is no single "AI search" to optimise for.

KEY IDEA

Reddit is 24% of Perplexity's citations and 0.1% of Gemini's. Any strategy built on "get mentioned on Reddit to win AI search" is therefore correct for one engine and irrelevant for another. **Identify which engines your buyers actually use before choosing tactics.**

Engine	Retrieval style	Source bias	Best used by you as
ChatGPT	Selective — retrieves when needed	Wikipedia, Reddit, mainstream web	Volume reality check
Perplexity	Always retrieves, always cites	Reddit and social very heavy	Diagnostic testing ground
Gemini	Always retrieves	Established web, YouTube, Wikipedia	Google-ecosystem proxy
Claude	Retrieves on demand	Documentation, primary sources	B2B and technical signal
Copilot	Retrieves, Bing-backed	Bing index, corporate sources	B2B distribution check

There is one more finding worth holding onto, because it reframes the whole competitive picture. Analysis across 200 million prompts found that **even the most-cited domain on any given platform rarely exceeds 5% of total citations.** AI search is a long tail with a few outliers, not a winner-takes-all market.

That is encouraging. There is no incumbent occupying the answer the way a single site can occupy position one.

IN PRACTICE

Building your engine priority list

Do not chase all five. Work out which two matter:

1. **Ask your customers.** In your next ten sales calls or support conversations, ask: "How did you first come across us?" and "Do you use ChatGPT or similar tools when researching this kind of purchase?" Ten conversations gives you more signal than any industry report.
2. **Check your referral data.** In analytics, look for referrals from `chatgpt.com`, `perplexity.ai`, `gemini.google.com`, `copilot.microsoft.com`. Volumes will be small, but the ratios are informative.
3. **Weight by fit.** B2B and technical: Copilot and Claude over-index. Consumer and lifestyle: ChatGPT and Perplexity. Anything where Google already dominates your acquisition: Gemini and AI Mode.

Pick two. Track those properly rather than all five badly.

COMMON MISTAKE

Do not build a strategy around Perplexity because it is the easiest to observe. Its transparency makes it a superb diagnostic tool and its actual user base is a fraction of ChatGPT's. Measure where you can see; prioritise where your buyers are.

CASE STUDY · KETTLEWORTH

Kettleworth asked ten customers on sales and support calls how they had researched their purchase. Six mentioned ChatGPT. Two mentioned Reddit specifically. Nobody mentioned Perplexity or Claude.

Their referral data agreed: `chatgpt.com` sent about eight times more sessions than every other assistant combined, though all the numbers were small.

They chose ChatGPT as their primary tracked engine and Perplexity as their secondary — not because customers used Perplexity, but because it always cites and therefore showed them which sources the retrieval systems trusted in their category. One engine for the truth about their market, one for the working.

CHAPTER SUMMARY

- The five engines differ enough that "optimise for AI" is a category error
- ChatGPT has the volume; Perplexity has the transparency; Gemini shares Google's infrastructure; Claude and Copilot skew professional
- Reddit is 24% of Perplexity citations and 0.1% of Gemini's — source strategy must be engine-specific
- No single domain exceeds ~5% of citations on any platform: this is a long-tail market with real openings
- Always-retrieve engines (Perplexity, Gemini) always give you a shot; selective ones (ChatGPT, Claude) decide, so record whether each answer was retrieved or generated
- Definitional explainer content is the weakest category, because engines answer those without retrieving anything
- Choose two engines to track properly based on where your buyers actually are

CHECK YOURSELF

1. Why is "get mentioned on Reddit" good advice for one engine and near-useless for another?
2. What makes Perplexity the best diagnostic tool, and why should it not drive your strategy?
3. What is the difference between `GPTBot` and `OAI-SearchBot`, and why does confusing them matter?
4. What does the ~5% citation ceiling tell you about competitiveness in AI search?

CHAPTER 7

Who Gets Cited, and How to Judge a Claim

The evidence, its limits, and a method for filtering the noise

WHAT YOU'LL LEARN

- What the citation data across engines actually shows about which sources win
- Why user-generated content punches so far above its weight, and what that means if you are a brand
- A repeatable method for evaluating any AI-SEO claim you encounter
- How to run your own small tests when the published evidence does not cover your situation

This chapter does two jobs. It tells you what the current evidence says about who gets cited — and then it gives you the tool to keep evaluating claims yourself, long after these specific numbers have dated. The second job is the more valuable one.

What the citation studies show

Several large studies through 2025 and 2026 have analysed citation patterns at scale. The consistent findings:

User-generated content dominates far beyond its apparent quality. Reddit is the standout — roughly 24% of Perplexity's citations and over 5% of ChatGPT's responses. Across multiple engines aggregated, one Q1 2026 analysis put Reddit at around 40% of citation frequency, the highest concentration on any single domain. Wikipedia sits at around 12-13% of ChatGPT citations.

Reddit's share is also still climbing: the same analysis measured growth of roughly 73% between October 2025 and January 2026 across the categories it tracked.

But the distribution is long-tailed. Even the most-cited domain rarely exceeds 5% of total citations on a platform. There is a head, but it is a small head with a very long tail behind it.

Vertical matters more than any general rule. For consumer, lifestyle and experience-driven products, Reddit ranks highest. For B2B, professional services and enterprise software, LinkedIn outranks it. YouTube performs strongly across the board, particularly in Google's surfaces.

Why user-generated content wins

It is understanding the mechanism rather than just the fact, because the mechanism tells you what to do.

1. **It contains genuine first-hand experience.** "I used this for eight months and here is what broke" is information that exists nowhere in marketing copy. It maps directly onto the "Experience" in E-E-A-T (Chapter 15).
2. **It is structured as question-and-answer.** A Reddit thread is literally a question followed by attempted answers — the exact shape a retrieval system is looking for.
3. **It is dense with specifics.** Real prices, real model numbers, real failure modes.
4. **It is comparative and unhedged.** Users say "X is better than Y for Z", which brands almost never do about themselves.

> KEY IDEA
>
> The reason user-generated content wins is not that models love Reddit. It is that Reddit content is **specific, experiential, comparative and directly answer-shaped**. Brand content that adopts those four properties competes considerably better. That is the transferable lesson, and it is available to you.

A method for judging any claim

You will encounter a constant stream of confident assertions about what works. Here is a five-question filter. Run any claim through it.

The filter, and one claim run through it

ASK THIS OF ANY CLAIM	"ADDING llms.txt IMPROVES AI VISIBILITY"
1 **What was actually measured?** citation rate, position, frequency — all different	→ ✕ usually nothing; asserted, not tested
2 **On which engine, and more than one?** a finding on one engine is a claim about one engine	→ ? partial — only Perplexity claims to fetch it
3 **Retrieval, or use once retrieved?** the weaker claim is routinely reported as the stronger	→ ✕ neither has been demonstrated
4 **Does it survive competitors doing it?** effects that depend on scarcity erode as adoption rises	→ ✕ untested
5 **Is there a plausible mechanism?** it would need something to actually fetch the file	→ ✕ 97% never fetched — the mechanism is absent

Verdict: fails four of five. Costs little, does close to nothing – so it is a low priority, not a scandal.
Most claims you meet will fail question one, which is why it is question one.

Figure 7.1 – The five questions, and the same claim judged at each one. The worked example below walks through it in prose.

1. What was actually measured?

The most common failure. “Improves AI visibility by 40%” — visibility measured how? Citation rate? Position within the answer? Frequency of appearance across a prompt set? These are different things, and a gain in one can coincide with no change in the others.

2. On which engine, and was it tested on more than one?

A finding on Perplexity may not transfer to Gemini. If the study tested one engine, the claim is about one engine.

3. Was the source already retrieved, or did the tactic affect retrieval?

This is the distinction that invalidated most of the popular GEO advice. Improving how a passage is *used once it has been retrieved* is a much weaker claim than improving *whether it gets retrieved*. The original GEO paper measured the first; the industry reported it as the second.

4. Does it survive competitors doing it too?

The 2026 critical survey named this **competitive erosion**. If a tactic works because few people use it, its measured effect will shrink as adoption rises. Formatting tricks erode fast. Genuinely better content does not.

5. Is there a plausible mechanism?

Given Chapters 3 and 4, does the claim have a mechanism? “Adding an FAQ schema helps a passage be understood as answering a question” has one. “Publishing on a Tuesday improves AI citation” does not.

IN PRACTICE

Applying the filter to a live example

Claim: “Adding `llms.txt` to your site improves AI visibility.”

1. *What was measured?* Usually nothing. The claim is typically asserted, not tested.
2. *Which engine?* Only Perplexity has said it retrieves the file at all.
3. *Retrieval or usage?* Neither has been demonstrated.
4. *Survives competition?* Untested.
5. *Plausible mechanism?* It would need crawlers to fetch the file. An Ahrefs study of 137,000 sites found **97% of `llms.txt` files were never fetched**. In a 90-day window covering over 500 million AI bot visits, only 408 requests targeted `llms.txt`. Google has stated it does not support the file, with John Mueller comparing it to the long-discredited keywords meta tag.

Verdict: fails four of five tests. Costs little, does close to nothing. Not a scandal — just not a priority. Chapter 12 covers this in full.

NOTE

This filter judges claims made *about* GEO. The same scepticism applies one level down, to the individual claims inside an AI answer and whether the citation beside each one actually supports it — a citation can be topically relevant and silent on the sentence it sits next to. Chapter 44 is that method.

MYTH vs REALITY

"Following the checklist gets you cited"

The July 2026 survey of 45 studies found **no technique with a stable, longitudinal, cross-platform causal effect** on organic discoverability. Only topical relevance and context position replicated reliably.

That is not a reason for despair. It is a reason to invest in what was always going to work: being the best, clearest, most specific answer on a topic you have real authority in, rather than chasing formatting tactics with a half-life measured in months.

Running your own tests

The studies behind this chapter, with their identifiers, are listed in Chapter 61 if you want to read them yourself. Published evidence will never cover your specific market. You can generate your own, and doing so is one of the strongest differentiators available to a marketer right now.

A workable method:

1. **Build a prompt set.** Thirty to fifty real questions your customers ask, phrased as they would phrase them. Fix this list; do not change it between measurements or the comparison is meaningless.
2. **Establish a baseline.** Run every prompt on your two chosen engines. Record: are you mentioned, are you cited with a link, what position in the answer, which competitors appear. This is tedious and it is the whole basis of everything that follows.
3. **Change one thing.** Restructure one cluster of pages. Do not simultaneously restructure, add schema, and run a PR campaign, or you will learn nothing.
4. **Wait properly.** Four to eight weeks. Re-crawling and re-indexing take time.
5. **Re-measure identically.** Same prompts, same engines, same recording method.
6. **Be honest about confounds.** A core update, a competitor's campaign, or a seasonal shift can all produce the change you are attributing to your work.

COMMON MISTAKE

Run every prompt in a logged-out, incognito session, and preferably from a consistent location. Personalisation and conversation memory will otherwise contaminate the results — the engine remembers you asked about your own brand last week and treats you differently.

CASE STUDY · KETTLEWORTH

Kettleworth's baseline turned up an uncomfortable pattern. For *"is the Kettleworth K3 any good"*, three of four engines answered primarily from a single Reddit thread — nine comments, two of them complaining about static cling.

Kettleworth had 4,100 reviews averaging 4.6. None were being cited.

Reading the Reddit thread against their own product page explained it. The Reddit comments named specific conditions ("in a dry flat in winter"), gave duration of ownership, and compared directly against two competitors. The product page said the K3 offered "precision engineering and exceptional consistency".

One of those is an answer to a question. The other is an adjective.

CHAPTER SUMMARY

- User-generated content dominates citations because it is specific, experiential, comparative and answer-shaped — properties you can adopt
- Citation distribution is long-tailed; no domain owns a category the way a top-ranking page could
- Vertical determines platform: Reddit for consumer, LinkedIn for B2B, YouTube broadly
- Filter every claim through: what was measured, which engine, retrieval or usage, does it survive competition, is there a mechanism
- Your own prompt-set testing beats published studies for your specific market, provided you change one variable at a time

CHECK YOURSELF

1. Give the four properties of Reddit content that make it citable, and say how a brand page could adopt each.
2. Apply the five-question filter to: "Publishing original research increases AI citations."
3. Why must your prompt set stay fixed between measurements?
4. Why does the five-question filter matter more than any specific tactic in this book?

PART III

Technical Foundations

Gates one and two. Unglamorous, frequently broken, and the reason a great deal of good content is invisible.

CHAPTER 8

Crawlability and the AI Crawlers

Who is allowed to fetch your site, and what you actually decided when you blocked them

WHAT YOU'LL LEARN

- Every AI crawler that matters, what it feeds, and the consequence of blocking each
- Why "block the AI bots" is a decision with two very different halves
- How to read your server logs to see what is really happening
- A defensible default policy you can take to a nervous stakeholder

This is gate one. If a crawler cannot fetch your page, nothing else in this book applies to you. It is also the area where the most damage is currently being done by well-intentioned decisions taken without understanding the mechanism.

The two jobs crawlers do

AI companies run **two categories** of crawler, and they do completely different things:

- **Training crawlers** collect text to train future models. Content gathered here influences what a model generally "knows" — including background associations about your brand. The feedback loop is years long and unmeasurable.
- **Search / retrieval crawlers** fetch pages so they can be cited in answers *right now*. This is the inference-time retrieval from Chapter 3. This is what produces citations, referrals and visibility this quarter.

Not all AI bots do the same job – blocking is not one decision

CRAWLER	WHAT IT FEEDS	IF YOU BLOCK IT
GPTBot	Trains OpenAI models	Block = no training use
OAI-SearchBot	Powers ChatGPT Search	Block = invisible in ChatGPT
ClaudeBot	Trains Anthropic models	Block = no training use
Claude-SearchBot	Powers Claude search	Block = invisible in Claude
PerplexityBot	Powers Perplexity answers	Block = invisible there
Google-Extended	Gemini training only	Does NOT affect AI Overviews

Blocking the training bot but allowing the search bot is a coherent position. Blocking everything is usually not.

Figure 8.1 – The crawlers that matter and what blocking each actually costs you.

KEY IDEA

Blocking a **training** crawler protects your content from being absorbed into a model's parameters. Blocking a **search** crawler makes you invisible in that assistant's answers.

These are entirely different decisions with entirely different costs. Sites that "blocked the AI bots" in one sweep have frequently opted themselves out of visibility while believing they were only opting out of training.

The crawlers, precisely

Crawler	Operator	Job	Blocking it means
`GPTBot`	OpenAI	Model training	Not used for training. No effect on ChatGPT search
`OAI-SearchBot`	OpenAI	ChatGPT search retrieval	**Invisible in ChatGPT search results**
`ChatGPT-User`	OpenAI	Live fetch when a user shares a link	Users cannot pull your page into a chat
`ClaudeBot`	Anthropic	Model training	Not used for training
`Claude-SearchBot`	Anthropic	Claude search retrieval	**Invisible in Claude's answers**
`PerplexityBot`	Perplexity	Indexing for answers	**Invisible in Perplexity**
`Google-Extended`	Google	Gemini training only	Not used for Gemini training. **No effect on AI Overviews**
`Googlebot`	Google	Everything Google	Invisible in Google entirely, including AI Overviews
`Bingbot`	Microsoft	Bing, feeds Copilot	Invisible in Bing and Copilot

COMMON MISTAKE

The `Google-Extended` misunderstanding is the most expensive single error in this chapter. It controls Gemini **training**. It does **not** remove you from AI Overviews, because AI Overviews retrieve from the ordinary Google index via `Googlebot`. There is no way to be in Google Search but out of AI Overviews. Many publishers added `Google-Extended` believing otherwise and got no protection from the thing they were worried about.

Reading `robots.txt`

The file lives at `yoursite.com/robots.txt`. The syntax is simple: name a crawler, then allow or disallow paths.

```
# Allow everything by default
User-agent: *
Allow: /

# Refuse training use, permit search retrieval
User-agent: GPTBot
Disallow: /

User-agent: OAI-SearchBot
Allow: /

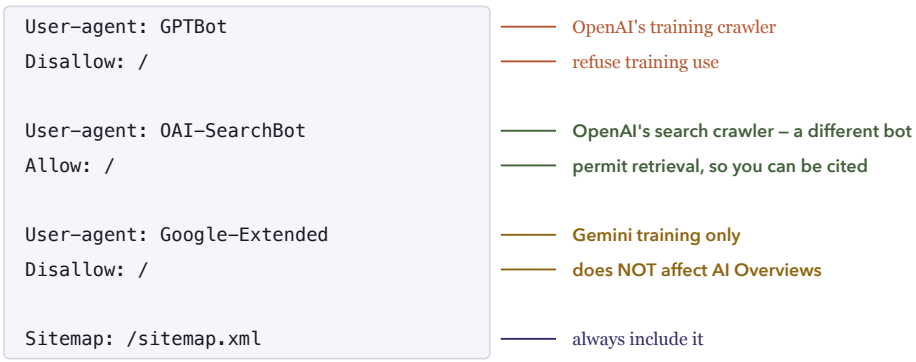
User-agent: ClaudeBot
Disallow: /

User-agent: Claude-SearchBot
Allow: /

User-agent: PerplexityBot
Allow: /

Sitemap: https://yoursite.com/sitemap.xml
```

A robots.txt that states a coherent position



Two OpenAI crawlers, two different decisions. Blocking the wrong one makes you invisible.

Figure 8.2 – The same file, annotated. Two OpenAI crawlers, two different decisions.

What each part does, in order.

Lines beginning `#` are comments. Crawlers ignore them entirely; they are for whoever opens this file next. Use them, because in six months you will not remember why you blocked something.

`User-agent:` opens a *block* — a set of rules belonging to one named crawler. Everything beneath it, until the next `User-agent:` line, applies to that crawler and to no other. The `*` is a wildcard meaning “any crawler that does not have a block of its own”.

`Disallow: /` refuses the whole site; `Allow: /` permits it. Paths are matched from the site root, so `Disallow: /admin` blocks every URL whose path starts with `/admin`, and a bare `/` matches everything.

The rule that catches people out is **precedence**. A crawler reads the block that names it and ignores the wildcard block completely — it does not combine the two. `GPTBot` above does not inherit `Allow: /` from the `*` block; it finds its own block, reads `Disallow: /`, and stops there. So the moment you write a named block, you must state everything you want that crawler to know, because the general rules no longer reach it.

`Sitemap:` is the exception to all of that. It sits outside every block and applies to all crawlers regardless of what came before it.

Now read the file again with that in mind. `GPTBot` is refused and `OAI-SearchBot` is permitted — two crawlers from the same company given deliberately opposite answers. The first collects text to train future models; the second fetches pages to answer a question someone is asking right now. Refusing one and permitting the other is the entire position, expressed in two blocks: *do not train on my work, but do cite me when answering questions.*

> COMMON MISTAKE
>
> `robots.txt` is a request, not a lock. Well-behaved crawlers honour it; nothing enforces it. It also hides nothing — the file is public, so `Disallow: /secret-launch/` announces to the world that the directory exists. If something must not be reachable, put it behind authentication, not behind a `Disallow`.

For most commercial sites the configuration above is the sensible default, and it is defensible in a meeting.

> IN PRACTICE
>
> **The audit that takes ten minutes**
>
> 1. Open `yoursite.com/robots.txt` in a browser. Read every line.
> 2. For each `Disallow`, ask: which crawler, and is it a training or a search bot?
> 3. Check for a blanket `User-agent: * / Disallow: /` — this happens by accident when a staging configuration ships to production, and it is catastrophic.
> 4. Confirm the `Sitemap:` line is present and the URL loads.
> 5. In Google Search Console, use the URL Inspection tool on three important pages and confirm each says it is indexed.
>
> If you do nothing else technical from this book, do this. Sites lose the entire game at this step more often than anyone expects. The full version of this checklist, and every other one in the book, is collected in Chapter 58.

Beyond robots.txt: the blocks you did not know about

`robots.txt` is voluntary and visible. The blocks that actually catch people out are elsewhere:

- **CDN and WAF bot rules.** Cloudflare, Fastly and Akamai all ship bot-management features, and several have added AI-crawler blocking as a one-click default. Your infrastructure team may have enabled it without telling marketing. **Check this specifically** — it is now the single most common cause of unexplained AI invisibility.
- **Rate limiting.** Aggressive limits cause crawlers to back off and fetch a fraction of your site.
- **Geographic or IP blocking.** Crawlers fetch from particular ranges; if those are blocked, you are invisible regardless of `robots.txt`.
- **Login walls and paywalls.** Content behind them is not retrievable. That may be exactly what you want, but be deliberate.
- **Aggressive cookie or consent interstitials.** If the crawler receives a consent wall instead of content, that wall is what gets indexed.

IN PRACTICE

Confirming what actually reached your server

`robots.txt` tells you your *intent*. Server logs tell you the *reality*. Ask whoever manages your hosting for the last 30 days of access logs, then filter for the user agents in the table above.

What you are looking for:

- **Which AI crawlers visited at all.** If `OAI-SearchBot` never appears, you are not in ChatGPT search, whatever `robots.txt` says.
- **What status codes they received.** A wall of 403s means something is blocking them above the application layer — almost always the CDN.
- **Which pages they fetched.** If they only ever fetch your homepage, your internal linking is not leading them inward.

If you cannot get logs, Cloudflare and most CDN dashboards expose the same information under bot analytics.

Should you block AI crawlers at all?

An honest framing rather than a slogan. The case for blocking training crawlers is real: your content is being used to build a commercial product you are not paid for, and which may reduce your traffic. Large publishers with licensing leverage have a genuine strategic reason to withhold.

The case against blocking search crawlers is much stronger for almost everyone else. If your business depends on being discovered, removing yourself from the surfaces where discovery increasingly happens is a serious cost paid to prevent a mostly symbolic harm.

KEY IDEA

Default recommendation for a commercial site that needs to be found: **block training crawlers if you feel strongly, allow every search crawler.** Revisit if your licensing position changes.

The exception is a publisher whose product *is* the content and who has, or plausibly could have, a licensing relationship. There, withholding is a negotiating position rather than a technical decision, and it belongs to the commercial team, not to SEO.

CASE STUDY · KETTLEWORTH

Kettleworth's `robots.txt` allowed every AI crawler. They were still absent from Perplexity entirely.

The server logs settled it in ten minutes. `PerplexityBot` and `OAI-SearchBot` were both arriving, and both receiving 403 responses. Their Cloudflare configuration had AI bot blocking enabled, switched on eleven months earlier as part of a default security profile by a contractor who had since left.

Nobody in marketing knew. Nothing in Search Console showed it. `robots.txt` said one thing and the infrastructure did another.

Turning it off took four minutes. First Perplexity citation: nineteen days later.

CHAPTER SUMMARY

- Training crawlers and search crawlers are different; blocking them has entirely different consequences
- `Google-Extended` controls Gemini training only and does nothing about AI Overviews
- CDN bot-management defaults now block AI crawlers silently and are the leading cause of unexplained invisibility
- Server logs are the only source of truth about what actually reached you
- A defensible default: refuse training, permit retrieval

CHECK YOURSELF

1. Your CEO says "block the AI bots, they're stealing our content." What two questions do you ask before acting?
2. What exactly does adding `Google-Extended: Disallow` achieve, and what does it not?
3. `robots.txt` allows `PerplexityBot` but you never appear in Perplexity. Name three things to check.
4. Why are server logs more trustworthy than `robots.txt` when diagnosing this?

CHAPTER 9

Rendering, JavaScript and What Bots Can Read

The invisible-content problem that AI search made much worse

WHAT YOU'LL LEARN

- What rendering is and why it decides whether your content exists
- Why AI crawlers are much less capable here than Googlebot, and what that means
- How to check what a bot actually sees on your pages
- The specific patterns that make content unreadable, and their fixes

A page can be crawled successfully and still be empty as far as the crawler is concerned. This chapter is about that gap, and it has become sharply more important because AI crawlers are far less capable than Googlebot at closing it.

What rendering means

When a browser loads a page, it fetches the HTML, then runs any JavaScript, which may fetch more data and build most of the visible content. What you see is the *rendered* result, which can be very different from the raw HTML that arrived.

Modern JavaScript frameworks — React, Vue, Angular — commonly ship a nearly empty HTML shell and construct everything in the browser. For a user this is fine. For a crawler it is a problem, because running JavaScript is expensive.

Googlebot renders JavaScript, in a second pass that can lag the initial crawl by anything from hours to weeks.

Most AI crawlers do not render JavaScript at all. `GPTBot`, `ClaudeBot`, `PerplexityBot` and their siblings overwhelmingly fetch the raw HTML and read what is in it. If your content is not in that HTML, it does not exist to them.

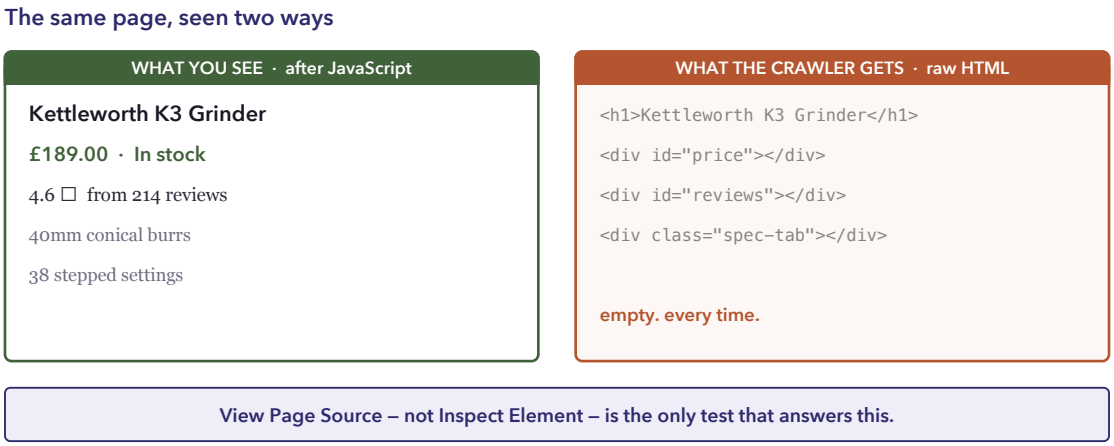


Figure 9.1 – A product page as a person sees it, and as a crawler receives it.

> KEY IDEA
>
> Googlebot will eventually see your JavaScript-rendered content. AI crawlers will not. A site that ranks perfectly well in Google can be **completely invisible** to every AI assistant for this reason alone, and nothing in the SEO reporting will show it.

How to check what a bot sees

Two methods, neither requiring any coding.

Method 1: view the source, not the page

In any browser, right-click and choose **View Page Source** (not "Inspect"). This shows the raw HTML as delivered. Use your browser's find function to search for a distinctive sentence from the middle of your article.

- **Found it?** The content is in the HTML. AI crawlers can read it.
- **Not found?** It is being built by JavaScript. AI crawlers cannot see it.

Do this on a product page, a blog post and your homepage. It takes two minutes and answers the question definitively.

Method 2: fetch the page the way a bot does

If you have terminal access, this shows exactly what arrives before any JavaScript runs:

```
curl -A "Mozilla/5.0 (compatible; GPTBot/1.0)" \
  https://yoursite.com/your-page | head -200
```

If the output contains navigation and a shell but none of your article text, you have confirmed the problem.

IN PRACTICE

The three-page rendering audit

Run View Page Source on:

1. **Your highest-value commercial page** — search for the price, the product name and the main benefit statement
2. **Your best-performing blog post** — search for a sentence from the middle
3. **A page with reviews, comments or user content** — these are the most commonly JavaScript-injected, and the most citable

Record for each: is the main content present, is the price present, are reviews present. If any answer is no, you have found a concrete, fixable reason for AI invisibility, and one that most competitors have not checked either.

The specific patterns that break things

Pattern	Why it breaks	The fix
Client-side-only rendering	No content in the HTML	Server-side rendering or static generation
Content behind “Load more”	Only appears on click	Render the first screenful in HTML; paginate with real links
Tabbed content	Hidden tabs often not in HTML	Include all tab content in the HTML, hide with CSS
Infinite scroll	No URLs for later items	Provide paginated URLs alongside
Reviews via a third-party widget	Loaded in an iframe or by script	Also render review text server-side
Prices set by script	Price absent from HTML	Include the price in the HTML and in structured data
Text inside images	Not readable as text	Real text, with the image as decoration
Content requiring interaction	Crawlers do not click	Expose it at a URL

COMMON MISTAKE

The reviews case deserves particular attention. Review content is exactly what AI assistants want — specific, experiential, comparative, and it is very commonly delivered by a third-party JavaScript widget that no AI crawler can read. Sites with hundreds of excellent reviews are frequently invisible for “is *[product]* any good”, which is precisely the question they should own.

What to ask a developer for

You do not need to implement this. You need to ask for the right thing, in language that will be understood.

> *"We need the main content of these page types to be present in the initial HTML response, before JavaScript executes. Several important crawlers don't execute JavaScript, so anything rendered client-side is invisible to them. Can we use server-side rendering or static generation for these templates?"*

That sentence is precise, is technically correct, and will get a useful answer. Expect the response to be about framework capability and build complexity; both are legitimate, and the fix is usually a configuration change rather than a rewrite.

CASE STUDY · KETTLEWORTH

With crawlers finally reaching the site, Kettleworth ran View Page Source on three page types.

The buying guides were fine — server-rendered, all text present. The product pages were not. The price was injected by script, so the HTML contained an empty element. The specification table lived in a tab that only populated on click. The reviews came from a third-party widget in an iframe.

To an AI crawler, a Kettleworth product page consisted of a title, a breadcrumb and a marketing paragraph.

The developer's estimate to server-render prices and specifications was three days. It had been costing them the entire product half of their catalogue for two years.

CHAPTER SUMMARY

- Rendering is the gap between the HTML delivered and the page a user sees
- Googlebot renders JavaScript on a delay; most AI crawlers do not render it at all
- View Page Source and search for your own sentences — the definitive two-minute test
- Reviews, prices and tabbed content are the most commonly invisible, and the most commercially damaging
- Ask developers for content in the initial HTML response via server-side rendering

CHECK YOURSELF

1. What is the practical difference between View Page Source and Inspect Element, and why does it matter here?
2. Why can a page rank well in Google and be entirely invisible to ChatGPT?
3. Which three content types are most commonly JavaScript-injected, and why is each commercially costly?
4. Write the request you would send a developer, in one paragraph.

CHAPTER 10

Architecture and Internal Linking

Making a site legible to a machine that reads passages

WHAT YOU'LL LEARN

- Why site structure still matters when retrieval works at passage level
- How internal links direct both crawlers and meaning
- Building topic clusters that map onto query fan-out
- URL and heading structure as machine-readable context

It is tempting to conclude that if retrieval is passage-level, site architecture no longer matters. The opposite is true: architecture is how a machine works out what your passages are *about* and how authoritative they are.

What internal links actually do

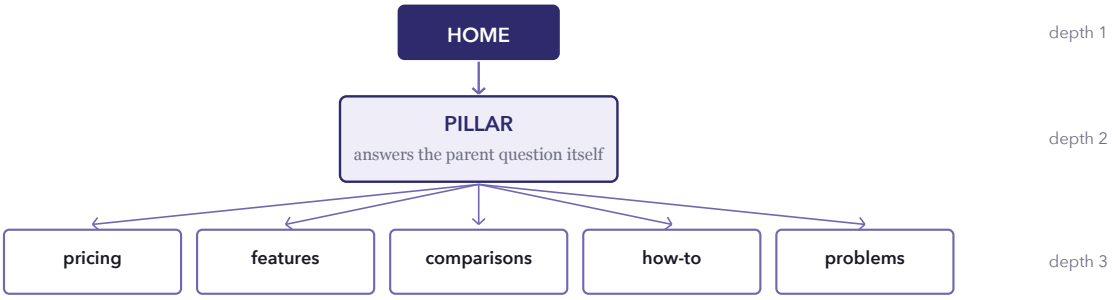
Three jobs at once:

1. **Discovery.** A crawler finds pages by following links. A page nothing links to is found only via a sitemap, and is treated as correspondingly unimportant.
2. **Standing distribution.** Authority earned by one page flows through internal links to others. Your homepage usually has the most; the pages it links to inherit some.
3. **Semantic context.** This is the one that matters most for AI search. The anchor text and the surrounding sentence tell the system what the target page is about. A link reading “our guide to agency CRM pricing” is a much stronger signal than one reading “click here”.

KEY IDEA

Internal links are the cheapest, most controllable ranking and retrieval lever you have. You own every one of them. No outreach, no budget, no approval from anyone outside your team.

Click depth and the topic cluster



Links run both ways, with descriptive anchor text.
Beyond depth four, pages are crawled less and treated as less important.

Figure 10.1 – Click depth and the hub-and-spoke cluster.

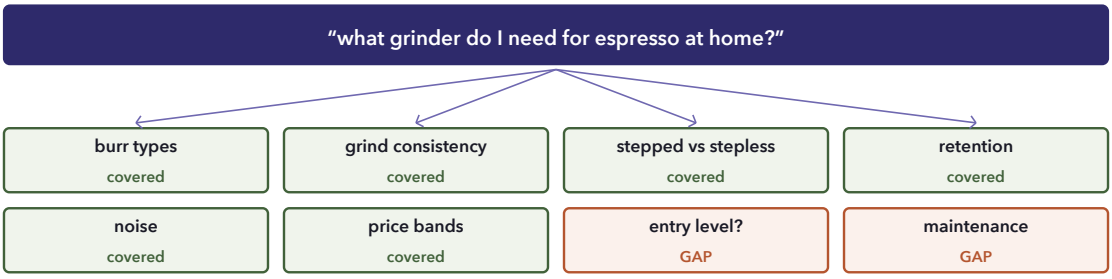
Topic clusters, and why fan-out makes them work harder

The hub-and-spoke model — one comprehensive pillar page on a broad topic, linked bidirectionally with focused pages on sub-topics — predates AI search. Query fan-out makes it substantially more valuable.

Recall from Chapter 4 that one question becomes eight to twelve sub-queries. A well-built cluster gives you a strong candidate passage for many of those sub-queries at once, rather than one page competing for all of them.

The mechanism behind this is entity association, which Chapter 14 covers in full — see Figure 14.1 for how a brand becomes connected to a topic across independent sources.

One question, eight sub-queries, one cluster



The gaps are the content plan – derived from how retrieval works, not from search volume.

Figure 10.2 – The fan-out turned into a content plan. The gaps are the brief.

IN PRACTICE

Building a cluster from the fan-out

Work backwards from how the system retrieves:

1. Choose a commercially important question: *"how do agencies price retainers"*
2. Generate the fan-out: ask a model for the 12 sub-questions it would research. You will get things like: what is a retainer, typical retainer sizes, hourly versus value pricing, how to raise retainer fees, retainer contract terms, when retainers go wrong.
3. Map each sub-question to a page. Some exist; some do not.
4. **Write the pillar page to answer the parent question directly** — not as a table of contents that fobs the reader off onto other pages. It must stand alone.
5. **Write each spoke to answer its sub-question completely**, again standing alone.
6. Link every spoke to the pillar and the pillar to every spoke, using descriptive anchor text.
7. Cross-link spokes where genuinely related.

The result is a structure where a fan-out across this topic finds you repeatedly rather than once.

COMMON MISTAKE

The failure mode is the pillar page that answers nothing. Pages that say "There are many factors in retainer pricing. Read our guide to hourly billing, our guide to value pricing…" are useless in AI search: no passage on them answers anything, so nothing is retrievable. **Every page must be independently useful**, including hubs.

Content modelling: deciding what your pages *are*

Before architecture comes a question most marketing teams never ask explicitly: what **types** of page does this site have, and what does each type always contain?

A content model is that answer written down. For a typical B2B site it might be five types — service page, case study, guide, author profile, comparison — each with required fields. A guide always has: a question-shaped title, a first-sentence answer, a named author, a published and updated date, a specific figure or example, and internal links to its cluster.

Why this matters for retrieval:

- **Consistency is a signal.** When every guide carries an author and a date, the whole content type becomes legible rather than each page being a special case.
- **It makes structured data automatable.** A defined type maps cleanly onto a schema type, so markup is generated rather than hand-written per page.

- **It stops drift.** Without a model, twenty writers over four years produce twenty formats, and the decay audit in Chapter 17 becomes archaeology.
- **It makes gaps visible.** "We have no comparison pages" is a much more actionable finding than "our content feels thin."

IN PRACTICE

Writing your content model in an hour

1. List every distinct kind of page on your site. Most sites have four to seven; if you have twenty, you have drift rather than variety.
2. For each, write the fields it must always have. Be strict — a field that is optional is not part of the model.
3. Map each type to a schema type: guide › `Article`, service › `Service` or `WebPage`, case study › `Article`, profile › `Person`.
4. Check ten existing pages of each type against the model. The mismatches are your backlog.
5. Put the model in the brief template (Chapter 59) so new content is born compliant.

URLs, headings and machine-readable context

URLs

Keep them short, readable and stable. `/blog/agency-retainer-pricing` beats `/blog/2024/11/post-4471?ref=nav`. URLs are weak ranking signals but strong *context* signals — they appear in citations, and a readable URL communicates what a page is about to both humans and machines.

Above all, **do not change them without redirects**. A URL change without a 301 discards everything the old URL had accumulated.

Heading structure

Headings do more work in AI search than they did in classical SEO, because they define the boundaries of the chunks discussed in Chapter 4. A chunking system uses headings as natural split points.

The practical rules:

- **One `H1` per page**, describing what the page is
- `H2` **for each major section**, phrased as far as possible as *the question that section answers*
- `H3` **for sub-sections**, in a genuine hierarchy — never skip levels for styling reasons
- **Each section under a heading should stand alone**

That third rule about phrasing is the highest-leverage change in this chapter. "Pricing" is a weak `H2`. "How much does an agency retainer cost?" is a strong one, because it matches the

shape of a sub-query directly.

Chapter 13 shows the effect of this in full — Figure 13.1 contrasts the same information structured two ways, only one of which can be lifted into an answer.

IN PRACTICE

The heading rewrite

Open your five most important pages and list every H2 and H3. For each, ask: *is this a question a real person would type or say?*

- "Features" › "What can *[product]* actually do?"
- "Pricing" › "How much does *[product]* cost?"
- "About us" › "Who is *[company]* and who do they work with?"
- "Benefits" › "Why choose *[product]* over *[main alternative]*?"

Then ensure the first sentence under each heading answers the question directly. This is perhaps two hours of work per page and it addresses gates three, four and five simultaneously.

Depth and orphans

Two structural problems worth checking:

Click depth is how many links from the homepage a page sits. Anything beyond four is crawled less and treated as less important. If a commercially important page is six clicks deep, restructure.

Orphan pages have no internal links at all. They are usually accidental — pages made for a campaign, or left behind after a navigation change. They are invisible to discovery. A crawl tool will list them; so will comparing your sitemap against your internal link graph.

CASE STUDY · KETTLEWORTH

Kettleworth built their first proper cluster around espresso grinding.

The pillar answered *"what grinder do you need for espresso at home?"* directly and completely — not as a table of contents. Around it sat eight spokes, each answering one sub-question from Chapter 4's fan-out, each linked to and from the pillar with descriptive anchor text.

The hardest part was cultural. Their previous pillar page had deliberately withheld the recommendation to drive clicks to product pages. Making it answer the question directly felt like giving something away.

Within two months the pillar was being cited for three distinct sub-queries. It also converted better, because the people arriving had been given a straight answer.

CHAPTER SUMMARY

- Internal links do three jobs: discovery, authority distribution and semantic context
- Topic clusters map directly onto query fan-out, giving you multiple candidate passages per question
- Every page must be independently useful — hub pages that only point elsewhere retrieve nothing
- Phrase headings as the questions they answer; this aligns chunk boundaries with sub-queries
- Check click depth and orphan pages; both are common and both are quietly expensive

CHECK YOURSELF

1. Why does site architecture still matter if retrieval works at passage level?
2. What is wrong with a pillar page that mainly links to other pages?
3. Rewrite these headings as questions: “Integrations”, “Our Process”, “Support”.
4. Why is descriptive anchor text more valuable in AI search than it was previously?

CHAPTER 11

Structured Data and Schema

Telling machines what your content means, without hoping they infer it

WHAT YOU'LL LEARN

- What structured data is and what it genuinely does
- The schema types that carry weight, and the ones that are decoration
- How structured data supports E-E-A-T and citation
- Honest limits: what schema will not fix

Structured data is a way of stating facts about your page in a format machines parse unambiguously, rather than leaving them to infer those facts from prose. It is one of the few genuinely technical levers where the evidence is reasonably good — and one where the claims still outrun it.

What it is

Schema.org provides a shared vocabulary. You embed it in your page, usually as a block of JSON in the HTML, and it states things explicitly: this is a product, its price is £49, it has 214 reviews averaging 4.6, the author is Jane Smith who is a cardiologist.

```
{
  "@context": "https://schema.org",
  "@type": "Product",
  "name": "Acme Field Notebook",
  "description": "A pocket notebook for site surveys.",
  "brand": { "@type": "Brand", "name": "Acme" },
  "offers": {
    "@type": "Offer",
    "price": "12.50",
    "priceCurrency": "GBP",
    "availability": "https://schema.org/InStock"
  },
  "aggregateRating": {
    "@type": "AggregateRating",
    "ratingValue": "4.6",
    "reviewCount": "214"
  }
}
```

You do not need to write this by hand. Most content management systems generate it via a plugin. What you need is to know **which types matter** and to check that what is generated is accurate.

What it actually does

Three things, in descending order of confidence:

1. **It enables rich results.** Star ratings, prices, FAQ dropdowns and recipe cards in search results are driven by structured data. This is well established and directly observable.
2. **It disambiguates entities.** By stating that your `Organization` has a particular `sameAs` link to a Wikipedia page or LinkedIn profile, you connect your brand to a known entity rather than leaving a system to guess. This matters considerably for the entity work in Chapter 14.
3. **It assists machine understanding and verification.** Both Google and Microsoft have confirmed structured data helps their AI systems understand and verify content. Structured data acts as a stated source of truth that a system can check prose against.

KEY IDEA

Structured data does not make weak content rank or get cited. It makes **good content easier to understand, verify and reuse**. Think of it as removing ambiguity, not adding authority.

MYTH vs REALITY

"Schema markup gives a 2.5× higher chance of appearing in AI answers"

You will see figures like this, and like "40% more AI Overview appearances with complete schema". Run them through the Chapter 7 filter:

- *What was measured?* Usually a correlation between sites having schema and sites appearing in AI results.
- *Is there a confound?* Enormously. Sites with complete, accurate structured data are overwhelmingly sites with competent technical teams, better content and stronger authority. The schema is a **marker** of quality operations, not necessarily the cause of the visibility.
- *Plausible mechanism?* Yes, partially — disambiguation genuinely helps.

Sensible conclusion: implement it because it is cheap, well-supported and removes ambiguity. Do not expect it to be transformative on its own, and be sceptical of the multiplier claims.

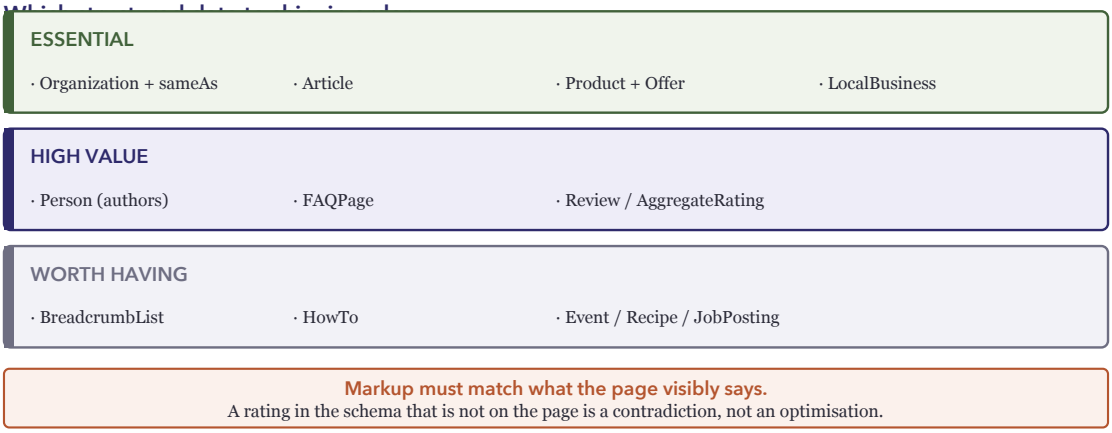


Figure 11.1 — What to ship, in order, and the rule that governs all of it.

The types that carry weight

Schema type	Use it for	Priority
`Organization`	Your company identity, `sameAs` links	**Essential**
`Article` / `BlogPosting`	Editorial content, with author and dates	**Essential**
`Product` + `Offer`	Anything with a price	**Essential** if selling
`FAQPage`	Genuine question-and-answer sections	High
`Person`	Author profiles, credentials	High
`BreadcrumbList`	Site hierarchy	Medium
`HowTo`	Step-by-step instructions	Medium
`LocalBusiness`	Physical premises	Essential if local
`Review` / `AggregateRating`	Genuine reviews you host	High
`Event`, `Recipe`, `JobPosting`	Self-explanatory	As relevant

The `Organization` block is the one people skip

It is the most under-implemented and among the most valuable, because it is how you assert your identity in machine-readable form:

```
{
  "@context": "https://schema.org",
  "@type": "Organization",
  "name": "Acme Consulting",
  "url": "https://acme.example",
  "logo": "https://acme.example/logo.png",
```

```
  "description": "Operations consulting for creative agencies.",
  "foundingDate": "2019-03-01",
  "sameAs": [
    "https://www.linkedin.com/company/acme-consulting",
    "https://en.wikipedia.org/wiki/Acme_Consulting",
    "https://www.crunchbase.com/organization/acme-consulting"
  ]
}
```

The `sameAs` array is doing the important work: it connects your site to independent profiles, which is how a machine confirms you are a real, identifiable entity rather than an unknown string. Chapter 14 develops this.

COMMON MISTAKE

Structured data must match what is visible on the page. Marking up a rating of 4.8 that does not appear anywhere on the page, or an `Offer` price that differs from the displayed price, is a guideline violation that can cost you rich results entirely. It is also, in the AI context, actively counterproductive: you are feeding a verification system a contradiction.

IN PRACTICE

Validating what you have

1. Go to Google's **Rich Results Test** and enter one URL from each page template.
2. Note which types are detected and which errors and warnings appear.
3. Fix errors — these prevent the markup working. Warnings are usually optional fields worth filling in.
4. Repeat on **schema.org's own validator**, which checks the vocabulary more strictly than Google's tool.
5. In Search Console, open the **Enhancements** reports for site-wide errors.

Then check the `Organization` block specifically. On most sites it is either missing, or present with no `sameAs` array — the highest-value quick fix in this chapter.

CASE STUDY · KETTLEWORTH

Kettleworth's structured data was better than average and missing the piece that mattered most: they had no `Organization` markup at all.

They added it — name, URL, logo, founding date, a plain description, and a `sameAs` array linking to their Companies House record, LinkedIn page and Crunchbase profile.

It took forty minutes. Six weeks later, the "coffee shop" confusion from Chapter 3 had disappeared from three of four assistants.

Correlation, not proof — they had also corrected the directory listing in the same period. But it cost forty minutes, it removed a genuine ambiguity, and the wrong answer stopped appearing.

CHAPTER SUMMARY

- Structured data states facts about your page unambiguously rather than leaving them to inference
- It enables rich results, disambiguates entities and assists machine verification
- `Organization` with `sameAs`, `Article` with author, and `Product` with `Offer` are the priorities
- It cannot rescue weak content — it clarifies good content
- Markup must match visible content, or it becomes a liability

CHECK YOURSELF

1. What are the three things structured data does, in order of how confident you should be?
2. Why should you be sceptical of "schema gives 2.5× more AI citations"?
3. What is the `sameAs` array for, and why does it matter for entity recognition?
4. Why is marking up a rating that is not visible on the page a bad idea for two separate reasons?

CHAPTER 12

llms.txt and the Emerging Standards

A short chapter, because the honest answer is short

WHAT YOU'LL LEARN

- What `llms.txt` proposes and why it was appealing
- What the adoption and crawl data actually shows
- How to think about emerging standards generally
- What to do about it — a defensible, low-cost position

This chapter exists because you will be asked about `llms.txt`, probably by someone who read that it is essential. The evidence is unusually clear, and knowing it will save you from spending real effort on something that currently does very little.

The proposal

`llms.txt` is a proposed file at the root of your site — by analogy with `robots.txt` — offering a curated, Markdown-formatted guide to your most important content. The pitch: rather than making a model crawl and parse your whole site, hand it a clean summary of what matters and how you would like to be described.

It is a genuinely sensible idea. Curation on top of access control. The problem is not the concept.

What the data shows

BY THE NUMBERS

- **97%** of `llms.txt` files are never fetched, per an Ahrefs study of 137,000 sites
- In a 90-day monitoring window covering over **500 million AI bot visits**, only **408** requests targeted `llms.txt` directly — a negligible fraction
- Adoption sits at roughly **10.1%** of 300,000 domains studied by SE Ranking, rising to ~28% among technically sophisticated sites
- **Google does not support it and has said it does not plan to.** John Mueller compared it to the keywords meta tag — the classic example of a signal that was proposed, adopted, ignored, and eventually became a spam marker
- **Adoption among the top 1,000 sites** was 8.7% by June 2026 — lower than the 10% across the broader web, which suggests the largest publishers are not persuaded either
- **No major provider** — OpenAI, Google, Anthropic, Meta, Perplexity or Mistral — has publicly stated that its production search systems read or act on the file

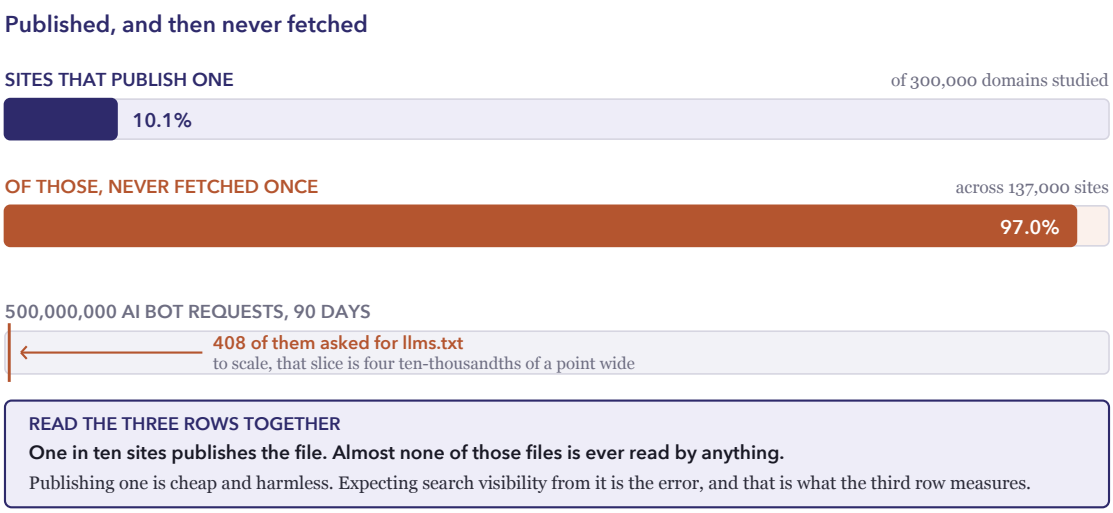


Figure 12.1 – Adoption, fetches, and the scale of the gap between them.

KEY IDEA

There is one place `llms.txt` genuinely works, and it is not search. **Coding assistants and IDE agents do fetch it** — Cursor, Claude Code, Copilot, Windsurf — along with some MCP servers and in-product AI assistants. A developer-facing company with an API has a real reason to publish one: it helps an agent decide which of your endpoints to read.

That is a documentation decision, not a GEO one. If you sell to developers, ship it for that reason. If you do not, the evidence above stands.

The gap between adoption and usage is the whole story. Roughly one site in ten has implemented it. Almost no crawler reads it.

MYTH vs REALITY

"You need llms.txt for AI visibility"

Claim: implementing `llms.txt` improves how AI systems see your site.

Evidence: the crawl figures above. Every major AI crawler skips the file and fetches HTML directly; only Perplexity confirms retrieving it.

Verdict: implementing it is close to harmless and close to useless. It is not a scam — it is a reasonable proposal that has not been adopted by the systems that would need to adopt it for it to matter.

What to do: if it takes an hour, fine, do it for the Perplexity case and optionality. If it would take a sprint, do literally anything else in Parts III to V instead.

How to think about emerging standards

`llms.txt` is a useful case study, because more proposals will follow. The test for any of them:

1. **Who has committed to reading it?** Not "who might" — who has stated they do. A standard is only as real as its consumers.
2. **Can you verify consumption?** Server logs will tell you whether anything fetches the file. This is checkable, and almost nobody checks.
3. **What is the cost of being wrong in each direction?** Implementing something useless costs an hour. Missing something that becomes universal costs more. The asymmetry usually favours cheap implementation, but only when it is cheap.
4. **Is it displacing real work?** This is the real risk. A team that spends a sprint on `llms.txt` while its CDN blocks `OAI-SearchBot` has actively harmed itself by working on the wrong layer.

COMMON MISTAKE

The genuine danger of standards like this is **displacement**. The work in Chapters 8, 9 and 13 is unglamorous and definitely matters. A novel file with a memorable name is more interesting to talk about in a meeting. Do not let the interesting thing crowd out the effective thing.

IN PRACTICE

A defensible position for a stakeholder

When asked "should we implement `llms.txt`?", the answer that holds up:

> *"It's cheap, so I've added it, but I want to be straight that the data doesn't support treating it as a priority. Studies show around 97% of these files are never fetched, and Google has said it doesn't*

> *support them. Only Perplexity confirms reading it.*
>
> >
>
> > *The things that measurably affect whether AI assistants can see us are: whether our CDN is blocking their crawlers, whether our content is in the HTML rather than built by JavaScript, and whether our pages answer questions in liftable passages. I've prioritised those. I'll keep monitoring* `llms.txt` *adoption and revisit if the engines start fetching it."*

That answer demonstrates you have read the evidence, made a proportionate decision, and know what actually matters. It is a considerably better look than either enthusiasm or dismissal.

CASE STUDY · KETTLEWORTH

Kettleworth had already implemented `llms.txt`. It was, in fact, the first thing they had done — six weeks earlier, after their founder read an article calling it essential. It had consumed most of a developer sprint, because they built tooling to generate it automatically from the CMS.

During the same six weeks, their CDN was returning 403 to every AI crawler that asked.

Their server logs showed `/llms.txt` had been fetched four times in three months, twice by monitoring tools they ran themselves.

This is displacement, in one company, costing one sprint. The file is still there. It does no harm.

CHAPTER SUMMARY

- `llms.txt` proposes a curated Markdown guide for AI systems — a sensible idea
- 97% of these files are never fetched; only 408 requests across 500 million bot visits targeted one
- Google does not support it; only Perplexity confirms retrieving it
- Implement it if it is an hour's work; never at the expense of crawler access, rendering or content structure
- Judge future standards by who has committed to reading them, and verify in your own logs

CHECK YOURSELF

1. What does `llms.txt` propose, and why is the idea reasonable?
2. What does the crawl data show, and what is the single most telling number?
3. What is “displacement” and why is it the real risk here?
4. Draft your own two-sentence answer to a CEO who read that `llms.txt` is essential.

PART IV

Content

Gates three, four and five. This is where AI search is actually won, and where most of your time should go.

CHAPTER 13

Writing for Retrieval

The craft of building passages a machine can lift

WHAT YOU'LL LEARN

- The single test that determines whether a passage is retrievable
- The answer-first pattern, and how to apply it without writing badly
- Specificity as a retrieval signal — why vagueness is now a technical fault
- How to restructure existing content without damaging what already works

If you only apply one chapter of this book, apply this one. Technical work removes obstacles; this is the work that actually earns citations. It is also, conveniently, work you can do yourself without a developer, a budget or anyone's permission.

The test

Everything in this chapter reduces to one question, applied to every section of everything you publish:

> *If a machine lifted this section out, showed it to a stranger with no other context, and nothing else from the page — would it make complete sense, and would it answer something?*

That is the test. It is not a metaphor for the test; it is mechanically what happens, as Chapter 4 described. A chunking system splits your page at heading boundaries, and a reranker evaluates the resulting passage on its own merits.

KEY IDEA

Self-containment is the property. A passage that begins "As we saw above…", "This is why the third factor matters…" or "Building on that…" has failed. It cannot be lifted, so it will not be cited, however well written it is.

The answer-first pattern

The most common structural fault in marketing content is burying the answer. Writers were trained — often correctly, for humans — to build context first and deliver the payoff at the end. In retrieval, this is fatal: the chunk containing the build-up gets evaluated, finds no answer in it, and loses.

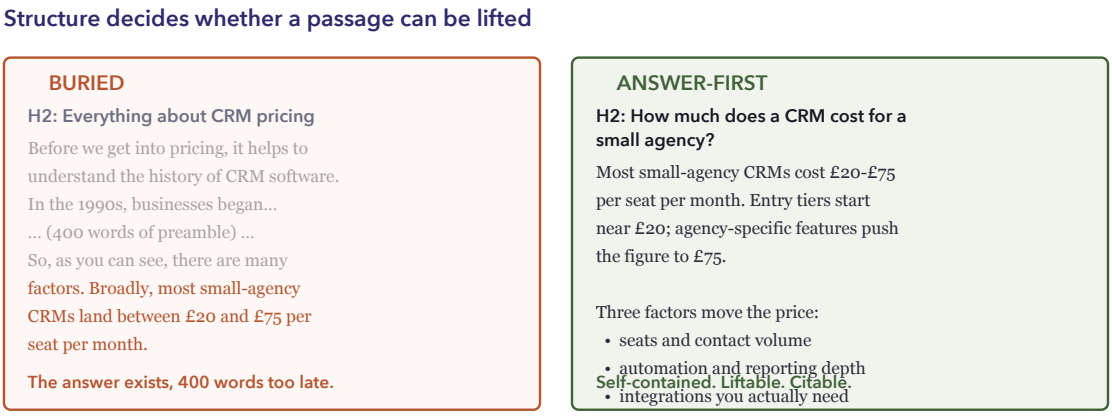


Figure 13.1 – Identical information, structured two ways. Only one is liftable.

The pattern:

1. **Heading = the question**, phrased as a person would ask it
2. **First sentence = the direct answer**, complete and specific
3. **Following sentences = the qualification, nuance and context**
4. **Then examples, caveats and detail**

Note this is not the same as writing shallowly. All the nuance stays. It moves *after* the answer instead of before it.

IN PRACTICE

The rewrite, worked

Before — a real pattern from a consultancy site:

> *### Our approach to pricing Every business is different, and we believe in tailoring our engagement to your specific needs. Over the years we've worked with organisations of all sizes, and one thing we've learned is that a one-size-fits-all approach rarely serves clients well. That's why we start every relationship with a discovery conversation...*

This passage answers nothing. Retrieved on its own, it tells a reader precisely one thing: that this company will not tell you what it costs.

After:

> *### How much does an engagement cost? Most of our engagements run between £4,000 and £12,000 per month, over an initial six-month term. Smaller diagnostic projects start at £6,500 as a one-off.*
>
> *The range depends on three things: the number of workstreams running at once, whether you need us embedded on site, and the seniority of the team assigned. A single-workstream retainer with fortnightly remote sessions sits at the bottom of that range; a three-workstream programme with a weekly on-site day sits at the top.*
>
> *We confirm the exact figure after a discovery call, but you should not need a call to know whether we're in your budget.*

Same company, same policy, same nuance. The second version can be retrieved and cited for "how much do operations consultants cost", and the first cannot.

Specificity is now a technical requirement

Vague copy has always been weak marketing. In retrieval it is a technical fault, because a passage containing no specific information is a poor match for any specific question.

Vague	Specific
"Affordable pricing"	"£12.50 per user per month"
"Fast delivery"	"Dispatched same day if ordered before 3pm"
"Works with your existing tools"	"Native integrations with Slack, HubSpot, Xero and Notion"
"Trusted by leading brands"	"Used by 340 agencies across the UK and Ireland"
"Significant improvement"	"Cut average response time from 14 hours to 3"
"Suitable for small businesses"	"Designed for teams of 3-25 people"

The right-hand column is retrievable. The left is not — there is no question for which "affordable pricing" is a good answer.

COMMON MISTAKE

The overcorrection is to bolt numbers onto everything, including places where you do not have them, or to invent precision. This is worse than vagueness: AI systems cross-check claims across sources, and contradictions damage the trust signals in Chapter 15. If you do not have a number, say what you do know precisely: "pricing depends on seat count; most customers land between X and Y" is specific and honest.

Formatting that helps

Some formatting genuinely assists chunking and extraction. Some is cargo cult. The distinction:

Genuinely helpful

- **Question-shaped headings** — they align chunk boundaries with sub-queries
- **Short paragraphs** — two to four sentences; dense blocks are harder to extract cleanly
- **Lists for genuinely enumerable things** — steps, options, criteria
- **Tables for comparisons** — machine-readable and directly liftable
- **Definitions stated plainly** — "X is Y" is a highly citable sentence shape
- **A summary near the top** for long pieces

Cargo cult

- Bolding random phrases for "emphasis signals"
- Adding an FAQ section that repeats the page content in question form
- Chopping everything to a uniform 150 words
- Inserting statistics unrelated to the argument because "statistics improve GEO"
- Keyword variants stuffed into headings

MYTH vs REALITY

"Adding an FAQ section to every page improves AI visibility"

FAQ sections help when they contain **genuine questions with genuine answers not covered elsewhere on the page**. They do nothing when they restate the page's content in question form — you are giving the retrieval system two near-duplicate passages, which can dilute rather than concentrate relevance.

Better: rather than appending an FAQ, restructure the page itself so its H2s *are* the questions. Same benefit, no duplication, better for readers.

Restructuring without breaking what works

If a page currently ranks and converts, you should be careful. The safe sequence:

1. **Do not change the URL.** Ever, without a redirect.
2. **Do not cut content.** Restructure it. Removing text that is currently earning rankings is the main way this goes wrong.
3. **Reorder within sections** so answers come first. This is usually a matter of moving the last sentence of a paragraph to the front.
4. **Rewrite headings** as questions.
5. **Add specifics** where you were vague — this is usually additive.
6. **Note the date** and watch rankings for four weeks.

IN PRACTICE

The 90-minute page upgrade

For one important page:

1. **20 min** — List every heading. Rewrite each as the question it answers.
2. **30 min** — For each section, move the direct answer to the first sentence.
3. **20 min** — Find every vague claim and replace it with a specific one, or state honestly what varies and why.
4. **10 min** — Read each section in isolation. Fix any that reference "above", "below" or "as mentioned".
5. **10 min** — Add one comparison table or one clearly enumerated list where the content warrants it.

Then leave it alone for four weeks. The page-structure checklist in Chapter 58 covers the same ground in a form you can tick off. Do this to your ten most commercially important pages and you will have done more for your AI visibility than any amount of schema or `llms.txt`.

CASE STUDY · KETTLEWORTH

Kettleworth rewrote their grinder guide using the 90-minute method. Nothing was deleted; the word count rose slightly.

"Getting Started" became "Do you need a burr grinder for espresso, or will a blade grinder do?" — answered in the first sentence: *No. Espresso requires a burr grinder, because blade grinders produce inconsistent particle sizes that cause channelling.*

"A Note on Burrs" became "What is the difference between conical and flat burrs?"

Eight headings, rewritten as questions, each with the answer moved to the first sentence.

Five weeks later the page was being cited for four distinct sub-queries. It had ranked in the same position throughout.

CHAPTER SUMMARY

- The test: does this section make complete sense and answer something when lifted out alone?
- Answer first, then qualify — the nuance stays, it just moves after the answer
- Specificity is now a technical requirement, not just good writing
- Question-shaped headings align chunk boundaries with sub-queries
- Restructure rather than rewrite; never change URLs without redirects

CHECK YOURSELF

1. State the self-containment test in your own words.
2. Take this heading and first sentence and fix them: “Our Services — We offer a comprehensive range of solutions tailored to your needs.”
3. Why can adding an FAQ section actively dilute relevance?
4. Why is inventing precision worse than admitting vagueness?

CHAPTER 14

Entities and Topical Authority

Being a known thing, connected to a known subject

WHAT YOU'LL LEARN

- What an entity is and why machines think in entities rather than keywords
- How to establish your brand as a recognised entity
- What topical authority means mechanically, and how to build it
- Why consistency across the web matters more than any single page

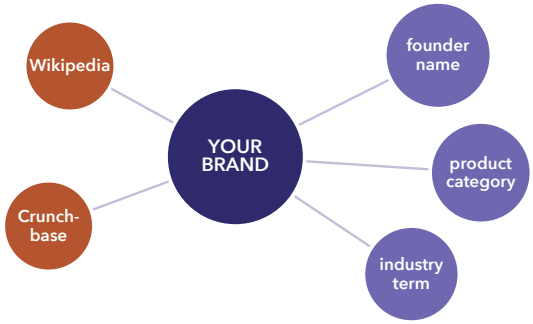
Chapter 3 explained that models work with meaning rather than strings. This chapter is the practical consequence: search systems maintain a model of *things in the world* — people, companies, products, concepts — and the relationships between them. Being a recognised thing is a prerequisite for being an answer.

What an entity is

An entity is a specific, identifiable thing that exists independently of the words used to describe it. "Apple" the company is an entity. "Apple" the fruit is a different entity. The string is identical; the entities are unrelated.

Search systems resolve strings to entities using context and, crucially, corroboration from multiple independent sources. This is why the `sameAs` structured data from Chapter 11 matters: it explicitly connects your string to known identifiers.

Entities: how machines connect your brand to a topic



WHY IT MATTERS

A model answers about topics, not keywords. If your brand is not connected to the topic across independent sources, you are not a candidate answer — however well your page is written.

Figure 14.1 – Your brand becomes an entity by being consistently connected to the same attributes across independent sources.

KEY IDEA

If a system cannot resolve your brand name to a specific entity, it cannot confidently associate you with a topic, and it will not offer you as an answer. **Entity recognition is a precondition, not an optimisation.**

How to become a recognised entity

The mechanism is corroboration: the same facts about you, stated consistently, across sources the system already trusts.

1. Consistency of your core facts

Your company name, founding date, location, leadership and description should be **identical** everywhere they appear. Not similar — identical. Variation ("Acme Ltd", "Acme Consulting", "Acme Consulting Limited", "ACME") fragments the signal across what may be treated as several entities.

Write down the canonical version of each fact once, and use it everywhere.

2. Presence on the corroborating sources

These are the sources that entity systems lean on. In rough order of weight:

Source	Why it matters	Difficulty
Wikipedia	Heavily weighted; strong entity anchor	Very hard, and requires genuine notability
Wikidata	The structured layer beneath Wikipedia	Moderate
LinkedIn company page	Universally referenced, especially B2B	Easy
Crunchbase	Standard for company data	Easy
Companies House / registry	Authoritative primary record	Automatic
Industry directories	Category association	Easy
Google Business Profile	Essential if you have premises	Easy

> **COMMON MISTAKE**
>
> Do not attempt to create a Wikipedia article about your own company. Wikipedia's conflict-of-interest rules are strictly enforced, the article will very likely be deleted, and the attempt can attract lasting negative attention. Wikipedia presence is a *consequence* of genuine notability — sustained independent coverage — not something you can manufacture. Wikidata, by contrast, is a legitimate and much lower-barrier place to establish structured facts.

3. Named people with real profiles

Entity systems connect organisations to people. A company whose founder, key staff and authors are themselves recognised entities — with consistent profiles, bylines and independent mentions — is far more legible than one presenting as a faceless brand.

This is one of the highest-leverage moves available to a small business, because it is entirely within your control and most competitors have not done it. Give every author a real profile page with credentials, link it from every article, mark it up with `Person` schema, and connect it to their LinkedIn via `sameAs`.

Topical authority, mechanically

"Topical authority" is used loosely. Mechanically, it means: **across the corpus, is this source consistently associated with this topic?**

That association is built from:

- **Depth of coverage** — you address the topic and the concepts adjacent to it, not one page in isolation
- **Consistency over time** — publishing on it repeatedly, not once
- **External corroboration** — others discussing you in connection with the topic
- **Internal coherence** — your site's structure clearly organises around it

This is why the topic cluster from Chapter 10 works, and why a scattergun content calendar does not. Ten articles across ten unrelated topics build authority in none of them. Ten articles covering one topic thoroughly build authority in one.

IN PRACTICE

The topical authority audit

1. List every page you have published in the last 18 months.
2. Assign each to a topic. Be strict — "marketing" is not a topic; "email deliverability for e-commerce" is.
3. Count pages per topic and sort.
4. Look at the tail: topics with one or two pages.

For each tail topic, decide honestly: **commit or cut.** Either plan six to ten more pieces to build genuine depth, or accept you will not be an authority there and stop publishing into it.

Most content calendars fail this audit badly. Choosing two topics to own beats being adjacent to twelve.

MYTH vs REALITY

"You need to publish constantly to build topical authority"

Volume without coherence builds nothing. A site publishing weekly across scattered topics accumulates pages, not authority.

The mechanism rewards **depth within a topic** and **corroboration from outside**. Twelve well-researched pieces covering one subject completely, plus mentions from others in that subject, beats fifty scattered posts, and takes less time.

The consistency check

A short exercise that reliably finds problems:

1. Search your exact brand name on Google. Read the first two pages.
2. Note every description of what you do. Are they consistent?
3. Check your LinkedIn, Crunchbase, directory listings and your own About page side by side.
4. Ask ChatGPT, Claude and Perplexity "what does *[brand]* do?" and compare to your intended description.

Divergence here is extremely common — an outdated description on a directory from four years ago, a LinkedIn page describing a pivoted-away-from product, and it directly produces the wrong answers models give about you.

CASE STUDY · KETTLEWORTH

There is a Kettleworth in Cheshire that makes garden furniture. For two years, assistants had been blending the two.

The fix was unglamorous and mostly clerical. They standardised on "Kettleworth Coffee" everywhere their name appeared — website, LinkedIn, Crunchbase, eleven directories, and the footer of every invoice. They added founding date and location consistently. They created Wikidata entries for the company and for their head of buying.

They also gave every guide a named author with a profile page, credentials, and a `sameAs` link to LinkedIn — replacing the byline that had said "Kettleworth Team".

The garden furniture confusion took about four months to clear.

CHAPTER SUMMARY

- Systems resolve strings to entities; unrecognised entities cannot be confidently offered as answers
- Consistency of core facts across independent sources is the primary mechanism
- Named people with real profiles are a high-leverage, under-used entity signal
- Topical authority means consistent association with a topic — depth beats breadth
- Audit for divergent descriptions across the web; they directly produce wrong AI answers

CHECK YOURSELF

1. Why can a system fail to offer you as an answer even when your content is excellent?
2. Why is creating your own Wikipedia article a bad idea, and what is the better alternative?
3. Your content covers 14 topics with two pages each. What does the audit tell you to do?
4. How does author profile work connect to entity recognition?

CHAPTER 15

E-E-A-T, Trust and Why It Now Matters More

The framework behind quality judgements, and how it translates to AI search

WHAT YOU'LL LEARN

- What E-E-A-T is, and what it is not
- Why the first E — Experience — is the one most brands can actually win
- How trust signals translate into retrieval and citation
- Concrete signals you can add, in priority order

E-E-A-T stands for Experience, Expertise, Authoritativeness and Trust. It is not a ranking factor — it is a framework Google's human quality raters use to evaluate results, which then informs how systems are trained. That distinction matters, but the practical implication is the same: content demonstrating these properties performs better, and in AI search the effect strengthens.

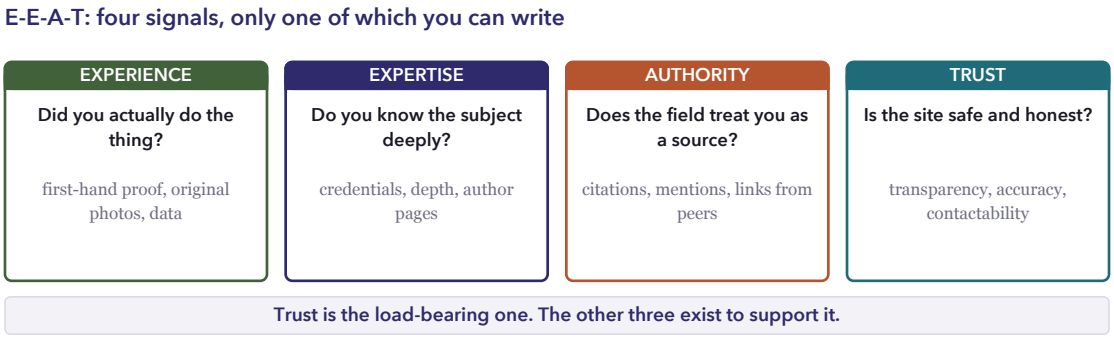


Figure 15.1 – Four signals, one of which you can create directly and three you mostly earn.

Why it matters more in AI search

Two reasons, both mechanical.

First, hallucination risk makes source quality expensive. A system that cites a wrong source produces a wrong answer with a name attached to it. That is far more damaging than a wrong link on a results page, so retrieval systems weight source reliability heavily.

Second, the systems increasingly cross-check. When several retrieved passages disagree, something must adjudicate. Corroboration across trusted sources is how. Content that agrees with the established record, from a source with standing, survives that process.

Experience – the winnable one

The first E was added in 2022 and is the most under-exploited. It asks a simple question: **did the author actually do the thing?**

This is where most brand content fails and where user-generated content wins, as Chapter 7 established. A review written by someone who used a product for six months contains information that cannot be synthesised from marketing copy: what broke, what surprised them, what they would do differently.

KEY IDEA

Experience is the one E you can create directly and quickly. Expertise takes credentials, authority takes years, trust takes consistency, but you can add genuine first-hand experience to your content this week, and almost nobody in your competitive set has.

IN PRACTICE

Injecting genuine experience

Concrete moves, in ascending order of effort:

- **Original photographs.** Not stock. Your actual office, product, process, whiteboard. Stock imagery signals the opposite of experience.
- **Specific numbers from your own work.** "In 34 client engagements, the median implementation took 11 weeks."
- **Named failure modes.** "This approach does not work when the team is under six people, and here is what we saw when we tried."
- **Process detail only a practitioner would know.** The step everyone gets wrong, the tool that breaks, the question the client always asks in week three.
- **First-person accounts with dates and context.** "When we migrated a 40,000-SKU catalogue in March, the thing that cost us two weeks was…"
- **Original data.** Covered fully in Chapter 16.

Test: could a competent writer with no experience of your field have produced this paragraph from desk research? If yes, it demonstrates no experience.

Expertise

Who wrote this, and are they qualified? Practically:

- **Named authors.** Content attributed to "Admin" or the company name demonstrates nothing.
- **Real author pages** with credentials, relevant history and links to independent profiles.
- `Person` **schema** connecting the author to those profiles.
- **Consistency** — the same author writing repeatedly in the same subject area builds a recognisable expert entity.

For YMYL topics — Your Money or Your Life, meaning health, finance, legal and safety — expertise requirements are considerably stricter, and content without demonstrable professional credentials struggles badly. If you operate in these areas, credentialed authorship is not optional.

Building author entities deliberately

Expertise attaches to people, and people are entities (Chapter 14). That makes author development one of the few authority levers entirely inside your control.

For each person who writes for you:

- **A real profile page** at a stable URL, with credentials, relevant history, and what they specifically know about
- `Person` **schema** with `sameAs` pointing to LinkedIn and any professional register
- **A consistent byline** — the same name form everywhere, never "Admin" or the company name
- **Bylines on external work too.** A guest article or podcast appearance under the same name compounds the same entity
- **Subject consistency.** One author writing repeatedly about one area becomes recognisable as an expert in it; the same author writing about six areas becomes recognisable as nothing

> KEY IDEA
>
> A named author with a real profile, writing consistently in one subject, is worth more than an anonymous byline on better prose. The person is the entity that accumulates authority — and if they leave, the pages they wrote keep the credential while the organisation keeps the corroboration.

Authoritativeness

Do others in the field treat you as a source? This is largely earned rather than built, through the work in Chapters 20 to 21: citations, mentions, links, speaking, being quoted.

The AI-search-specific point: **unlinked mentions now carry much of the weight that once required a link**, because a model reading text learns the association regardless of markup. This meaningfully lowers the barrier — being talked about is nearly as good as being linked to.

Trust

The load-bearing one. Google has been explicit that Trust is the most important member of the family; the other three exist to support it. Signals:

- **Transparency.** Real address, real phone number, named people, company registration.
- **Accuracy.** Claims that survive checking. Corrections when you get things wrong.
- **Currency.** Visible, honest dates. Reviewed-on dates for evergreen content.
- **Honest commercial disclosure.** Affiliate relationships, sponsorship, limitations of your own product.
- **Security and basic hygiene.** HTTPS, working forms, no broken checkout.
- **Balance.** Content acknowledging what your product is *not* good for is more trusted than content that does not.

COMMON MISTAKE

The single fastest way to damage trust signals with AI systems is **contradicting yourself across sources**. A price on your site that differs from your pricing page, a headcount that differs from LinkedIn, a founding date that differs from Companies House. Cross-checking systems find these, and unresolvable contradictions push a source down. Run the consistency check from Chapter 14 seriously.

IN PRACTICE

The trust audit

Score your site honestly, one point each:

1. Every article has a named author with a profile page
2. Author pages include relevant credentials and external profile links
3. A physical address and working phone number are on the site
4. An About page names actual people with photographs
5. Content includes visible published and updated dates
6. Commercial relationships are disclosed
7. At least some content states where your product is *not* the right choice
8. Core facts match across your site, LinkedIn and directories
9. HTTPS everywhere, no mixed-content warnings
10. A visible, working way to report an error

Below 6 out of 10, this is your highest-priority content work — ahead of anything in Chapter 13.

CASE STUDY · KETTLEWORTH

Kettleworth's guides had been written by a freelance copywriter who had never used any of the equipment. They read well and demonstrated nothing.

The change was to route every guide through the people who actually handled the products. Their head of buying dictated notes; the copywriter shaped them.

What appeared in the rewritten guides: photographs of their own workshop bench rather than manufacturer renders. The observation that one popular grinder's adjustment collar loosens after about four months of daily use. The fact that they had returned an entire batch of one model in 2024, and why.

That last paragraph — an admission that cost them a supplier relationship — became their most-cited passage of the year.

CHAPTER SUMMARY

- E-E-A-T is a quality framework, not a ranking factor, but the practical implication is the same
- It matters more in AI search because wrong citations are costly and systems cross-check
- Experience is the winnable one: original photos, real numbers, named failures, practitioner detail
- Unlinked brand mentions now carry much of the authority weight that used to require links
- Trust is load-bearing; self-contradiction across sources is the fastest way to lose it

CHECK YOURSELF

1. Why is E-E-A-T more consequential in AI search than in classical search?
2. Why is Experience the most winnable of the four, and how would you add it this week?
3. What is YMYL and how does it change the expertise bar?
4. Give three examples of self-contradiction across sources that would damage trust.

CHAPTER 16

Citable Assets

Making the thing that other people have to reference

WHAT YOU'LL LEARN

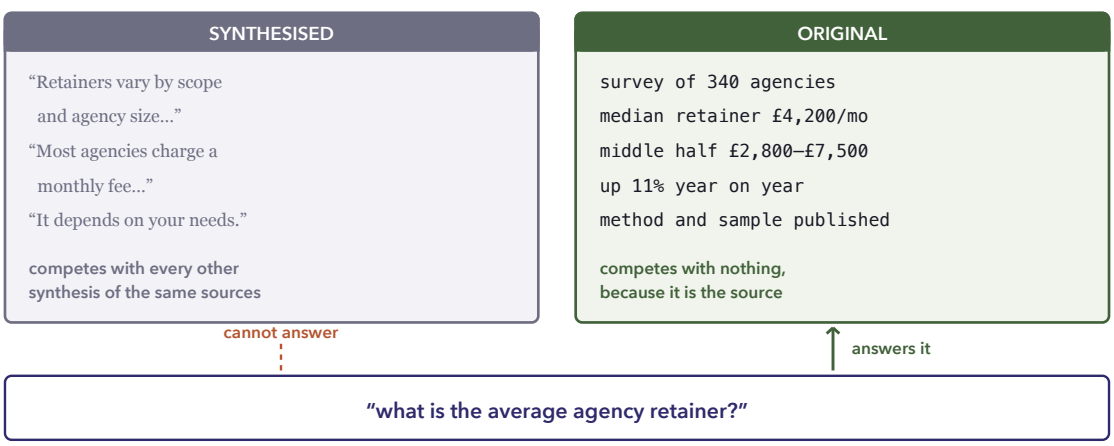
- Why original data outperforms every other content type for citation
- The formats that get cited most, and why
- How to produce original research without a research budget
- Making an asset maximally easy to cite

There is one content type that consistently earns citations across every engine, is defensible against competitive erosion, and compounds over time: **information that exists nowhere else**. If you produce the only source for a fact, anyone answering a question involving that fact must reference you.

Why original data wins

Chapter 4 explained reranking. A passage containing a specific, unique, verifiable fact is an extremely strong match when a question needs that fact, and there is no substitute source competing with it.

Compare two pages about agency retainers. The first synthesises what other people have written. The second reports a survey of 340 agencies with median retainer values, distributions and year-on-year change. When a model needs “the average agency retainer”, only the second can supply it.



Only one of these pages contains the number, so only one of them can be retrieved to answer with.
That is the whole argument for original data, and it is the one content position competitive erosion cannot reach.

Figure 16.1 – The same topic, written two ways, and the question only one of them can answer.

> **KEY IDEA**
>
> Synthesised content competes with every other synthesis of the same sources. Original data competes with nothing, because it is the source. This is the only content strategy with a genuine moat.

Formats that earn citations

Format	Why it gets cited	Effort
Original survey	Produces statistics nobody else has	Medium
Proprietary platform data	Unique by definition	Low if you have it
Systematic testing	Comparative results with method	Medium
Longitudinal tracking	Trends emerge over time; compounds	Low per instalment
Aggregated public data	New analysis of scattered sources	Medium
Definitive glossary	Becomes the definition others cite	Low
Genuine case study with numbers	Specific outcomes with context	Low
Calculator or tool	Referenced repeatedly	High

Producing research without a budget

The barrier is much lower than people assume.

Survey your own audience

Two hundred responses is enough to report medians and distributions honestly. Your email list, LinkedIn network and any community you participate in will supply that. Ten well-chosen questions, three weeks in the field.

Be transparent about the method — sample size, how respondents were recruited, when it ran. Stated limitations increase citability rather than reducing it, because they make the data verifiable and let others use it responsibly.

Report your own operational data

If you run any platform, service or process at volume, you are sitting on data nobody else has. Anonymise, aggregate, and publish.

Examples: average time-to-resolution across support tickets; the distribution of project sizes; the most common reason customers churn; seasonal demand patterns. All interesting, all already in your systems.

Test things systematically

Buy the five leading products in your category and test them against consistent criteria. Document the method. This is how review sites earn citations, and it is available to anyone with a modest budget and patience.

Track something over time

Pick one measurable thing in your industry and measure it quarterly. The first instalment is unremarkable; the fourth is the only longitudinal source in your field. This is the highest-compounding option on the list and the one most people abandon after two rounds.

IN PRACTICE

Designing a citable survey

1. **Ten questions maximum.** Completion rate collapses beyond that.
2. **Ask for numbers, not opinions.** "What is your monthly retainer?" produces citable data; "How satisfied are you with retainers?" does not.
3. **Include one question nobody has asked before.** This is what makes it uniquely citable.
4. **Collect segment data** — company size, sector, region, so you can report cuts and produce many citable facts from one survey.
5. **Publish the method openly**, including sample size and recruitment.
6. **Give every finding its own heading**, phrased as the question it answers. This turns one report into fifteen retrievable passages.
7. **Include a plain-text summary table** of headline figures near the top.
8. **State a clear citation line**: "Source: Acme Agency Retainer Survey 2026, n=340."

Making an asset easy to cite

Produce the data, then remove every obstacle between it and someone quoting it:

- **State each finding as a complete sentence.** "The median UK agency retainer in 2026 was £4,200 per month." That sentence is liftable exactly as written.
- **Put key figures in text, not only in charts.** An image of a graph is invisible; the number must exist as text.
- **Use a table for headline figures.** Machine-readable and directly extractable.
- **Give it a stable, memorable URL** you will not change.
- **Date it explicitly** and keep the year in the title.
- **Provide a citation line** people can copy.
- **Do not gate it.** A form-gated PDF is invisible to every retrieval system. If you need leads, gate a supplementary cut and publish the headline findings openly.

COMMON MISTAKE

Gating research behind a form is the single most common way organisations waste this work. The asset's value in AI search comes entirely from being freely retrievable. A gated report earns no citations, no mentions and no authority — it collects a few email addresses in exchange for the entire strategic benefit.

CASE STUDY · KETTLEWORTH

Kettleworth owned a grind-consistency test rig they used for buying decisions. It had never occurred to them to publish anything from it.

They tested 24 grinders under identical conditions, measured particle distribution, and published the full dataset: method, sample sizes, the models that performed worse than their price suggested, and the limitations of their own rig.

They stated each finding as a complete sentence under a question-shaped heading, gave the tables plain text alongside the charts, and did not gate it.

Within four months it was cited by two trade publications, one manufacturer, and — the outcome they cared about — by all four assistants when asked about grind consistency.

CHAPTER SUMMARY

- Original data is the only content type with a genuine moat — it competes with nothing
- Surveys of your own audience, your operational data, systematic testing and longitudinal tracking are all achievable without a research budget
- Longitudinal tracking compounds the fastest and is most often abandoned
- State findings as complete sentences, put figures in text, and give each its own question-shaped heading
- Never gate it: gating trades the entire strategic benefit for a handful of email addresses

CHECK YOURSELF

1. Why does original data resist competitive erosion when formatting tactics do not?
2. You have no research budget. Name three ways to produce original data this quarter.
3. Why does stating your survey's limitations make it more citable, not less?
4. Explain to a demand-gen colleague why gating the report defeats its purpose.

CHAPTER 17

Content Decay and the Refresh Discipline

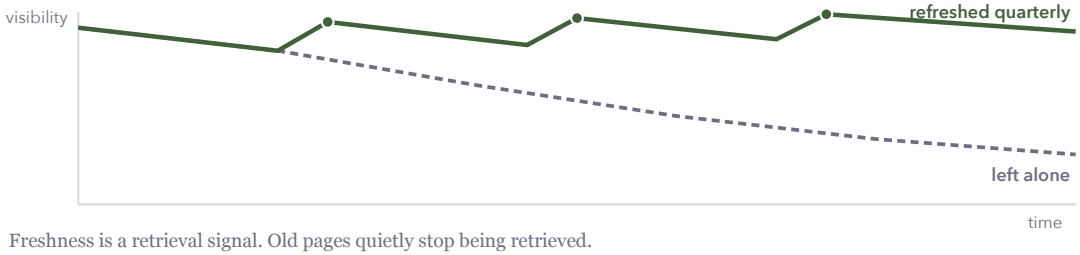
Why publishing less and maintaining more usually wins

WHAT YOU'LL LEARN

- What content decay is and why AI search accelerates it
- How freshness functions as a retrieval signal
- A workable audit and refresh cycle
- When to update, when to consolidate, and when to delete

Most content teams are organised around publishing. The higher-return activity, for almost every established site, is maintenance — and AI search has sharpened this considerably, because freshness is weighted heavily at the retrieval stage.

Why refreshing beats publishing



Freshness is a retrieval signal. Old pages quietly stop being retrieved.

Figure 17.1 – Two identical pages, one maintained. The gap compounds.

Why decay happens faster now

A page loses visibility over time for reasons that have little to do with its quality:

- **The facts age.** Prices, tool names, statistics and screenshots go stale.
- **Competitors publish newer material** covering the same ground.
- **The question itself changes.** What people mean by a query shifts.
- **Freshness weighting.** Retrieval systems favour recent sources for anything time-sensitive, and they interpret "time-sensitive" broadly.
- **Contradiction risk.** As Chapter 15 noted, out-of-date claims that contradict current sources actively push you down rather than merely being neutral.

That last point is what makes decay more dangerous than it used to be. An old page is not just less useful; it can be a liability that undermines the trust signals for your whole domain.

> KEY IDEA
>
> In classical search, an outdated page slowly lost rankings. In AI search, an outdated page can actively contradict current information and reduce your credibility as a source. **Decay is now a risk, not just a missed opportunity.**

The audit

Run this quarterly. It takes about half a day for a site of a few hundred pages.

1. **Export every URL with its traffic and impressions** for the last 12 months, plus the previous 12 for comparison.
2. **Sort by decline** — the largest absolute drops first.
3. **Add a "last updated" column.**
4. **Classify each declining page** into one of four buckets.

Bucket	Criteria	Action
Refresh	Still relevant, still gets traffic, facts have aged	Update thoroughly
Consolidate	Overlaps with two or more other pages	Merge into one, redirect the rest
Rewrite	Topic matters, execution is weak	Rebuild from scratch at the same URL
Delete	No traffic, no relevance, no links	Remove, redirect to nearest relevant page

IN PRACTICE

What a proper refresh includes

Changing the date is not a refresh, and systems detect the difference. A real refresh:

1. **Verify every fact.** Prices, statistics, product names, regulations. Update or remove anything you cannot confirm.
2. **Replace stale examples.** Tools that no longer exist, screenshots of old interfaces.
3. **Restructure using Chapter 13.** Most older content predates the answer-first discipline; this is the moment to fix it.
4. **Add what has emerged since.** New considerations, new alternatives, new evidence.
5. **Improve specificity.** Replace vague claims with concrete ones.
6. **Check internal links** in both directions — from this page and to it.
7. **Update the structured data**, including `dateModified`.
8. **Record what changed** in a visible note if the change is material. This is a genuine trust signal.

A refresh done properly takes about as long as writing a new piece and reliably outperforms it, because the page already has standing, links and history.

Consolidation

Sites accumulate near-duplicates: four posts about roughly the same thing, written over five years by different people. Each is mediocre; each competes with the others; none is authoritative.

Merging them into one comprehensive page and redirecting the rest concentrates the signals and usually produces a page substantially stronger than any of the originals. It also removes the ambiguity a retrieval system faces when deciding which of your four similar passages to use.

COMMON MISTAKE

When consolidating, always **redirect** the removed URLs to the surviving one. Deleting them outright discards every link and every accumulated signal. A 301 redirect passes most of that value to the merged page. This is the most common and most costly mistake in content consolidation.

The publishing ratio

A useful default for an established site: **spend at least as much time maintaining as publishing.** Many teams find a 60:40 split towards maintenance beats 90:10 towards new

content. A refreshed page starts from existing authority; a new page starts from nothing.

MYTH vs REALITY

"You need to publish frequently to stay visible"

Publishing frequency is a proxy that got mistaken for a cause. What actually helps is a corpus that is current, coherent and thorough.

A site publishing twice a month while properly maintaining its existing 200 pages outperforms a site publishing twice a week onto a decaying archive. The second site has more pages; the first has more *retrievable* pages.

CASE STUDY · KETTLEWORTH

Kettleworth's decay audit covered 340 published pages. The classification was sobering: 61 to refresh, 44 to consolidate into 12, 9 to rewrite, and 78 to delete.

The consolidation was the largest single gain. They had four separate posts about milk frothing, written between 2019 and 2024 by three different people. Each was mediocre; each competed with the others.

Merged into one properly structured page — with 301 redirects from all four URLs — the result outperformed the best of the originals within seven weeks.

They also fixed the habit that produced the problem: no new post now gets commissioned without checking whether an existing page should be extended instead.

CHAPTER SUMMARY

- Content decays through ageing facts, newer competitors, shifting intent and freshness weighting
- Outdated content is now a liability, not merely a lost opportunity, because contradiction damages trust
- Audit quarterly, classify into refresh, consolidate, rewrite or delete
- A real refresh verifies facts, restructures and adds substance — changing the date does nothing
- Always redirect when consolidating; deleting discards accumulated value

CHECK YOURSELF

1. Why is an outdated page more dangerous in AI search than in classical search?
2. What are the four buckets in the decay audit, and what triggers each?
3. Why does a refreshed page usually outperform an equivalent new page?
4. What is the single most costly mistake when consolidating overlapping content?

CHAPTER 18

International and Multilingual AI Search

Why everything in this book behaves differently outside English

WHAT YOU'LL LEARN

- Why non-English AI search is measurably weaker, and why that is an opportunity rather than a problem
- How citation sources differ by language, and why Reddit advice does not travel
- Translation versus transcreation, and why machine-translated content retrieves badly
- How `hreflang`, entity signals and measurement all change across markets

Almost everything published about AI search assumes English. That assumption is doing quiet damage to anyone operating elsewhere, because several of the mechanisms in this book behave differently — and in one respect, far more favourably — outside the Anglophone web.

Why non-English AI search is weaker

A language model's competence in a language is a function of how much of that language it saw in training. English dominates the training corpus of every major model by a very large margin. The consequence is not that models are bad at German or Spanish — they are fluent, but that the **supporting machinery is thinner**:

- **Entity resolution is weaker.** There are fewer independent sources connecting a Spanish company to its topic, so the corroboration mechanism from Chapter 14 has less to work with.
- **The retrieval corpus is smaller.** Fewer high-quality pages exist per topic per language, so the candidate pool for any sub-query is shallower.
- **Fact density is lower.** Specialist information is disproportionately published in English first, and sometimes only.
- **Translation artefacts creep in.** Models sometimes answer non-English questions using English sources, translating on the fly — which means an English page can be cited for a German query, and a German page can be passed over.

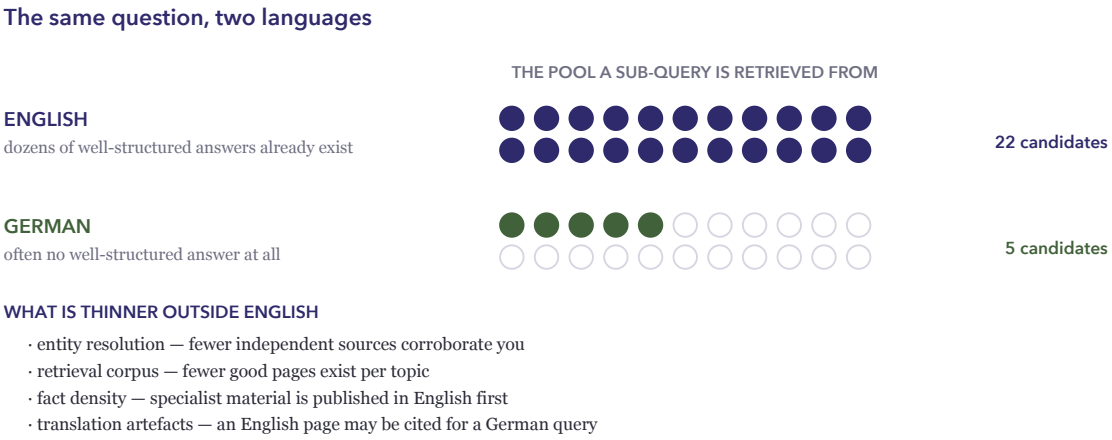


Figure 18.1 – Why the non-English corpus is thinner, and why a thinner corpus is an opportunity.

> KEY IDEA
>
> Thinner corpora mean **less competition**. In English, being the best answer to a specific question is hard. In German, Dutch, Polish or Portuguese, the same question often has no well-structured answer at all. This is the single largest under-exploited opportunity in the book.

Citation sources do not travel

Chapter 22's platform advice is Anglophone advice. Reddit's dominance — 24% of Perplexity citations — is a fact about the English-language web, not about retrieval systems.

Market	What tends to carry weight instead
Germany	Sector forums, `gutefrage`, trade association sites, `Stiftung Warentest`-style testing bodies, German Wikipedia
France	Specialist forums, `Que Choisir`-style consumer bodies, French Wikipedia
Netherlands	Highly active niche communities, Dutch review platforms
Spain / LatAm	YouTube disproportionately strong, regional forums
Japan	Domestic platforms dominate; Western social sources barely register
Nordics	English-language sources often cited for domestic queries

The practical instruction is simple and almost nobody follows it: **run the Chapter 29 source-concentration analysis separately for each language you operate in.** The

list of third-party domains the engines rely on will look nothing like the English one.

> COMMON MISTAKE
>
> Do not assume your English platform strategy transfers. A German-market team that invests six months in Reddit because an English blog recommended it will have spent six months on a platform with a fraction of the domestic weight — while the trade publications that genuinely feed German AI answers go untouched.

Translation versus transcreation

This is where most international content fails, and the reason is mechanical rather than aesthetic.

Retrieval matches **the way people actually ask questions** in a language. Machine translation preserves meaning but not idiom, and it certainly does not preserve how a native speaker phrases a query. A German user does not ask the literal translation of "how much does a coffee grinder cost" — they use different compound constructions, different framing, often a more specific noun.

Three failure modes, in ascending severity:

1. **Literal translation of headings.** Your question-shaped H2s become questions nobody asks. The alignment with sub-queries that Chapter 13 depends on is lost entirely.
2. **Translated examples and figures.** Prices in the wrong currency, regulations from the wrong jurisdiction, product names that do not exist in that market. These read as errors and damage trust signals.
3. **Translated first-hand experience.** The "Experience" pillar from Chapter 15 does not survive translation — an account of testing something in Manchester, rendered into German, demonstrates no experience of the German market.

IN PRACTICE

Transcreating a page properly

For each page you take into a new market:

1. **Re-derive the questions.** Do not translate your headings. Ask a native speaker, or a model, prompted in that language — how someone in that market would ask this. The phrasing will differ more than you expect.
2. **Replace the specifics.** Local prices, local currency, local suppliers, local regulations, local units.
3. **Replace the examples** with ones from that market. This is the step everyone skips and it is the one that carries the experience signal.
4. **Have a native speaker edit for idiom.** Machine translation produces text that is correct and subtly foreign; retrieval systems and readers both notice.
5. **Keep the structure.** Answer-first and self-containment are language-independent. Only the content changes.

A properly transcreated page takes perhaps three times as long as a translated one and performs several times better. Translating fifty pages badly is worse than transcreating fifteen well.

The technical layer

`hreflang`

`hreflang` is a set of tags telling search engines which language and region each version of a page serves, so the right one is shown to the right user. It is well supported by Google and does real work in classical search.

Its status in AI retrieval is less clear, and you should be honest about that. What is safe to say: `hreflang` helps the classical index serve the correct version, and since the generative layer retrieves from that index, getting it right removes a source of confusion. What is not established: that it directly influences which language version a model cites.

The practical rules remain:

- Every version links to every other version, including itself
- Use language-only codes (`de`) unless you genuinely serve different content by region (`de-AT` versus `de-DE`)
- Include an `x-default` for unmatched users
- Never point `hreflang` at a redirecting or non-indexable URL

URL structure

Country-code domains (`example.de`), subdirectories (`example.com/de/`) or subdomains (`de.example.com`). Subdirectories are usually the right default: they inherit the parent domain's accumulated authority, which matters more when your entity signals in the new

market are weak.

A country-code domain signals local commitment strongly but starts from zero authority — a serious cost given how much of retrieval candidacy depends on the classical layer.

Local entity signals

Chapter 14's entity work must be repeated per market:

- A **local-language Wikipedia entry** is a distinct entity anchor from the English one
- **Local business registries** — for Germany, the `Handelsregister`; for France, `Infogreffe`
- **Local directories and trade association memberships**
- **A local address and phone number**, if you genuinely have them
- **Local-language `sameAs` profiles**

Measuring across markets

Everything in Chapter 29 applies per language, separately:

- **Separate prompt sets**, written natively rather than translated
- **Separate baselines**, since your share of model will differ enormously by market
- **VPN or location-consistent sampling** — answers vary by region, and running German prompts from a UK IP gives you a result that describes nobody
- **Separate source-concentration lists**

> COMMON MISTAKE
>
> Running your English prompt set through machine translation to create a German one produces a prompt set that measures nothing useful. The questions will not be the questions Germans ask, so your share of model against them tells you how you perform on queries nobody types.

CASE STUDY · KETTLEWORTH

Kettleworth's German launch began the way most do: 60 of their best English guides pushed through machine translation, `hreflang` correctly implemented, live in three weeks.

Six months later their German share of model was 4%. The UK set, measured the same way, was at 29%, and the two numbers are not comparable except in the way that matters here: each is measured against its own natively written prompt set, in its own market. The baseline showed why. Their translated heading "Wie viel kostet eine Kaffeemühle?" was a reasonable sentence that almost nobody searched; German buyers were asking about `Mahlgradverstellung` — grind adjustment mechanics — a distinction English-language buyers rarely lead with. Their prices were in pounds. Their "we tested this in our Leeds workshop" experience passages translated into something that read as foreign and demonstrated nothing about German conditions.

They cut the German site from 60 pages to 14, transcreated properly with a native-speaking coffee professional, replaced every example with German suppliers and euro pricing, and added a `Handelsregister`-linked German entity block.

German share of model at the next quarterly measurement, against the same 30 German prompts: 19%. Fourteen good pages beat sixty translated ones.

CHAPTER SUMMARY

- Non-English AI search has thinner corpora, weaker entity resolution and less competition — the opportunity is larger, not smaller
- Citation sources are market-specific; Reddit advice is Anglophone advice and does not transfer
- Transcreate rather than translate: re-derive the questions, replace the specifics, replace the examples
- `hreflang` helps the classical index and therefore retrieval candidacy; its direct effect on citation is unestablished
- Measure each market separately with natively written prompts and location-consistent sampling

CHECK YOURSELF

1. Why does thinner training data in a language create an opportunity rather than only a disadvantage?
2. What are the three failure modes of translation, and which destroys the E-E-A-T experience signal?
3. Why is a machine-translated prompt set worse than useless for measurement?
4. When would a country-code domain be the wrong choice for a new market?

CHAPTER 19

Multimodal GEO

Images, video, audio and voice — making non-text retrievable

WHAT YOU'LL LEARN

- Why everything that is not text has to become text before it can be retrieved
- How to make images, video and audio citable, in the order that pays best
- What changes when the answer is spoken rather than shown — and why voice is winner-takes-all
- Knowledge panels and the knowledge graph: what they are and what you can actually influence

Retrieval systems work on text. That single fact governs everything in this chapter. A photograph, a chart, a video and a podcast are all invisible to a retrieval pipeline until something converts them into words. The quality of that conversion is entirely under your control, which makes it one of the most neglected easy wins in the discipline.

Everything becomes text before it can be retrieved

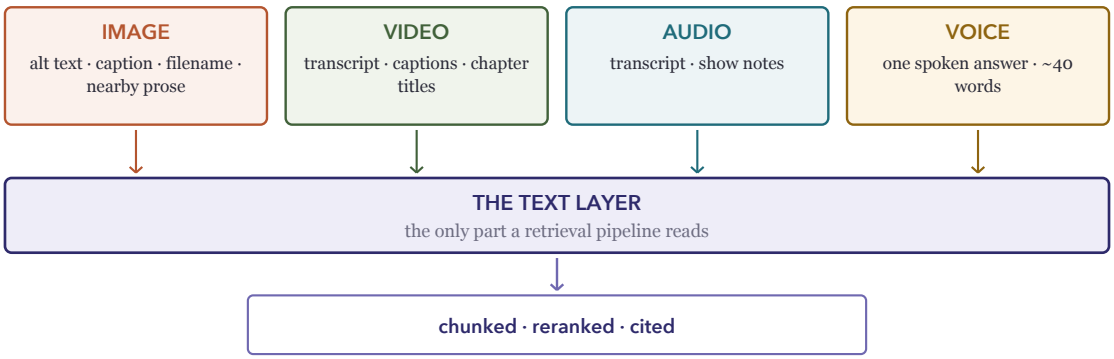


Figure 19.1 — Four media, one requirement. Everything becomes text before it can be retrieved.

The general rule

Modern models can genuinely see images and hear audio. Vision-capable models describe photographs accurately, and speech models transcribe well. But there is a gap between *can* and *does*: running vision over every image on the web at crawl time is expensive, and retrieval pipelines overwhelmingly still index the **text near and about** a piece of media rather than the media itself.

> **KEY IDEA**
>
> Assume your media will be understood only through its text layer — alt text, captions, transcripts, filenames and surrounding prose. If the text layer is missing, the media is invisible, however good it is.

Images

In descending order of return:

1. Alt text that describes, not labels

Alt text exists for screen readers first, and that purpose produces exactly what retrieval wants: a plain description of what the image shows.

Weak	Strong
`alt="grinder"`	`alt="Burr chamber of the Kettleworth K3 with the retention gap visible"`
`alt="chart"`	`alt="Grind consistency by price band, showing the £150–200 band outperforming £200–300"`
`alt="team photo"`	`alt="Our buying team testing grinders on the workshop bench in Leeds"`

The right-hand column can be retrieved for a question. The left cannot.

2. The sentence next to the image

Retrieval reads the paragraph around an image as context for it. An image dropped between two unrelated paragraphs is orphaned; one introduced by a sentence that says what it shows and why is anchored.

3. Real captions

A visible caption serves readers and retrieval at once. It is the single highest-value place to put the specific claim the image supports.

4. Filenames and structured data

`kettleworth-k3-burr-chamber.jpg` beats `IMG_4417.jpg`. `ImageObject` schema with a `caption` and `contentUrl` makes the relationship explicit. Both are cheap; neither is transformative on its own.

> COMMON MISTAKE
>
> Charts and infographics are the worst offenders. A chart carrying your original data — precisely the citable asset Chapter 17 argues for — is frequently published as an image with `alt="chart"` and no figures in the text. The data is then unreachable. **Always put the numbers in text as well**, in a table or a sentence. The image is for humans; the text is for retrieval.

Video

Chapter 23 makes the core point: the transcript is the retrievable asset, not the video. To extend it:

- **Upload your own captions.** Auto-generated captions mangle product names, jargon and figures — exactly the specific terms you want matched.
- **Structure the spoken content answer-first.** The Chapter 14 pattern applies to speech. If the answer arrives at 4:20 after four minutes of preamble, the retrievable passage is four minutes of preamble.
- **Use descriptive chapter markers.** They function as headings, and headings define chunk boundaries.
- **Publish a written version.** A transcript or article on your own site captures value that would otherwise sit entirely on a platform you do not own.
- `VideoObject` **schema** with `transcript` and `description` makes the text layer explicit.

Audio and podcasts

The same logic, with one difference: podcasts are rarely accompanied by a text version, so the gap between what was said and what is retrievable is enormous.

If you appear on a podcast — which Chapter 22 recommends as a mention-earning tactic — ask whether a transcript will be published. If not, publish your own on your site: *"I spoke to X about Y — here is the transcript and the three points that matter."* You capture the retrievable text from an appearance that would otherwise leave no trace.

Voice, and why it is different

Voice assistants change one thing fundamentally: **they return a single answer.** A screen shows a summary and three citations; a speaker reads one response. There is no second place.

This produces different behaviour worth planning for:

- **Queries are longer and more conversational.** People speak in full sentences.
- **Local intent is over-represented.** Much voice use is on the move.

- **Answers are shorter.** A spoken answer is one or two sentences, so the passage that wins is one that resolves the question in about forty words.
- **Follow-ups are common**, which means conversational context carries between turns.

IN PRACTICE

Writing a passage that survives being read aloud

Take an answer-first section from Chapter 14 and apply two extra constraints:

1. **Can the first sentence stand alone as a spoken answer?** Read it out. If it needs a table, a list or a preceding sentence to make sense, it will not be selected.
2. **Is it under about forty words?** Longer answers get truncated or passed over.

Before: "Grind size depends on brew method, dose, and the specific machine — see the table below for our recommended starting points."

After: "For espresso, start at a fine grind, around 200 to 300 microns. Adjust finer if the shot runs fast and coarser if it chokes. The table below gives starting points for other brew methods."

The first sentence now works alone, aloud, and the detail still follows for readers.

Knowledge panels and the knowledge graph

A **knowledge graph** is a structured store of entities and the relationships between them — the machinery behind Chapter 15's entity work. A **knowledge panel** is the visible box of facts a search engine shows about a recognised entity.

What you can influence:

- **Corroborated facts across independent sources** — the mechanism from Chapter 15, unchanged
- **Wikidata**, which is directly editable and feeds several graphs
- `Organization` **and** `Person` **schema with** `sameAs`, which states your identifiers explicitly
- **Google Business Profile**, if you have premises (Chapter 30)

What you cannot: create a panel on demand, or edit one directly beyond suggesting corrections. Panels are a *consequence* of entity recognition, not a lever on it.

MYTH vs REALITY

"Optimising images means compressing them"

Image compression is a page-speed concern, and page speed is a weak tie-breaker (Chapter 2). It has essentially nothing to do with whether an image can be retrieved.

The retrieval question is entirely about the text layer. A perfectly optimised 40KB WebP with `alt="image"` is invisible; a heavy PNG with a precise description and a real caption is not.

CASE STUDY · KETTLEWORTH

Kettleworth's grind-consistency study — their most citable asset — existed almost entirely as charts.

The full dataset had been published as eleven images exported from a spreadsheet, each with `alt` text that read `chart1.png` through `chart11.png`. The written article around them described the method and the conclusion but contained almost none of the actual figures.

The consequence showed up in the baseline: assistants credited the study when asked *whether* Kettleworth had tested grinders, but could not answer *what the results were* without hedging.

The fix was a morning's work. Every chart gained a descriptive alt text and a visible caption stating its headline finding. Each chart was accompanied by a plain HTML table of the same numbers. The eleven figures went into the article text as sentences.

Three months later the study was being quoted with specific numbers by three of four engines, including one that named the £150-200 band finding directly. The data had been there all along, sealed inside pictures.

CHAPTER SUMMARY

- Retrieval works on text; media is understood through alt text, captions, transcripts and surrounding prose
- Alt text should describe what the image shows, not label it — and charts must always have their numbers in text as well
- Transcripts are the retrievable asset for video and audio; upload your own captions rather than relying on auto-generation
- Voice returns a single answer, so the winning passage resolves the question in roughly forty spoken words
- Knowledge panels are a consequence of entity recognition, not a lever; influence them through Wikidata, schema and corroboration

CHECK YOURSELF

1. Why is a chart carrying original data one of the most commonly wasted assets in AI search?
2. Rewrite this alt text: `alt="dashboard"` for a screenshot showing organic traffic falling 34% year on year.
3. What are the two extra constraints a passage must meet to work as a spoken answer?
4. What can you actually influence about a knowledge panel, and what can you not?

PART V

Authority Beyond Your Site

Most of what determines whether you are cited is written by other people, somewhere else.

CHAPTER 20

Links, Mentions and What Changed

Why the unlinked mention became nearly as valuable as the link

WHAT YOU'LL LEARN

- What links still do, and what they have stopped being sufficient for
- Why unlinked brand mentions carry real weight in AI search
- How co-occurrence builds topical association
- A realistic view of what link building is now worth

Link building consumed a large share of SEO effort for twenty years, on sound reasoning: links were the strongest available proxy for reputation. That reasoning still holds, but a second mechanism now runs alongside it, and it changes where your effort should go.

What links still do

Everything from Chapter 2 remains true. Links are discovery paths and endorsement signals; they feed the classical index; the classical index feeds the generative layer. A site with no links from anywhere will struggle to rank, will therefore struggle to be retrieved, and will therefore struggle to be cited.

So: links still matter. What has changed is that they are no longer the *only* way to build the association between your brand and a topic.

Why mentions now carry weight

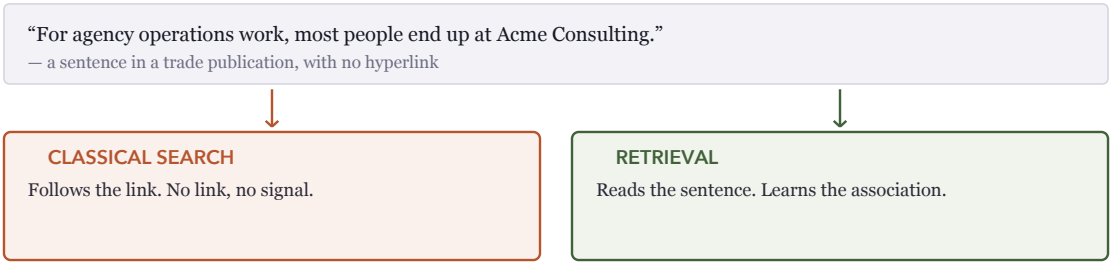
The mechanism is straightforward once you have Chapter 3. A language model reading text learns associations from the text itself. If a hundred articles say “Acme Consulting specialises in agency operations”, the model learns that association whether or not a single one of them contains a hyperlink.

A hyperlink is a machine-readable pointer. A model does not require one to understand a sentence.

KEY IDEA

Classical search needed the link. Retrieval systems need the sentence. An unlinked mention of your brand alongside a topic builds the topical association almost as effectively as a linked one, and unlinked mentions are considerably easier to earn.

Why an unlinked mention still counts



A hyperlink is a machine-readable pointer. Understanding a sentence does not require one.

Figure 20.1 – One sentence, two systems. Only one of them needs the hyperlink.

This is liberating for anyone who has tried to do link outreach. The journalist who quotes you but does not link, the podcast that names you in the show notes, the Reddit commenter who says "we use Acme for this" — all previously worth close to nothing for SEO, all now doing real work.

Co-occurrence

The related concept is **co-occurrence**: how often your brand name appears in proximity to a topic across the corpus. This is the raw material of topical association.

Practically, you want your brand name and your core topic appearing in the same sentence, paragraph and page as often as possible, across as many independent sources as possible. Not stuffed — genuinely, in context, in text other people wrote.

IN PRACTICE

Auditing your co-occurrence

1. Search Google for `"your brand name"` in quotes and read the first five pages of results.
2. For each result, note: what topic is your brand associated with in that text?
3. Tally the topics.
4. Compare against the topic you *want* to own.

The common findings are instructive. Either your brand co-occurs mostly with generic terms (“marketing”, “software”) rather than your specific niche, or it co-occurs with a product line you have moved away from, or it barely appears outside your own properties.

Each of those is a different problem with a different fix, and you cannot know which you have without looking.

What link building is worth now

An honest assessment:

Still worth doing: earning links from relevant, genuinely authoritative sources in your field. These pass standing, drive referral traffic, and come bundled with a mention.

Worth much less than it was: volume link building. Directory submissions, low-quality guest posts, link exchanges. These were always weak; they are now weak *and* they do not produce the contextual mentions that matter.

Newly valuable: anything that gets you talked about in text, whether or not it links. Podcast appearances, expert quotes, community participation, being named in comparison articles, industry surveys that list you.

MYTH vs REALITY

“Links don’t matter any more in AI search”

An overcorrection. Links feed the classical index, the classical index supplies retrieval candidates, and link-derived authority is a documented input to reranking.

The accurate statement: **links remain necessary but are no longer sufficient, and mentions have become a parallel route to the same goal.** Budget accordingly, but do not stop.

Where to redirect effort

If you previously ran a link-building programme, the highest-return reallocation is usually:

From	To
Guest post volume	Fewer, better placements in genuinely read publications
Directory submissions	Structured profiles that support entity recognition (Ch. 14)
Link exchanges	Podcast appearances and expert commentary
Broken-link building	Original data that others cite unprompted (Ch. 16)
Anchor text optimisation	Ensuring your brand co-occurs with your topic in the text

CASE STUDY · KETTLEWORTH

Kettleworth's link building had been a monthly retainer producing guest posts on marginal blogs. They cancelled it and redirected the budget.

Their head of buying began appearing on coffee industry podcasts — four in six months. None linked to the site. All named the company alongside the topic, in text, in show notes and transcripts.

They also audited their co-occurrence and found the real problem: outside their own site, "Kettleworth" appeared mostly alongside generic terms like "coffee equipment" rather than the specialist grinding and extraction topics they wanted to own.

The podcasts fixed that faster than two years of guest posts had.

CHAPTER SUMMARY

- Links still feed the classical index and still matter for retrieval candidacy
- Unlinked mentions now build topical association nearly as well, because models read sentences
- Co-occurrence — brand and topic appearing together across independent sources — is the underlying mechanism
- Volume link building has lost most of its remaining value; contextual mentions have gained
- Audit what your brand currently co-occurs with; it is rarely what you assume

CHECK YOURSELF

1. Explain, using the mechanism from Chapter 3, why an unlinked mention has value.
2. What is co-occurrence and how would you measure yours?
3. Why is “links don’t matter any more” wrong?
4. Name three activities that produce mentions without necessarily producing links.

CHAPTER 21

Share of Model and Brand Presence

Being the answer, not just being in the index

WHAT YOU'LL LEARN

- What share of model means and why it is becoming the headline metric
- How brand strength influences generative answers
- The compounding relationship between brand search and AI visibility
- Practical ways to increase the frequency with which you are named

In classical search you competed for position. In generative search you compete for **inclusion in the answer** — and the natural metric is what proportion of relevant answers name you. This chapter is about that metric and how to move it.

Share of model, defined

Share of model — also called share of voice in AI, or answer share — is the proportion of relevant generated answers in which your brand is mentioned or cited.

To take an arbitrary illustration: ask fifty questions your customers would ask, and if your brand appears in twelve of the answers, your share of model on that prompt set is 24%. It is a simple, robust and honest metric, and it maps onto how the systems behave better than rank tracking does. Chapter 29 covers measuring it properly.

KEY IDEA

Rank measures where you sit on a list. **Share of model measures how often you are part of the answer.** The second is closer to what actually drives commercial outcome, because in AI search there is no list for the customer to scan.

How brand strength feeds generative answers

Three mechanisms, worth separating because they need different work:

1. Training-data presence

If your brand appeared often enough, in enough contexts, before the model's training cutoff, the model has learned associations about you and can name you without retrieving anything. This is why established brands appear in answers even when no source is cited.

You cannot optimise this directly. It is the slow accumulation of being written about, and it argues for patience and consistency rather than campaigns.

2. Retrieval-time presence

Your pages, or pages mentioning you, are retrieved and inform the answer. This is the fast, controllable route, and it is what Parts III and IV are about.

3. Corroboration weight

When several sources agree that you are a leading option in a category, that agreement itself becomes a signal. A model that retrieves five comparison articles, four of which list you, treats you as an established option. One that retrieves five and finds you in none does not.

This is why **being included in third-party comparison and "best of" content is disproportionately valuable**. Those pages are retrieval magnets for exactly the commercial queries you care about, and inclusion in them directly changes what answers say.

The pages the engines actually read about your category

THIRD-PARTY SOURCE	ENGINES CITING IT	YOUR PRESENCE
Category round-up A	■■■■	● listed – price wrong
Review platform	■■■■	● listed – accurate
Reddit thread	■■■	● absent
Trade publication	■■■	● listed – outdated spec
Comparison blog	■■	● absent
Competitor's guide	■■	● absent

Correcting one wrong price on a heavily cited page is an email. It changes every answer.

Figure 21.1 – The output of a source-concentration analysis: which third-party pages the engines lean on, and how you appear on each.

IN PRACTICE

The comparison-article campaign

This is one of the highest-return activities in the whole book, and it is systematic rather than creative.

1. **Find the retrieval set.** Ask each engine your five most commercially important questions ("best X for Y", "X alternatives", "X vs Z"). Record every cited source.
2. **Compile the list.** You will typically find 15-40 distinct third-party pages that the engines lean on for your category.
3. **Check your presence.** For each: are you listed? Accurately? With current information?
4. **Prioritise.** Pages cited by multiple engines matter most.
5. **Approach each appropriately.** Many are editorial round-ups whose authors accept corrections and additions. Some are review platforms where you can claim a profile. Some are community threads where participation is appropriate and self-promotion is not.
6. **Fix inaccuracies first.** An outdated price or a wrong feature list on a heavily cited page is doing active damage, and correcting it is usually a two-line email.

The inaccuracy fix alone often produces a measurable change in how you are described, for perhaps an hour of work.

The brand search feedback loop

Branded search volume and AI visibility reinforce each other:

1. You appear in an AI answer
2. Some readers search your brand name to investigate
3. Rising branded search strengthens brand entity signals
4. Stronger entity signals make you more likely to appear in answers

This is why the widely cited finding that AI-Overview citation produces a ~15% lift in branded search matters beyond the immediate traffic — it feeds back into the mechanism that produced it.

The practical implication: **branded search volume is one of the best available proxy metrics for AI visibility**, and unlike most AI-search metrics, it is free and reliable. Chapter 30 builds on this.

COMMON MISTAKE

Share of model is only meaningful against a **fixed prompt set**. If you change the questions between measurements, the number moves for reasons unrelated to your work. Write the prompts down, version them, and change them only deliberately — noting when you did.

CASE STUDY · KETTLEWORTH

Kettleworth listed every third-party page the engines cited for their category. Thirty-one pages. They appeared on nine of them, and on four of those the information was wrong.

One heavily cited round-up listed their flagship grinder at a price they had abandoned in 2023 and described it as discontinued.

A polite email with current information got it corrected in six days.

That single correction changed how three assistants described their product line. Total effort: one email, twenty minutes to find the right contact. It remains the highest-return thing they did all year, and it was available to them the entire time.

CHAPTER SUMMARY

- Share of model is the proportion of relevant answers naming you — the natural headline metric for AI search
- Brand influences answers through training presence, retrieval presence and corroboration weight
- Inclusion in third-party comparison content is disproportionately valuable and highly actionable
- Fixing inaccuracies on heavily cited third-party pages is among the cheapest wins available
- Branded search volume is a reliable free proxy, and it feeds back into AI visibility

CHECK YOURSELF

1. Define share of model and explain why it beats rank as a headline metric.
2. Which of the three brand mechanisms can you influence fastest, and which not at all?
3. Why are third-party comparison articles such high-value targets?
4. Why must the prompt set stay fixed?

CHAPTER 22

The Platforms That Feed the Models

Reddit, Wikipedia, YouTube and LinkedIn — and how to be present without being unwelcome

WHAT YOU'LL LEARN

- Why a handful of platforms carry disproportionate citation weight
- How to participate in each legitimately
- The specific ways brands get this wrong and damage themselves
- Which platforms matter for which business types

Chapter 7 established that user-generated platforms dominate citations. This chapter is about what to actually do with that, including the substantial risk of doing it badly.

Reddit

The single most cited domain on some engines — around 24% of Perplexity's citations, over 5% of ChatGPT's responses, and around 0.1% of Gemini's. That variance is the first thing to internalise.

Why it wins

As Chapter 7 set out: genuine first-hand experience, explicit question-and-answer structure, dense specifics, and unhedged comparison. Reddit contains what marketing copy structurally cannot.

How to participate

Reddit communities are hostile to marketing, and correctly so. The rules that work:

- **Participate as a person, with a real account with history.** A new account posting about your product will be removed and may get your domain banned across large parts of the site.
- **Contribute far more than you promote.** A useful ratio is thirty helpful comments before one that mentions your product.
- **Disclose affiliation every time.** "Disclosure: I work at Acme" — Reddit is far more tolerant of disclosed affiliation than of discovered affiliation.

- **Answer the question asked.** Even when the honest answer is that a competitor suits them better. This is the behaviour that builds standing, and it is also the behaviour that gets quoted.
- **Find the right subreddits** — usually small, specific professional communities rather than large general ones.

> COMMON MISTAKE
>
> Do not pay for Reddit mentions, run sockpuppet accounts, or brief an agency to "seed positive sentiment". Reddit users detect this reliably, moderators ban domains rather than just accounts, and the resulting threads — which are highly citable — become permanent negative material about your brand appearing in AI answers. This is one of the few tactics in this book that can leave you materially worse off than doing nothing.

Wikipedia

Around 12-13% of ChatGPT citations and a foundational entity source. Also the platform where brands most reliably embarrass themselves.

The rule: do not edit articles about your own organisation. Conflict-of-interest editing is against policy, is detected routinely, and generates lasting negative attention. There is a proper process — declare your conflict on the article's talk page and request changes, and it is the only route you should use.

What you can legitimately do:

- **Correct factual errors via the talk page**, with sources.
- **Build genuine notability** through sustained independent coverage. Wikipedia presence follows notability; it cannot precede it.
- **Contribute to Wikidata**, which has a much lower barrier and directly supports entity recognition (Chapter 14).
- **Improve articles about your subject area** — not your company — as a genuine contributor.

YouTube

Strong citation performance across engines, and particularly in Google surfaces where it leads alongside Wikipedia and Fandom. Two reasons: Google owns it, and video transcripts are richly indexed text.

The practical point most people miss: **the transcript is the retrievable asset**. A video with an accurate, well-structured transcript is a text document that can be chunked and cited like any other.

- Upload accurate captions rather than relying on auto-generation
- Structure videos so spoken content follows the answer-first pattern from Chapter 13

- Use chapters with descriptive, question-shaped titles
- Put a full written summary in the description

LinkedIn

Outranks Reddit for B2B, professional services and enterprise software citations, per the Semrush prompt study. If you sell to businesses, this is your equivalent of Reddit.

- **Personal profiles outperform company pages** substantially
- Long-form posts with genuine expertise get indexed and cited
- Company pages matter for entity recognition regardless of engagement
- Employee content aggregates into brand association

Which platform for which business

Business type	Priority platforms
Consumer product	Reddit, YouTube, review platforms
B2B software	LinkedIn, G2/Capterra, industry publications
Professional services	LinkedIn, industry bodies, YouTube
Local service	Google Business Profile, local directories, Reddit locals
Publisher / media	Wikipedia, YouTube, industry citation
E-commerce	Reddit, YouTube, review platforms, comparison sites

IN PRACTICE

Choosing where to invest

Do not spread across all of them. Instead:

1. Run your ten most commercially important questions through your two priority engines.
2. Tally which platforms appear in the citations.
3. Pick the top two and commit properly for six months.

Two platforms done properly beats six done superficially, and on community platforms, superficial participation is actively counterproductive.

CASE STUDY · KETTLEWORTH

Kettleworth's head of buying started participating in two subreddits, from her own account, with her employer in her flair.

The rule she set herself was that she would answer the question actually asked, including when the answer was that a competitor's product suited the person better. She did this perhaps thirty times before mentioning a Kettleworth product once.

Nine months in, a long comment she wrote explaining burr alignment tolerances — technical, specific, clearly affiliated, and useful — became the most-cited Reddit source in their category.

It took nine months to produce one citation. That citation is still being retrieved.

Nine months, thirty answers, one citation



THE RATIO THAT MAKES IT WORK

Roughly thirty helpful comments before one that mentions your product — and disclosure every single time.

Superficial participation is worse than none at all.

Moderators ban domains rather than accounts, and the thread about it is itself highly citable.

Figure 22.1 – What earning one community citation actually costs, and why the ratio is the point.

CHAPTER SUMMARY

- A small number of platforms carry disproportionate citation weight, and the mix varies enormously by engine
- Reddit rewards genuine, disclosed, generous participation and punishes manipulation severely
- Never edit your own Wikipedia article; use the talk page, and use Wikidata instead
- YouTube's citable asset is the transcript, not the video
- LinkedIn is the B2B equivalent of Reddit; personal profiles outperform company pages

CHECK YOURSELF

1. Why does Reddit dominate on Perplexity and barely register on Gemini?
2. What is the correct procedure for fixing an error in your company's Wikipedia article?
3. Why is a video's transcript the part that matters for citation?
4. Why is superficial participation on community platforms worse than no participation?

CHAPTER 23

Digital PR for AI Visibility

Earning the mentions that shape what models say

WHAT YOU'LL LEARN

- How PR objectives change when the audience includes retrieval systems
- What makes a story both press-worthy and citation-worthy
- Working with journalists in a way that produces durable mentions
- Measuring PR against AI visibility rather than clip counts

Public relations was always about being talked about. In AI search, being talked about has a direct mechanical effect on what systems say about you. This makes PR more measurable and more strategically central than it has been for years — provided you change what you optimise for.

What changes

Traditional PR optimises for reach and prestige: the biggest publication, the largest audience, the most prominent placement.

PR for AI visibility optimises for **corroboration**: your brand associated with your topic, in text, across many independent sources that retrieval systems trust.

These often point in different directions. A single feature in a national newspaper delivers enormous reach. Twenty mentions across specialist trade publications delivers far better corroboration, because the association is repeated across independent sources in the relevant topical context.

KEY IDEA

For AI visibility, **breadth of consistent association beats depth of individual placement.** Twenty trade mentions saying the same thing about you outperform one national feature saying it once.

What makes a story citable

A story earns press coverage if it is newsworthy. It earns *citations* if it contains something a future answer will need. The overlap is:

- **Original data** (Chapter 16) — journalists need statistics; models need sources for them
- **A specific, quotable claim** attributed to a named person
- **A number** that becomes the reference figure for a question
- **A named phenomenon** — if you name a trend and it sticks, you become the source for it
- **A clear, extractable definition** of something people ask about

IN PRACTICE

Building a data-led PR asset

The reliable pattern, which serves both audiences at once:

1. **Run a survey** of your audience (method in Chapter 16).
2. **Find the counter-intuitive finding.** Journalists need a story, not a table.
3. **Write the press release around one headline number**, with two supporting figures.
4. **Publish the full data openly on your own site**, ungated, with each finding under a question-shaped heading.
5. **Offer named expert commentary** — a real person with a title who can be quoted.
6. **Pitch to trade publications first**, then broader outlets. Trade coverage is easier to earn and better for topical association.
7. **When coverage appears, check what it says about you** and correct errors immediately.

The result: press coverage, an owned citable asset, and consistent co-occurrence of your brand with your topic across many sources.

Working with journalists

Practical points that specifically help AI visibility:

- **Ask for your descriptor, not just your name.** "Acme Consulting, which specialises in agency operations" builds the topical association; "Acme Consulting" alone does much less.
- **Provide a boilerplate description** and use it consistently everywhere. Consistency is the entity mechanism from Chapter 14.
- **Offer a named spokesperson with a title.** Quotes attributed to a person with credentials carry expertise signals; quotes attributed to "a spokesperson" do not.
- **Do not insist on a link.** It is nice to have, but refusing an unlinked mention now means refusing something genuinely valuable.
- **Respond to journalist requests quickly.** Services connecting sources to reporters are a reliable route to consistent mentions, and speed is the main differentiator.

COMMON MISTAKE

Never supply a statistic you cannot substantiate, and never let a rounded figure drift upward between tellings. Retrieval systems cross-check across sources; a number that appears as “30%” in one article and “over a third” in another and “nearly 40%” in a third creates exactly the contradiction pattern that damages trust signals (Chapter 15).

Measuring PR against AI visibility

Clip counts and reach estimates tell you little about whether the work moved retrieval. Better measures:

Measure	How
Mention volume with topical context	Track brand + topic co-occurrence over time
Share of model change	Fixed prompt set, before and after (Ch. 23)
Description accuracy	Ask engines what you do; compare to intent
Citation of your data asset	Track who references your figures
Branded search volume	Search Console and trend data

CASE STUDY · KETTLEWORTH

Kettleworth pitched their grind-consistency study to trade publications rather than national press.

Eleven covered it. None had circulation worth mentioning. All of them described Kettleworth as “the Leeds specialty coffee equipment retailer”, because that phrase was in the press release and they asked for it specifically, rather than just their name.

Eleven independent sources, in the relevant topical context, saying the same thing about them.

Their share of model on grinding queries roughly doubled over the following quarter. They also got one national mention, which produced a spike of traffic and, as far as they could measure, no lasting change in how assistants described them.

CHAPTER SUMMARY

- PR for AI visibility optimises for corroboration — consistent association across many independent sources
- Breadth of trade coverage beats depth of a single national placement
- Data-led stories serve journalists and retrieval systems simultaneously
- Ask for your descriptor alongside your name; consistency is the entity mechanism
- Measure share of model and description accuracy, not clip counts

CHECK YOURSELF

1. Why might twenty trade mentions be worth more than one national feature here?
2. What five properties make a story citable as well as newsworthy?
3. Why should you not insist on a link?
4. Why is an inconsistently reported statistic actively harmful?

CHAPTER 24

When the Answer Is Wrong About You

Diagnosing and correcting what assistants say, and what to do when it is damaging

WHAT YOU'LL LEARN

- The five reasons an assistant says something false about your brand, and how to tell which you have
- How to trace a wrong answer back to the source that caused it
- The correction ladder, cheapest first — and what is genuinely unfixable
- A crisis playbook for when the wrong thing is also damaging

Every organisation that starts measuring AI visibility discovers the same thing within an afternoon: the assistants say things about them that are not true. Prices from two years ago, products they discontinued, a competitor's feature attributed to them, occasionally something worse. This chapter is the repair manual.

Five causes, five different fixes

Diagnosis matters, because the remedies do not overlap.

Cause	What it looks like	Where the fix is
Stale source	Old but plausible facts — a former price, a discontinued line	The source page, not yours
Entity confusion	Facts belonging to a similarly named organisation	Chapter 15 — entity signals
Your own contradiction	Different figures in different places on your own site	Your site
A loud negative	One critical thread quoted as representative	Corroboration, not suppression
Hallucination	A specific claim with no traceable source at all	Ungrounded — hardest to fix

KEY IDEA

Never start fixing before you know which of the five you have. The most common error is rewriting your own pages when the problem is a third-party listing you had forgotten existed — effort spent where the answer was not coming from.

Tracing the wrong answer

The method takes about twenty minutes per claim.

1. **Reproduce it.** Ask the same question three times, logged out, in fresh sessions. If it appears once in three, it is unstable rather than settled — note it and move on.
2. **Ask for the source.** *"Where did you get that?"* Retrieval-based engines will usually name a page. This is what Perplexity is best for (Chapter 6).
3. **Open the source and find the exact sentence.** Nine times in ten it is there, stated plainly, on a page you did not know was being read.
4. **If no source is offered**, ask the same question on an always-retrieve engine. If that one cites something, you have your source. If none of them can, you are probably looking at a hallucination or a training-data association.
5. **Record it** in the accuracy column of your tracker (Chapter 29). Wrong answers recur, and you want the history.

IN PRACTICE

Worked trace

Claim: an assistant states your product costs £149. It costs £189.

1. Reproduce — appears in three of three. Settled, not noise.
2. Ask for the source — it names a marketplace listing.
3. Open it — the listing shows £149 and was last edited in 2024.
4. Check your own site — £189 in the HTML, correct.

Diagnosis: stale source. **Fix:** update or remove the marketplace listing. Nothing on your own site needed changing, and rewriting your pricing page would have achieved precisely nothing.

The correction ladder

Work upward. Each rung costs more than the last.

1. Fix your own contradictions first

Free, fast, and a precondition for everything above it. Run the consistency check from Chapter 15: name, description, founding date, pricing, headcount, across your site,

LinkedIn, Crunchbase, directories and the public register. A cross-checking system that finds you disagreeing with yourself will not resolve in your favour.

2. Correct the third-party source

Usually an email. Most editorial round-ups, review platforms and directories accept corrections readily — you are offering accuracy, not asking a favour. Chapter 22's source-concentration list tells you which ones matter.

3. Publish the canonical answer yourself

If a question keeps getting answered wrongly, publish the correct answer in a form built to be retrieved: question-shaped heading, answer in the first sentence, specific and dated. *"How much does the K3 cost? The K3 is £189 as of August 2026. It was £149 during a 2024 promotion that has ended."*

Stating the superseded fact explicitly is deliberate. It gives a retrieval system something that reconciles the contradiction rather than adding a third competing claim.

4. Build corroboration

If one source says the wrong thing and nothing says the right one, the wrong one wins by default. Several independent sources stating the correct fact resolve it. This is Chapter 22's work applied to a specific problem.

5. Use the platform's feedback route

Every major assistant has one — thumbs-down, a report link, a feedback form. Effectiveness is unclear and unacknowledged, which is why it is the last rung rather than the first. Use it, log it, expect nothing quickly.

> **COMMON MISTAKE**
>
> Do not attempt to suppress genuine criticism. If an assistant reports a real weakness in your product, the answer is not a correction campaign — it is either fixing the product or addressing the criticism openly. Manufacturing contradicting content to bury a true statement is the manipulation Chapter 56 warns about, and it fails in the same way: the attempt itself becomes citable.

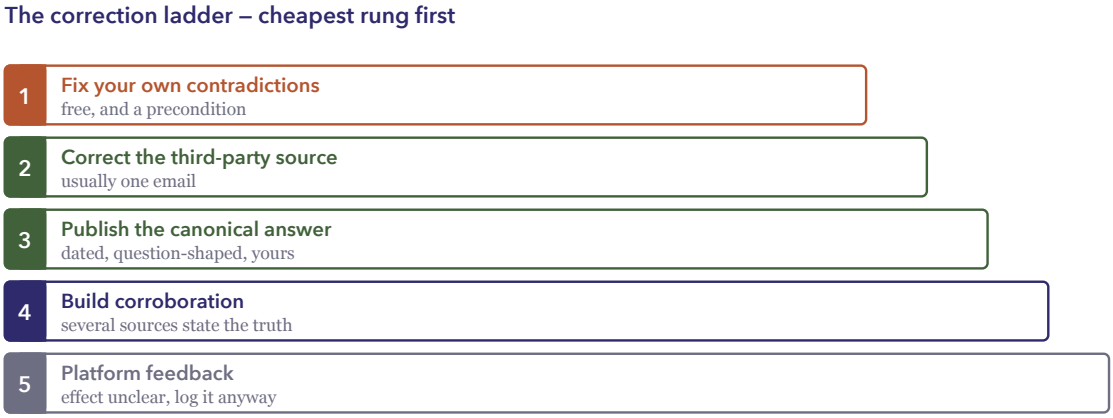


Figure 24.1 – Work upward. Most wrong answers are fixed on the first two rungs.

When it is damaging: the crisis playbook

Occasionally the wrong answer is not a stale price but something that costs you customers — a safety claim, a legal implication, a confusion with a company in trouble.

1. **Establish the blast radius within the hour.** All four engines, five phrasings each. Is it one engine or all of them? Every phrasing or one? A single-engine, single-phrasing problem is not a crisis.
2. **Trace the source.** The method above. Damaging claims almost always have one traceable origin.
3. **Deal with the source directly**, with escalation proportionate to the harm. Most publishers correct genuine factual errors quickly when asked plainly.
4. **Publish your own unambiguous statement**, dated, structured to be retrieved, on a URL you control.
5. **Tell sales and support the same day.** They will be asked about it before your fix propagates, and they need the correct answer and a link.
6. **Re-measure weekly.** Correction propagation is measured in weeks, not days — the source must be re-crawled and the change must outweigh the accumulated prior signal.

> KEY IDEA
>
> Expect **four to eight weeks** for a corrected source to change what assistants say, and longer where the wrong claim is widely repeated. Say this to stakeholders on day one. The pressure to “fix it today” produces panicked over-correction that makes the record messier.

What you cannot fix

Being straight about the limits saves a lot of wasted effort:

- **A model's training-data associations** cannot be edited. If a model learned something before its cutoff, retrieval can override it in grounded answers but not in ungrounded ones.
- **Genuinely true criticism** is not a correction problem.
- **Pure hallucinations with no source** have nothing to correct. They often decay on their own; sometimes publishing a clear canonical answer gives the system something better to reach for.
- **Another company's reputation**, when you share a name. Chapter 15's entity work reduces the confusion but rarely eliminates it.

CASE STUDY · KETTLEWORTH

Kettleworth's damaging one arrived in March. Asked whether their grinders were safe for commercial use, two of four assistants said the K5 had been "subject to a recall."

There had been no recall. The trace took forty minutes: a forum post from 2023 described Kettleworth "recalling a batch" — which was true, and was the batch return described in Chapter 16, handled with the supplier and never a safety matter. A later aggregator page had summarised that post as a product recall, and that page was being cited.

They did four things in a day. Emailed the aggregator with the facts and the supplier correspondence. Published a dated page titled "Has the Kettleworth K5 ever been recalled?" answering plainly — no, and here is what the 2023 batch return actually was. Briefed sales and support with the link. Logged it in the tracker.

The aggregator corrected in nine days. The claim persisted in one engine for a further five weeks, then stopped. Their own page is now the cited source for the question — which means the honest version, including the batch return, is what gets repeated.

CHAPTER SUMMARY

- Five causes — stale source, entity confusion, self-contradiction, a loud negative, hallucination — with five different fixes; diagnose before acting
- Trace every claim to its source before touching your own site; the answer usually is not coming from you
- Work the correction ladder upward: your contradictions, the third-party source, your own canonical answer, corroboration, then platform feedback
- Do not suppress true criticism; the attempt becomes citable
- Corrections take four to eight weeks to propagate — set that expectation on day one

CHECK YOURSELF

1. Name the five causes of a wrong answer and the fix each one calls for.
2. An assistant quotes a price you abandoned two years ago. What do you check first, and what would rewriting your pricing page achieve?
3. Why does a corrective page state the superseded fact explicitly rather than ignoring it?
4. What are the four things that genuinely cannot be fixed, and why is saying so early useful?

PART VI

Commerce and Place

Selling through a surface that answers before you get a chance to, what to do with the visitor a citation earns you, and what changes when the answer depends on where someone is standing.

CHAPTER 25

E-commerce and Product Discovery

When the thing being retrieved is a product, not a paragraph

WHAT YOU'LL LEARN

- How AI surfaces select products, and why it is not the same as selecting passages
- Product feeds, structured data and the machine-readable layer that decides eligibility
- Why reviews are your most valuable and most commonly invisible asset
- How to win the comparison and "best X" queries that decide category purchases

Everything so far has treated the retrievable unit as a passage of text. For anyone selling products, a second unit exists alongside it — the product record — and it is governed by a different pipeline with different rules. Missing this is why e-commerce sites with excellent content still lose category queries.

Two pipelines, not one

A product can reach an AI answer by two routes:

Through content. Your review, guide or comparison page is retrieved and cited like any other passage. Everything in Parts III to V applies unchanged.

Through the product graph. Structured product data — from your feed, your schema, or a marketplace listing — is used to name and compare actual products, with prices and availability. This route bypasses your prose entirely.

KEY IDEA

A model answering *"what's the best burr grinder under £200"* is doing two things at once: retrieving **passages** that discuss grinders, and consulting **product records** for what exists at that price. You can win one and lose the other, and most retailers do exactly that.

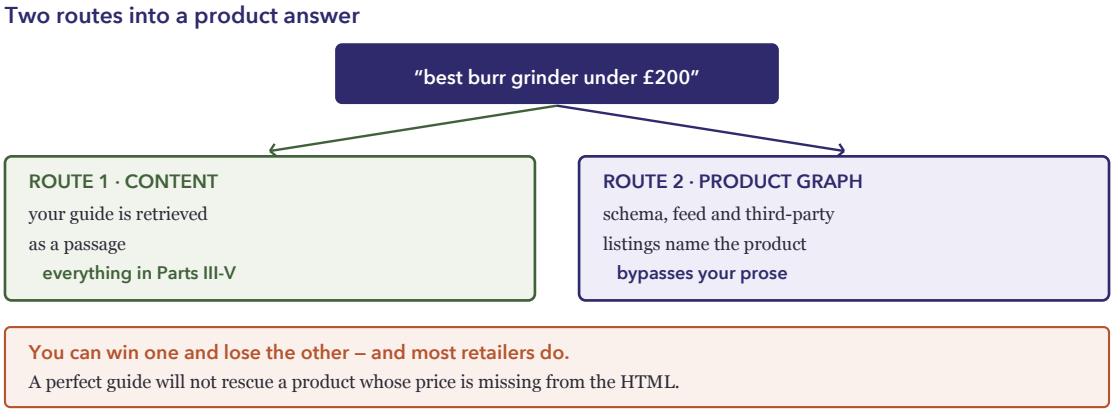


Figure 25.1 – Two routes into a product answer. Winning one does not win the other.

The machine-readable layer

Three overlapping sources feed the product graph. You want to be correct and consistent across all three, because contradictions between them are resolved against you.

1. Product structured data

Chapter 11's `Product` and `Offer` markup is the foundation. For e-commerce it is not optional and it must be complete:

```
{
  "@context": "https://schema.org",
  "@type": "Product",
  "name": "Kettleworth K3 Burr Grinder",
  "sku": "KW-K3-BLK",
  "gtin13": "5060123456789",
  "brand": { "@type": "Brand", "name": "Kettleworth" },
  "description": "40mm conical burr grinder with 38 stepped settings.",
  "offers": {
    "@type": "Offer",
    "price": "189.00",
    "priceCurrency": "GBP",
    "availability": "https://schema.org/InStock",
    "priceValidUntil": "2026-12-31",
    "shippingDetails": { "@type": "OfferShippingDetails" },
    "hasMerchantReturnPolicy": { "@type": "MerchantReturnPolicy" }
  },
  "aggregateRating": {
    "@type": "AggregateRating",
    "ratingValue": "4.6", "reviewCount": "214"
  }
}
```

How to read that, if you have never read JSON before.

It is a set of labelled facts nested inside one another. Each line is a label, a colon, and a value. Curly braces group labels into one thing, and a label whose value is itself in braces means "this fact is a thing in its own right, with facts of its own".

Four labels do the structural work:

- `@context` names the vocabulary in use. `schema.org` is the shared dictionary that search engines and assistants all read, so this line is what makes everything below it mean anything to them. It is identical on every block you will ever write.
- `@type` says what kind of thing is being described. `Product` at the top means "everything inside these braces describes one product".
- **The plain labels** — `name`, `sku`, `description` — are facts about that product.
- **The nested blocks** are separate things attached to it. `brand` is a `Brand`. `offers` is an `Offer`: price, currency, availability. `aggregateRating` summarises reviews. Each carries its own `@type` because each is a distinct kind of object rather than a bare value.

So the block states, in a form a machine can act on without interpretation: *this is a Product called the Kettleworth K3, made by the Brand Kettleworth, sold under an Offer of £189 and in stock, rated 4.6 across 214 reviews.* Nothing has to be inferred from your page layout, your fonts, or where the price happens to sit — which is the entire point, and why a price rendered only in JavaScript (Chapter 9) is invisible while this is not.

> **COMMON MISTAKE**
>
> Do not invent values to fill the fields. `aggregateRating` in particular must reflect reviews that genuinely exist and are visible on the page. Fabricated ratings are a policy violation that can cost you rich results altogether, and the downside vastly outweighs the upside.

The fields that matter most and are most often missing: `gtin` (the barcode number — this is how systems match your listing to the same product elsewhere), `sku`, **shipping and returns**, and `priceValidUntil`.

`gtin` deserves particular emphasis. Without it, a system cannot confidently tell that your "K3 Burr Grinder" and a marketplace's "Kettleworth K3" are the same object. With it, your listing joins a product record that already has reviews, price history and comparisons attached.

2. Merchant feeds

If you run Google Shopping, your Merchant Center feed is a second, richer description of your catalogue. Feed quality affects eligibility for Google's shopping surfaces, which increasingly include AI-driven ones. Keep the feed and the on-page schema consistent — divergence between them is a common and quietly expensive error.

3. Marketplace and retailer listings

Your products almost certainly exist on other sites. Those listings feed the same product graph, and if their data is wrong — old price, wrong specification, discontinued variant — that wrong data is competing with yours.

IN PRACTICE

The product data audit

For your twenty best-selling products:

1. **Check the schema** with the Rich Results Test. Note missing `gtin`, `sku`, shipping and returns.
2. **View Page Source** and confirm the price is in the HTML, not injected by script (Chapter 9). This is the single most common e-commerce failure.
3. **Search the exact product name** and list every third-party page selling or describing it.
4. **Compare the data** on each: price, specification, availability. Note contradictions.
5. **Ask four engines** "how much is *[exact product name]*" and record what each says.

Step 5 usually produces the shock. Retailers routinely discover that assistants are quoting a price they abandoned eighteen months ago, sourced from a stale marketplace listing they forgot existed.

Reviews: the most valuable invisible asset

Chapter 7 established why user-generated content dominates citations — specific, experiential, comparative, answer-shaped. Product reviews are all four, and you are sitting on hundreds of them.

They are also, on the majority of e-commerce sites, completely invisible to AI crawlers, because they are delivered by a third-party JavaScript widget that renders in the browser (Chapter 9).

COMMON MISTAKE

This is the highest-value, lowest-effort fix available to most retailers. A site with 4,000 genuine reviews that no AI crawler can read is losing exactly the queries it should own — *"is the K3 any good"*, *"K3 problems"*, *"K3 vs Baratza"* — to a Reddit thread with nine comments.

Ask your reviews provider for a server-side rendering option. Most now have one. If they do not, render a summary of review text into the HTML yourself alongside the widget.

Beyond making them readable:

- **Surface the specifics.** "Works great!" is not citable. Prompt reviewers with questions — how long have you owned it, what did you use it for, what would you change, and the resulting text becomes retrievable.
- **Keep negative reviews.** A page with only five-star reviews reads as unreliable to a cross-checking system, and gives it nothing to answer "what are the downsides" with.

Someone will answer that question; better it is you.

- **Answer reviews publicly**, especially critical ones. The reply is text too, and it demonstrates the trust signals from Chapter 15.

Winning the comparison query

Category purchases are decided by *"best X for Y"* and *"X vs Z"*. These trigger heavy fan-out and lean on third-party comparison content, as Chapter 21 covered.

The specific e-commerce moves:

1. **Publish genuine comparisons that include competitors**, and say plainly when a competitor is better for a particular use. This feels wrong and it works: comparative, unhedged content is what gets cited, and the alternative is that only your competitors' comparisons exist.
2. **Build "best for *[use case]*" pages** rather than "best" pages. The fan-out generates use-case sub-queries; a page answering *"best grinder for espresso on a budget"* has far less competition than one answering *"best grinder"*.
3. **State the trade-off explicitly.** "The K3 is quieter but slower than the K5" is a liftable comparative sentence. A specification table alone is not.
4. **Fix your presence in third-party comparisons.** Chapter 21's campaign, applied to category round-ups.

> CASE STUDY · KETTLEWORTH
>
> Kettleworth had 4,100 reviews averaging 4.6 across their catalogue, and none of them existed as far as any AI crawler was concerned — their reviews platform rendered entirely client-side.
>
> The consequence showed up in the baseline. Asked *"is the Kettleworth K3 any good"*, three of four engines answered from a single Reddit thread in which two users complained about static cling. Asked about price, two engines quoted £149 — a promotional price from a marketplace listing Kettleworth had stopped maintaining in 2024.
>
> The fixes took a fortnight: server-side rendering enabled on the reviews platform, `gtin` and shipping data added to the product schema, and the stale marketplace listings either updated or removed.
>
> Six weeks later, the same question returned Kettleworth's own review summary in three of four engines, including the static-cling complaint alongside 4,100 ratings in context. Price accuracy went from 50% to 100%.
>
> Nothing was written. The content had existed all along; it had simply been unreadable.

CHAPTER SUMMARY

- Products reach answers by two routes — retrieved passages and the product graph, and you can win one while losing the other
- `gtin` is how systems match your listing to the same product elsewhere; its absence isolates you from the product record
- Keep schema, merchant feed and third-party listings consistent; contradictions resolve against you
- Reviews are your most citable asset and are usually invisible because of client-side rendering
- Keep negative reviews and publish honest comparisons — comparative, unhedged content is what gets cited

CHECK YOURSELF

1. Describe the two routes a product can take into a generated answer.
2. Why does a missing `gtin` matter more than a missing description?
3. Why is a page with only five-star reviews weaker than one with a realistic spread?
4. Why does “best grinder for espresso on a budget” beat “best grinder” as a page?

CHAPTER 26

After the Citation

The visitor who arrives already half-answered

WHAT YOU'LL LEARN

- Why a visitor from an AI assistant arrives in a different state from a search visitor
- What that changes about landing pages, messaging and calls to action
- The specific mistake of re-answering a question the model already answered
- How to measure and improve this without clean attribution

Everything up to here has worked towards being cited. This chapter starts the moment that succeeds — because the visitor who clicks through from an AI answer is not the visitor you spent twenty years optimising for, and treating them the same wastes most of what the citation was worth.

A different arrival state

Two visitors, two states

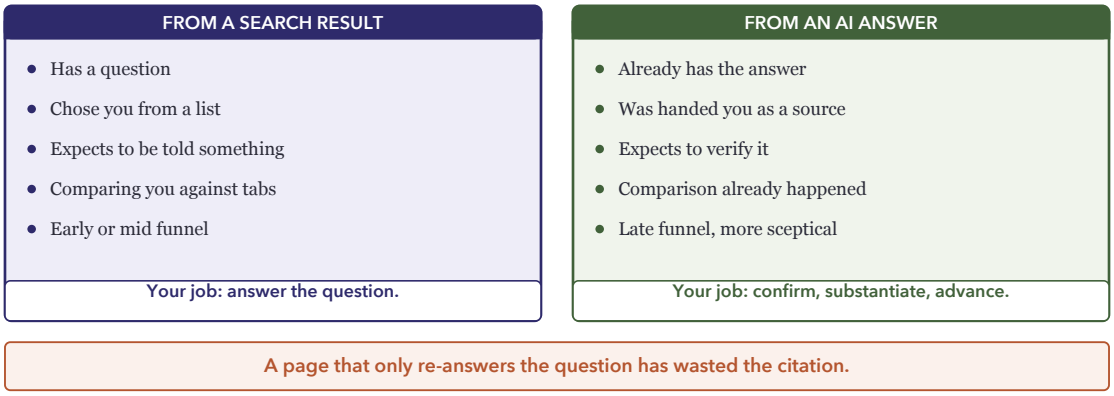


Figure 26.1 – Two visitors, two states. The AI-referred visitor arrives with the answer already given.

A visitor from a classical search result arrives **seeking an answer**. They have a question, they picked your link from a list, and they expect your page to tell them something.

A visitor from an AI answer arrives **verifying an answer they already have**. The model has already told them what a grinder costs, or which consultancy suits their size, or what the trade-offs are. They clicked because they wanted to check the source, get more detail, or decide whether to trust it.

That difference produces four practical consequences:

1. **They are further down the funnel.** They have done their comparison inside the assistant. This is the mechanism behind the widely reported finding that AI-referred traffic converts at much higher rates — it is not that the traffic is magical, it is that the early stages happened elsewhere.
2. **They are more sceptical, not less.** They are checking a claim. Trust signals matter disproportionately on arrival.
3. **They arrive deep, not on your homepage.** They land on the specific page that was cited, often mid-site, with no context about who you are.
4. **They have limited patience for repetition.** They already know the answer. Being told it again reads as padding.

> KEY IDEA
>
> The AI-referred visitor has already had their question answered. Your page's job is no longer to answer it — it is to **confirm, substantiate and advance**. A page that simply restates what the assistant said has wasted the citation.

What the page needs to do

Confirm quickly

Within the first screen, the visitor should see that this page is indeed about the thing they were told, and that the specific fact they came to check is here. If the model said your grinder is £189, the number £189 should be visible immediately — not after a scroll, not behind a variant selector.

Substantiate

They came to verify. Give them the evidence the assistant summarised away: the method behind your data, the sample size, the caveats, the photographs, the named author with credentials. This is where the E-E-A-T investment from Chapter 15 pays off commercially rather than just mechanically.

Establish who you are, fast

They landed deep with no context. A brief, concrete statement of who you are and why you are credible on this — not a mission statement — needs to be present on every page, not just the homepage.

Advance the decision

The call to action should match a late-funnel state. "Learn more" and "download our guide" are answers to questions this visitor has already resolved. What they need is the next concrete step: see the price, check availability, book the call, start the trial, ask a specific question.

IN PRACTICE

Auditing a cited page for the AI visitor

Take a page you know is being cited. Open it as if you had just been told about it by an assistant.

1. **Within five seconds**, can you confirm the specific fact you came to check?
2. **Is there evidence** for that fact, beyond the assertion — data, method, photograph, named author?
3. **Can you tell who this company is** and why they would know, without navigating away?
4. **Is the next step appropriate** for someone who has already decided this is roughly right, or is it aimed at someone just starting to look?
5. **Does the page repeat** at length what an assistant would already have summarised?

Most cited pages fail on 1, 3 and 4. They were written as destinations for early-stage searchers and are being used as verification pages by late-stage buyers.

The re-answering mistake

The most common error is subtler than a bad call to action. A page optimised hard for retrieval — answer-first, self-contained, specific — is by construction the page an assistant can summarise most completely. Which means the visitor arriving from that summary has already read the substance of your page.

The resolution is not to make the page worse at retrieval. It is to make sure the page contains **more than the answer**:

- The answer, stated cleanly at the top, for retrieval
- The evidence underneath it, which the summary omits
- The edge cases and exceptions, which summaries always drop
- The specific application to the reader's situation
- The route to acting on it

A page built that way serves both audiences. The retrieval system takes the top; the human who arrives gets the rest.

COMMON MISTAKE

Do not respond to this by gating the detail or holding back the answer to force a click. It removes you from the answer entirely — you lose the citation and the visitor. The correct move is to be fully citable *and* to have more depth available than can be summarised.

NOTE

If you sell B2B

The arrival state is the same; the timescale is not. A B2B visitor arriving from an AI answer is still verifying rather than seeking, but they may be one of five people who will touch the decision, and the gap between citation and purchase can be months.

Three adjustments:

- **Substantiate for a sceptic who must convince colleagues.** The evidence on the page needs to be forwardable: a figure with a method, a named author, something they can paste into an internal thread.
- **Make the next step low-commitment.** "Book a demo" is a large ask of someone still building an internal case. A specification sheet, a pricing range or a comparison they can circulate does more.
- **Expect no attributable click at all.** The person who read the AI answer may never visit; the person who visits may arrive weeks later via a colleague's link. Chapter 30's proxy measures matter more here than anywhere else in the book.

Measuring it

Attribution is poor (Chapter 30), but you are not helpless:

- **Segment by landing page.** Pages you know are cited can be compared against comparable non-cited pages for bounce rate, scroll depth and conversion.
- **Segment by referrer** where assistants pass one, accepting that it undercounts.
- **Watch conversion rate on cited pages** specifically. If citation is working, it should rise as the traffic mix shifts towards verifiers.
- **Ask.** The "how did you hear about us" field from Chapter 28, plus one further question on enquiry forms: *what did you already know about us before you got in touch?* The answers tell you what the assistants are saying.

CASE STUDY · KETTLEWORTH

Kettleworth's grinder comparison page started getting cited in October. Traffic to it fell 22%, and revenue attributable to it rose 40%.

The page had been built for early-stage researchers: a long explanation of burr types, then a comparison table, then product links at the bottom. The visitors now arriving had already been told which grinder suited them; they wanted to check the claim and buy.

Three changes, none of which touched the retrievable top section. The specific model recommendations moved above the fold, with prices and stock visible. A two-line "who we are and why we test this" block went onto every guide. And the bottom-of-page "read our other guides" links were replaced with a direct add-to-basket for the recommended models.

Conversion rate on that page went from 1.1% to 3.4%. The citation had been working for two months; the page had been wasting it.

CHAPTER SUMMARY

- AI-referred visitors arrive verifying an answer, not seeking one — further down the funnel and more sceptical
- Their job for your page is confirm, substantiate, establish credibility, advance — not answer
- They land deep with no context, so credibility and identity must be present on every page
- The re-answering mistake: a page perfectly optimised for retrieval has already been summarised; it needs depth beyond the summary
- Never gate the answer to force a click — you lose the citation and the visitor

CHECK YOURSELF

1. In what state does an AI-referred visitor arrive, and how does it differ from a search visitor?
2. Why does a page perfectly optimised for retrieval risk disappointing the visitor it earns?
3. What is wrong with "download our guide" as the call to action on a cited page?
4. Why is gating the detail the wrong response to the re-answering problem?

CHAPTER 27

Local and Service-Area GEO

When the answer depends on where the person is standing

WHAT YOU'LL LEARN

- Why local queries behave differently from everything else in this book
- The three signal families that decide local visibility, and their relative weight
- How assistants answer "near me" when they have no map to show
- A practical local GEO routine for a single site or a hundred

Local is the exception to several rules in this book, and if you work for a business with premises or a service area, it is probably the most commercially important chapter here. It is also the one place where AI search has so far changed *less* than people expected — for reasons worth understanding.

Why local is different

Three structural differences:

The answer set is bounded by geography. For "best CRM" the candidate pool is global. For "emergency plumber in Leeds" it is a few dozen businesses, and the ranking factors are dominated by proximity, opening hours and whether anyone can verify the business exists.

AI Overviews trigger far less. As Chapter 5 established, generative summaries appear heavily on informational queries and lightly on local and transactional ones. If your business lives on "near me" searches, your exposure is genuinely low — and calculating that, rather than assuming, is the first thing to do.

Verification carries unusual weight. A local answer that sends someone to a business that has closed is a much worse failure than a slightly wrong statistic. Systems compensate by leaning hard on verified, structured, corroborated local data.

KEY IDEA

If you are local, run the exposure calculation from Chapter 5 before doing anything else in this book. Many local businesses discover that under 15% of their impressions sit on AI-Overview-triggering queries — in which case classical local SEO remains the priority and this chapter is a monitoring job, not a transformation.

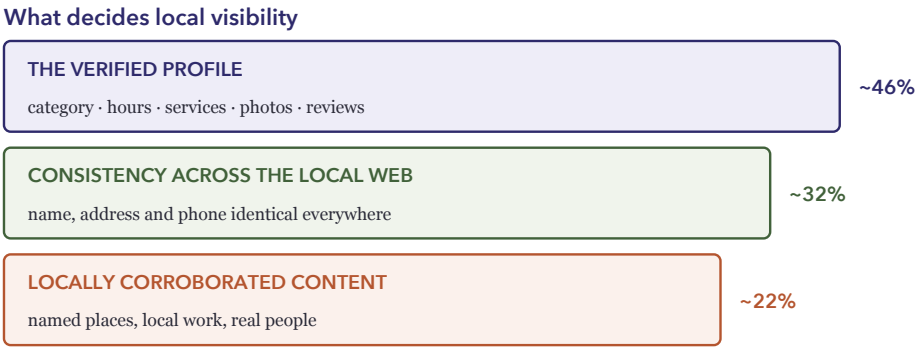


Figure 27.1 – The three signal families behind local visibility, and roughly how much each is worth.

The three signal families

1. The verified profile

Your **Google Business Profile** — and its equivalents on Apple Business Connect and Bing Places — is the single highest-leverage local asset. It is a structured, verified record, which is exactly what a system wants when the cost of being wrong is high.

What actually matters on it, roughly in order:

- **Correct primary category**, chosen precisely rather than aspirationally
- **Opening hours**, including holidays, kept current
- **Complete service or product listings**
- **Photographs**, refreshed periodically
- **Reviews**, and replies to them
- **Posts and updates**, which signal an active business
- **Attributes** — accessibility, payment methods, service options

2. Consistency across the local web

The old **NAP** discipline — Name, Address, Phone — identical everywhere. This is entity consistency (Chapter 15) applied to local, and it matters more here because the verification bar is higher. One transposed digit across forty directories is a real problem.

Where consistency counts most: the major aggregators, industry directories for your trade, local chambers and associations, and your own site's contact page and schema.

3. Locally corroborated content

Content that demonstrates you actually operate where you claim: named neighbourhoods, local landmarks, case studies from local clients, staff who are visibly local, involvement in local events. This is the "Experience" pillar (Chapter 16) with a geographic dimension.

COMMON MISTAKE

The classic failure is the doorway page — a template repeated for fifty towns with the place name swapped. It fails on every axis this book cares about: no self-contained answer, no specific information, no experience signal, and a strong duplicate-content problem. Systems have penalised this pattern for over a decade and generative retrieval finds it even less useful.

If you serve fifty towns, write genuinely about the five that matter most and let a single well-structured service-area page cover the rest.

How assistants answer local questions

Without a map to show, an assistant does one of three things:

1. **Refuses to be specific** — "I'd recommend searching for plumbers in your area", the most common response for genuinely local intent
2. **Names a few businesses** it can verify, usually those with strong structured presence and review corroboration
3. **Describes how to choose** rather than choosing, which is the shape most assistants default to for high-stakes local decisions

The opportunity sits in the second and third. Being nameable requires verified structured data. Being useful in the third case means having content that answers *"how do I choose a plumber in Leeds"* — which is an informational query, does trigger AI Overviews, and is where local businesses can win visibility that pure "near me" optimisation never reaches.

IN PRACTICE

The local GEO routine

Monthly, about ninety minutes:

1. **Verify the profile.** Hours, categories, services, phone. Check it renders correctly on both desktop and mobile.
2. **Run five local prompts** on your two tracked engines: "best *[trade]* in *[town]*", "how do I choose a *[trade]*", "*[your business]* opening hours", "is *[your business]* any good", "*[competitor]* alternatives in *[town]*".
3. **Check what it says about your hours and location.** This is the most commonly wrong local fact and the most commercially damaging.
4. **Reply to every review**, positive and negative. Replies are text, and text is retrievable.
5. **Add one piece of locally specific content** — a job you did, a local question answered, a neighbourhood guide.

Quarterly, add a NAP consistency sweep across your top twenty citations.

Multi-location businesses

If you run ten or more locations, two additional rules:

- **One page per location, each genuinely distinct.** Different staff, different photographs, different local detail, different reviews. If two location pages differ only by place name, merge them into a service-area page.
- **Structured data per location**, with `LocalBusiness` carrying the individual address, hours and coordinates — not the head-office details repeated.

CASE STUDY · KETTLEWORTH

Kettleworth had a Leeds workshop that they treated as a warehouse. It had no Google Business Profile, and their contact page listed a PO box.

This mattered more than they realised. Asked *"where can I try a coffee grinder before buying in Leeds"* — a question with obvious commercial intent — assistants named two competitors with shops. Kettleworth went unmentioned, despite running monthly open testing sessions at the workshop, and having done so for years.

They claimed the profile and set the category to "Coffee machine supplier", listing the open sessions as an attribute. The real address went on the contact page with `LocalBusiness` schema. And they wrote one page: *"Can you try a grinder before buying?"* — answering yes, here is when, here is what happens, here is how to book.

Within two months they were being named for the original question. The open sessions had existed for three years; nothing about the business changed. It had simply never been stated anywhere a machine could read.

CHAPTER SUMMARY

- Local behaves differently: bounded answer sets, far lower AI Overview exposure, and an unusually high verification bar
- Calculate your exposure before investing — many local businesses have a monitoring job rather than a transformation
- Three signal families: the verified profile, NAP consistency across the local web, and locally corroborated content
- Doorway pages fail on every axis; write properly about the few places that matter
- The reachable opportunity is often the informational local query — "how do I choose a X in Y" — not the "near me" search itself

CHECK YOURSELF

1. Give three reasons local queries behave differently from the rest of this book.
2. Why is verification weighted more heavily for local answers than for informational ones?
3. What is wrong with fifty near-identical town pages, in the terms this book uses?
4. Why is “how do I choose a plumber in Leeds” a better GEO target than “plumber near me”?

PART VII

Measurement and Proof

Where most practitioners are weakest — including how to show that your work, rather than the weather, caused the change.

CHAPTER 28

The Measurement Problem

Why the old numbers stopped working, and what honesty looks like here

WHAT YOU'LL LEARN

- Exactly which measurement capabilities you lost and why
- The four dimensions of visibility, and which are observable
- Why proxy metrics are legitimate rather than an admission of defeat
- How to talk about uncertainty without losing credibility

Everything in Parts III to V is doable. This part is where the discipline gets genuinely hard, because the feedback loop that made SEO a measurable practice has partly broken — and pretending otherwise is the fastest way to lose a stakeholder's trust when they eventually notice.

What you lost

In classical SEO you could observe: your position for a query, impressions and clicks per query, which pages received them, and how that changed over time. Rank tracking was possible because ranks existed and were public.

In AI search:

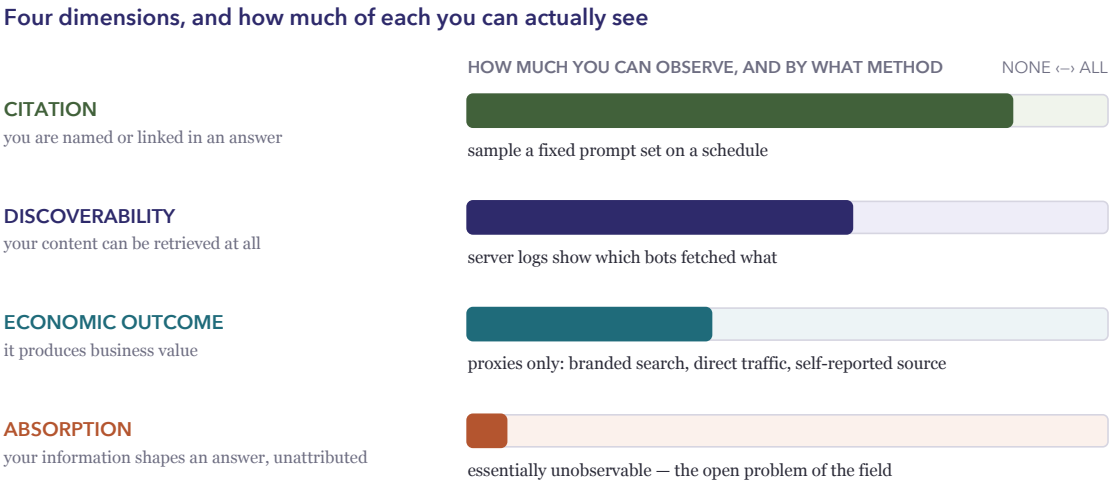
- **There is no rank.** A passage is retrieved or not.
- **There is no impression count.** Nobody reports how many generated answers mentioned you.
- **There is often no click.** The user got their answer.
- **Answers vary between users.** Personalisation, conversation history and non-determinism mean two people asking the same question get different answers.
- **Referral data is patchy.** Some assistants pass a referrer; some do not; some strip it.

KEY IDEA

The honest position: **you cannot measure AI search visibility directly and completely. You can sample it, proxy it, and triangulate it.** Everyone claiming a precise figure is either sampling and not saying so, or selling something.

The four dimensions

The July 2026 critical survey's **visibility vector** is the most useful framework available, because it separates things routinely conflated:



SAY THE BOTTOM ROW OUT LOUD, EARLY

Your content can be read, used, and shape what millions are told — with no citation, no click and no line in any report.
That is a property of the medium, not a failure of your reporting. Stakeholders take it far better before they discover it themselves.

Figure 28.1 – The four dimensions, ordered by how much of each you can observe, and by what method.

Read the figure top to bottom and you have the shape of the whole discipline. **Citation** is close to fully observable, because you can ask the engines yourself on a schedule and count what comes back. **Discoverability** is observable at one remove: server logs tell you which bots fetched which pages, which is a fact about access rather than about answers. **Economic outcome** you reach only through proxies. And **absorption** — your material shaping an answer that never names you — is barely observable at all.

The ordering matters more than the individual rows, because it tells you where to spend effort. Measurement work at the top of the figure returns real information for the hours you put in. Work at the bottom returns estimates at best, and the sensible response is to stop trying to measure it precisely and start saying out loud that it exists.

Absorption is the uncomfortable one and deserves to be stated plainly to stakeholders early, before someone else discovers it. Your content can be read, used, and materially shape what millions of people are told about your category, while producing no citation, no click, and no line in any report.

That is not a reporting failure. It is a property of the medium.

Why proxies are legitimate

There is a temptation to treat proxy metrics as second-class. They are not. Large parts of marketing have always been measured by proxy — brand tracking, share of voice, media

reach. Search was briefly, unusually, directly measurable, and that was the exception rather than the norm.

The proxies that work here:

- **Branded search volume.** Rises when more people encounter your name. Free, reliable, and directly connected to the feedback loop in Chapter 21.
- **Direct traffic.** Rises when people who learned your name type it in.
- **Share of model against a fixed prompt set.** Sampled, but consistent and comparable over time.
- **Referral traffic from assistant domains.** Undercounts badly, but the *trend* is informative.
- **Conversion rate of organic traffic.** Should rise as pre-qualified visitors replace browsers.
- **"How did you hear about us?"** on forms. Unfashionable, unreasonably effective.

IN PRACTICE

Adding the one question that fixes a lot

Add an open-text "How did you hear about us?" field to your enquiry forms, and instruct sales to ask on every call.

Within a quarter you will have qualitative data that no analytics platform can give you: people saying "ChatGPT recommended you", "I asked Claude for options and you came up", "someone on Reddit mentioned you".

This is the cheapest, most robust AI-attribution instrument available, it works when referrers are stripped, and almost nobody does it.

Talking about uncertainty

The instinct is to hide uncertainty, because uncertainty sounds like incompetence. It does the opposite. A stakeholder who is given a precise number that later proves wrong stops trusting you; one given an honest range with a stated method trusts you more.

Some phrasings that hold up:

> *"We sample 50 fixed prompts monthly across two engines. Our share of model went from 18% to 26% — that is nine answers to thirteen, on one run each. The sample is consistent, so it is worth watching, but four answers is inside the run-to-run variation we measured. I will call it a trend if it holds next month." (Illustrative figures, not Kettleworth's.)*

> *"We can't measure how often our content shapes an answer without being credited. What we can measure is branded search, which is up 22% over the period, and that usually moves when visibility moves."*

> *"I'd rather give you a range than a false decimal. The honest answer is somewhere between X and Y, and here's why the spread exists."*

COMMON MISTAKE

Do not report a single-engine, single-day sample as though it were a stable metric. Answers vary between runs even for the same prompt on the same engine. Always sample multiple prompts, note the date, and report the trend rather than the point.

KEY IDEA

The arithmetic matters more than it looks. A move from 18% to 26% across 50 prompts is a move from nine mentions to thirteen — **four answers**. Assistants are non-deterministic, so four answers can change without anything about your visibility changing. The next chapter's repeated-run protocol is what tells you whether four is inside your noise or outside it, and until you have measured that, "the trend is real" is a sentence you cannot support.

CASE STUDY · KETTLEWORTH

Kettleworth's founder asked the reasonable question: how much revenue had this produced?

The honest answer was that nobody could fully know. Referral data showed 1,200 sessions from assistant domains that quarter, which was certainly an undercount. Branded search was up 31%. Direct traffic was up 19%.

What they could not measure was absorption — how often their grind-consistency data shaped an answer without being credited. Given that two trade publications now cited it, the number was plainly not zero.

They set the expectation in writing, in the first report, before anyone asked: we sample, we do not census, and some of our influence will never appear in any dashboard.

CHAPTER SUMMARY

- Rank, impressions and often clicks are gone; direct complete measurement is not available
- The four dimensions — discoverability, citation, absorption, economic outcome — have very different observability
- Absorption is real, significant and largely unmeasurable; say so early
- Proxy metrics are legitimate; branded search and "how did you hear about us" are the strongest
- Report ranges and trends with stated methods, never false precision

CHECK YOURSELF

1. Name the four dimensions of the visibility vector and say which you can observe.
2. What is absorption, and why must you raise it with stakeholders proactively?
3. Why is branded search volume a good proxy specifically?
4. Rewrite this to be defensible: "Our AI visibility is 34%."

CHAPTER 29

Building a Measurement System

A practical, repeatable method you can run yourself

WHAT YOU'LL LEARN

- How to build a prompt set that represents real demand
- A monthly measurement routine that takes a few hours
- What to record, and what the records tell you
- When paid tools earn their cost, and when they do not

This chapter is a procedure. Follow it and you will have better AI visibility measurement than the large majority of organisations, including many with substantial budgets — because most have not built a baseline at all.

What to measure when the click disappears

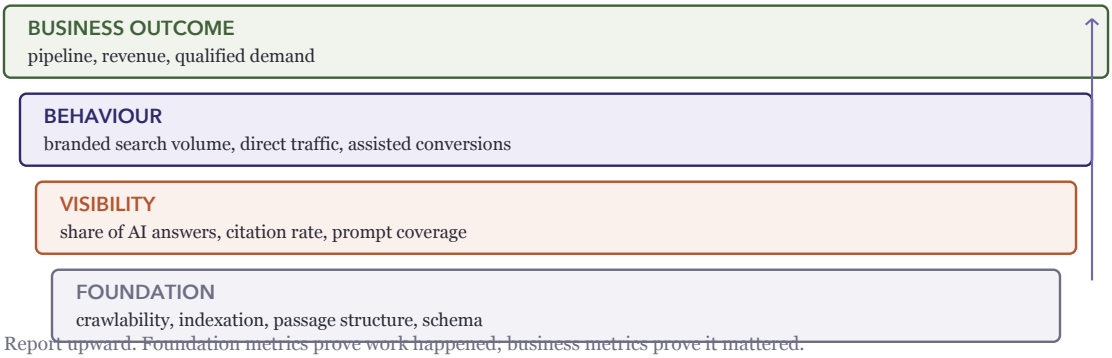


Figure 29.1 – Four layers of measurement. Report upward.

Step 1: Build the prompt set

Your prompt set is the foundation. Everything else compares against it, so it must be stable and representative.

Aim for 40-60 prompts spread across the buying journey:

Category	Share	Example
Problem-aware	25%	“why are our support response times slipping”
Solution-aware	25%	“what software helps agencies track retainers”

Category	Share	Example
Comparison	25%	"*[competitor]* alternatives for small agencies"
Branded	15%	"is *[your brand]* any good", "*[brand]* pricing"
Adjacent	10%	"how do agencies usually price retainers"

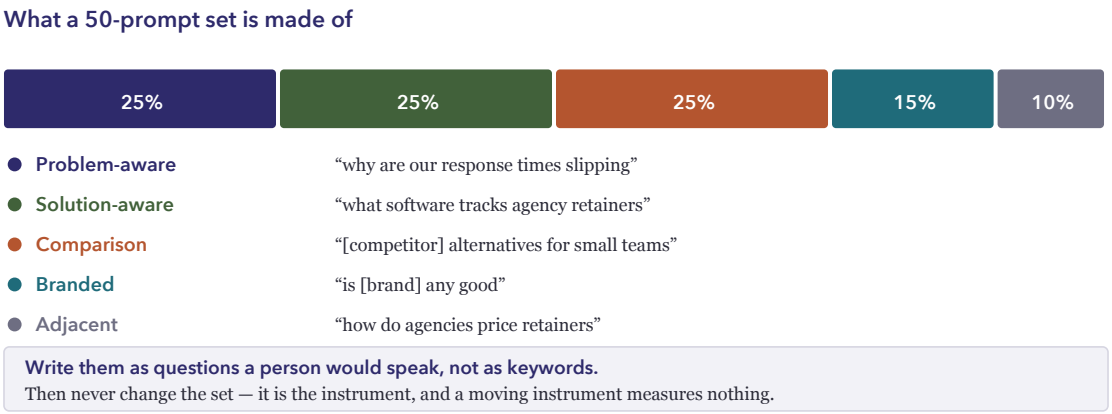


Figure 29.2 – The composition of a 50-prompt set, and why it never changes once fixed.

Sources for the prompts:

- Sales and support conversations — the actual words customers use
- Your Search Console queries, rewritten as a person would speak them
- Ask a model: "what would someone ask before buying *[category]*?"
- Competitor comparison queries
- Objection-shaped questions: "is *[your category]* worth it for small teams"

> COMMON MISTAKE
>
> Write prompts as **natural language questions**, not keywords. "CRM pricing" is a search query. "How much should a five-person agency expect to pay for a CRM?" is a prompt. The systems behave differently for each, and measuring the wrong one gives you a number that does not correspond to how anyone actually uses these tools.

NOTE

If you sell B2B

Adjust the prompt-set mix. Consumer buyers ask about the product; business buyers ask about the *decision* — risk, integration, procurement, and what happens when it goes wrong.

Add a sixth category, roughly 15% of the set, drawn from questions like:

- "is *[category]* worth it for a team of 30"
- "what goes wrong with *[category]* implementations"
- "how long does *[category]* take to implement"
- "what should be in a *[category]* contract"
- "*[competitor]* vs *[competitor]* for regulated industries"

These retrieve differently from consumer-style questions, and they are where B2B purchases are decided. They are also, in most markets, comprehensively unanswered — which is the opportunity.

Step 2: Establish the baseline

Run every prompt on your two priority engines. For each, record:

Field	What to record
Date	The run date
Engine	Which assistant
Prompt	Verbatim
Brand mentioned	Yes / no
Brand cited with link	Yes / no
Position in answer	First third / middle / last
Sentiment	Positive / neutral / negative
Competitors named	List them
Sources cited	List the domains
Accuracy	Was anything said about you wrong?

A spreadsheet is entirely sufficient. Fifty prompts across two engines is a hundred rows and takes about three hours the first time, and rather less thereafter. Chapter 59 has the column layout ready to copy, along with the monthly note and quarterly report templates.

IN PRACTICE

The rules that keep the data honest

1. **Always use a logged-out, incognito session.** Conversation memory otherwise contaminates results — the engine remembers you asked about your own brand and behaves differently.
2. **Use a consistent location.** Answers vary by region.
3. **Ask each prompt in a fresh conversation.** Prior turns influence subsequent answers.
4. **Record the answer verbatim** in a column, not just your classification. You will want to re-read these later.
5. **Run on the same day of the month.** Reduces variance from engine updates.
6. **Do not change the prompt set** without noting the version and date.

Step 3: Compute the metrics

From that spreadsheet:

- **Share of model** = prompts mentioning you ÷ total prompts
- **Citation rate** = prompts citing you with a link ÷ total prompts
- **Competitive share** = your mentions ÷ mentions of your top three competitors
- **Accuracy rate** = correct statements about you ÷ total statements about you
- **Source concentration** = which third-party domains keep appearing (this becomes your Chapter 21 target list)

The last one is quietly the most actionable output of the whole exercise. It tells you exactly which third-party pages the engines rely on for your category, which is a work list you cannot get any other way.

Step 4: The monthly routine

Roughly two to three hours per month:

1. Re-run the prompt set on both engines (90 min)
2. Update the metrics (20 min)
3. Note any answer that is newly wrong about you (15 min)
4. Note any new third-party source that has entered the citation set (15 min)
5. Write three sentences on what changed and why you think so (10 min)

That last step matters. Numbers without interpretation get ignored; a short written narrative each month is what turns this into something a stakeholder reads.

Step 4b: How many runs, and when a movement counts

One response per prompt per engine is the cheapest possible design and it is what most people run. It is also the design in which a handful of stochastic answer changes look exactly like progress.

Pick a tier and record which one you are on, because the tier determines what you are allowed to claim:

Tier	Protocol	What it supports
Solo	25 prompts × 2 engines × 2 runs	Direction over a quarter
Standard	50 prompts × 2 engines × 3 runs	Month-on-month movement, with a stated range
Scaled	Automated sampling, multiple markets and modes	Movement by segment

With more than one run per prompt, record the **repeat number** alongside everything else, then report two things instead of one: the majority result across repeats, and the spread between the best and worst run.

Establishing your noise floor. Before your first comparison, run the whole set twice on the same day with nothing changed. The difference between those two runs is your noise floor — the amount your number moves when nothing about the world has. Most sets land somewhere between four and eight percentage points, which is more than most reported "improvements".

> KEY IDEA
>
> **Do not call a one-period movement a trend until it survives repeated runs or appears across two or more measurement periods.** This is the single rule that separates a defensible measurement practice from a monthly exercise in reading noise, and it costs you nothing but a month of patience.

Step 4c: Recording the configuration, not just the engine

"ChatGPT" is not a measurement condition. The same product produces materially different citation behaviour depending on the model, the mode and the account, and a set sampled across a silent product change is not comparable with itself.

Record these fields on every run. Where the product does not reveal one, the value is `UNKNOWN` — not blank, and not a guess.

```
product:           model:             mode:
search enabled:    research mode:     account type:
```

```
logged in:          memory on:          country:
language:           device:             run timestamp (with offset):
```

The timestamp needs its offset. A "monthly" comparison in which half the runs happened in a different timezone is comparing different news cycles, and that is not a small effect in fast-moving categories.

> NOTE
>
> This is the level of detail that feels excessive until the first time a number moves and someone asks why. Then it is the difference between "the product switched to a new default model on the eleventh" and a shrug. Chapter 48 covers preserving these records; this is about capturing them in the first place.

Step 5: Supporting instrumentation

Alongside the prompt sampling:

- **Server log monitoring** for AI crawler access — confirms discoverability (Chapter 8)
- **Branded search volume** from Search Console, monthly
- **Referral traffic** from assistant domains, as a segment in analytics
- **"How did you hear about us?"** responses, tallied
- **Organic conversion rate**, watched for the pre-qualification effect

Do you need a paid tool?

There is now a substantial market: Profound, Peec AI, Otterly, Ahrefs Brand Radar, Semrush's AI toolkit, Conductor and others. They automate the sampling described above at much larger scale.

They are worth it when: you need to track hundreds of prompts, you need multi-market or multi-language coverage, you need to demonstrate rigour to a board, or the analyst time saved exceeds the licence cost.

They are not necessary when: you are one person, tracking one market, with 50 prompts. Your spreadsheet gives you the same signal, plus a much better understanding of what the numbers mean, because you read every answer.

> KEY IDEA
>
> Build the manual system first, even if you intend to buy a tool. Three months of reading answers yourself teaches you things about how your category is described that no dashboard will surface, and it means you can tell when a tool's numbers are wrong.

CASE STUDY · KETTLEWORTH

Kettleworth's prompt set ran to 48 questions, written from sales call recordings rather than keyword tools.

The difference mattered. Their keyword research said "best coffee grinder". Their customers were asking things like *"I've got a Gaggia Classic and my espresso tastes sour, is it the grinder?"* — a question no keyword tool had ever shown them, and one they had no content for at all.

They ran the set on the first Monday of each month, logged out, in fresh sessions, and recorded every answer verbatim.

The verbatim column turned out to be the most useful thing they built. Reading back six months of answers showed the shift more convincingly than any metric.

CHAPTER SUMMARY

- A fixed prompt set of 40-60 natural-language questions across the buying journey is the foundation
- Run logged-out, in fresh conversations, from a consistent location, on a consistent schedule
- Record verbatim answers, not just classifications
- Share of model, citation rate, competitive share, accuracy and source concentration are the core metrics
- Source concentration is the most actionable output; it names your third-party target list
- Build manually first; buy tools for scale, not for understanding

CHECK YOURSELF

1. Why must prompts be natural-language questions rather than keywords?
2. Why run logged-out and in fresh conversations?
3. Which metric produces a direct work list, and what is that list?
4. When does a paid AI visibility tool genuinely earn its cost?

CHAPTER 30

Attribution and Proving Value

Making a defensible business case when the data is incomplete

WHAT YOU'LL LEARN

- How to frame AI search value when clicks undercount it
- Building a business case that survives scrutiny
- Which numbers to lead with for which audience
- What to do when asked for a figure you honestly cannot produce

This chapter is about the conversation rather than the tooling, because that is where careers in this discipline are made or lost. The work in Parts III to V will produce results. Whether anyone believes it did is a separate problem, and it is solved with framing rather than data.

How an AI-driven sale gets credited to something else

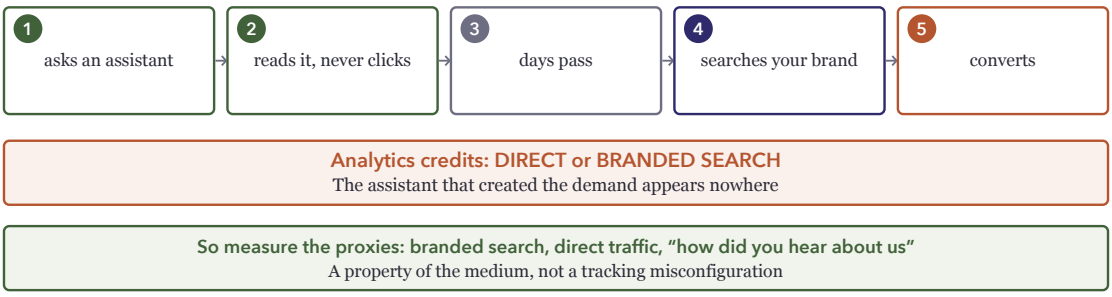


Figure 30.1 – Five steps from AI answer to conversion, and where the credit lands.

The attribution gap

The mechanics of why this is hard:

1. Someone asks ChatGPT for options and is told about you
2. They do not click; they read the answer
3. Days later they search your brand name directly
4. They arrive as **direct** or **branded organic** traffic
5. They convert

Every analytics platform attributes that conversion to direct or branded search. The AI answer that created the demand is invisible. This is not a tracking configuration problem; the information genuinely does not reach you.

> KEY IDEA
>
> AI search behaves like a **demand-generation channel** — it creates awareness that converts later through other paths — while being measured with **demand-capture tooling** built to track last clicks. The mismatch is structural, and the framing you use should acknowledge it rather than fight it.

Building the case

The argument that survives scrutiny has four parts.

1. Exposure

Establish that your audience is using these tools. Use your own "how did you hear about us" data and your sales team's observations, not an industry statistic. Internal evidence is far more persuasive than a report from a vendor.

2. Visibility movement

Show share of model rising against a fixed prompt set, with the method stated. This is the direct evidence that your work changed something.

3. Proxy corroboration

Show branded search volume and direct traffic moving in the same direction over the same period. Two independent measures moving together is much stronger than either alone.

4. Honest limitation

State plainly what you cannot prove. Doing this *before* being challenged converts a vulnerability into a demonstration of rigour.

IN PRACTICE

The quarterly report structure

One page. Numbers in the order a stakeholder cares about them.

What changed commercially Organic conversions, branded search volume, direct traffic — with period comparisons.

What we did Three to five specific pieces of work, named.

What moved in AI visibility Share of model before and after, method stated in one line. Two example answers, quoted verbatim — one where you now appear and previously did not.

What we cannot measure One short paragraph on absorption and the attribution gap. Same wording each quarter.

What we are doing next, and why Three priorities tied to the gaps the measurement found.

The verbatim answer quotes do disproportionate work. A stakeholder who reads ChatGPT recommending their company, when it did not three months ago, understands the value immediately in a way no chart achieves.

Leading with the right number

Audience	Lead with	Because
CEO / founder	Branded search and direct traffic growth	Proxies for demand, which is what they think about
CFO	Organic conversion rate and cost per acquisition	Efficiency, not volume
Sales leader	"How did you hear about us" responses	Their team hears it first-hand
Marketing peers	Share of model, competitive share	Comparable, competitive framing
Board	One trend line plus one verbatim answer	Time-constrained; needs the gist

When you are asked for a number you do not have

You will be asked "how much traffic is AI sending us?" and the honest answer is that nobody can fully know. What works:

> *"Referral data shows about 400 sessions last month, but that materially undercounts it — most assistants don't pass a referrer, and most people don't click through at all; they read the answer and search for us later.*

>

> *The better indicators are that branded search is up 22% and 14 people told us on forms this quarter that they found us through an AI assistant. I'd rather give you those two real numbers than one fake one."*

That answer is honest, specific, and demonstrates you understand the medium. It is a considerably better position than producing a confident figure that collapses under a follow-up question.

COMMON MISTAKE

Do not construct an "AI-influenced revenue" figure by applying assumed multipliers to referral data. It will be challenged, the assumptions will not survive, and you will lose credibility on the parts of your reporting that were sound. Report what you measured and label the rest as unmeasured.

CASE STUDY · KETTLEWORTH

Kettleworth's quarterly report was one page.

It led with commercial numbers: organic conversions up 22%, branded search up 31%, direct traffic up 19%. Then three named pieces of work. Then share of model on the UK prompt set — 48 fixed questions, ChatGPT and Perplexity, logged out, run monthly — moving from 14% at baseline to 29%, with that method stated in a single line beneath it.

Then the part that landed. Two verbatim answers to the same question, three months apart. In the first, ChatGPT recommended three competitors. In the second, it recommended two competitors and Kettleworth, citing their grind-consistency study.

The founder read the two quotes and approved the following year's budget. He did not ask about the chart.

CHAPTER SUMMARY

- AI search creates demand that converts through direct and branded paths, making last-click attribution structurally blind to it
- Build the case from exposure, visibility movement, proxy corroboration and honest limitation
- Verbatim answer quotes persuade stakeholders more effectively than any chart
- Lead with different metrics for different audiences
- Never fabricate an influenced-revenue figure; state what is unmeasured

CHECK YOURSELF

1. Trace the five steps by which an AI-driven conversion gets attributed to direct traffic.
2. Why is stating your limitations before being challenged a strength?
3. Why do verbatim answer quotes work so well with senior stakeholders?
4. Draft your answer to “how much revenue has AI search generated for us?”

CHAPTER 31

Testing and Experiment Design

Establishing that your work caused the change

WHAT YOU'LL LEARN

- Why the usual A/B test does not work in AI search, and what replaces it
- How to design a change you can actually attribute
- The confounds that will ruin your result, and how to bound them
- What sample size means when the population is a prompt set

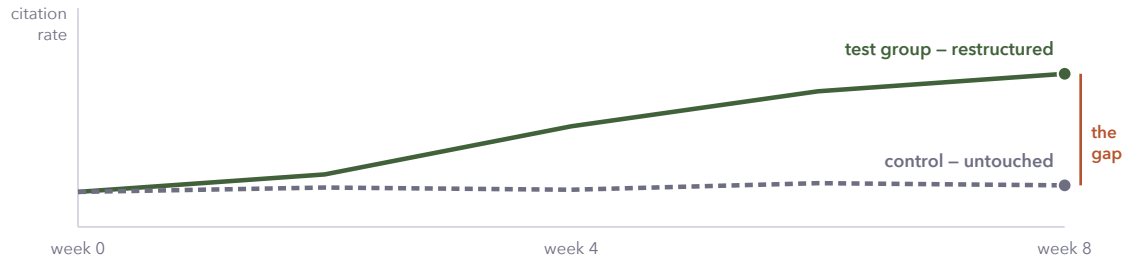
Chapter 7 tells you to distrust industry claims and measure your own results. This chapter is how to measure them in a way that survives a sceptical question — because "we changed some pages and share of model went up" is not evidence, and a stakeholder who has been burned once will say so.

Why you cannot A/B test AI search

Classical CRO splits traffic: half see variant A, half see B, and the difference is the effect. AI search offers no such split. There is one index, one set of engines, and one version of your page. You cannot show variant B to half the retrieval systems.

What you can do instead is **quasi-experimental**: change one thing for one group of pages, leave a comparable group alone, and compare the trajectories.

The control group is what makes it evidence



If both lines rise together, something else moved them — and you have learned that too.

Write the prediction, and what would falsify it, before week 0.

Figure 31.1 – The test group moves, the control group does not. The gap between them is your evidence.

KEY IDEA

The control group is what turns an anecdote into evidence. Without one you cannot separate your work from a core update, a seasonal shift, or a competitor's campaign — and those will move your numbers more than your restructuring does.

Designing a change you can attribute

1. Pick a coherent test group

Twenty to fifty pages that are similar to each other: same template, same topic area, similar current performance. Similar matters more than large.

2. Pick a control group that matches it

Another twenty to fifty pages, as similar as you can make them, that you will **not** touch for the duration. This is the discipline people abandon first, usually because the control pages are commercially important and leaving them alone feels wasteful. Leave them alone anyway — for eight weeks.

3. Change exactly one thing

Restructure answer-first, *or* add schema, *or* rewrite headings as questions. Not all three. If you change three things and the number moves, you have learned that something worked.

4. Write the prediction down first

Before you start, write it down: *"I expect share of model on the 15 prompts covering these pages to rise from X to Y within eight weeks."* Recording the prediction stops the result being reinterpreted afterwards. That reinterpretation is the commonest way marketing tests deceive the people running them.

5. Wait eight weeks

Four at absolute minimum. Re-crawling, re-indexing and re-evaluation all take time, and a two-week read is noise.

IN PRACTICE

A worked test design

Hypothesis: rewriting headings as questions increases citation rate.

- **Test group:** 24 guide pages in the espresso cluster
- **Control group:** 24 guide pages in the brewing cluster — same template, similar traffic, untouched
- **Change:** headings rewritten as questions; first sentence under each answers it. Nothing else — no schema, no new content, no internal linking changes.
- **Metric:** citation rate on the 18 prompts that map to those two clusters, measured monthly, logged out, both engines
- **Prediction, recorded before starting:** test group citation rate rises from 11% to at least 18%; control stays within 3 points of 11%
- **Duration:** eight weeks
- **What would falsify it:** both groups rise together (something else caused it), or the test group rises less than 3 points (the change did not do much)

That last line is the one that makes it a test rather than a demonstration.

The confounds that will ruin it

Confound	What it does	How to bound it
Core update	Moves everything at once	Note update dates; control group absorbs it
Seasonality	Category demand shifts	Compare against the same period last year
Competitor activity	Their gain is your loss	Track competitor mentions in the same prompts
Engine change	Retrieval behaviour shifts	Track both engines; a single-engine move is suspicious
Your own other work	PR, links, new content	Freeze other activity on both groups
Sampling noise	Answers vary between runs	Larger prompt set; report trends not points

COMMON MISTAKE

The freeze in the last-but-one row is the hardest to hold and the most often broken. If a PR campaign lands mid-test and mentions your test pages, the test is over — record that honestly and start again rather than reporting the result as if the campaign had not happened.

Sample size, when the population is a prompt set

You are not sampling users; you are sampling questions. That changes the arithmetic in a useful way.

A prompt set of 50 with a citation rate around 20% has roughly ±11 points of noise at 95% confidence — which means a move from 20% to 26% is **not** a detectable result. To see a six-point change reliably you need somewhere near 150-200 prompts, or you need to accept that you are reading a direction rather than a measurement.

Two practical consequences:

1. **Small prompt sets can only detect large effects.** With 50 prompts, look for changes of 15 points or more, and describe anything smaller as suggestive.
2. **Consistency beats precision.** The same 50 prompts run monthly for a year produce a trend line that is far more informative than any single comparison, because the noise is stable while the signal accumulates.

MYTH vs REALITY

"We restructured and traffic went up 30%, so restructuring works"

No control group, no recorded prediction, no isolated variable, and no stated period. This describes something that happened; it does not establish what caused it.

Say instead: *"We restructured 24 pages. Their citation rate rose from 11% to 21% while an untouched comparison set moved from 11% to 12% over the same eight weeks. No core update landed in the window. That is reasonable evidence the restructuring did it."*

The second version is longer, and it is the one that survives being questioned.

CASE STUDY · KETTLEWORTH

Kettleworth's first test was a failure worth reporting.

They restructured 20 grinder pages and predicted citation rate would rise from 9% to 18% in eight weeks. It rose to 19%. They were about to declare it when someone checked the control group: their brewing pages, untouched, had gone from 10% to 17%.

Something else had moved both. Checking the dates, an engine update had landed in week three.

The test told them nothing about restructuring — but it did tell them their measurement system worked, and it stopped them attributing a platform-wide shift to their own work in a board report. They re-ran the test the following quarter with the same design. That time the test group rose 12 points and the control moved 2.

The second result is the one they reported, and it was believed, because they had already shown they would report the first.

CHAPTER SUMMARY

- AI search cannot be A/B tested; use a quasi-experiment with a matched, untouched control group
- Change exactly one thing, and write the prediction and the falsifier down before you start
- Eight weeks minimum; four is the floor and two is noise
- Six confounds will ruin a test — the hardest to hold is freezing your other activity
- A 50-prompt set only detects large effects; consistency over a year beats precision in any single comparison

CHECK YOURSELF

1. Why can you not A/B test AI search, and what replaces it?
2. What does the control group actually protect you from? Name three things.
3. Why record the prediction and the falsifier before starting?
4. With a 50-prompt set, a metric moves from 20% to 26%. What can you honestly say?

PART VIII

Practice

The toolkit, the audit, the plan and the working week. Turning all of it into something you could start on Monday.

CHAPTER 32

The Audit Playbook

A complete diagnostic, in the order that finds problems fastest

WHAT YOU'LL LEARN

- A full AI-search audit you can run on any site
- The order that surfaces the biggest problems first
- How to turn findings into a prioritised list rather than a list of complaints
- What a finished audit document should contain

This chapter is a procedure to work through. It is ordered deliberately: each stage is cheaper than the one after it, and an early failure makes the later stages irrelevant. Do not reorder it — checking content quality before checking whether crawlers can reach the site wastes days.

Stage 0: Establish the baseline (half a day)

Before diagnosing anything, capture the current state. Without this you cannot demonstrate improvement.

1. Build the prompt set (Chapter 29)
2. Run it on your two priority engines and record everything
3. Note branded search volume and direct traffic for the last 12 months
4. Screenshot three current answers about your brand

Stage 1: Access (one hour, catches the most catastrophic problems)

IN PRACTICE

The access checklist

- [ ] `robots.txt` loads and contains no accidental blanket `Disallow: /`
- [ ] `OAI-SearchBot`, `Claude-SearchBot` and `PerplexityBot` are not blocked
- [ ] You know whether `GPTBot` / `ClaudeBot` are blocked and that this was deliberate
- [ ] `Google-Extended` status is understood (it does **not** affect AI Overviews)
- [ ] CDN / WAF bot management checked for AI-crawler blocking — **ask explicitly**
- [ ] Server logs show AI crawlers visiting in the last 30 days
- [ ] Those visits return 200, not 403 or 429
- [ ] Sitemap exists, is referenced in `robots.txt`, and loads
- [ ] Key pages report as indexed in Search Console

If anything fails here, stop and fix it. Nothing downstream matters while the crawlers cannot get in. This stage alone resolves a surprising share of "we're invisible in AI" problems.

Stage 2: Readability (one hour)

- [ ] View Page Source on a commercial page — is the main content in the HTML?
- [ ] Is the **price** in the HTML?
- [ ] Are **reviews** in the HTML, or in a JavaScript widget?
- [ ] View Page Source on a blog post — is the article text present?
- [ ] Is anything important inside an image, an iframe or behind a click?
- [ ] Does a consent banner block content for a non-consenting visitor?

Stage 3: Structure (half a day)

For your twenty most important pages:

- [ ] Does each `H2` read as a question a person would ask?
- [ ] Does the first sentence under each heading answer it directly?
- [ ] Does each section stand alone without "as we saw above"?
- [ ] Are sections roughly 130-200 words where the content allows?
- [ ] Are vague claims replaced with specifics?
- [ ] Is there a comparison table where the content warrants one?
- [ ] Are published and updated dates visible?
- [ ] Is there a named author with a profile page?

Stage 4: Entity and trust (two hours)

- [ ] `Organization` schema present, with a populated `sameAs` array
- [ ] Company name, description and founding facts identical across site, LinkedIn, Crunchbase and directories
- [ ] Author pages exist, with credentials and `Person` schema
- [ ] Physical address and phone number present
- [ ] Ask four engines "what does *[brand]* do?" — is the answer right?
- [ ] Search your brand name and read two pages of results — what are you associated with?

Stage 5: Content coverage (half a day)

- [ ] Generate the fan-out for your five most important questions (Chapter 4)
- [ ] Map each sub-question to a page that answers it
- [ ] List the sub-questions with no page — these are your gaps
- [ ] Run the topical authority audit (Chapter 14) — commit or cut on tail topics
- [ ] Run the decay audit (Chapter 17) — classify declining pages

Stage 6: Off-site (two hours)

- [ ] From your baseline, list every third-party domain the engines cite for your category
- [ ] Check your presence and accuracy on each
- [ ] Identify the two platforms most relevant to your business type
- [ ] Note where competitors appear and you do not

Migrations and replatforming

Sooner or later someone will propose a redesign, a CMS change or a domain move. These are the highest-risk events in this discipline, because they can undo years of accumulated signal in an afternoon.

The non-negotiables, in order:

1. **Crawl and record the current state before anything changes.** Every URL, its title, its status code, its structured data. This is the only way to prove what was lost.
2. **Map every old URL to a new one.** One to one, no bulk redirects to the homepage. Anything without a destination is a decision, not an oversight.
3. **301, never 302.** A temporary redirect tells engines to keep the old URL, which is precisely wrong.
4. **Preserve the content, not just the URL.** A redirect to a page that no longer contains the answer loses the passage even though the link resolves.
5. **Carry the structured data across.** It is the most commonly dropped item in a replatform.

6. **Re-check crawler access on the new platform.** New infrastructure means new CDN defaults — the Chapter 8 failure mode, freshly reintroduced.
7. **Re-run the prompt set the week before and four weeks after.** This is the only measurement that will tell you what the migration cost.

> COMMON MISTAKE
>
> Expect a dip. Even a well-executed migration usually costs visibility for four to eight weeks while the new URLs are re-crawled and re-evaluated. Say so in advance and in writing. A predicted dip is a managed risk; an unpredicted one is a crisis with your name on it.

Turning findings into priorities

An audit that lists forty problems achieves nothing. Score each finding on two axes — impact and effort, and work the top-left quadrant first.



The four in the top-left were three days of work and most of the problem.

Figure 32.1 – Scoring findings by impact and effort, with Kettleworth's audit as the example.

From finding to verified

A priority order tells you what to do next. It says nothing about whether it worked, and an audit that cannot close its own rows becomes a document people stop opening.

Give every action seven fields:

Field	Example
Finding	OAI-SearchBot receives 403
Mechanism	The crawler cannot retrieve any page

Field	Example
Action	Change the CDN bot rule
Owner	Infrastructure
Definition of done	That user agent receives HTTP 200
Immediate verification	A log line showing the 200
Measurement delay	Two to eight weeks before answers move

Two of those do work the others cannot. **Definition of done** stops "we looked at it" counting as completion. **Immediate verification** separates the thing you can confirm today from the thing you have to wait for — which is what keeps a stakeholder patient through the eight weeks in between.

A few more, in the same shape:

- *Render reviews server-side* › done when review sentences appear in View Page Source.
- *Add Organization schema* › done when the Rich Results Test passes **and** the markup matches the visible facts.
- *Rewrite a passage* › done when it survives the isolation test, checked before publication rather than after.
- *Correct a third-party comparison page* › done when the page visibly changes, then rechecked after recrawl.
- *Publish original research* › done when the method and the headline findings are reachable without a form.

> NOTE
>
> Add two more fields for anything touching infrastructure: **failure response** (what you check if the fix does not take — rate limits, a second rule, an IP block) and **rollback** (how to restore the previous state). Both are cheap to write down in advance and expensive to work out at the point you need them.

IN PRACTICE

The audit document

Six sections, no more:

1. **Executive summary** — three sentences: current state, biggest problem, what it will take
2. **Baseline** — share of model, cited/not, one verbatim answer
3. **Critical findings** — anything failing at Stage 1 or 2, with impact stated in plain terms
4. **Prioritised actions** — the matrix above, with named owners and dates
5. **What we will measure** — the metrics and the cadence
6. **Appendix** — full checklists and raw data

Chapter 58 holds every checklist in this book in one place, and Chapter 59 the document templates, including the audit write-up itself.

Write section 1 last, and write it for someone who will read only that.

CASE STUDY · KETTLEWORTH

Kettleworth's audit produced 40 findings. Presented as a list, it would have achieved nothing.

Scored on impact and effort, the picture was clear. Four items sat in high-impact, low-effort: the Cloudflare block, the missing `Organization` schema, the four wrong third-party listings, and the price rendering. Together they represented perhaps three days of work and addressed the majority of the problem.

The restructuring programme and the original research went into "plan this quarter". Eleven findings went into "do not do", including a proposed replatform that had been circulating for a year.

They did the four things first. Everything else got easier afterwards.

CHAPTER SUMMARY

- Audit in cost order: access, readability, structure, entity, coverage, off-site
- A Stage 1 failure makes everything downstream irrelevant — check CDN bot rules explicitly
- The most common catastrophic finding is infrastructure blocking AI crawlers without marketing's knowledge
- Score findings by impact and effort; act on high-impact, low-effort first
- The audit document is six sections and the summary is written last

CHECK YOURSELF

1. Why is the audit ordered this way, and what happens if you reorder it?
2. Which single check catches the most catastrophic and most common problem?
3. What belongs in “do not do”, and why is naming that explicitly useful?
4. Why write the executive summary last?

CHAPTER 33

The GEO Toolkit

What to actually use, what it tells you, and what to ignore

WHAT YOU'LL LEARN

- The minimum viable stack — what you can do with free tools alone
- What each category of paid tool genuinely adds, and when it starts to pay
- How to assemble a working stack rather than a collection of subscriptions
- What to automate, what to keep manual, and why the distinction matters

A traditional marketer arriving in this discipline usually asks "what tools do I need?" before "what am I trying to find out?", which is the wrong order and produces expensive stacks that answer nothing. This chapter goes the other way: here is what you need to know, and here is the cheapest thing that tells you.

The stack by tier – almost everything decisive is free

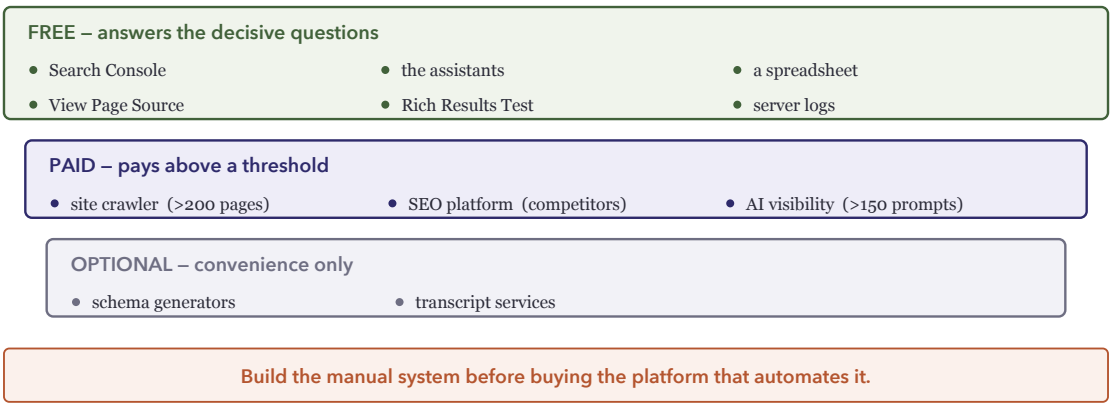


Figure 33.1 – The stack by tier. Almost everything decisive sits in the free layer.

The free tier does most of the work

You can run everything in Chapters 29 and 30 with these:

Tool	What it answers	Cost
Google Search Console	What you rank for, what is indexed, crawl behaviour, branded search volume	Free

Tool	What it answers	Cost
The assistants themselves	What is actually said about you — the prompt set	Free tier adequate
A spreadsheet	The tracker, the baseline, every metric in Chapter 29	Free
View Page Source	Whether a crawler can see your content (Chapter 9)	Free
Rich Results Test	Whether your structured data parses	Free
Your server logs	Which AI crawlers actually visited, and what they got	Free — ask hosting
Google Analytics 4	Landing-page behaviour, referral segments, conversion	Free

KEY IDEA

The two most decisive diagnostics in this book — whether AI crawlers are being blocked, and what assistants actually say about you — both cost nothing. Anyone telling you that GEO requires a platform before you can start is selling something.

Search Console, specifically

It is the most under-used tool most marketers already have. Four reports carry nearly all the value:

- **Performance** — queries, impressions, clicks. Filter to branded queries for the proxy metric from Chapter 28.
- **Pages / Indexing** — which URLs are indexed and which were crawled and rejected. "Crawled — currently not indexed" is the engine telling you it read your page and declined.
- **Crawl stats** (under Settings) — response codes, response time, and what Googlebot spent its budget on. This is where crawl waste shows up.
- **Enhancements** — structured data errors, site-wide.

What paid tools add

Crawlers — Screaming Frog, Sitebulb

Fetch your whole site the way a bot does and list what is broken: missing alt text, thin pages, redirect chains, orphan pages, missing schema. **Worth it above roughly 200 pages**, where manual checking stops being feasible. Screaming Frog is free to 500 URLs, which covers many sites entirely.

SEO platforms — Ahrefs, Semrush

Backlinks, competitor visibility, keyword data, rank tracking, and increasingly AI-answer tracking. **Worth it when you need competitive intelligence** you cannot get from your own data. If you only need your own performance, Search Console tells you more, more accurately.

AI visibility platforms — Profound, Peec, Otterly, Brand Radar, Conductor

Automate the prompt sampling from Chapter 29 at a scale a person cannot: hundreds of prompts, multiple engines, multiple markets, daily. **Worth it when the analyst time saved exceeds the licence**, which is usually somewhere above 150 tracked prompts or two markets.

> COMMON MISTAKE
>
> Build the manual system first, even if you intend to buy. Three months of reading answers yourself teaches you what your category sounds like — and, more usefully, it means you can tell when a dashboard is wrong. Teams that start with a platform learn the numbers without learning the market, and cannot sanity-check either.

Content and structured data helpers

Schema generators, hreflang validators, transcript services. Individually minor, collectively a real time saving. None of them decide anything.

Assembling a stack rather than a subscription pile

Work backwards from the question:

Question you need answered	Cheapest thing that answers it
Can crawlers reach us?	Server logs, `robots.txt`, CDN settings — free
Can they read the page?	View Page Source — free
Are we indexed?	Search Console — free
What do assistants say?	Ask them, logged out — free
Which third parties do they cite?	Your own baseline notes — free
Is our schema valid?	Rich Results Test — free
What is broken across 5,000 pages?	A crawler — paid above 500 URLs
How are competitors doing?	An SEO platform — paid
What happens across 300 prompts daily?	An AI visibility platform — paid

Anything that does not map to a question you are actually asking is not part of your stack.

Automation: what to and what not to

Automate the repetitive and verifiable: running the prompt set, pulling Search Console data into a sheet, crawling for broken structured data, flagging pages whose traffic dropped, generating schema from a template.

Keep manual: reading the answers, judging whether a claim is true, deciding what a finding means, choosing what to write. Chapter 34's division holds here — transformation can be automated, judgement cannot.

IN PRACTICE

A first automation, with no coding

The single highest-return automation is getting your prompt-set results into a sheet without retyping them.

1. Create the tracker with the columns from Chapter 59.
2. Add a column called **Answer verbatim**.
3. As you run each prompt, paste the full answer into that column rather than summarising it.
4. Use your spreadsheet's search to count mentions across the month: a formula counting cells containing your brand name gives you share of model without manual tallying.

That is not glamorous, and it removes the step people abandon the tracker over. If you later learn enough Python or spreadsheet scripting to call an engine's API directly, the structure is already right — you are only replacing the paste.

What an API is, and whether you need one

An **API** is a way for one program to ask another program for something without a person clicking. Most assistants and SEO platforms have one, which means your prompt set could be run automatically overnight rather than by hand.

You do not need this to do the work in this book. It becomes worth learning when you are running more than about a hundred prompts a month, at which point the manual method costs more hours than the learning does. It is also the single most useful technical skill a marketer can add in this field, because it converts a repetitive task into a scheduled one.

CASE STUDY · KETTLEWORTH

Kettleworth's stack, at the point their programme was working, cost £0 for the first four months.

Search Console for indexing and branded search. A spreadsheet for the 48-prompt tracker. View Page Source for the rendering audit that found their invisible prices. Their hosting dashboard for the log check that found the Cloudflare block. Free-tier assistants for the monthly run.

They added Screaming Frog in month five, when the catalogue audit outgrew manual checking. They looked at an AI visibility platform in month eight and did not buy: at 48 prompts and one market, the licence would have cost more than the two hours a month it replaced.

They bought it a year later, when the German set took them to 78 prompts across two markets and the manual run stopped fitting into a morning. By then they knew exactly what to ask it for.

CHAPTER SUMMARY

- The free tier answers almost every decisive question, including the two that matter most
- Search Console's four key reports are the most under-used asset most marketers already own
- Crawlers pay above ~200 pages; SEO platforms pay for competitive intelligence; AI visibility platforms pay above ~150 prompts or two markets
- Build the manual system before buying a platform, or you will not be able to tell when it is wrong
- Automate transformation and collection; keep reading, judging and deciding manual

CHECK YOURSELF

1. Which two diagnostics in this book are the most decisive, and what do they cost?
2. What does "Crawled — currently not indexed" in Search Console actually mean?
3. At what point does an AI visibility platform start to pay for itself?
4. Why build the manual tracker before buying a tool that automates it?

CHAPTER 34

Using AI to Do the Work

Where models genuinely help, where they quietly damage you

WHAT YOU'LL LEARN

- The tasks where AI assistance is straightforwardly good
- The tasks where it produces convincing, useless output
- How to use models without producing the generic content that fails retrieval
- A working process for AI-assisted content that keeps the parts that matter human

There is an irony worth naming: the same technology creating this discipline is also the main threat to doing it well. Models make it trivially easy to produce content that is fluent, plausible, and exactly the kind of undifferentiated material that retrieval systems have no reason to cite.

The dividing line is whether the average is the right answer

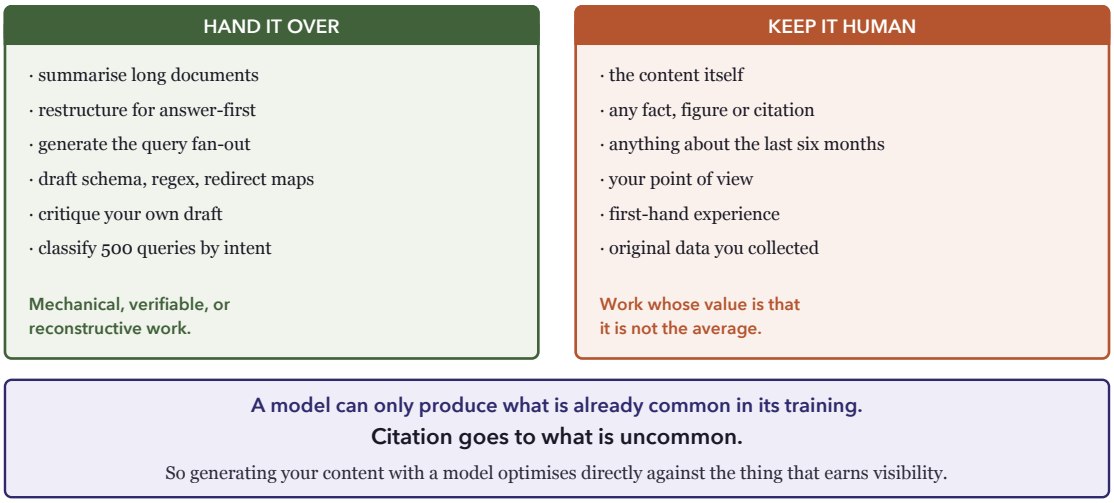


Figure 34.1 – What to hand a model, what to keep, and the single rule that sorts them.

Where models genuinely help

Research and synthesis

Summarising long documents, extracting themes from customer interviews, comparing sources. Reliable, and a real time saving.

Structure and reformatting

"Restructure this section so the answer comes first" is a task models do well and quickly. Applying Chapter 13's pattern across fifty pages is exactly the kind of mechanical transformation worth automating.

Generating the fan-out

Asking a model what sub-questions it would research is not a workaround — it is asking the same class of system that performs the decomposition. This is one of the most legitimate uses in the whole book, provided you are clear that it produces a *plausible* decomposition rather than the one some other assistant used. Plan from it; do not report from it.

First-draft technical work

Schema markup, regex, spreadsheet formulas, redirect maps. Verify the output, but the drafting is faster.

Critique

"What is missing from this article?" and "what would a sceptical expert say about this claim?" produce useful, honest answers. Models are better critics than authors.

Scale classification

Sorting 500 queries by intent, tagging 200 pages by topic. Tedious, mechanical, and well within capability.

Where they quietly damage you

Producing the content itself

A model writing from its training produces the *average* of what has been written on a topic. Average is precisely what does not get cited. Everything in Chapters 13 to 16 — specificity, first-hand experience, original data, genuine expertise — is by definition absent from an average.

> KEY IDEA
>
> A model can only produce what is already common in its training. **Citation goes to what is uncommon.** Using a model to generate your content is optimising directly against the thing that earns you visibility.

Facts and figures

Models confabulate statistics with total confidence, including plausible-looking sources that do not exist. Every number must be verified against a primary source. This is not optional and it is where most AI-assisted content quietly fails.

Anything requiring current information

Training cutoffs mean a model's unaided knowledge is months old. In a field moving as fast as this one, that is disqualifying.

Your point of view

Models produce balanced, hedged, consensus positions. A distinctive argument is a competitive asset, and it will not come from a model.

> COMMON MISTAKE
>
> The failure mode is not that AI content gets penalised — Google has been explicit that it judges quality, not production method. The failure mode is that **AI-generated content is generic, and generic content does not get retrieved**. You will not be punished. You will simply be ignored, which produces the same outcome more quietly.

Detection, disclosure and the quality bar

Three questions come up constantly and are worth settling.

Will AI-generated content be detected and penalised?

Google has been explicit that it judges content by quality, not by how it was produced. There is no penalty for AI assistance as such. Detection tools exist, and they are unreliable in both directions — they flag careful human writing as machine-made and miss heavily edited machine writing. Do not make decisions based on their output, yours or a competitor's.

The real risk is the one from earlier in this chapter: not detection, but irrelevance. Average content is not penalised. It is simply never the best passage for anything.

Should you disclose AI assistance?

It depends where you operate and what you publish. Some jurisdictions are moving towards disclosure requirements for synthetic content, and some sectors — health, finance, legal — hold a higher bar regardless of law. A defensible default: disclose where the reader's decision could be affected by how the content was made, keep an internal record of human involvement for everything, and follow your organisation's policy where one exists.

What is the quality bar, concretely?

Not "does this read well" — models clear that easily. Use these instead:

Test	Fails if
Uncommonness	A competent writer with no access to your organisation could have produced it
Verifiability	Any figure cannot be traced to a primary source you have actually opened
Experience	Nothing in it required having done the thing
Position	It hedges everywhere and commits nowhere
Retrievability	No section answers a specific question when read alone

KEY IDEA

A piece that passes all five is worth publishing whoever or whatever drafted it. A piece that fails two or more is not worth publishing even if a person wrote every word. The production method is not the question; these five are.

A working process

The division that works: models handle transformation, humans supply substance.

IN PRACTICE

The AI-assisted content workflow

1. **Human: choose the topic and the angle.** What do you know that others do not? This cannot be delegated.
2. **AI: generate the fan-out.** "What 12 sub-questions would you research to answer this?"
3. **Human: fill the gaps only you can.** Your data, your client examples, your failure modes, your numbers.
4. **AI: draft the connective structure.** Transitions, definitions of standard concepts, background sections.
5. **Human: write every section containing experience or judgement.** Non-negotiable.
6. **AI: restructure for retrieval.** "Rewrite each section so the first sentence answers the heading."
7. **Human: verify every fact.** Every statistic against a primary source. Every claim about a product against the product.
8. **AI: critique.** "What would an expert say is missing or wrong here?"
9. **Human: address the critique and add the point of view.**
10. **Human: read each section in isolation.** Apply the Chapter 13 test.

Steps 1, 3, 5, 7, 9 and 10 are human. That is the ratio that produces content worth citing.

Prompting well enough

Three techniques cover most of what you need:

- **Give it a role and an audience.** "You are a technical SEO reviewing this for a non-technical marketing director" produces materially better output than an unframed request.
- **Give it an example of what good looks like.** One or two examples of the output format you want beats any amount of description.

- **Ask it to critique before it produces.** "Before drafting, tell me what information you'd need to do this well." This surfaces the gaps you need to fill yourself.

CASE STUDY · KETTLEWORTH

Kettleworth used models heavily and wrote nothing with them.

The fan-out generation, the heading rewrites across 60 pages, the first-draft schema, the classification of 340 pages by topic — all delegated, all mechanical, all faster.

The grind-consistency study, the workshop observations, the admission about the returned batch — all human, because no model had access to any of it.

They did try generating three guides. The output was fluent, plausible and indistinguishable from every other guide on the subject. None was ever cited. They were not penalised; they were simply never the best passage for anything, because nothing in them was uncommon.

CHAPTER SUMMARY

- Models are good at research, restructuring, fan-out generation, technical drafting, critique and classification
- They are bad at producing substance, generating facts, anything current, and having a point of view
- Model output is the average of the corpus; citation goes to the uncommon
- The failure mode is being ignored, not penalised
- Humans own topic, experience, verification and judgement; models own transformation

CHECK YOURSELF

1. Why does AI-generated content fail specifically at getting cited?
2. Which six steps of the workflow must stay human, and why each?
3. Why is asking a model to generate the fan-out particularly legitimate?
4. What does "you will not be punished, you will be ignored" mean practically?

CHAPTER 35

The Prompt Library

Prompts that do real work, organised by task

WHAT YOU'LL LEARN

- A working set of prompts for research, auditing, content and measurement
- How to adapt them to your own context
- Why prompts should be saved and versioned rather than improvised

These are written to be used. Replace anything in square brackets. Save the ones that work in a document you keep — improvising the same prompt badly every time is the most common waste of effort in AI-assisted marketing.

Research and planning

Generate the query fan-out

```
You are a search retrieval system. For the question below, list the 12
sub-queries you would run to gather information before answering.

Give only the sub-queries, one per line, phrased as a person would
type them.

Question: [your target question]
```

Why it is shaped this way. Assigning the model a role — *you are a search retrieval system* — makes it reason about how a machine decomposes a question rather than how a person would rephrase it, which is the thing you actually need. Asking for the sub-queries *only* stops it wrapping them in commentary you then have to strip out. Chapter 4 explains what fan-out is; this makes it visible for your own topic.

Build a prompt set for measurement

```
I sell [product/service] to [audience]. They typically [context about
their situation].

Generate 40 questions this audience would genuinely ask an AI assistant
while researching this purchase. Spread them across:
- 10 problem-aware (they have the problem, don't know solutions exist)
- 10 solution-aware (they know the category, comparing approaches)
- 10 comparison (evaluating specific options)
```

```
- 5 branded (about [my brand] specifically)
- 5 adjacent (about the wider topic)

Write them as full natural-language questions, not keywords.
```

Why it is shaped this way. Left alone, models write keyword-shaped queries because that is what most training text about SEO looks like. Naming the buying stages forces coverage across the journey instead of a cluster of near-duplicates, and asking for natural speech is what makes the set measure the systems people actually use (Chapter 29).

Find content gaps

```
Here are the H2 headings from my page on [topic]:
[paste headings]

Here are the 12 sub-questions a retrieval system would generate for
"[target question]":
[paste fan-out]

Which sub-questions does my page fail to answer in a self-contained
section? List only the gaps, ranked by how commercially important
they seem for [my audience].
```

Why it is shaped this way. Ranking by commercial importance rather than volume is the whole point. A model will happily list forty gaps; the constraint turns a list into a decision. Note that it is being asked to judge *your* audience, so the more specific that bracket, the more useful the ranking.

Auditing and improving content

The retrieval test

```
Below is one section from a web page. Read it in complete isolation,
as if you had never seen the rest of the page.

1. What question does it answer?
2. Would the answer make sense to someone who saw only this?
3. What context does it assume that isn't present?
4. Rewrite it so the first sentence answers the question directly and
   nothing depends on surrounding content. Keep all the nuance.

Section:
[paste]
```

Why it is shaped this way. *In complete isolation* is doing the work. It reproduces the condition a retrieved passage is actually judged under — no surrounding page, no preceding paragraph, no navigation (Chapter 13). Question 3 is the one that finds real defects: assumed context is invisible to the person who wrote it and fatal to a lifted chunk.

Find vagueness

```
Review the text below. List every claim that is vague, unquantified or
could be said by any competitor.

For each, show:
- the vague phrase
- what specific information would replace it
- the question I'd need to answer internally to supply that

Do not rewrite. Just identify.

Text:
[paste]
```

Why it is shaped this way. Vagueness is easy to feel and hard to see in your own writing. Asking for specific unsupported claims rather than a general critique produces a list you can act on line by line, and it surfaces exactly the sentences that make a passage unciteable — a claim with no number, date or source attached is one an engine has no reason to prefer over anyone else's.

Rewrite headings as questions

```
Below are the headings from my page. Rewrite each as the natural
question a person would ask an AI assistant that this section should
answer. Keep them short. Do not add keywords artificially.

[paste headings]
```

Why it is shaped this way. Headings are the strongest single signal of what a passage is about, and question-shaped headings match the shape of what people actually ask. Doing this as a batch rewrite rather than page by page keeps the phrasing consistent across a site, which matters more than any individual heading.

Sceptical review

```
You are a sceptical subject-matter expert in [field] reviewing this
draft for publication.

Identify:
1. Claims made without evidence
2. Statements that are true but too general to be useful
3. Anything a practitioner would know is wrong or oversimplified
4. What a genuine expert would have included that is missing

Be direct. Do not compliment the writing.

Draft:
[paste]
```

Why it is shaped this way. Models default to agreement, which makes them poor reviewers unless instructed otherwise. Casting it as a sceptical expert *in your field* changes both the stance and the vocabulary it argues with. Treat the output as a list of objections to answer in the copy, not as a verdict.

Measurement

Analyse a competitor's citation advantage

```
I asked [engine] "[question]" and it cited these sources:
[paste sources and the relevant passages]

For each cited passage, identify:
1. What specific question it answered
2. What made it more citable than a general article on the topic
3. What information it contained that a competitor could not simply
   restate

Then tell me what my content would need to contain to compete.
```

Why it is shaped this way. It asks what is *different*, not what is good — a comparison produces something actionable where an assessment produces flattery. Feeding it the actual cited sources rather than your impression of them is what keeps the answer grounded in evidence you gathered (Chapter 39).

Classify queries by intent

```
Classify each query below as: informational, commercial-investigation,
transactional, navigational, or local.

Then flag which are likely to trigger an AI Overview and why.

Output as a table. Queries:
[paste]
```

Why it is shaped this way. Intent decides the format of a winning answer, so classifying before writing prevents the common error of answering a comparison question with a definition. Batch classification also makes the boundaries visible: the queries the model hesitates over are usually the genuinely ambiguous ones worth handling twice.

Technical

Draft structured data

```
Generate JSON-LD structured data for this page.

Page type: [article / product / organisation]
URL: [url]
Content summary: [summary]
```

```
Author: [name, credentials]
Key facts: [facts]

Use only properties that correspond to content genuinely visible on
the page. Flag any property you'd recommend that I haven't given you
information for.
```

Why it is shaped this way. Models are reliable at the *shape* of schema and unreliable at its truth — they will cheerfully invent a rating. So the prompt asks for structure from facts you supply, and the output must be validated and fact-checked before it goes near a live page (Chapter 11).

Explain a technical finding

```
Explain the following technical SEO issue to a marketing director with
no technical background.

Cover: what it means, why it matters commercially, what happens if we
don't fix it, and what I should ask the development team for.

Keep it under 200 words and avoid jargon.

Issue: [paste]
```

Why it is shaped this way. Naming the audience and the decision they need to make is what turns an explanation into a recommendation. Without it you get an accurate paragraph nobody can act on, which is the usual failure mode when technical findings reach non-technical stakeholders (Chapter 38).

Strategy

Pressure-test a plan

```
Here is my 90-day AI search plan:
[paste]

Act as a sceptical head of marketing. Identify:
1. Which activities have the weakest evidence behind them
2. What I'm assuming that might not hold
3. What I've prioritised that won't matter for six months
4. What's missing that would matter sooner

Rank by how much each would change the plan.
```

Why it is shaped this way. Asking what would have to be true is more productive than asking whether the plan is good, because it converts optimism into checkable assumptions. The ones you cannot verify are your real risks, and they belong in the plan as such rather than being discovered in month three.

What good output looks like

Prompts are only half the skill; recognising a weak result is the other half. Two worked examples.

The fan-out generator

> *Note before you use it: the output is a hypothetical decomposition, not a record of any queries actually used by another assistant. Label it INFERRED — Chapter 42.*

Weak output — restatements of the question in different words:

```
what is the best project management tool
best project management software
top project management tools
project management tool comparison
```

Strong output — different sub-questions, several of which you would not have thought of:

```
what features do agencies need that generic PM tools lack
how much do agency-focused PM tools cost per seat
do agencies need time tracking built in or separate
how do PM tools handle client access and permissions
what is the migration cost of switching PM tools
which PM tools integrate with common agency accounting software
```

If you get the first kind, the model has not understood that you want the *research plan*, not synonyms. Add: *"Each sub-query must seek different information, not rephrase the question."*

The sceptical review

Weak output — praise with hedged suggestions: *"This is a strong piece with clear structure. You might consider adding more examples."* Useless.

Strong output — specific, uncomfortable and checkable:

```
- "Most agencies charge £4–8k" — no source, no sample, and it
  contradicts the £4–12k range you give in the pricing section.
- The section on retainer creep describes the symptom but never
  says what causes it. A practitioner would name scope ambiguity
  in the original SOW.
- No mention of what happens when a retainer client wants to
  pause. This is the single most common question in the category
  and its absence is conspicuous.
```

If you are not slightly stung by the output, the prompt has not worked. Re-run it with: *"Assume the draft is not good enough to publish. Explain why."*

COMMON MISTAKE

Never paste confidential client data, unreleased financials or personal data into a consumer AI tool. Check your organisation's policy first. Where a business-tier agreement exists, data handling is usually different, but confirm rather than assume.

CHAPTER SUMMARY

- Save and version the prompts that work; improvising the same prompt repeatedly wastes effort
- The highest-value prompts are the fan-out generator, the retrieval test and the sceptical review
- Ask models to identify and critique more often than to produce
- Adapt the bracketed sections properly — generic inputs produce generic outputs
- Check your data-handling policy before pasting anything sensitive

CHECK YOURSELF

1. Which three prompts in this chapter would you use before writing a single word of a new page, and in what order?
2. Why does the sceptical-review prompt explicitly instruct the model not to compliment the writing?
3. Rewrite the vagueness prompt so it works on a product page rather than an article.
4. What is the risk of reusing these prompts verbatim without adapting the bracketed sections?

CHAPTER 36

Strategy and the First 90 Days

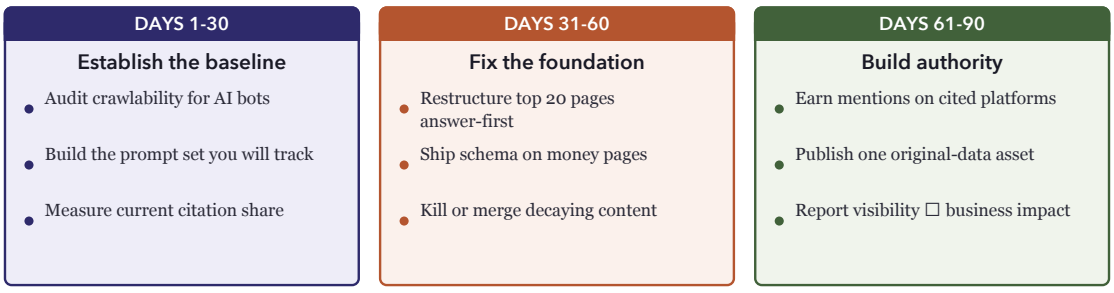
Sequencing the work so it compounds

WHAT YOU'LL LEARN

- How to sequence AI search work so early wins fund later ones
- A concrete 90-day plan
- How to decide what not to do
- Adapting the plan to your situation

Everything in this book is doable. Doing it in the wrong order wastes months. This chapter is about sequence.

A 90-day plan that survives contact with a real job



Baseline first. Without it you cannot prove anything you do next actually worked.

Figure 36.1 – Ninety days, structured so each phase makes the next possible.

The sequencing principle

Three rules determine the order:

1. **Measure before you change anything.** Without a baseline you cannot demonstrate value, and you will be asked to.
2. **Fix access before content.** Perfect content behind a blocked crawler is worth nothing.
3. **Do the cheap high-impact work first.** It buys you the credibility and time to do the expensive work.

Days 1-30: Baseline and access

Goal: know where you stand and remove the blockers.

Week	Work
1	Build the 40-60 prompt set. Run the baseline on two engines. Record everything.
2	Stage 1 and 2 of the audit — access and readability. Fix anything critical immediately.
3	Entity audit. `Organization` schema, `sameAs`, consistency across profiles. Correct divergent descriptions.
4	Map your citation source set. List every third-party domain the engines rely on. Check your accuracy on each.

By day 30 you should have: a baseline you can compare against, confirmation that crawlers can reach you, consistent entity facts, and a list of third-party pages to fix.

> KEY IDEA
>
> If you achieve nothing else in month one, achieve the **baseline** and the **crawler access check**. The first makes everything measurable; the second is where the largest single failure usually hides.

Days 31-60: Structure and foundation

Goal: make your existing content retrievable.

Week	Work
5-6	Restructure your top 20 pages: question headings, answer-first, specificity. Roughly 90 minutes each.
7	Fix the third-party inaccuracies found in week 4. Mostly short emails.
8	Run the decay audit. Refresh the three biggest decliners properly; consolidate obvious duplicates.

By day 60 you should have: your most commercially important content restructured for retrieval, and your description corrected on the pages the engines actually read.

Days 61-90: Authority and proof

Goal: build things that compound, and prove the work mattered.

Week	Work
9	Choose two platforms (Chapter 22). Begin genuine participation.
10-11	Produce one original-data asset. A survey of your own audience is the realistic option.

Week	Work
12	Re-run the prompt set. Compare to baseline. Write the quarterly report.

By day 90 you should have: a measurable change in share of model, one citable asset that will keep earning, and a report that makes the case.

Deciding what not to do

The plan above deliberately excludes several popular activities. Being able to say why is part of the job.

Excluded	Why
`llms.txt` implementation	97% never fetched; hour of work at most, not a priority
Tracking all five engines	Two done properly beats five done badly
Buying a visibility tool in month one	Build the manual system first; you will not understand the numbers otherwise
Volume content production	Restructuring existing pages returns more, faster
Chasing every AI feature launch	Most do not survive; the mechanics in Part I do
Blocking AI crawlers	Unless you have a licensing position, this costs more than it protects

Adapting to your situation

If you are one person at a small company: halve the prompt set to 25. Skip the original research in month three and do a single strong case study with real numbers instead. Everything else holds.

If you are at a large organisation: the constraint is not effort, it is coordination. Spend week 1 identifying who controls `robots.txt`, the CDN, and the CMS templates. The access check may take a month to action rather than a day.

If you are in a YMYL field (health, finance, legal): move the E-E-A-T work from Chapter 15 into month one. Credentialed authorship is a precondition rather than an enhancement in these categories.

If you have no existing content: invert months two and three. Build a small, deep topic cluster on one subject rather than restructuring content you do not have.

CASE STUDY · KETTLEWORTH

Kettleworth's actual first 90 days did not go to plan, and the way it failed is instructive.

They spent month one as prescribed: baseline, access audit, entity work. Month two went to restructuring, and stalled, because the developer resource for server-side rendering was not available until week nine.

Rather than wait, they moved the content restructuring forward and the technical work back. The guides were fully restructured by day 60; the product page rendering did not ship until day 80.

Both worked. The sequencing principle that mattered was measure-first, not the exact order of everything after it. The baseline was what let them prove either had done anything.

CHAPTER SUMMARY

- Measure first, fix access second, improve content third, build authority fourth
- Month one: baseline and access. Month two: structure. Month three: authority and proof
- The baseline and the crawler check are the two things you cannot skip
- Explicitly decide what not to do, and be able to justify each exclusion
- Adapt by team size, organisational complexity, sector risk and existing content

CHECK YOURSELF

1. Why must measurement precede any change?
2. What are the two non-negotiables in month one?
3. Give three activities the plan excludes and the reason for each.
4. How would you adapt this plan for a two-person company with no content?

CHAPTER 37

The GEO Practitioner's Week

What the job actually is, hour by hour

WHAT YOU'LL LEARN

- What a GEO specialist does daily, weekly, monthly and quarterly
- How to allocate time so the important work does not lose to the urgent
- How the role changes over the first year
- Which skills to build in which order, whether you are solo or in a team

Chapter 36 gives you ninety days. This chapter gives you the steady state after them — because the most common failure in this discipline is not doing the wrong things, it is doing the right things once and never again.

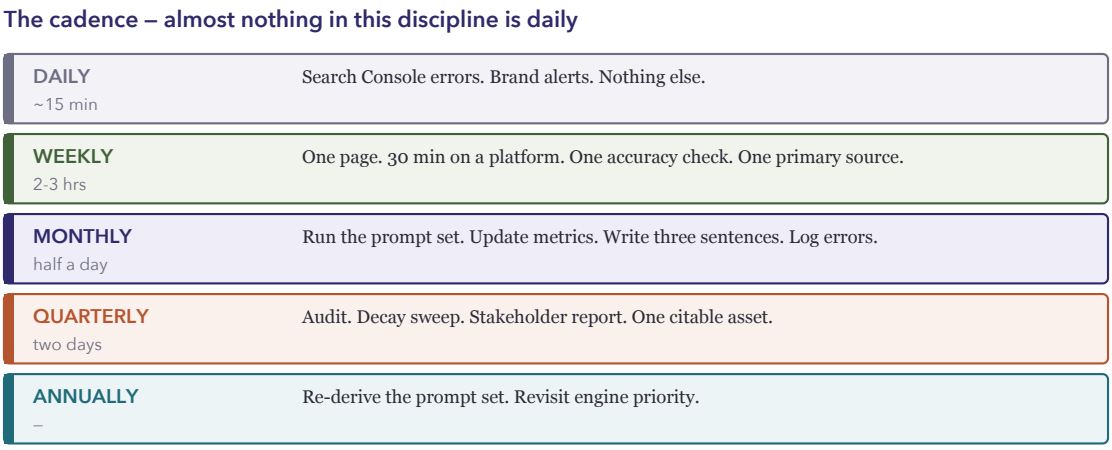


Figure 37.1 – The cadence. Almost nothing in this discipline is daily, and that is the point.

The cadence

Daily — about fifteen minutes

Almost nothing. Resist the pull to check.

- Scan Search Console for indexing errors or crawl anomalies
- Check any alerting you have on brand mentions
- Nothing else

Daily measurement of AI visibility is noise. Answers vary between runs; a daily reading tells you about variance, not performance.

Weekly — two to three hours

- **Publish or refresh one thing.** One properly restructured page beats three rushed ones.
- **Spend thirty minutes on a platform** — the community participation from Chapter 23. This is the item that compounds and the first that gets dropped.
- **Check one third-party source** from your citation list for accuracy.
- **Read one primary source** — an engine's own documentation, a study, an update note. Not a blog summarising one.

Monthly — half a day

- **Run the prompt set** (Chapter 30). Ninety minutes.
- **Update the metrics** and write the three-sentence narrative. Thirty minutes.
- **Log accuracy issues** — anything newly wrong about you (Chapter 24).
- **Note new entrants** to the citation source set.
- **Review the content decay list** (Chapter 18) and pick next month's refreshes.

Quarterly — two days

- **Full audit** (Chapter 32), or at least stages 1 and 2
- **Decay audit** across all content, with the refresh/consolidate/rewrite/delete classification
- **The stakeholder report** (Chapter 31)
- **One original-data asset** shipped or substantially advanced (Chapter 17)
- **Re-check the technical foundation** — crawler access drifts, CDN rules change, developers ship things

Annually

- Re-derive the prompt set from current sales conversations. Customer language moves.
- Revisit the engine priority. Market share shifts.
- Re-read your own oldest content with fresh eyes.

Allocating the time

For a full-time role, a defensible split:

Share	Work
40%	Content — writing, restructuring, refreshing
20%	Measurement and reporting

Share	Work
15%	Technical — audits, fixes, working with developers
15%	Off-site — platforms, PR, third-party corrections
10%	Learning and testing

KEY IDEA

The 10% for learning is the first thing to disappear under pressure and the one that determines whether you are still employable in this field in three years. Defend it in the calendar, not in intention.

How the year unfolds

Months 1-3. Diagnostic and foundational. You are finding out what is broken and unblocking it. Visible results are mostly technical fixes and a baseline nobody had.

Months 4-6. Structural. Restructuring existing content, fixing third-party inaccuracies, building the first cluster properly. Share of model starts moving. This is when stakeholders begin to believe you.

Months 7-9. Compounding. Original research, platform participation and PR start returning. The work you did in months 4-6 is being cited.

Months 10-12. Systematising. What was manual becomes routine or automated. You are running experiments (Chapter 35) rather than only executing.

COMMON MISTAKE

Expect a plateau somewhere around month five. The technical fixes are done, their gains have been banked, and the content and authority work has not yet landed. This is where programmes get cancelled by people who expected a straight line.

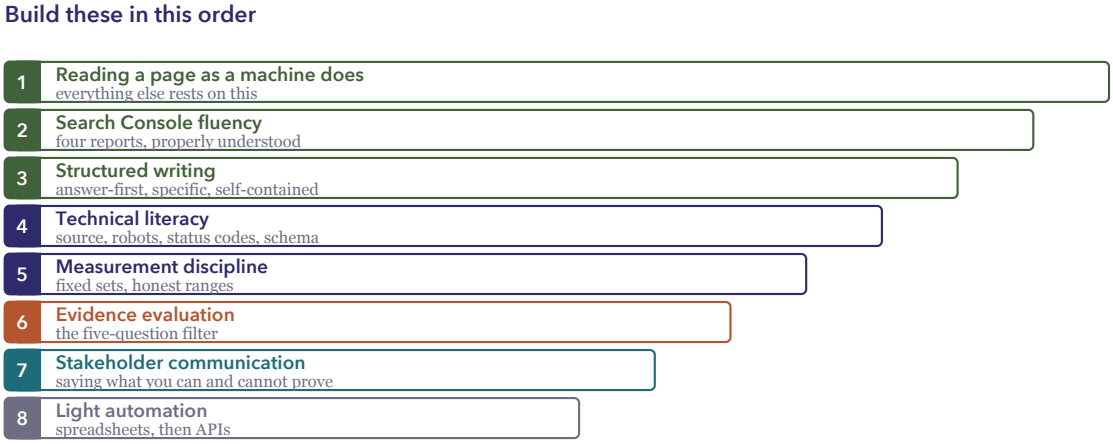
Say it will happen, in advance, in writing. A predicted plateau is a sign your model of the work is right. An unpredicted one looks like failure.

Solo versus in a team

Solo. Halve the prompt set to 25. Drop to one tracked engine. Do the weekly platform participation and the monthly measurement without fail, and let the quarterly audit slip to twice a year if it must. The two things never to drop: the baseline measurement and the crawler access check.

In a team. The work splits along a natural seam — technical foundation, content craft, off-site authority, measurement. The seam that causes trouble is measurement, because it is nobody's favourite and everybody's evidence. Make it one named person's job.

Agency-side. You will be asked for monthly reporting on a discipline whose signal takes eight weeks to move. Set the cadence at quarterly for outcomes and monthly for activity plus leading indicators, and say why in the first meeting rather than the fourth.



Tool expertise is absent deliberately: tools change, and every skill here transfers.

Figure 37.2 – The order that compounds. Note what is absent.

Skills, in the order worth building them

1. **Reading a page as a machine does** — the Chapter 14 test. Everything else rests on it.
2. **Search Console fluency.** Four reports, properly understood.
3. **Structured writing.** Answer-first, specific, self-contained. The single highest-return craft skill.
4. **Basic technical literacy.** View Page Source, `robots.txt`, what a status code means, what schema is.
5. **Measurement discipline.** Fixed prompt sets, honest ranges, stated methods.
6. **Evidence evaluation.** The Chapter 7 filter. This is what stops you wasting a year.
7. **Stakeholder communication.** Chapter 38.
8. **Light automation.** Spreadsheets first, then APIs if the volume justifies it.

Notice what is not on the list: tool expertise. Tools change; every skill above transfers.

CASE STUDY · KETTLEWORTH

Kettleworth's SEO manager kept a note of where her hours actually went in month seven, having assumed she was content-led.

The real split was 22% content, 31% reporting and chasing numbers, 18% technical firefighting, 9% off-site, and 20% meetings about AI that produced no work. Almost the inverse of the intended shape.

Three changes. The monthly measurement was templated so it took ninety minutes rather than most of a week. The recurring "AI strategy" meeting was replaced with the one-page quarterly report. And Friday mornings were blocked for content and defended.

By month ten the split was 38% content, 17% reporting, 14% technical, 16% off-site, 15% learning and testing. The programme's results improved, and the only thing that had changed was where the hours went.

CHAPTER SUMMARY

- Daily work is fifteen minutes and mostly monitoring; daily measurement is noise
- Weekly: one page, thirty minutes of platform participation, one accuracy check, one primary source
- Monthly: the prompt set and the narrative. Quarterly: audit, decay, report, one citable asset
- Defend the 10% for learning in the calendar; it decides whether you are employable in three years
- Predict the month-five plateau in writing, before it happens
- Build skills in order: reading as a machine, Search Console, structured writing, technical literacy, measurement, evidence evaluation, communication, automation

CHECK YOURSELF

1. Why is daily AI visibility measurement a mistake?
2. Which weekly item compounds most and is dropped first?
3. What causes the month-five plateau, and why announce it in advance?
4. Why is tool expertise absent from the skills list?

CHAPTER 38

Working with Stakeholders

The conversations that decide whether any of this happens

WHAT YOU'LL LEARN

- How to present AI search work to people who have heard a lot of hype
- Handling the most common objections and misconceptions
- Setting expectations that survive contact with reality
- Positioning yourself credibly in a field full of overclaiming

Technical competence is not what usually blocks this work. What blocks it is a stakeholder who has read a frightening article, or an enthusiastic one, and has arrived at a position you now have to work with.

The four stakeholders you will meet

Four people, four different questions under the same words

THE PANICKER "AI killed search — what are we doing?" wants urgency	Validate, then quantify it for this business specifically. A real number calms faster than reassurance.
THE DISMISSER "It is a fad, and not in our industry." wants nothing to change	Do not argue. Ask the four engines what the company does, and read the answer out verbatim.
THE TACTICIAN "Why have we not shipped llms.txt?" wants visible activity	Show the evaluation, the proportionate call, and what you did instead with the same hour.
THE MEASURER "What is the ROI, exactly?" wants proof	Name the limits first, then show the proxies and the trend. Credibility here comes from the caveats.

All four are asking the same underlying question: is this real, and is it being handled?
They differ only in what evidence they will accept as an answer.

Figure 38.1 — Four stakeholders, what each is actually asking for, and the reply that lands.

The panicker

Has read that AI killed SEO and wants to know what you are doing about it. Wants urgency.

Approach: validate the concern, then quantify it specifically for your business. "You're right that this is significant. I've measured our exposure — 34% of our impressions are on queries that trigger an AI Overview. Here's the plan for those, and here's why the other 66% needs different work." Specificity converts panic into confidence.

The dismisser

Thinks it is a fad, or that it does not affect their industry. Wants to change nothing.

Approach: do not argue. Show them a verbatim answer. Ask the four engines what your company does and what your competitors offer, and put the answers in front of them. If a competitor is being recommended and they are not, the argument makes itself. If nothing is wrong, they were right, and you have learned something cheaply.

The tactician

Has read a list of tactics and wants to know why you have not implemented `llms.txt`. Wants activity.

Approach: the Chapter 12 answer. Show you have evaluated it, made a proportionate call, and prioritised things with better evidence. Demonstrating a *method* for evaluating claims is more reassuring than any individual answer.

The measurer

Wants a number, wants it monthly, and wants it to go up. Reasonable, and the hardest to satisfy honestly.

Approach: give them share of model against a fixed prompt set, with the sampling method stated every time. Set the absorption expectation early (Chapter 28). Never invent precision.

The objections, answered

IN PRACTICE

Objections and honest answers

"Isn't this just SEO with a new name?" Substantially, yes — the foundations are the same, and anyone telling you to throw out your SEO is wrong. What is new is that retrieval works at passage level rather than page level, which changes how we structure content, and that measurement works by sampling rather than rank tracking.

"How do we know any of this works?" Honestly, the published evidence for specific tactics is weak — a 2026 survey of 45 studies found almost nothing with a proven cross-platform effect. That is why we measure our own results against a fixed prompt set rather than trusting industry claims. What we do have strong evidence for is that crawler access, content structure and entity consistency affect whether we can be retrieved at all.

"Our competitors aren't doing this." That is the argument for doing it now. Citations are long-tailed — no domain holds more than about 5% on any platform, so there is no incumbent to displace yet.

"Can't we just block the AI companies?" We can block training crawlers, and that is a legitimate position. Blocking search crawlers removes us from the answers our customers are reading. Those are different decisions and I'd like to take them separately.

"When will we see results?" Access fixes show up in weeks. Content restructuring takes four to eight weeks to be re-crawled and re-evaluated. Authority work takes two quarters. I'll report monthly so you can see the trend rather than waiting for an outcome.

Briefing an agency or freelancer

If the work is going outside, you need to buy the right thing. Most GEO proposals are SEO proposals with the words changed.

Ask these five questions. The answers separate practitioners from vendors:

1. *"How will you measure whether this worked?"* — you want a fixed prompt set and a stated method, not "increased AI visibility".
2. *"What would you check first?"* — crawler access and rendering. If the first answer is content or `llms.txt`, they have not read the evidence.
3. *"Which tactics do you think have weak evidence?"* — anyone who says none is selling. A good answer names two or three specifics.
4. *"Show me a control group from previous work."* — Chapter 31. Vanishingly few will have one, and the ones who do are worth talking to.
5. *"What will you not do?"* — a practitioner has exclusions. A vendor has a package.

What to put in the brief: your exposure figure (Chapter 5), your baseline if you have one, the two engines that matter for your market, the business metric you are actually trying to

move, and what is out of scope.

What to reject: guaranteed citation counts, "AI-optimised content" priced per word, anything involving seeded reviews or community accounts, and monthly reporting on a metric that moves quarterly.

Setting expectations that survive

Three things to say early, in writing, before you are asked:

1. **"I will be sampling, not measuring."** Explain the method once, properly, and repeat the one-line version every report.
2. **"Some of our influence will be invisible."** Introduce absorption before someone else discovers the attribution gap.
3. **"Some of what the industry recommends has weak evidence, and I will tell you which."** This is the sentence that makes you the reliable source in the room.

> COMMON MISTAKE
>
> The largest career risk in this field is overclaiming. The people making the boldest promises about AI search in 2026 will spend 2027 explaining why they did not materialise. Being the person who was consistently honest about uncertainty, and who measured properly — is a considerably better position eighteen months out.

Positioning yourself

Practical advice for making this a career asset rather than a task:

- **Run the measurement even if nobody asked.** Being the only person with a baseline makes you the person who owns the topic.
- **Write up what you find internally.** A monthly note on how your category is being described is read more widely than you expect.
- **Be specific about evidence quality.** "This is well supported" and "this is plausible but untested" said clearly builds more trust than confident uniformity.
- **Keep the verbatim answers.** A file of screenshots showing what engines said about your company over time is compelling in a way no chart is.

CASE STUDY · KETTLEWORTH

Kettleworth's board contained one panicker and one dismisser.

The panicker had read that AI was ending search and wanted a strategy immediately. He got the 61% exposure figure and a plan addressed specifically to those queries, and stopped panicking, because the number was smaller and more tractable than the article had implied.

The dismisser thought it was a fad. He got a printout of four assistants recommending three competitors in response to the question his own customers asked most. He stopped dismissing.

Neither was convinced by argument. Both were convinced by their own data, which had been available to anyone who bothered to collect it.

CHAPTER SUMMARY

- Match your approach to the stakeholder: panicker needs specificity, dismisser needs evidence, tactician needs method, measurer needs honest metrics
- Set the sampling and absorption expectations in writing before being asked
- Distinguish clearly between well-supported and merely plausible — this is what makes you trusted
- Overclaiming is the main career risk in this field
- Running measurement nobody asked for is the fastest route to owning the topic

CHECK YOURSELF

1. How would you handle a CEO who read that AI has killed organic traffic?
2. What is the honest answer to "how do we know any of this works?"
3. Which two expectations should you set in writing before being asked?
4. Why is being explicit about weak evidence a credibility gain rather than a loss?

PART IX

Instrumentation

Measuring citation properly, building the instrument yourself, and watching what the crawlers actually do. The part where you stop guessing.

CHAPTER 39

Measuring Citations in Practice

What counts as a citation, how to find the passage that was lifted, and the three numbers worth reporting

WHAT YOU'LL LEARN

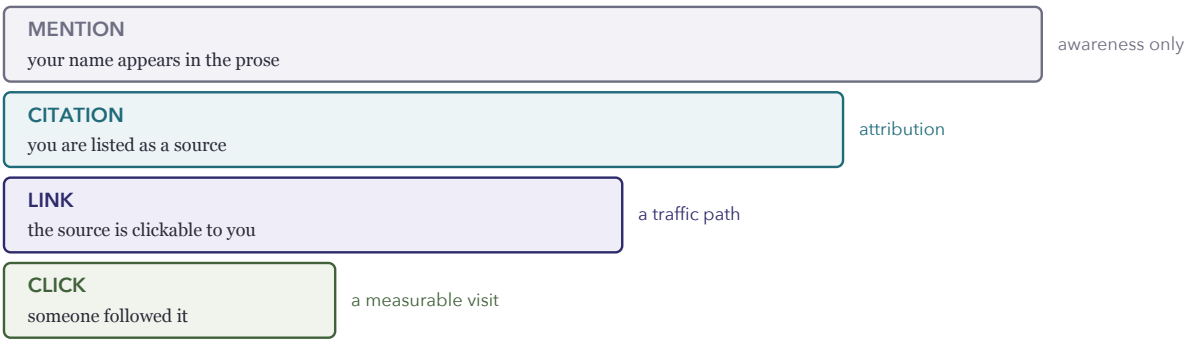
- The difference between a mention, a citation and a link, and why conflating them flatters your numbers
- How to find which specific passage of yours was used
- How to compute share of model, citation rate and source concentration from raw records
- What a defensible measurement claim sounds like, and what an indefensible one sounds like

Chapter 29 gave you the routine: a fixed prompt set, run monthly, logged out, recorded in a spreadsheet. This chapter is about what you do with the answers once you have them. It is the difference between having data and having evidence.

Three different things people call a citation

Most reporting on AI visibility collapses three distinct outcomes into one word. They are worth different amounts and they are earned in different ways.

Four different things, one word



Width shows roughly how often each happens. Most tools report the top rung and let you infer the bottom one.

Figure 39.1 — Four outcomes, decreasing in frequency and increasing in value. Most tools report the first and imply the last.

Outcome	What happened	What it is worth
Mention	Your brand name appears in the answer text	Awareness. No traffic. The model may be drawing on training data, not retrieval
Citation	Your page is listed as a source	Attribution. The reader can see you are the origin
Link	The citation is clickable and points at you	Attribution plus a traffic path
Click	Someone followed it	A visit, measurable in analytics

The gap between mention and citation matters more than it looks. A mention with no citation usually means the model is reciting something it absorbed during training — which means it may be months out of date, and you cannot correct it by publishing. A citation means retrieval happened *this time*, from a live document. Only the second is a system you can influence this quarter.

> COMMON MISTAKE
>
> When a vendor reports your "AI visibility" as a single percentage, ask which of the four rows it counts. Several tools count mentions. Mentions are the easiest number to move and the least connected to anything a finance director recognises as value.

Recording an answer so it can be re-examined

The single most common failure in this work is recording a classification and throwing away the evidence. Six months later you can see that a number moved but not why, and you cannot re-score the old answers against a definition you have since sharpened.

Save, for every prompt and engine:

- **The answer verbatim.** All of it, including hedging and caveats.
- **The cited sources**, as full URLs, in the order given.
- **The date and the engine**, including the model name if it is shown.
- **A screenshot** where the interface shows something the text does not — a carousel, a product panel, an inline card.

IN PRACTICE

Why verbatim matters more than the classification

Kettleworth's first three months of records had a `Brand mentioned: yes/no` column and nothing else. When they later wanted to know whether they were being described as *"a specialist supplier"* or *"one of several options"*, the records could not answer it. They had thrown away the sentence and kept the tick.

They re-ran the set with a verbatim column. The wording turned out to be the finding: they were consistently described as expensive, sourced from a two-year-old forum thread they had never seen.

Finding the passage that was lifted

When an engine cites you, it retrieved a *passage*, not your page. Knowing which passage is the most directly actionable thing in this entire discipline, because it tells you what a working chunk looks like on your own site — and you can then write more of them deliberately.

The method is comparison. Take the sentence or two in the answer that carries your attribution, then find the closest matching text on the cited page.

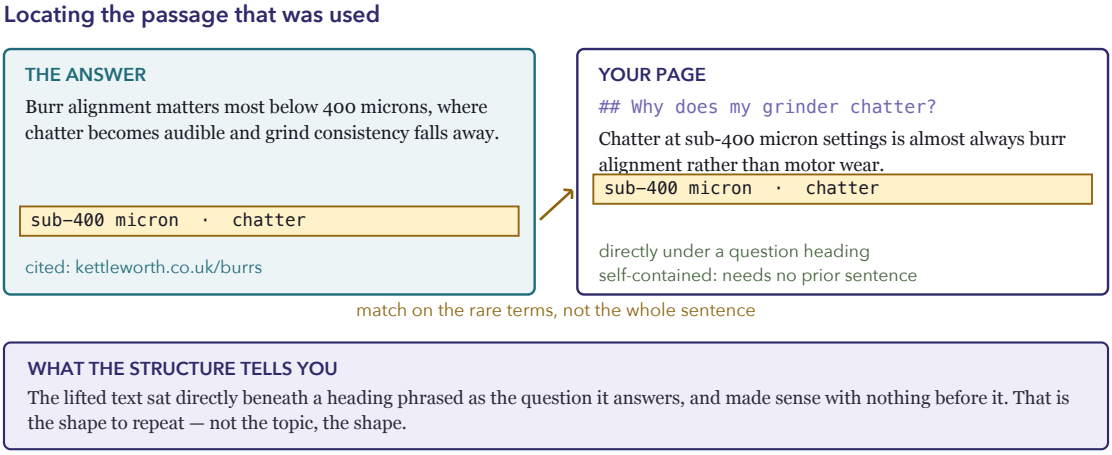


Figure 39.2 – Locating the lifted passage by overlap, and what the surrounding structure tells you.

Work through it in this order:

1. **Isolate the claim.** Which specific assertion in the answer is attributed to you? Usually one sentence, sometimes a number.
2. **Open the cited URL** and search the page for the distinctive nouns from that sentence. Not the whole sentence — models paraphrase — but the rare words: a product name, a figure, an unusual pairing of terms.
3. **Note where it sits.** Under which heading? How far down? Is it inside a list, a table, a paragraph?

4. **Note what precedes it.** Retrieval works on passages that make sense alone (Chapter 13). If the lifted text sits directly under a heading that asks the question the passage answers, that is the pattern to repeat.

After a dozen of these you will have an inventory of what your own site's citable passages look like. That inventory is worth more than any tool's dashboard, because it is specific to how your content is written.

> NOTE
>
> **When you cannot find the passage**
>
> Two innocent explanations. The engine may have retrieved a cached older version of the page — check whether you have edited it recently. Or the citation may be *decorative*: some systems attach a plausible source to a claim assembled from several documents. If you repeatedly cannot locate the text, you are probably looking at the second, which is worth knowing and worth not celebrating.

The three numbers worth reporting

From the records, three metrics carry most of the signal. Compute them the same way every month or the trend is meaningless.

Share of model. Of the prompts where an answer named *any* provider in your category, in what share were you one of them?

```
share of model = prompts naming you / prompts naming anyone in category
```

The denominator matters. Dividing by all prompts, including those the engine answered generically without naming anyone, understates you and makes the number jumpy. Dividing by prompts where naming happened isolates the competitive question: when the engine picks names, does it pick yours?

Citation rate. Of the prompts where you were named, in what share were you cited as a source?

```
citation rate = prompts citing you / prompts naming you
```

A high share of model with a low citation rate is a specific diagnosis: the model knows you from training but is not retrieving you now. That is a crawlability, freshness or structure problem, and Part III is the fix.

Source concentration. Across every answer in the set, count the domains cited and rank them.

```
concentration = citations for a domain / total citations across the set
```

This is the most immediately useful output in the book. The top of that list is not a curiosity; it is your outreach target list, in priority order, derived from evidence rather than intuition.

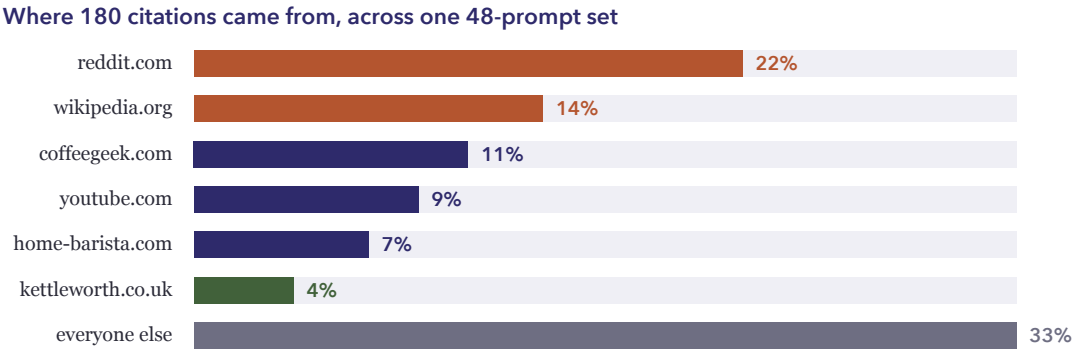


Figure 39.3 – A worked source-concentration count, and the work list it produces.

> **KEY IDEA**
>
> Source concentration converts measurement into a task list. Every domain above a few per cent of citations in your category is a page you should be accurate on, present on, or cited by. Chapter 23 covers how; this is how you decide *where*.

What a defensible claim sounds like

Measurement in this discipline is sampling, not census. Nobody can tell you your "true" visibility, because the systems are non-deterministic, personalised and changing. What you can do is sample consistently and be honest about what the sample supports.

Indefensible	Defensible
"We have 34% AI visibility"	"In our 50-prompt set, on two engines, run logged out on 3 March, we were named in 34% of answers that named anyone"
"AI search sends us 12% of traffic"	"Referral traffic from assistant domains was 12% of sessions; this undercounts, because most assistant traffic arrives without a referrer"
"We improved 8 points this month"	"We improved 8 points month over month; the set is 50 prompts, so 8 points is four prompts, which is within the range we have seen fluctuate without any change on our side"

That last row is the discipline that separates this from theatre. **Work out your noise floor before you report a movement.** Run the same set twice in one week with no changes in between. Whatever it moves by is noise. Anything smaller than that is not a result.

BY THE NUMBERS

Establishing your own noise floor

Run the full set on a Monday and again on the Thursday of the same week, changing nothing. Compare share of model.

Most practitioners who do this are surprised: run-to-run variation of several percentage points on a 50-prompt set is common, because a difference of one or two prompts is already 2-4 points. That is not a flaw in your method. It is the resolution of the instrument, and knowing it stops you reporting weather as climate.

CHAPTER SUMMARY

- Mention, citation, link and click are four different outcomes; report which one you mean
- Record answers verbatim, with full source URLs — a classification alone cannot be re-examined
- Find the lifted passage by matching distinctive terms, then note the structure around it
- Share of model, citation rate and source concentration are the three numbers worth reporting
- Source concentration produces a prioritised outreach list directly from evidence
- Establish your noise floor by running the set twice with no changes, and never report a movement smaller than it

CHECK YOURSELF

1. Why is a mention without a citation a weaker signal than a citation?
2. Why is share of model computed against prompts that named someone, rather than all prompts?
3. What does a high share of model with a low citation rate tell you to go and fix?
4. How do you establish the noise floor of your measurement, and why does it matter?

CHAPTER 40

Building a Citation Tracker

A working tool you can build in an afternoon, with no coding background and nothing to install

WHAT YOU'LL LEARN

- Why a custom tracker is worth building even though tools exist
- How to build one in a spreadsheet, step by step, with the code supplied
- How to schedule it, control what it costs, and keep the data honest
- What it cannot do, and when to stop building and buy

This chapter contains code. You do not need to understand programming to use it — every part is explained in plain language, and the whole thing runs inside Google Sheets with nothing to install. If you would rather not, Chapter 39's manual method produces the same data. This is the same job, done in three minutes a month instead of three hours.

Why build rather than buy

The tools listed in Chapter 33 are competent and getting better. Build your own anyway, at least once, for three reasons.

You learn what the numbers are made of. A dashboard showing "visibility: 41%" is a claim about a method you cannot inspect. When you have written the counting yourself, you know exactly which prompts, which engine, which definition of *named*.

It costs almost nothing. A 50-prompt set on two engines is 100 API calls a month. At the per-call prices typical of 2026 that is small change, against a monthly licence for a platform.

It fits your question. Commercial tools measure what is common across their customers. If your category has a peculiarity — a regulatory term, a comparison that matters only in your market — you can measure it directly.

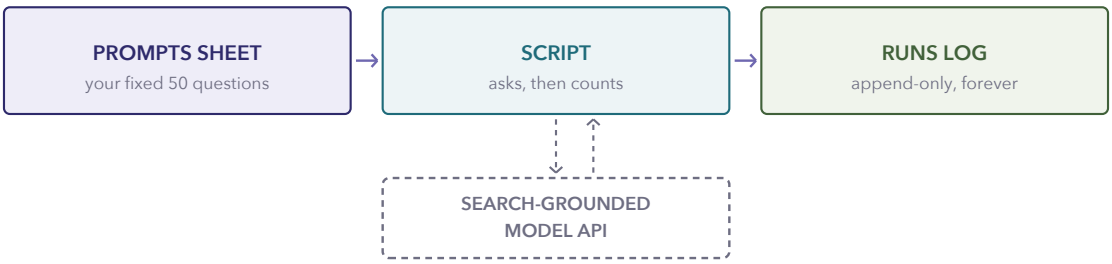
NOTE

What you are trading away

Honesty about the other side: a built tool has no historical database, no competitor benchmarking you did not set up, no multi-market coverage unless you build it, and no support. If you need hundreds of prompts across several languages, buy. This chapter is for the practitioner with one market and a real question.

What you are building

The whole tool



One function talks to the vendor. Everything else is yours, and survives a change of engine.
The log is the asset: three months of rows is what turns a snapshot into a trend.

Figure 40.1 — The whole tool. A sheet holds the questions, a script asks them, a second sheet keeps every answer.

Three parts:

- A **Prompts** sheet: your fixed prompt set, one per row.
- A **script** that reads each prompt, sends it to a search-grounded model, and records what comes back.
- A **Runs** sheet: one row per prompt per run, appended forever, never edited.

The Runs sheet is the asset. Everything else can be rebuilt; the accumulated history cannot.

Step 1: Set up the sheets

Create a new Google Sheet. Make two tabs, named exactly `Prompts` and `Runs`.

In `Prompts`, put a header row and your prompt set beneath it:

```
     A                                          B
1    prompt                                     category
2    what should a 5-person agency pay for a    solution-aware
     CRM?
3    best CRM for small agencies                comparison
```

In `Runs`, one header row only. The script writes everything below it:

```
A1: run_date   B1: engine   C1: prompt   D1: answer
E1: cited_domains   F1: brand_named   G1: brand_cited
```

> **COMMON MISTAKE**
>
> Name the tabs exactly `Prompts` and `Runs`, lower-case spelling as shown for the headers. The script looks them up by name, and a trailing space in a tab name is the single most common reason this does not work first time.

Step 2: Add the script

In the sheet, open **Extensions › Apps Script**. Delete whatever is in the editor and paste the following. Then read the explanation underneath — it is worth understanding what you are running.

```
// ---- settings you change ----------------------------------------
const BRAND = "Kettleworth";            // your brand, as written
const KEY_NAME = "SEARCH_API_KEY";      // label for the stored key
const ENGINE = "perplexity";            // what the log should call it
const ENDPOINT = "https://api.perplexity.ai/chat/completions";

// ---- the part that talks to the model ---------------------------
// The only vendor-specific function. Check the current API docs for
// the endpoint, the request shape, and where citations appear in
// the response. Those change; the rest of this script does not care.
function askEngine(prompt) {
  const props = PropertiesService.getScriptProperties();
  const key = props.getProperty(KEY_NAME);
  if (!key) throw new Error("No API key stored. See step 3.");

  const response = UrlFetchApp.fetch(ENDPOINT, {
    method: "post",
    contentType: "application/json",
    headers: { Authorization: "Bearer " + key },
    payload: JSON.stringify({
      model: "sonar",
      messages: [{ role: "user", content: prompt }]
    }),
    muteHttpExceptions: true
  });

  const code = response.getResponseCode();
  if (code !== 200) {
    return { answer: "ERROR " + code, citations: [] };
  }

  const body = JSON.parse(response.getContentText());
  return {
    answer: body.choices[0].message.content,
    citations: body.citations || []
```

```
  };
}

// ---- the part that does the counting ---------------------------
function domainsOf(urls) {
  const out = [];
  urls.forEach(function (u) {
    try {
      out.push(new URL(u).hostname.replace(/^www\./, ""));
    } catch (e) { /* not a usable URL; skip it */ }
  });
  return out;
}

function contains(haystack, needle) {
  return haystack.toLowerCase().indexOf(needle.toLowerCase()) !== -1;
}

// ---- the part that runs the whole set --------------------------
function runPromptSet() {
  const book = SpreadsheetApp.getActiveSpreadsheet();
  const rows = book.getSheetByName("Prompts")
                   .getDataRange().getValues().slice(1);
  const runs = book.getSheetByName("Runs");
  const today = new Date();

  rows.forEach(function (row) {
    const prompt = row[0];
    if (!prompt) return;

    const result = askEngine(prompt);
    const answer = result.answer || "";
    const domains = domainsOf(result.citations);

    // "Named" = the brand appears in the prose the reader sees.
    const named = contains(answer, BRAND);

    // "Cited" = one of the sources is our own site. Change this if
    // your brand name and your domain are not the same word.
    const cited = domains.some(function (d) {
      return contains(d, BRAND);
    });

    // Joined with a bare comma, no space: the counting formula in step 5
    // splits on "," and a stray space would make " reddit.com" and
    // "reddit.com" count as two different domains.
    runs.appendRow([today, ENGINE, prompt, answer,
                    domains.join(","), named, cited]);

    Utilities.sleep(1200);  // polite, and stays under rate limits
  });
}
```

What each part does, in order.

The block at the top is settings. `BRAND` is the string searched for in answers. `KEY_NAME` is a label, not the key itself — the key is stored separately in step 3 so it never appears in the file.

`askEngine` is the only function that knows anything about a particular vendor. It sends the prompt and hands back two things: the answer text and a list of source URLs. If the request fails it returns an error string rather than stopping the run, so one bad response does not lose you the other forty-nine. **This is the function to change if you switch engines**, and the only one.

`domainsOf` turns full URLs into bare domains, so `https://www.example.com/guide/x` becomes `example.com`. Stripping `www.` matters: otherwise the same site counts as two when you tally concentration.

`runPromptSet` is the job. It reads the prompts, asks each one, works out whether you were named and whether you were cited, and appends a row. `Utilities.sleep(1200)` waits 1.2 seconds between calls — this keeps you inside rate limits and is simply good manners toward an API you do not own.

> NOTE
>
> **Why `named` and `cited` are computed separately**
>
> They are the two rows of Figure 39.1 that get confused most often. Keeping them as separate columns means you can compute citation rate later without re-reading every answer — and if you decide your definition of *named* was too loose, the verbatim answer is in column D and you can re-score without re-running.

Step 3: Store the key safely

Never paste an API key into a spreadsheet cell or into the script itself. Anyone you share the sheet with would get it, and shared sheets outlive the people who made them.

In the Apps Script editor: **Project Settings › Script Properties › Add script property**. Name it `SEARCH_API_KEY` and paste the key as the value.

> COMMON MISTAKE
>
> If you have already pasted a key somewhere visible, treat it as compromised: revoke it in the provider's dashboard and issue a new one. Keys in shared documents are one of the more common small security incidents in marketing teams.

Step 4: Run it, then schedule it

Run `runPromptSet` once from the editor. Google will ask you to authorise the script — it is your own script, asking for permission to edit your own sheet and to make outbound requests. Watch the `Runs` tab fill.

Once a manual run works, schedule it: in Apps Script, **Triggers › Add Trigger** › choose `runPromptSet`, time-driven, month timer, and pick a day. The first Monday of the month is a reasonable convention, and matching Chapter 29's manual cadence means the two data sets are comparable.

> IN PRACTICE
>
> **Check the first run by hand**
>
> Before trusting anything, take three rows from `Runs` and repeat those prompts yourself in a browser, logged out. You are checking that the answers are recognisably similar and that the citation list matches.
>
> They will not be identical — these systems are non-deterministic, and the API may not use exactly the same retrieval path as the consumer interface. What you are looking for is whether the tool is measuring the same *kind* of thing. If the API returns answers with no citations at all while the browser shows several, your request is not asking for a grounded search and the whole log is measuring the wrong thing.

Step 5: Turn rows into the three numbers

Add a third tab, `Summary`, and use ordinary spreadsheet formulas. No more code.

```
Share of model   =COUNTIFS(Runs!F:F,TRUE)/COUNTA(Runs!C2:C)
Citation rate    =COUNTIFS(Runs!G:G,TRUE)/COUNTIFS(Runs!F:F,TRUE)
```

The two above are ordinary counting. `COUNTIFS` counts rows meeting a condition, `COUNTA` counts non-empty cells, and each formula is one divided by the other — exactly the definitions from Chapter 39, written in spreadsheet.

Source concentration takes more work, because column E holds a *list* in each cell — `reddit.com, wikipedia.org, coffeegeek.com` — and you need to count across all of those lists at once. The formula looks alarming; it is four simple steps nested inside one another:

```
=QUERY(FLATTEN(SPLIT(TEXTJOIN(",",TRUE,Runs!E2:E),",")),
       "select Col1, count(Col1) group by Col1 order by count(Col1) desc",0)
```

Read it inside out, because that is the order it runs in:

1. `TEXTJOIN(",",TRUE,Runs!E2:E)` glues every cell in column E into one enormous comma-separated string. `TRUE` tells it to skip empty cells, so blank rows do not produce stray commas.
2. `SPLIT(..., ",")` chops that string back apart at every comma — giving one domain per cell, now spread across a single wide row.
3. `FLATTEN(...)` tips that row on its side into a single tall column, because the next step can only count down a column.

4. `QUERY(..., "select Col1, count(Col1) group by Col1 order by count(Col1) desc", 0)` counts them. In plain English: *take column one, group the identical values together, count each group, and sort with the biggest first*. The final `0` says the data has no header row of its own.

The result is two columns — domain, and how many times it was cited — sorted with the most-cited first. That is the ranked table from Chapter 39: your outreach list, regenerated every month, for free.

> NOTE
>
> **If the counts look wrong**
>
> The usual cause is a separator mismatch: if column E holds `reddit.com, wikipedia.org` with a space after the comma, splitting on `","` leaves `wikipedia.org` with a leading space, and the same domain gets counted twice under two spellings. This is why the script joins with a bare comma. If you edit either side, change both — they only have to agree with each other.

> KEY IDEA
>
> The moment this stops being a toy and becomes an instrument is the moment you have three months of rows. One run tells you where you stand. Three tells you which direction you are moving, and whether the movement is bigger than the noise floor you established in Chapter 39.

What this will not do

Stated plainly, so you do not discover it later:

- **It measures an API, not the product a person uses.** Consumer interfaces personalise, use conversation history and sometimes different retrieval. Treat the tool as an index, not a replica.
- **It cannot see Google AI Overviews.** There is no supported API for them. Overviews remain a manual check, which is one reason Chapter 29's routine does not disappear entirely.
- **It has no opinion about quality.** Being named is not the same as being described well. The verbatim column exists so a human can read for tone, and a human should, monthly.
- **It will drift.** Vendors change models and response shapes. When the numbers move sharply with no plausible cause, suspect the tool before believing the finding.

CASE STUDY · KETTLEWORTH

Kettleworth built this in an afternoon and ran it for a year.

The tracker itself produced no insight in month one that the manual method had not already given them. Its value showed up in month seven, when share of model fell 11 points in a single run. Because they had the verbatim answers, they could see immediately that the fall was not competitive: the engine had started answering grinder questions generically, naming no brands at all, in about a third of the set.

Their share of *named* answers was in fact unchanged. Had they been tracking only a headline percentage from a dashboard, they would have spent a quarter fixing a problem they did not have.

CHAPTER SUMMARY

- Build once, even if you later buy: you learn what the numbers are made of
- Three parts — a prompt sheet, a script, an append-only log; the log is the asset
- Isolate the vendor-specific call in one function so switching engines is a small change
- Store the API key in script properties, never in the sheet
- Compute share of model, citation rate and concentration with ordinary formulas
- The tool measures an API, cannot see AI Overviews, and has no view on quality — read the answers

CHECK YOURSELF

1. Why is `askEngine` kept as a separate function from `runPromptSet`?
2. Why must the API key live in script properties rather than a cell?
3. Why does the script record the verbatim answer when it already records true/false columns?
4. Name two things this tracker cannot measure, and what you do about each.

CHAPTER 41

Watching the Crawlers

Server logs as evidence: getting them, reading them, verifying them, and connecting crawl to citation

WHAT YOU'LL LEARN

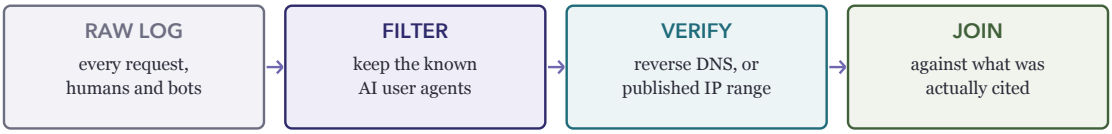
- How to obtain logs when you do not run the server
- How to read a log line, field by field
- How to verify a crawler is genuine rather than something wearing its name
- How to connect crawl activity to citation outcomes, and the limits of that connection

Chapter 8 established that server logs are the only source of truth about which crawlers actually reached you. This chapter is the practical skill: obtaining them, reading them and proving what they show. It is the supply side of measurement — Chapters 39 and 40 measure what engines *say*, this measures what they *fetch*.

Getting hold of the logs

You probably do not administer the server. The request is small and the people who can grant it are used to it.

From raw log to evidence you can defend



MOST OF THE VOLUME DROPS OUT AT THE FIRST TWO STEPS

THE FIELD THAT ANSWERS GATE ONE

```
203.0.113.42 ... "GET /guides/burrs" 403 ... "OAI-SearchBot/1.0"
```

A 403 here is invisible from every other vantage point. robots.txt shows intent; this shows what happened.

Figure 41.1 – From raw access log to defensible evidence. Most of the work is filtering and verification.

Where your site runs	What to ask for, and of whom
Traditional host / VPS	"The last 30 days of access logs" — hosting support or whoever deploys

Where your site runs	What to ask for, and of whom
Cloudflare in front	Bot analytics in the dashboard; Logpush for the raw article
Vercel, Netlify, similar	Log drains, or the platform's analytics with bot traffic shown
A CMS you do not control	Ask the vendor for crawler access reports; some expose them, some do not
No access at all	Fall back to the indirect evidence at the end of this chapter

Ask for **raw access logs** rather than a summary. Summaries are aggregated by the tool's idea of what matters, and AI crawlers are frequently bundled into "other bots".

Reading a log line

A standard access log line looks like this. It is one long line in the file; it is wrapped here to fit the page. It is intimidating for about ninety seconds.

```
203.0.113.42 - - [03/Mar/2026:11:14:52 +0000]
  "GET /guides/grinder-burrs HTTP/1.1" 200 18342 "-"
  "Mozilla/5.0 (compatible; OAI-SearchBot/1.0;
   +https://openai.com/searchbot)"
```

Read it in pieces:

Piece	Meaning
`203.0.113.42`	The IP address that made the request
`[03/Mar/2026:11:14:52 +0000]`	When, with timezone
`"GET /guides/grinder-burrs"`	What was requested
`200`	The response code your server gave
`18342`	Bytes sent
`"OAI-SearchBot/1.0..."`	The user agent — who the requester claims to be

Four of these carry the whole story for our purposes: **who claimed to come, what they asked for, what they got, and when.**

The response code is where problems surface. `200` means served. `403` means your server refused. `404` means the page is gone. A crawler receiving `403` across the board is the exact failure from Chapter 8's case, and it is invisible from every other vantage point.

IN PRACTICE

The first pass, in a spreadsheet

You do not need special software. Open the log in a spreadsheet, split on spaces, then:

1. **Filter the user-agent column** for `GPTBot`, `OAI-SearchBot`, `ChatGPT-User`, `ClaudeBot`, `Claude-SearchBot`, `PerplexityBot`, `Google-Extended`, `CCBot`.
2. **Pivot by user agent and response code.** You are looking for any bot whose responses are not overwhelmingly `200`.
3. **Pivot by user agent and URL.** This tells you which pages are actually being fetched, which is often not the pages you assumed.

Fifteen minutes, and it answers the only technical question that matters at gate one.

Verifying a crawler is genuine

A user agent is a self-declaration. Anyone can send a request claiming to be `GPTBot`, and plenty of scrapers do, precisely because sites allow it. If you are going to make claims to a stakeholder about who is crawling you, verify.

The standard method is **reverse DNS, then forward confirmation**:

1. Take the IP address from the log line.
2. Look up its reverse DNS name.
3. Look that name back up to an IP address.
4. Confirm the result matches the original IP.

A genuine crawler resolves to a hostname on the operator's own domain, and that hostname resolves back to the same address. An impostor fails at step 2 or step 4.

```
# on any Mac or Linux terminal
host 203.0.113.42
# -> ... domain name pointer crawl-203-0-113-42.openai.com

host crawl-203-0-113-42.openai.com
# -> crawl-203-0-113-42.openai.com has address 203.0.113.42
```

The two agree, so the request genuinely came from the operator.

NOTE

The easier route

Most major operators now publish their crawler IP ranges as a file you can download. Checking an address against a published range is quicker than reverse DNS and does not require a terminal. Search for the operator's name plus "crawler IP ranges"; treat the published list as authoritative over anything a third party republishes.

COMMON MISTAKE

Do not skip verification and then present crawl volumes to a stakeholder. “We had 4,000 AI crawler visits” is a claim that collapses the moment someone asks whether they were real, and a meaningful share of traffic claiming to be an AI crawler is not.

Connecting crawl to citation

This is the join that makes the two halves of measurement one system. You have, from Chapter 40, a log of which pages were cited. You have, from this chapter, a log of which pages were fetched. Put them side by side.

Four cases, each with a different meaning:

Crawled	Cited	What it means	What to do
Yes	Yes	Working as intended	Study the passage (Chapter 39), write more like it
Yes	No	Reachable but not chosen	A structure or authority problem, not an access one — Parts IV and V
No	Yes	Cited from training data or a third party	Check whether the claim is current; you cannot update it by publishing
No	No	Not in the game	Gate one. Fix access before anything else

KEY IDEA

The *crawled but not cited* row is where most established sites actually sit, and it is routinely misdiagnosed as a crawlability problem because that is the easier thing to check. If the logs show `200` responses across your key pages and you are still not cited, stop optimising access. The problem is upstream of everything technical: what the passage says and whether anyone else corroborates it.

When you cannot get logs at all

Sometimes access is genuinely impossible. The indirect evidence, in descending order of usefulness:

- **A canary page.** Publish a page with a distinctive, otherwise-unused phrase, link to it from somewhere crawlable, and periodically ask engines a question that phrase would answer. If the phrase comes back, something fetched and retained it.
- **CDN bot analytics.** Less granular than raw logs but usually enough to answer “are they arriving and are they getting errors”.

- **Live retrieval checks.** Ask an engine about a page you published yesterday. A correct answer proves retrieval more directly than any log.
- `robots.txt` **request monitoring.** Some platforms log requests for that one file even when they do not expose general logs, and every well-behaved crawler asks for it first.

CASE STUDY · KETTLEWORTH

Kettleworth could not get raw logs — their site was on a managed platform whose support politely declined.

They used a canary. A short page on burr alignment contained the phrase *"burr chatter at sub-400 micron settings"*, which appeared nowhere else on the web. Eleven days after publishing, an assistant asked about grinder noise returned that phrase almost verbatim and cited the page.

That single observation told them what a month of arguing with support had not: retrieval was working, the platform was not blocking anything, and their problem was elsewhere. The canary cost them forty minutes.

CHAPTER SUMMARY

- Ask for raw access logs, not summaries, which bury AI crawlers under "other bots"
- Four fields carry the story: who claimed to come, what they asked for, what they got, when
- Response codes are where failures surface; a crawler receiving 403 is invisible from anywhere else
- Verify crawlers by reverse DNS with forward confirmation, or against published IP ranges
- Join crawl records to citation records; the four-case table converts the join into a diagnosis
- Without logs, a canary page proves retrieval faster than arguing for access

CHECK YOURSELF

1. Why ask for raw access logs rather than the analytics summary?
2. What does a `403` in the response-code column of an AI crawler's rows tell you?
3. Describe reverse DNS verification, and what an impostor fails.
4. Your key pages return `200` to `OAI-SearchBot` and you are still never cited. Which two parts of this book address that, and why is it not a crawlability problem?

PART X

Forensics

Why one source was chosen and another was not — and how to say what you know, what you reconstructed, and what you cannot see. The part that makes your findings defensible.

CHAPTER 42

Evidence Status: Actual, Inferred, Unknown

The one discipline that separates a professional finding from a plausible story

WHAT YOU'LL LEARN

- Why a model's explanation of its own behaviour is not a system log
- The three labels every process claim in your reporting should carry
- How to apply them to the techniques already in this book
- What to do when a stakeholder wants certainty you do not have

This is the shortest chapter in Part X and the most important. Everything that follows — forensics, claim validation, lineage, counterfactuals — depends on one habit: knowing, and saying, which class of evidence each statement rests on. Without it, the rest of this part produces confident-sounding fiction.

The failure this prevents

Ask a model why it cited a particular source and it will tell you. The answer will be fluent, specific and plausible. It will cite relevance, authority, freshness. It will read exactly like an explanation.

It is not one. It is a *new generation*, produced after the fact, by a system that has no privileged access to the reasoning that produced the earlier answer. The model is not reading a log. It is writing the kind of text that tends to follow a question like yours.

COMMON MISTAKE

A model's retrospective account of its own retrieval is generated text, not telemetry. It is often *reasonable*. It is never *evidence*. The distinction survives no matter how detailed and confident the account sounds — detail and confidence are properties of fluent generation, not of accuracy.

The same trap sits inside a technique this book already recommended. Chapter 4 suggests asking a model what sub-questions it would research, as a way of generating your own fan-out. That is a good content-planning move. It tells you nothing whatsoever about the

sub-queries some other assistant actually issued when it answered a real user. One is a hypothesis you generated. The other is a fact you cannot see.

The three labels

Every claim you make about how an answer was produced falls into one of three classes. Label it.

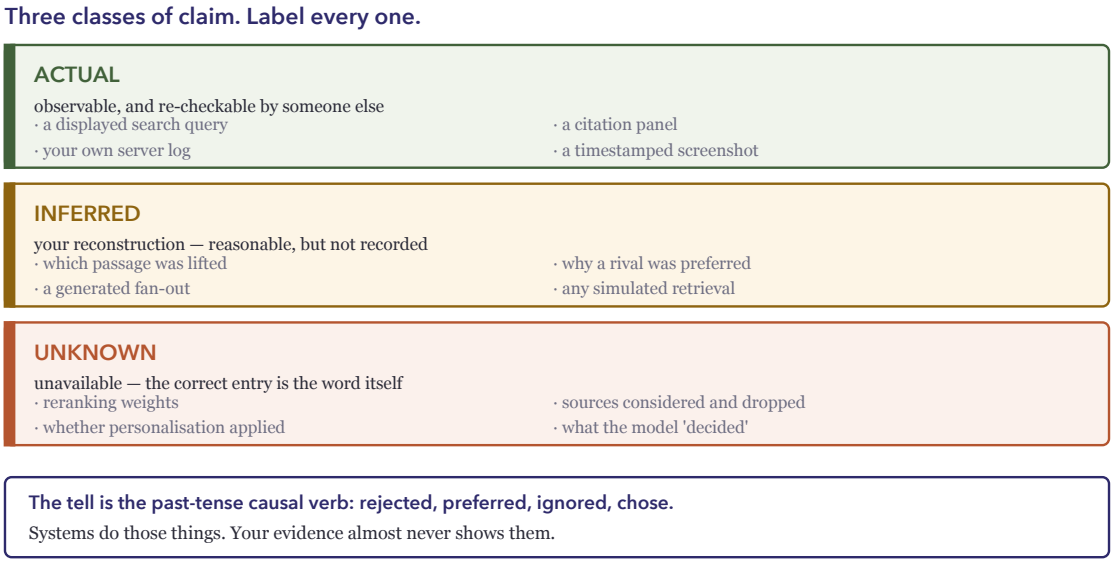


Figure 42.1 – The three classes, what belongs in each, and the failure that follows from mixing them.

ACTUAL. Directly observable, and re-checkable by someone else. A visible search query in a product that shows one. A tool-call record. The list of sources in a citation panel. Your own server log line. A screenshot with a timestamp. If a sceptical colleague could open the same artefact and see the same thing, it is actual.

INFERRED. A reconstruction you built, which is reasonable but was not recorded. "The passage that won was probably the comparison table, because it is the only text on any cited page that answers the question directly." That may well be right. It is still your reasoning, not the system's record.

UNKNOWN. Unavailable, and not to be guessed. Reranking weights. Which relevant sources were considered and rejected. Whether personalisation affected the answer. The correct entry here is the word UNKNOWN, not a plausible sentence.

> KEY IDEA
>
> The purpose of the labels is not modesty. It is that **inferences can be revised and actuals cannot**. If you record which is which, a later finding that contradicts your reconstruction costs you one line in a document. If you recorded everything as fact, it costs you your credibility with the person who believed you.

Applying it to what you already do

Nothing in the earlier chapters becomes wrong. Several things become labelled.

What you do	Evidence status	Why
Recording an answer verbatim	ACTUAL	You have the text
Listing the sources shown in a citation panel	ACTUAL, within the panel's scope	The panel is a record, but it shows what was cited, not what was retrieved
Reading a search query the product displays	ACTUAL	It is displayed
Asking a model to generate a fan-out	INFERRED	A hypothesis about sub-questions, not a record of any
Identifying which passage of yours was lifted	INFERRED unless quoted verbatim	Word-for-word overlap is strong; paraphrase is a judgement
Explaining why a competitor was preferred	INFERRED	Your analysis of observable differences
Stating which sources were considered and dropped	UNKNOWN	No product exposes this
Stating the reranking weights	UNKNOWN	Not disclosed by anyone
Server log showing a 403 to OAI-SearchBot	ACTUAL	Your own infrastructure recorded it

IN PRACTICE

Relabelling one existing report

Take the last piece of AI-visibility analysis you wrote or received. Go through it sentence by sentence and mark each process claim A, I or U in the margin.

Two things usually happen. First, the proportion of I is higher than anyone expected — commonly more than half. Second, at least one sentence that reads as fact turns out to be U dressed as I: a claim about something nobody can observe, stated in the past tense.

Rewrite those sentences. "Perplexity rejected our page because it lacked structured data" becomes "Our page was not cited. It lacks structured data, which the cited pages have — but the selection reason is not observable." The second sentence is less satisfying and considerably more useful, because it does not stop the investigation at a guess.

Saying it out loud

Stakeholders do not, in practice, punish honest uncertainty. They punish being wrong after sounding certain. The phrasing that works:

> *"Two things here are observed and two are my reading of the evidence. Observed: we are not cited on any of the twelve buying-intent prompts, and our server returns 403 to two AI crawlers. My reading: the 403 is the larger cause, because the pages that do get cited are not obviously better. I can prove the first pair today. The second I can only test by fixing the block and remeasuring."*

That paragraph contains a diagnosis, a confidence level, and a falsification plan. It is more persuasive than certainty, because it tells the listener exactly what would change your mind.

NOTE

Chapter 38 covers stakeholder conversations in general. The specific move here is narrower: never let an inference cross into a report wearing the grammar of an observation. The tell is the past-tense causal verb — *rejected*, *preferred*, *ignored*, *chose*. Systems do these things; your evidence rarely shows them.

CHAPTER SUMMARY

- A model's account of its own retrieval is generated text, not a log, however detailed it sounds
- Label every process claim ACTUAL (observable and re-checkable), INFERRED (your reconstruction) or UNKNOWN (unavailable, not to be guessed)
- The generated fan-out is a planning tool and an INFERRED artefact — never a record of another assistant's queries
- Inferences can be revised without damage; misreported facts cannot
- Report the label alongside the finding, and name what would change your mind

CHECK YOURSELF

1. Why is a model's explanation of why it cited a source not evidence, even when the explanation is accurate?
2. Which label applies to a list of sources taken from a citation panel, and what is the limit of what that panel proves?
3. A report says "Gemini preferred the competitor's guide because it was more recent." What is wrong with the sentence, and how would you rewrite it?
4. What practical advantage does an INFERRED label give you three months later?

CHAPTER 43

Citation Decision Forensics

Reconstructing why one source was used and another was not

WHAT YOU'LL LEARN

- The difference between improving your content and investigating a decision
- A thirteen-part audit structure for a single answer
- Which stages are observable, which are reconstructed, and which are closed to you
- How to turn one forensic audit into a prioritised action list

Chapter 39 taught you to trace an answer back to the passages it used. That exercise begins with the sources that were visibly cited — which is to say, it begins after most of the decision has already happened. This chapter starts earlier and asks a harder question: of everything that could have been used, why this?

Why the earlier exercise is not enough

Tracing a citation tells you what won. It does not tell you what it beat.

That matters because the two produce different work. “Their comparison table is more specific than ours” is a content note. “Their comparison table was one of eleven candidate pages, four of which were opened, and ours was never returned in the result set at all” is a diagnosis — and it points at retrieval, not at writing. You would have spent a month improving a page that was never in the running.

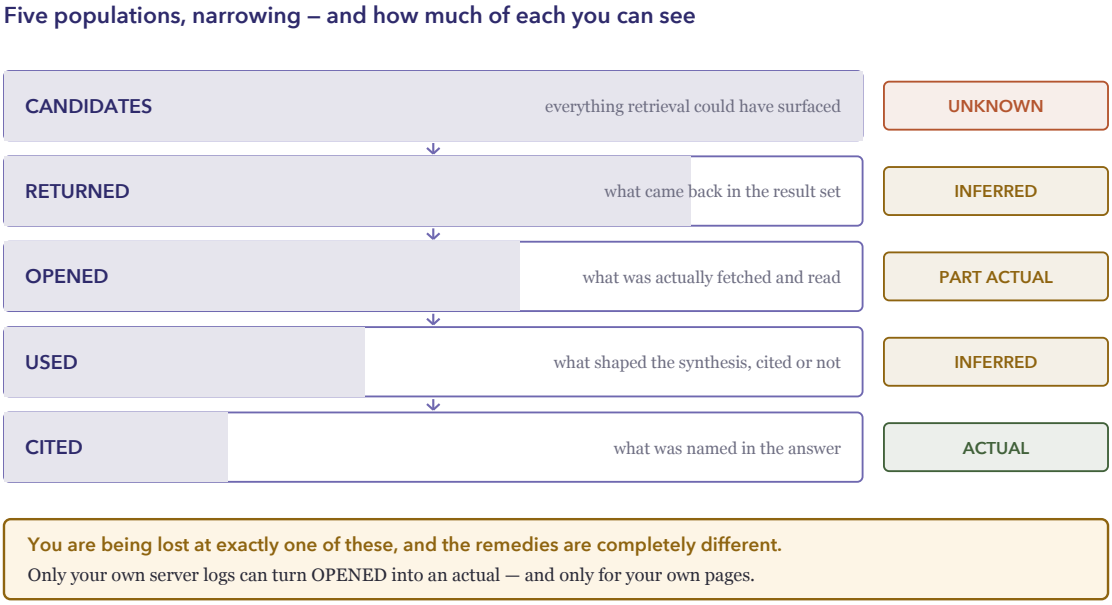


Figure 43.1 – The five populations a source passes through, and how much of each you can actually see.

The populations narrow at every step:

- **Candidates.** Everything the retrieval layer could plausibly have surfaced. Largely UNKNOWN, though you can approximate it by running the sub-queries yourself as ordinary searches.
- **Returned.** What came back in the result set. Sometimes visible in products that show their search steps; otherwise INFERRED.
- **Opened.** What was actually fetched and read. Occasionally visible; often UNKNOWN. Your own server logs can prove it for *your* pages, which is the one hard corner of this you own.
- **Used.** What contributed to the synthesis, cited or not. Almost always INFERRED, from verbatim overlap.
- **Cited.** What was named. ACTUAL.

> KEY IDEA
>
> Most GEO work optimises for the last box while the failure is in the first. A page that is never a candidate cannot be improved into a citation. Establish where in the funnel you are being lost before you decide what to fix — the remedies are completely different, and only one of them is writing.

The thirteen-part audit

This is the structure. It is deliberately long, because a shorter one lets you skip the parts that are uncomfortable to admit you cannot answer. Every section that you cannot fill in gets the word UNKNOWN, which is itself a finding.

1. Interpreted intent. What did the assistant appear to take the question to mean? Compare with what the user meant. A mismatch here explains more failures than any technical issue.

2. Citation-requiring claims. Break the answer into atomic claims. Mark which ones need external evidence and which are general knowledge. Only the first group is contestable.

3. Search and tool record. What the product shows about whether it searched, and with what queries. ACTUAL where displayed; UNKNOWN where not. Do not fill this in from a model's description of its own process.

4. Candidate set. Run the visible or reconstructed sub-queries as ordinary web searches. Record the top results. This is your best available proxy for what was in the pool, and it is INFERRED.

5. Returned, opened, used, cited. Place every source you can identify into the funnel above, with its status label.

6. Source quality assessment. Score each cited source on the dimensions in the next section.

7. Selection reasons. For each cited source, what observable property plausibly earned it the place. Evidence-based, and INFERRED.

8. Rejection reasons. For each strong source that was *not* cited — including yours — the most plausible reason, from a fixed list.

9. Claim-to-source map. Which citation supports which claim, and how well. This is Chapter 44.

10. Unsupported claims. Anything asserted with no citation attached, or with a citation that does not carry it.

11. Conflicting evidence. Where two cited sources disagree, and which one the answer followed.

12. Citation placement. Whether citations sit next to the claims they support or are pooled at the end. Pooled citations are much weaker evidence of what was actually used.

13. Unknowns. An explicit list. This section should never be empty, and a forensic report that has an empty one is not finished — it is overconfident.

Scoring a source, and scoring the one that lost

The point of a rubric is to stop you from concluding "theirs was just better." Score both the winner and the strongest source that lost, on the same dimensions.

Dimension	What you are asking
Direct relevance	Does a passage answer *this* sub-question, not the topic

Dimension	What you are asking
Primary status	Is it the origin of the fact, or reporting someone else's
Authority	Would a domain expert accept it as a reference
Freshness	Is the fact current, and does the page say when it was updated
Specificity	Numbers, models, dates, named conditions
Corroboration	Do independent sources agree
Geographic and language fit	Right market, natively written
Accessibility	Reachable, renderable, not paywalled or bot-blocked
Commercial bias	Does the source profit from the answer
Independence	Is it evidentially distinct from the others cited

Then record why the loser lost. A fixed vocabulary keeps this honest and makes the results countable across many audits:

`not returned` · `not opened` · `outdated` · `duplicate` · `secondary where a primary existed` · `paywalled or blocked` · `geographic mismatch` · `unsupported claim` · `commercially biased` · `contradicted by stronger evidence` · `unknown`

IN PRACTICE

One answer, end to end

Objective. Produce a completed forensic audit for a single commercially important answer.

Inputs. One prompt from your set that names a competitor and not you. One engine. Two hours.

Procedure.

1. Run the prompt logged out. Screenshot the answer and the source panel. Save the verbatim text.
2. List the atomic claims. Mark the ones needing evidence.
3. Record the search steps the product shows. If it shows none, write UNKNOWN — do not ask the model.
4. Run each sub-question you can identify as an ordinary web search. Record the top ten results each time. This is your candidate proxy.
5. Check your own server logs for that engine's crawler over the previous fourteen days. Did it fetch your page at all, and what status did it get? This converts one INFERRED cell into an ACTUAL one.
6. Score each cited source and your own best page on the ten dimensions.
7. Assign a rejection reason to your page from the fixed list.
8. Write the unknowns section.

Expected output. A two-page document ending in one sentence of the form: "We are being lost at the ____ stage, and the evidence for that is ____."

Definition of done. Someone else can read it and reach the same conclusion from the same artefacts.

What can go wrong. The commonest error is filling section 3 with a model's account of its own search. That single mistake invalidates sections 4 through 8, because everything downstream inherits the fiction.

Turning it into work

A forensic audit that ends in understanding has failed. It should end in a queue. The mapping is mechanical:

Where you are lost	What that means	What to do
Not a candidate	Not in the index for this sub-question, or blocked	Chapters 8, 9 and 10 — access and coverage
Returned but not opened	Title and snippet did not promise an answer	Question-shaped headings; front-loaded specifics

Where you are lost	What that means	What to do
Opened but not used	The passage did not stand alone	Chapter 13 — restructure, do not rewrite
Used but not cited	Absorption. Real, and largely unfixable	Chapter 28 — say so; measure the proxies
Cited, but for the wrong claim	Your page is being read as answering something else	Retitle the section; make the claim explicit

COMMON MISTAKE

Do not run this audit on twenty prompts before running it on one. The first audit takes two hours and teaches you what your category's answers are made of. The twentieth takes twenty minutes. Teams that start by automating skip the step where they learn to tell a good source set from a bad one, and then cannot sanity-check the automation.

CHAPTER SUMMARY

- Tracing a citation tells you what won; forensics tells you what it beat, and the two point at different work
- Sources pass through candidate, returned, opened, used and cited — and you are being lost at exactly one of them
- Score the strongest source that lost on the same rubric as the winner, and assign a rejection reason from a fixed list
- Every section you cannot fill gets UNKNOWN, and a report with no unknowns is not finished
- The audit ends in a queue, not an understanding

CHECK YOURSELF

1. Why can a page that is never a candidate not be improved into a citation?
2. Which of the five populations can your own server logs turn from inferred into actual, and only for which pages?
3. What is the practical difference in remedy between "returned but not opened" and "opened but not used"?
4. Why is a forensic report with an empty unknowns section a warning sign rather than a good result?

CHAPTER 44

Claim-Level Validation

Being cited and being correctly cited are different outcomes

WHAT YOU'LL LEARN

- Why a citation next to a sentence does not mean the source supports the sentence
- A six-level scale for how well evidence carries a claim
- The parts of a sentence that most often outrun the evidence attached to it
- What to do when the answer cites you for something you did not say

Chapter 7 gave you a filter for evaluating claims made *about* GEO. This chapter applies the same scepticism one level down: to the individual claims inside an AI answer, and to whether the citation sitting beside each one actually carries it. This matters commercially, because a citation that misrepresents you is worse than no citation.

The gap between adjacency and support

Citations in generated answers are placed, not proved. A superscript beside a sentence indicates that a source was involved somewhere in producing that region of text. It does not assert that the source states the sentence.

In practice the relationship takes one of six forms.

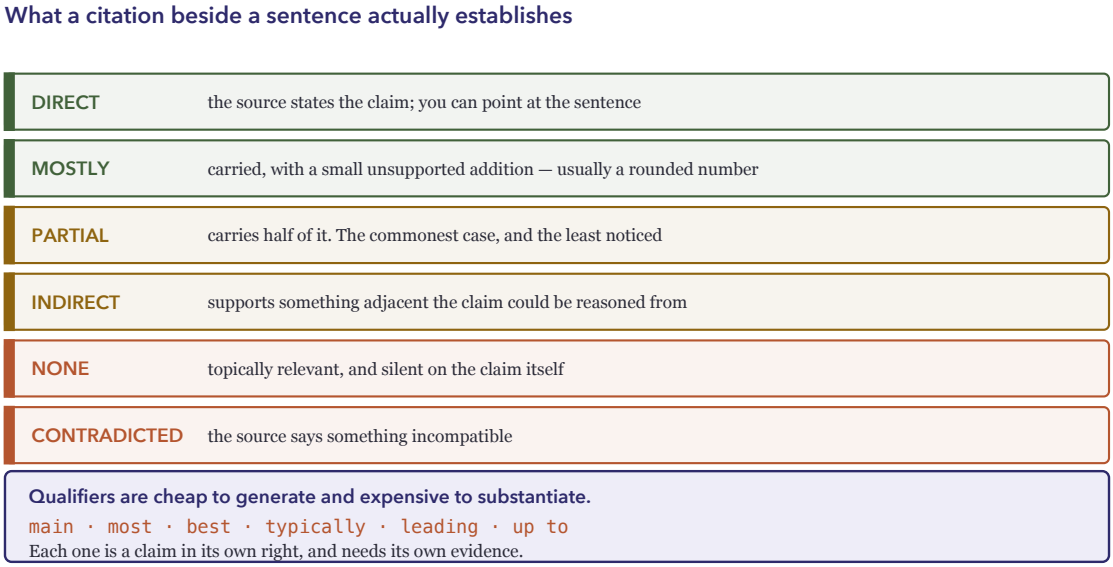


Figure 44.1 – Six relationships between a claim and the source cited beside it, from fully carried to actively contradicted.

DIRECT. The source states the claim, in substance, and you can point at the sentence.

MOSTLY SUPPORTED. The source carries the claim with a small unsupported addition — usually a rounded number or a tightened qualifier.

PARTIAL. The source carries half. "Burr alignment tolerance is the main cause of grind inconsistency in home grinders under £300" cites a page that establishes the tolerance figure but says nothing about the price band or about it being the main cause.

INDIRECT. The source supports something adjacent from which the claim could be reasoned, but does not state it.

NONE. The source is topically relevant and does not address the claim at all. Very common, and almost never noticed.

CONTRADICTED. The source says something incompatible. Rare, serious, and the strongest possible material for a correction request.

> KEY IDEA
>
> The commonest real-world case is **PARTIAL**, and the commonest cause is that the answer added a qualifier the source did not contain. Qualifiers are cheap to generate and expensive to substantiate: *main*, *most*, *best*, *typically*, *usually*, *leading*, *up to*. Each one is a claim in its own right and each one needs its own evidence.

What outruns the evidence

When you validate at claim level, the unsupported material clusters into predictable categories. Check for these specifically rather than reading for a general impression.

Category	Example of the overreach
Numbers	Source says 38–42; answer says 40
Dates	Source is undated; answer says “as of 2026”
Superlatives	Source says “a common cause”; answer says “the main cause”
Causal language	Source shows correlation; answer says “because”
Scope	Source studied UK retailers; answer says “retailers”
Recommendation	Source describes; answer advises
Comparison	Source describes one product; answer ranks it against another

The last two are the ones that damage brands. A source that neutrally describes your product becomes an answer that ranks it third. Nothing in the source supports the ranking; the ranking came from the model’s synthesis across sources, and the citation was attached to the descriptive part.

The hostile audit

Run this against answers that mention you — favourably as well as unfavourably. A flattering claim you cannot substantiate is a liability, because a competitor can dispute it and you will not be able to defend it.

IN PRACTICE

Auditing one answer at claim level

Objective. A claim-to-source map for one answer, with a support level on every row.

Inputs. One archived answer with its citations. Access to the cited pages. Forty minutes.

Procedure.

1. Split the answer into atomic claims — one assertion each. A sentence containing "and" is usually two claims.
2. For each claim, note the citation attached, or `none`.
3. Open the source. Find the passage that bears on the claim. Copy it verbatim into your record. If you cannot find one, that itself is the finding.
4. Classify: DIRECT, MOSTLY, PARTIAL, INDIRECT, NONE, CONTRADICTED.
5. For anything below MOSTLY, name the specific overreach from the table above.
6. Rewrite the claim so the evidence carries it. That rewrite is the correction you will send.

Expected output. A table of claims with support levels, and a shortlist of claims worth acting on.

Definition of done. Every row has a verbatim supporting passage or an explicit note that none exists.

How to verify it. Hand the map and the sources to a colleague without your classifications. Agreement above roughly eight rows in ten means your scale is being applied consistently; below that, your definitions have drifted.

When it is your page being misused

Two situations, two responses.

Your page is cited for a claim it does not make. Do not treat this as a win. It means the passage is being read as answering a question it does not answer. The fix is on your side: make the section's actual scope explicit, usually in the heading and the first sentence. "What burr alignment does to grind consistency" is being mistaken for "the main cause of grind inconsistency"; saying which one you are answering closes the gap.

Your page is cited accurately, and the surrounding answer is wrong. This is a synthesis error and you cannot fix it from your side. Record it, note the engine and date, and check whether the wrong material traces to another cited source — that source is where the correction work goes. Chapter 47 is the runbook.

COMMON MISTAKE

Do not send a correction request built on a PARTIAL classification without the verbatim passage in hand. Publishers respond well to "your article says 38–42 and the answer says 40, here is the sentence" and badly to "the AI is misrepresenting your article." One is a five-minute fix for them; the other is an argument.

CHAPTER SUMMARY

- A citation beside a claim shows involvement, not support; the two are routinely conflated
- Classify support as DIRECT, MOSTLY, PARTIAL, INDIRECT, NONE or CONTRADICTED, with a verbatim passage on every row
- Overreach clusters in numbers, dates, superlatives, causal language, scope, recommendation and comparison
- Being cited for a claim you did not make is a structural problem with your page, not a win
- Correction requests need the verbatim passage; without it you are asking a publisher to take your word

CHECK YOURSELF

1. What does a citation next to a sentence actually establish, and what does it not?
2. Why is PARTIAL the commonest classification in practice, and what usually causes it?
3. Your product page is cited for the claim that it is the best grinder under £300. The page makes no such claim. Why is this a problem, and where is the fix?
4. Why does a correction request need the verbatim source passage rather than a description of the error?

CHAPTER 45

Source Lineage and Independence

Ten cited domains are not ten pieces of evidence

WHAT YOU'LL LEARN

- Why source concentration and source independence are different measurements
- How to trace a cited page back to the evidence it actually rests on
- The five ways ten citations collapse into one original claim
- What to do when your category's entire evidence base traces to a single document

Chapter 39 made source concentration one of the three numbers worth reporting: which third-party domains the engines lean on for your category. It is a genuinely actionable metric. It is also, on its own, misleading — because counting domains counts publishers, and publishers are not evidence.

The collapse

An answer cites ten domains. It looks like a well-corroborated answer. Trace each one back and the picture changes:

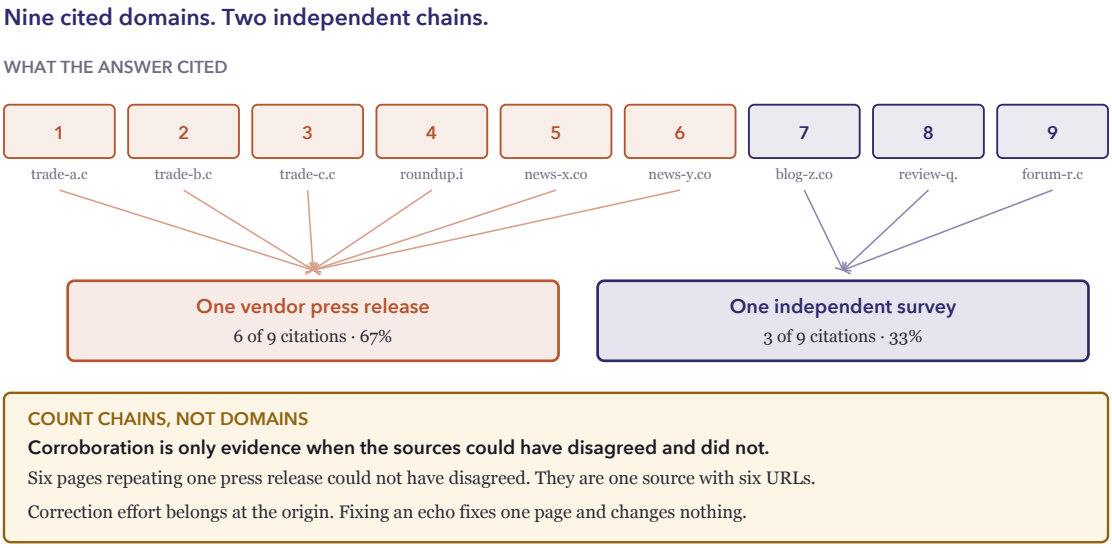


Figure 45.1 – Ten cited domains, traced to their origins. Two independent chains, not ten.

The five mechanisms that produce this:

- **Syndication.** The same article appears under several mastheads. Different domains, one newsroom, one set of facts.
- **Press-release propagation.** Six trade titles rewrite one announcement. The number in all six is the number the company chose to publish.
- **A single study.** Everyone cites the same survey. If the survey's method was weak, the weakness is now in ten places and looks like consensus.
- **Common ownership.** Several titles, one publisher group, one editorial standard, sometimes one shared research desk.
- **The upstream error.** A wrong specification on one marketplace listing gets copied into comparison articles, which are then cited independently of each other.

> KEY IDEA
>
> Corroboration is only evidence when the sources are **evidentially independent** — when they could have disagreed and did not. Ten pages repeating one press release could not have disagreed. They are one source with ten URLs, and treating them as ten is the single most common way a source-concentration report overstates how solid a category's evidence is.

Tracing a chain

For each cited page, record where the fact came from before it arrived there.

Field	What you record
Immediate source	The page that was cited
Original evidence	The document the fact actually comes from
Primary or secondary	Whether the cited page originated the fact
Ownership group	Who publishes it
Syndicated	Whether the text appears elsewhere verbatim
Evidence chain ID	Your own grouping — sources sharing an origin share an ID
Stakeholder	Whose interest the evidence serves: vendor, buyer, regulator, independent
Shares an origin with	The other source IDs in the same chain

Finding the original is usually three steps. Search a distinctive phrase from the cited page in quotation marks — syndication surfaces immediately. Follow the page's own outbound links to whatever it cites. Then check whether that document is itself reporting something else.

Then count differently:

- **Unique domains.** What the naive metric gave you.
- **Unique publishers.** Domains collapsed by ownership.
- **Independent evidence chains.** The number that matters.
- **Share that is primary.** How much of the answer rests on origin documents rather than on reporting.
- **Share dependent on the single largest chain.** Your concentration risk, properly measured.
- **Stakeholder mix.** Whose account of the category is being told.

IN PRACTICE

Lineage on one answer

Objective. Convert a domain count into a chain count.

Inputs. One archived answer with at least six citations. Ninety minutes.

Procedure.

1. List every cited URL.
2. For each, search one distinctive twelve-word phrase from it in quotes. Note every other domain carrying the same text.
3. Follow each page's own citations to the document the fact originates in.
4. Group sources sharing an origin into chains. Give each chain an ID.
5. Count: domains, publishers, chains, primary share, largest-chain share.
6. Record the stakeholder behind each chain.

Expected output. Two numbers you can say out loud: "The answer cites nine domains. It rests on three independent chains, and 61% of it traces to one vendor's own white paper."

What can go wrong. Stopping at the first outbound link. Many trade articles cite another trade article, which cites the release. Keep going until you reach a document that is not reporting anyone else.

What to do with the finding

The lineage picture points at cheaper work than the domain count does.

One chain dominates. Your effort belongs at the origin, not at the ten pages downstream. Correcting the original document propagates; correcting one of its echoes does not.

Every chain is vendor-supplied. The category has no independent evidence, which is an opening rather than a problem. Original data from you becomes a new chain — and Chapter 16 explains why a new chain is the only content position competitive erosion cannot reach.

Your own material is one chain appearing as five. Your press release, your blog post about the release, and three trade pickups are a single chain. Reporting it as five sources

overstates your position to your own stakeholders, which is the version of this error that damages you internally.

COMMON MISTAKE

A category whose evidence base is one chain is fragile in both directions. If the origin document is wrong, every assistant is confidently wrong in the same way. If it is corrected, the correction takes months to work through the downstream copies — and until it does, the engines will be citing pages that contradict the source they came from.

CHAPTER SUMMARY

- Counting cited domains counts publishers; counting evidence chains counts evidence
- Syndication, press releases, a single study, common ownership and copied errors all collapse many domains into one chain
- Corroboration only counts when the sources could have disagreed
- Report chains, primary share and largest-chain share alongside domain concentration
- Correction effort belongs at the origin of a chain, never at its echoes

CHECK YOURSELF

1. Why is “cited by ten domains” not evidence that a claim is well corroborated?
2. What single search technique surfaces syndication fastest, and what do you search for?
3. Your category’s evidence traces almost entirely to one vendor white paper. Why is that an opportunity, and what would you publish?
4. Why does reporting your own press coverage as five independent sources damage you rather than help you?

CHAPTER 46

Counterfactual Testing

Six controlled reruns that show whether your visibility is stable or lucky

WHAT YOU'LL LEARN

- Why one answer is an anecdote even when the prompt set is fixed
- Six reruns that isolate what a citation actually depends on
- How to tell a robust position from one resting on a single wording
- What each result tells you to do next

Chapter 31 taught intervention testing: change one thing, hold everything else, remeasure. That answers "did my change work?" This chapter answers a different question — "how fragile is the position I already have?" — and it needs a different kind of experiment, run against the answer rather than against your site.

The question these tests answer

You are cited for a query. Good. Now: is that citation a property of your content, or of one particular wording, one product mode, one source being temporarily absent?

The difference decides where your next month goes. A citation that survives paraphrase, mode changes and the removal of a rival is a position worth defending. A citation that appears only in one phrasing on one surface is noise you are about to report as a win.

Six reruns. What disappears tells you what you depend on.

PARAPHRASE five equivalent wordings	cited in 1 of 5?	→	a coincidence
MODE standard / search / research	search only?	→	no training presence
FRESHNESS only sources after [date]	you vanish?	→	cited by default
PRIMARY ONLY only first-party sources	you vanish?	→	a summary, not a source
EXCLUDE DOMAIN rebuild without the incumbent	you appear?	→	second in line
GEOGRAPHY answer for [country]	you vanish?	→	a sampling artefact

These probe the evidence landscape. They do not reconstruct the original decision.
Label every result INFERRED. Fresh conversation for every run, or the model converges on its own last answer.

Figure 46.1 — Six reruns, what each one holds constant, and what a change in the citation set tells you.

The six reruns

Each is run against the *answer*, by re-asking with an added constraint. You are not changing your site. You are probing what the answer depends on.

1. Prompt-paraphrase sensitivity. Ask five semantically equivalent versions of the same question. Compare cited domains, cited pages, named brands and recommendations across all five.

> *Interpretation. If your citation appears in one of five, you do not have a position; you have a coincidence. Three or more of five is a real foothold. Report the fraction, never the best run.*

2. Mode sensitivity. The same question in the product's standard mode, its search-enabled mode, and its research or deep mode.

> *Interpretation. Appearing only in the retrieval-heavy modes means you are winning retrieval but absent from training-era knowledge. Appearing only in the standard mode means the reverse, and no amount of content work this quarter will change it.*

3. Freshness control. "Answer using only sources published or updated after [date]."

> *Interpretation. If your citation disappears when recent sources are required, your page is being retrieved because nothing newer exists.*

> *That is a decaying position with a visible expiry.*

4. Primary-source-only. "Answer using only primary or first-party sources."

> *Interpretation. Reveals whether you are cited as an origin or as a convenient summary. Being dropped here is the strongest argument for Chapter 16's original-data work.*

5. Domain exclusion. "Answer this without using [dominant domain]. Rebuild the evidence base rather than removing the citation."

> *Interpretation. Shows who is next in line. If you appear only once the incumbent is removed, you are the second-best source and the gap is usually one specific missing fact.*

6. Geographic control. "Answer for a user in [country], preferring locally relevant sources."

> *Interpretation. Separates a genuinely international position from an artefact of where you happen to be sampling. Chapter 18 explains why the thinner market is usually the easier one.*

> **COMMON MISTAKE**
>
> These constraints change what the assistant does; they do not reveal what it did originally. A domain-exclusion rerun produces a new answer under a new instruction. It is a probe of the evidence landscape, not a reconstruction of the first decision. Label the results INFERRED, and never write "the model would have cited X instead" — write "with that domain excluded, X was cited."

Reading the pattern

One rerun rarely means much. The combination does.

Pattern across the six	What it usually means	Next move
Stable in all six	A genuine, defensible position	Protect it; check quarterly
Fails paraphrase only	One lucky wording	Broaden the page's question coverage

Pattern across the six	What it usually means	Next move
Fails freshness only	Cited by default, ageing out	Refresh with dated, current specifics
Fails primary-only	You are a summary, not a source	Publish original data
Appears only on exclusion	Second in line	Find the one fact the incumbent has
Fails geography	Sampling artefact	Re-baseline per market
Fails everything but one run	Not a position at all	Do not report it as visibility

IN PRACTICE

The robustness pass

Objective. Establish whether your three most valuable citations are stable.

Inputs. Three prompts where you are currently cited. Two engines. Half a day.

Procedure. For each prompt, run all six reruns on both engines, logged out, in fresh conversations. Record for every run: the constraint used, the full citation list, and whether you appeared.

Expected output. A grid of three prompts by six reruns, with a fraction in each cell.

Definition of done. You can state, for each of the three, the specific condition under which your citation disappears.

What can go wrong. Running the paraphrases in one conversation. The model carries the earlier answers as context and converges on its own previous output, which manufactures a stability that is not there. Fresh conversation every time.

NOTE

This suite is worth running once a quarter on your top prompts, not monthly on all of them. It is diagnostic, not a metric. The monthly number stays the one from Chapter 39; this tells you how much to trust it.

CHAPTER SUMMARY

- A fixed prompt set controls the question, not the fragility of the answer
- Six reruns — paraphrase, mode, freshness, primary-only, domain exclusion, geography — isolate what a citation depends on
- Report the fraction of paraphrases in which you appear, never the best single run
- The reruns probe the evidence landscape; they do not reconstruct the original decision, and must be labelled INFERRED
- Run the suite quarterly on your most valuable prompts, not monthly on all of them

CHECK YOURSELF

1. Why does a citation that appears in one of five paraphrases not count as a position?
2. What does it tell you if you are cited in standard mode but not in search-enabled mode?
3. Why must the paraphrase runs happen in separate conversations?
4. Why is “the model would have cited us instead” an indefensible sentence to write after a domain-exclusion run?

CHAPTER 47

The Correction Runbook

What to do in the hour after you find an assistant saying something wrong about you

WHAT YOU'LL LEARN

- A twelve-step procedure from discovery to closure
- Which of the five causes you are dealing with, and how to tell
- Templates for the three requests you will actually send
- What to record so the incident can be closed rather than merely forgotten

Chapter 24 explained why answers go wrong about a company and gave you the correction ladder. This chapter is the operational version: the procedure you follow, the evidence you capture, the emails you send and the log that tells you whether it worked. It exists because the discovery usually happens at an inconvenient moment and the instinct — publish a rebuttal on your own site — is almost always the wrong first move.

Twelve steps

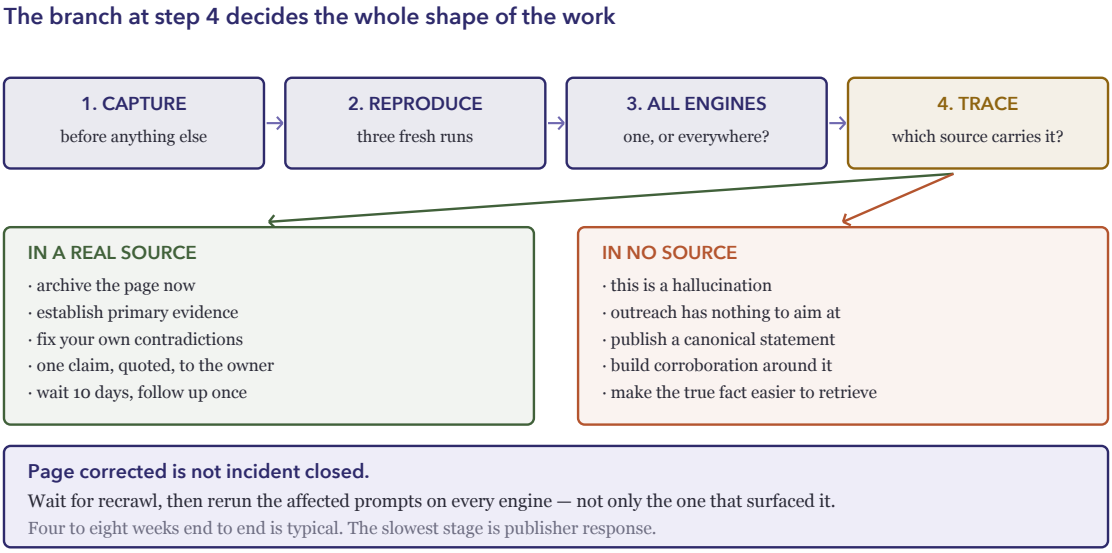


Figure 47.1 – From the wrong answer to a closed incident, with the branch that decides everything.

1. Capture before anything else. Screenshot the answer with the prompt visible. Save the verbatim text, the citation list, the engine, the mode, the date and your location. The

answer may not reproduce tomorrow, and an incident you cannot show is an incident you cannot escalate.

2. Reproduce. Run the same prompt twice more in fresh conversations. Note whether the error appears in all three. An error appearing once in three is a different — and lower — priority than one appearing every time.

3. Check the other engines. Same question, all four. An error confined to one engine points at one source that engine favours. An error present everywhere points at something upstream that everyone reads.

4. Trace to source. Which cited page carries the wrong fact? Use Chapter 44's claim-level method: find the verbatim passage. In the majority of cases the wrong fact is genuinely present in a real source and the assistant reported it faithfully.

5. Archive the source. Save a PDF or full-page capture of the offending page *now*, with the date. If the publisher corrects it later without a visible note, your evidence of what it said disappears with the edit.

6. Establish the correct evidence. What is the primary, first-party document that states the truth? If one does not exist on your own site in a citable form, that is your first fix regardless of anything else.

7. Classify the cause. Stale source, entity confusion, self-contradiction, a loud negative, or hallucination. The five have five different remedies and picking wrong wastes weeks.

8. Fix your side first. If your own site contradicts itself, or lacks the canonical statement, no amount of outreach will hold. This is also the only step entirely within your control.

9. Contact the owner. One page, one specific correction, with the evidence attached. Templates below.

10. Wait, then follow up once. Ten working days. One follow-up. Publishers correct factual errors at a high rate when the request is specific and low-effort; they ignore complaints.

11. Verify propagation. Page corrected is not incident closed. Wait for recrawl — check your own logs if the page is yours, or the cache date if not — then rerun the affected prompts on *every* engine, not only the one that surfaced it.

12. Close or escalate. Closed means the answer is corrected on the engines you checked, and you have the dated evidence. Escalate means the source will not correct, and you move to the next rung of Chapter 24's ladder: corroboration, then platform feedback.

KEY IDEA

The branch at step 4 decides the whole shape of the work. **If the wrong fact is in a real source, this is an outreach problem.** If it is in no source, this is a hallucination and outreach has nothing to aim at — your remedy is a canonical, well-structured first-party statement and enough corroboration that the correct fact is easier to retrieve than the invented one.

The three requests

A factual correction to a publisher. Short, specific, evidenced, and easy to act on.

> *Subject: Correction — [product name] specification in [article title]*
>
> *Hello [name],*
>
> *Your article "[title]" ([URL]), in the section on [section], states that [exact quoted claim]. The current figure is [correct fact].*
>
> *The source is our specification page: [URL], last updated [date]. I have attached the datasheet in case it is useful.*
>
> *This is currently being picked up by AI assistants answering questions about [category], which is why I am asking rather than leaving it.*
>
> *Happy to supply anything else you need.*
>
> *[Name, role, direct contact]*

Three properties make this work: one claim rather than a list, the exact quoted wording so the writer can find it in seconds, and a primary source they can verify without contacting you again.

An outdated marketplace or directory listing. Same shape, plus the identifiers the platform indexes on — listing ID, GTIN, registered company number. Most listing corrections fail because the request does not include the field the platform uses to find the record.

A review-platform profile. Correct the factual fields — category, pricing model, company details, integrations. Do not attempt to correct opinions, and do not respond to a factual-correction ticket with an argument about a review score; it converts a five-minute administrative fix into a dispute.

> COMMON MISTAKE
>
> Never ask a publisher to remove true criticism. It will not be removed, the request itself is frequently published, and that thread is highly citable. Chapter 24 makes the general case; this is the operational form of it. Correct facts, and only facts.

The incident log

Without a log, corrections become a folder of half-remembered emails. Thirteen fields close the loop.

Field	Purpose
Issue ID	Something to refer to
Incorrect claim	Exact wording as generated
Engine and mode	Where it appeared
Prompt and date	Reproduction instructions
Reproduced	How many of three runs
Source URL	Where the fact came from
Cause class	One of the five
Correct evidence	Your primary document
Owner contacted	Person and channel
Date contacted	Starts the ten-day clock
Page corrected	Date observed

Field	Purpose
Recrawl observed	Log line or cache date
Answer corrected	Date, and on which engines

IN PRACTICE

Running one incident to closure

Objective. Close a single real incident and time each stage, so you can set expectations on the next one.

Inputs. One prompt where an engine states something factually wrong about you. Two hours spread over six weeks.

Procedure. Steps 1 to 12 above, recording the date at every stage.

Expected output. A closed row in the incident log, and four real durations: discovery to contact, contact to page corrected, page corrected to recrawl, recrawl to answer corrected.

Definition of done. The prompt now returns a correct answer on at least the engine where the error was found, and you have dated evidence of both states.

Why the durations matter. The next time a stakeholder asks “how long will this take”, you will have your own numbers rather than the industry’s. Four to eight weeks end to end is typical; the slowest stage is almost always publisher response, not propagation.

CHAPTER SUMMARY

- Capture the answer before you do anything else; it may not reproduce
- Trace the wrong fact to a real source before assuming a hallucination — usually there is one
- Fix your own contradictions first, because outreach will not hold without a canonical statement
- Send one claim, quoted exactly, with a primary source attached; that is what gets corrected
- The incident is closed when the answer is corrected across engines and you have dated evidence, not when the email is sent

CHECK YOURSELF

1. Why is the first step capture rather than diagnosis?
2. What does it tell you if an error appears on one engine only, versus on all four?
3. Why does the branch at step 4 — is the wrong fact in a real source — change the entire remedy?
4. Why is “page corrected” not the same as “incident closed”, and what two things still have to happen?

CHAPTER 48

The Evidence Archive

Making your findings reproducible six months after you stop remembering them

WHAT YOU'LL LEARN

- Why AI-search evidence decays faster than the evidence in any other marketing discipline
- What to preserve for every audited answer, and in what structure
- The metadata that makes a run reproducible rather than merely recorded
- How much of this to do at each scale of operation

Everything in Part X produces findings that rest on artefacts. This chapter is about keeping the artefacts. It is unglamorous and it is the difference between an analysis you can defend in November and a spreadsheet of assertions you can no longer support.

Why this decays faster than you expect

In classical SEO the evidence persists. A ranking screenshot is corroborated by third-party trackers; the page you ranked is still there.

In AI search, seven things move underneath you:

- Answers are non-deterministic; the same prompt may not reproduce.
- Source pages change, often without a visible correction note.
- Publication dates change, sometimes silently, when a CMS re-stamps a page.
- URLs redirect, and the target may not be what was cited.
- Source panels are redesigned or removed between product versions.
- Models are replaced. The one that produced your baseline may be retired.
- Citations can point at dynamically assembled pages that no longer contain the passage.

KEY IDEA

Your finding is only as durable as its weakest artefact. An analysis that depends on a source page you did not archive has an expiry date set by someone else's CMS — and you will not be told when it passes.

What to preserve, per answer

One run, fully preserved

```
run-001/
  — prompt.txt          the exact question, character for character
  — answer.txt          verbatim text, so it is searchable
  — answer.png          the answer as displayed, with its source panel
  — metadata.json       product, model, mode, account, location, time + offset
  — sources/
      — S1.pdf          the page as it appeared — not as it appears now
      — S1.json         url, canonical, published + updated dates, retrieved-at
  — notes.md            what you concluded, and an explicit unknowns list
```

manifest.csv sits one level up: one row per run, metadata repeated.
That is what lets you filter and count without opening anything.

Archive the source when you capture, not when you decide it matters.
The days in between are exactly when a publisher quietly fixes the page you were going to quote.

Figure 48.1 – One run, fully preserved: what goes in each file and why each one is needed.

For each audited answer:

- Screenshot or PDF of the answer as displayed, including the source panel
- The verbatim answer text, separately, so it is searchable
- The exact prompt, character for character
- Timestamp with timezone, and the location the session ran from
- Product, model, mode, and whether search was enabled
- The citation list as URLs, not as domains
- A snapshot of each cited page — PDF or full HTML
- The canonical URL of each source, where it differs from the cited one
- Each source's publication and last-updated date as displayed
- The specific passage you believe was used, quoted
- Analyst notes: what you concluded and what you could not determine
- A note on whether anything captured contains personal data

The two most often skipped are the source snapshots and the timezone. The first is what fails you when a publisher edits quietly. The second is what makes a "monthly" comparison meaningless when half the runs happened in a different offset and the answers reflect a different news cycle.

A structure that survives handover

Flat folders of screenshots named `Screenshot 2026-08-07 at 14.22.11.png` are not an archive. The following is enough, and takes no tooling:

```
client/
  2026-08-baseline/
    manifest.csv          one row per run, all metadata
    run-001/
      prompt.txt
      answer.txt
      answer.png
      metadata.json       product, model, mode, time, location
      sources/
        S1.pdf            snapshot as it appeared
        S1.json           url, canonical, dates, retrieved-at
        S2.pdf
        S2.json
      notes.md            conclusions, and what stayed unknown
    run-002/
      ...
```

The `manifest.csv` at the top is what makes this usable: one row per run with the metadata repeated, so you can filter and count without opening anything. The per-run folders hold the evidence that backs each row.

IN PRACTICE

Archiving one baseline

Objective. Produce one fully archived run you could hand to someone else.

Inputs. One prompt, one engine, forty minutes.

Procedure.

1. Create the folder structure above for a single run.
2. Run the prompt logged out, in a fresh conversation.
3. Save the answer as text and as an image, including the sources panel.
4. Write `metadata.json` by hand: product, model if shown, mode, search enabled yes/no/unknown, account state, country, language, device, timestamp with offset.
5. For each cited source, save a PDF and a small JSON with the URL, the canonical URL, the displayed publication and update dates, and the time you retrieved it.
6. Write `notes.md`: what you concluded, and an explicit unknowns list.

Expected output. One `run-001` folder that a colleague could audit without asking you a single question.

Definition of done. Hand it to someone and ask them to state what the answer said, which sources it used, and one thing you could not determine. If they can, it is complete.

How much of this to do

The full structure is right for client work, regulated categories, and anything that might be disputed. For a solo practitioner tracking their own site it is too much, and the useful subset is small.

Scale	Preserve
Solo, own site	Verbatim answer, prompt, date with offset, citation URLs, one screenshot
In-house team	The above, plus product/model/mode metadata and snapshots of sources you intend to contact
Agency or client work	The full structure, because you will be asked to show your working
Regulated or disputed	The full structure plus content hashes, and archive on capture rather than on decision

COMMON MISTAKE

Capture the source snapshot at the moment you find the answer, not when you decide the finding matters. The gap between those two moments is typically several days, and it is precisely the window in which a publisher quietly fixes a page — leaving you asserting that it once said something you can no longer show.

CHAPTER SUMMARY

- AI-search evidence decays in seven distinct ways, most of them invisible while they happen
- Preserve the answer, the prompt, full run metadata, and a snapshot of every cited source
- A manifest row per run plus a folder per run is enough structure, and needs no tooling
- Timezone and source snapshots are the two most-skipped fields and the two that fail you first
- Archive on capture, not on decision, because pages get quietly corrected in between

CHECK YOURSELF

1. Name three ways the evidence behind an AI-visibility finding can decay without anyone telling you.
2. Why does a citation list of domains rather than full URLs make an archive much less useful?
3. What does the manifest file give you that the per-run folders do not?
4. Why is archiving a source at the moment of capture rather than at the moment of decision the more important habit?

PART XI

Playbooks

The exact files, markup, commands and sequences — including the parts that only apply to your industry, and the piece of work you should finish the book by producing.

CHAPTER 49

Copy-Paste Implementation

The actual files: robots.txt, the four schema blocks, and the log command

WHAT YOU'LL LEARN

- A complete `robots.txt` implementing the book's default policy, line by line
- The four structured-data blocks that carry nearly all the value, ready to adapt
- Where each one goes in the four platforms most readers are on
- One command that answers whether AI crawlers are reaching you

Earlier chapters told you what to do. This one is what to paste. Everything here has bracketed placeholders you replace and a plain-English explanation of what each part does, because pasting configuration you do not understand is how sites end up blocking the crawlers they meant to admit.

The robots.txt

Chapter 8 argued for one default: refuse training, permit retrieval. Here it is in full.

```
# Training crawlers — content used to train future models.
User-agent: GPTBot
User-agent: ClaudeBot
User-agent: Google-Extended
User-agent: Applebot-Extended
User-agent: meta-externalagent
User-agent: CCBot
Disallow: /

# Retrieval crawlers — fetch pages to cite in answers now.
User-agent: OAI-SearchBot
User-agent: ChatGPT-User
User-agent: PerplexityBot
User-agent: Perplexity-User
User-agent: Claude-SearchBot
User-agent: Applebot
Allow: /

# Everything else, including Googlebot and Bingbot.
User-agent: *
Allow: /
Disallow: /cart/
Disallow: /checkout/
```

```
Disallow: /*?sort=
Disallow: /*?filter=

Sitemap: https://[your-domain]/sitemap.xml
```

How this works. A `robots.txt` file is a set of blocks. Each block names one or more crawlers by their user-agent string and then gives rules. A crawler reads the file, finds the block that names it, and obeys only that block — so a crawler named in the first group never reaches the `User-agent: *` group at the bottom.

What each part does. The first group refuses the training crawlers: `Disallow: /` means "no part of this site." The second group admits the retrieval crawlers with `Allow: /`. The third is everyone else, with two kinds of exclusion: pages that should never be indexed at all (`/cart/`, `/checkout/`) and URL patterns that generate near-infinite duplicates. The `*` in `/*?sort=` is a wildcard, so that line means "any URL containing `?sort=`". Blocking those protects crawl budget, which Chapter 2 explained matters on large sites and not on small ones.

> COMMON MISTAKE
>
> Two failure modes account for most damage. **Google-Extended does not affect AI Overviews** — it controls Gemini training only, so blocking it does not remove you from Google's AI answers, and many people believe it does. And `robots.txt` **is a request, not a wall**: it states intent. Chapter 41's server logs state what happened. Publish this file, then check the logs, because a CDN rule can override every line of it silently.

The four schema blocks

Structured data states facts about a page in a form a machine does not have to infer. These four carry nearly all the value. Each goes inside a `<script type="application/ld+json">` tag in the page's HTML.

Organization — one per site, on the homepage. It is what entity resolution attaches to.

```
{
  "@context": "https://schema.org",
  "@type": "Organization",
  "@id": "https://[domain]/#organization",
  "name": "[Legal or trading name]",
  "url": "https://[domain]/",
  "logo": "https://[domain]/logo.png",
  "description": "[One sentence. The same one everywhere.]",
  "foundingDate": "[YYYY]",
  "sameAs": [
    "https://www.linkedin.com/company/[handle]",
    "https://en.wikipedia.org/wiki/[article]",
```

```
    "https://www.wikidata.org/wiki/[Qid]"
  ]
}
```

What the parts do. `@id` gives the organisation a stable identifier that other blocks on your site can point at, so a person and an article can both be tied to the same entity rather than to a name that happens to match. `sameAs` is the important one: it lists other places the same entity is described. That is the corroboration mechanism from Chapter 14, stated explicitly instead of left to be inferred.

Person — for each named author. Chapter 15 called this the most under-used entity signal available.

```
{
  "@context": "https://schema.org",
  "@type": "Person",
  "@id": "https://[domain]/authors/[slug]#person",
  "name": "[Full name]",
  "jobTitle": "[Title]",
  "worksFor": { "@id": "https://[domain]/#organization" },
  "knowsAbout": ["[topic]", "[topic]", "[topic]"],
  "sameAs": ["https://www.linkedin.com/in/[handle]"]
}
```

`worksFor` points at the `@id` you defined above rather than repeating the company name, which is what turns two separate facts into one connected graph. `knowsAbout` is where you state the topics this person is an authority on — three to five, matching what they actually publish about.

Article — on every substantial content page.

```
{
  "@context": "https://schema.org",
  "@type": "Article",
  "headline": "[The H1, verbatim]",
  "datePublished": "[YYYY-MM-DD]",
  "dateModified": "[YYYY-MM-DD]",
  "author": { "@id": "https://[domain]/authors/[slug]#person" },
  "publisher": { "@id": "https://[domain]/#organization" }
}
```

`dateModified` is the field that matters most here and the one most often wrong. It should change when the substance changes and not when the template does. A site that re-stamps every page nightly is telling engines that nothing on it is ever stable.

Product with Offer — for e-commerce. Chapter 25 explained why `gtin` is the field that connects your listing to the same product elsewhere.

```
{
  "@context": "https://schema.org",
  "@type": "Product",
  "name": "[Product name]",
  "gtin13": "[13-digit barcode]",
  "brand": { "@type": "Brand", "name": "[Brand]" },
  "offers": {
    "@type": "Offer",
    "price": "[189.00]",
    "priceCurrency": "GBP",
    "availability": "https://schema.org/InStock",
    "url": "https://[domain]/[product-url]"
  }
}
```

COMMON MISTAKE

Markup must match what a visitor sees. Schema stating a price of £189 on a page displaying £229 is not an optimisation; it is a discrepancy that damages trust in every fact on the page. Validate with the Rich Results Test, then read the page and check the numbers agree.

Where it goes, by platform

Platform	Schema	Rendering fix
WordPress	An SEO plugin handles Organization and Article; add Person and Product via a custom field or a code snippet in the theme header	Most themes render server-side already; the risk is a page builder that loads content by JavaScript
Shopify	`theme.liquid` for Organization; `product.liquid` for Product, using the Liquid variables the theme already has	Reviews are the usual casualty — most review apps inject after load. Ask the vendor for a server-rendered or metafield option
Webflow	Site Settings › Custom Code › Head for Organization; a page-level embed for the rest	Native content is server-rendered; CMS-driven collection lists usually are too
Custom or headless	In the server-rendered head	The whole question. If the framework hydrates on the client, the crawler sees the shell — this is the single highest-value thing to verify

> NOTE
>
> There is one universal rule underneath that table: **structured data injected by a tag manager is invisible to crawlers that do not run JavaScript**, which is most of the AI crawlers. It is a convenient deployment route and it defeats the purpose. Put it in the page.

The one command

If you have shell access to your logs, this answers the question that matters most in this book.

```
grep -Ei 'oai-searchbot|perplexitybot|claudebot|gptbot' access.log \
  | awk '{print $9, $12}' | sort | uniq -c | sort -rn | head -20
```

What it does, piece by piece. `grep -Ei` searches the log for lines containing any of those crawler names, case-insensitively — the `|` between them means "or". The `\` at the end of the first line just continues the command onto the next. `awk '{print $9, $12}'` prints the ninth and twelfth whitespace-separated fields of each matching line, which in a standard combined log format are the HTTP status code and the user agent. `sort | uniq -c` groups identical pairs and counts them, `sort -rn` puts the biggest counts first, and `head -20` shows the top twenty.

What you are looking for. A healthy result is mostly `200`. A column of `403` means a crawler is being refused — and Chapter 41 explained that this is invisible from every other vantage point. If the field positions are wrong for your log format, print the whole line first with `grep -Ei '...' access.log | head -3` and count the columns.

IN PRACTICE

The thirty-minute implementation pass

Objective. Ship the access policy and the entity block, and verify both.

Inputs. Access to `robots.txt`, the site head, and either logs or a hosting support ticket. Thirty minutes.

Procedure.

1. Publish the `robots.txt` above with your sitemap URL substituted.
2. Add the Organization block to the homepage head, with real `sameAs` URLs.
3. Run the Rich Results Test against the homepage.
4. Run the log command, or ask hosting for fourteen days of raw access logs filtered to those user agents.
5. Record the status-code distribution.

Definition of done. The test validates, and you have seen at least one `200` to a retrieval crawler in the logs.

What can go wrong. A CDN bot-management rule sitting above your `robots.txt`, refusing crawlers the file admits. The logs are the only place this shows up, which is why step 5 is not optional.

CHAPTER SUMMARY

- Refuse training crawlers, permit retrieval crawlers, and state the difference explicitly in `robots.txt`
- Google-Extended controls Gemini training only and does nothing about AI Overviews
- Organization, Person, Article and Product carry nearly all the structured-data value; connect them with `@id` rather than repeating names
- Schema deployed via a tag manager is invisible to crawlers that do not run JavaScript
- `robots.txt` states intent; the server log states what happened, and only one of them is evidence

CHECK YOURSELF

1. Why does a crawler named in the first `robots.txt` block never see the rules in the `User-agent: *` block?
2. What does `sameAs` do, and which mechanism from Chapter 14 does it make explicit?
3. Why is deploying structured data through a tag manager self-defeating for AI search?
4. In the log command, what are you hoping the status-code column shows, and what does a run of 403s mean?

CHAPTER 50

The Page Teardown

One page, before and after, with the markup and the scoring sheet

WHAT YOU'LL LEARN

- What a retrieval-hostile page actually looks like in markup
- The same page rewritten, with every change explained
- A rubric a writer can apply without understanding retrieval
- How to run this across a team so the results are consistent

Chapter 13 gave you the method: answer first, self-contained sections, question-shaped headings. This chapter applies it to one real page and shows the markup on both sides, because a method you cannot see applied is a method most writers will apply differently from each other.

The page before

A product page. It ranks. It is never cited.

```
<h1>Kettleworth K3 Grinder</h1>
<h2>Overview</h2>
<p>The K3 represents the culmination of our design philosophy,
bringing together precision engineering and thoughtful ergonomics
in a package designed for the discerning home barista. Building on
the success of the K2, it delivers a grinding experience that
punches well above its price point. Whether you are pulling
espresso or brewing filter, the K3 has been engineered to deliver
consistency where it matters. It uses 40mm conical burrs.</p>
<h2>Features</h2>
<ul><li>Precision engineered</li><li>Compact footprint</li></ul>
<h2>Specifications</h2>
<div id="spec-tab"></div>
```

Four separate problems, and only one of them is writing quality.

The headings are labels, not questions. “Overview”, “Features”, “Specifications” match no sub-query anyone will ever issue. Chapter 4’s fan-out turns “which grinder for espresso under £300” into questions like “what burr size does it use” and “is it consistent enough for espresso” — and nothing here signals that it answers either.

The answer is in sentence six. "It uses 40mm conical burrs" is the only retrievable fact in the paragraph and it arrives after five sentences of positioning. A chunking system that splits at heading boundaries produces one passage here, and its first hundred words say nothing.

The section cannot stand alone. "Building on the success of the K2" only means something if you already read the site. Lifted out, the passage refers to things it does not contain.

The specifications are not in the HTML. `<div id="spec-tab"></div>` is an empty container filled in by JavaScript. Chapter 9's test — View Page Source, search for your own numbers — returns nothing. Most AI crawlers see exactly what is printed above.

The page after

```
<h1>Kettleworth K3 Grinder</h1>

<h2>What burr type and size does the K3 use?</h2>
<p>The K3 uses 40mm conical burrs, hardened steel, with 38 stepped
grind settings. Conical burrs at this diameter are the usual choice
for home grinders handling both espresso and filter, because they
generate less heat at low speeds than flat burrs of equivalent
size.</p>

<h2>Is the K3 consistent enough for espresso?</h2>
<p>Yes, within a tolerance of 15 microns across the espresso range,
measured across 40 doses at setting 8. Grind consistency at
espresso fineness is where inexpensive grinders usually fail; the
K3 holds tolerance because the burr carrier is machined rather than
cast, which limits alignment drift.</p>

<h2>What does the Kettleworth K3 cost, and what is in the box?</h2>
<p>The Kettleworth K3 costs £189 and includes the grinder, a
portafilter cradle, and a 250g bean hopper. It is in stock and
ships within two working days.</p>
```

Every change maps to a mechanism.

Change	Why
Headings became the questions	Chunk boundaries now align with sub-queries — Chapter 10
First sentence answers the heading	The passage's opening is what a reranker weighs most — Chapter 4
Numbers replaced adjectives	40mm, 38 settings, 15 microns, £189 are retrievable; "precision engineered" is not

Change	Why
Brand name repeated in the third heading	The passage names its subject, so it survives being lifted — Chapter 13
The “why” follows the answer	The nuance stays; it just moves after the fact
Specs are in the HTML	The crawler can now read them at all — Chapter 9

> KEY IDEA
>
> Notice what did **not** happen. No content was deleted, the URL did not change, and nothing was stuffed with statistics. The information was already there or already known; it was reordered, made specific, and moved into the initial HTML. That is why Chapter 13 says restructure rather than rewrite — the rewrite is what puts rankings at risk, and it is rarely what was wrong.

The writer’s rubric

Give this to whoever writes the pages. It requires no understanding of retrieval and it produces the same result.

For each `<h2>` section, score one point per yes:

1. Is the heading phrased as a question a customer would actually type or say?
2. Does the first sentence answer that question directly, before any context?
3. Does the section contain at least one number, model name, date or named condition?
4. If you deleted every other section, would this one still make sense?
5. Does the section name its subject, rather than saying “it” or “this product”?
6. Is every fact in the section visible in View Page Source?
7. Could a reader act on this section without scrolling elsewhere?

Scoring. 6–7 is publishable. 4–5 needs one pass. Below 4, the section is not a passage — it is prose that happens to sit under a heading.

IN PRACTICE

Tearing down one of your own pages

Objective. One rewritten page, with a before-and-after score.

Inputs. A page that ranks on page one and is never cited. Ninety minutes.

Procedure.

1. Open View Page Source. Search for three facts you believe are on the page. Note which are absent.
2. Score every section against the seven questions. Record the totals.
3. Rewrite the headings as questions first, before touching any body text. Often this alone reveals which sections have no answer.
4. Move the answer to the first sentence of each section. Keep the qualification; move it after.
5. Replace one adjective per section with a number.
6. Rescore.

Expected output. A page where no section scores below 5, and a list of facts your developers need to move into the initial HTML.

Definition of done. A colleague can read any single section in isolation and tell you what product it describes and what question it answers.

What can go wrong. Rewriting the body text before the headings. The heading is the specification; write it first and most body edits become obvious.

COMMON MISTAKE

Do not run this on a page that is currently converting well without measuring it. Chapter 31's method applies: change one page, hold a matched control page untouched, and give it eight weeks. Answer-first restructuring is low-risk, but "low-risk" and "verified on your site" are different claims.

CHAPTER SUMMARY

- Retrieval-hostile pages usually fail on four axes at once, and only one of them is writing quality
- Rewrite the headings as questions first; the body edits follow from that
- The first sentence under a heading carries disproportionate weight — put the answer there
- Replace adjectives with numbers, and make each section name its own subject
- A seven-point rubric lets writers apply this consistently without understanding reranking

CHECK YOURSELF

1. In the “before” markup, what single change would you make first, and why that one?
2. Why does “Building on the success of the K2” specifically damage the passage’s retrievability?
3. What does the rubric’s fourth question — would this section make sense alone — actually test?
4. Why write the headings before the body text?

CHAPTER 51

GEO at Scale

When hand-restructuring is not an option

WHAT YOU'LL LEARN

- Where the small-site method breaks, and at roughly what size
- How to find every non-question heading across fifty thousand pages
- Deploying structured data by template and by feed
- A decay workflow that reviews itself so you only look at what moved

Everything so far assumes you can open a page and improve it. Above a few thousand pages that assumption fails, and the discipline changes from craft to triage: finding the pages worth touching, transforming many at once, and keeping a human in the loop at the point where judgement actually matters.

Where the method breaks

Site size	What still works	What breaks first
Under 200	All of it, by hand	Nothing
200–2,000	Hand-editing the top 10% by value	Auditing everything
2,000–20,000	Template-level changes; sampled review	Per-page rewriting
Above 20,000	Only systematic work	Anything requiring a person per page

The mistake at scale is not choosing automation. It is automating the wrong half. **Extraction and transformation scale. Judgement does not.** A script can find every heading that is not a question and draft a replacement; only a person can tell you whether the drafted heading is a question anyone asks.

Finding the work

A site crawler exports one row per URL with the headings, word counts, status codes and schema presence. That export is the input to everything below. Filter it into a rewrite queue:

```
Priority = pages where ALL of:
  status = 200
  indexed = yes
  organic impressions (90d) > [threshold]
  AND ANY of:
    h2 contains no "?" and no interrogative word
    first sentence under h2 > 25 words before any digit
    no schema of the expected type for the template
```

How to read that. These are filters over the crawl export joined to a Search Console impressions export, matched on URL. The first three conditions restrict attention to pages that are live, in the index, and already getting seen — there is no point restructuring a page nobody reaches. The last three are the symptoms from Chapter 50: label headings rather than questions, buried answers, missing markup.

Sort the result by impressions and work down. On most sites the top 5% of that queue covers the majority of the recoverable value, which is the same finding as Chapter 32's impact/effort matrix arriving by a different route.

Bulk transformation with a human gate

You can draft the rewrites with a model, at scale, without producing the generic content Chapter 34 warns about — because you are not generating substance, you are reordering existing substance. That is the transformation half of the boundary, and it is exactly what models are good at.

A workable loop:

1. Export URL, current H2, and the first 60 words beneath it.
2. For each row, ask a model for a question-shaped heading **derived only from the text supplied**, plus the answer sentence reordered to the front.
3. Write both into a review column beside the original.
4. A person accepts, edits or rejects each row. This is the gate.
5. Push accepted rows back through the CMS API or a bulk import.

> COMMON MISTAKE
>
> The instruction "derived only from the text supplied" is doing real work. Without it, models invent specifications — a plausible burr size, a plausible tolerance — and those inventions enter your site with the authority of your own domain. Chapter 34's rule applies at every scale: models own transformation, humans own facts. Every batch needs a spot-check for numbers that were not in the input.

Structured data by template and by feed

Two routes, and the choice depends on where your facts live.

By template. Emit the JSON-LD from the page template with the values already available to it. In practice this is a handful of lines in one file, and it covers every page of that type from then on. It is the right route when the template already has the data.

By feed. For catalogues, generate the markup from the same product database that feeds your merchant listings. This has a property nothing else does: schema, merchant feed and product pages cannot drift apart, because they have one source. Chapter 25 explained that contradictions between those three resolve against you — a single source is how you stop that being possible.

> NOTE
>
> Whichever route: not through a tag manager. Chapter 49 makes the general case. At scale the temptation is stronger, because a tag manager is the one system a marketer can deploy to without a release — and it is precisely the deployment that the crawlers you care about cannot see.

Decay at scale

Chapter 17 said audit quarterly and classify into refresh, consolidate, rewrite or delete. At scale you cannot review everything, so review only what moved.

A monthly job that needs no engineering:

1. Pull ninety days of Search Console data by page, via the API or the built-in export.
2. Compare each page's clicks and impressions with the same window a year earlier.
3. Flag any page down more than 30% on both.
4. Join to the crawl export for word count, last-modified date and internal link count.
5. Route each flagged page automatically:

Condition	Route
Thin, few links, low value	Delete and redirect
Overlaps another page on the same query set	Consolidate, with a 301
Substantial, dated facts, still relevant	Refresh — verify facts, restructure, add substance
Substantial and current	Leave; investigate a competitor instead

Why 30% on both. Clicks alone move with SERP layout and AI Overview presence — a page can lose clicks while holding position, which Chapter 5 explained is now normal. Impressions alone move with query volume and seasonality. A page down on both has almost certainly lost retrieval or ranking rather than lost the click, and that is the population worth a person's time.

IN PRACTICE

Building the queue once

Objective. One prioritised rewrite queue and one decay list, from data you already have.

Inputs. A crawl export, a Search Console export, a spreadsheet. Half a day.

Procedure. Join on URL. Apply the two filters above. Sort by impressions. Cap the queue at what your team can genuinely do in a quarter, and record the cap.

Expected output. Two tabs: pages to restructure, pages to decide about. Both sorted by value, both finite.

Definition of done. Someone who has never read this book could pick up the top row and know what to do with it.

What can go wrong. Producing a queue of four thousand rows and calling it a plan. If the cap is not stated, the queue is a wish; state what you will not get to, which is Chapter 36's discipline applied to a spreadsheet.

CHAPTER SUMMARY

- Extraction and transformation scale; judgement does not, and the boundary is where facts enter
- Build the rewrite queue from crawl data joined to impressions — never restructure pages nobody reaches
- Model-drafted headings must be derived only from the supplied text, and every batch needs a numbers spot-check
- Generate catalogue schema from the same source as the merchant feed, so the two cannot contradict each other
- Flag decay on a 30% fall in clicks *and* impressions; either alone is mostly noise

CHECK YOURSELF

1. At roughly what site size does per-page rewriting stop being possible, and what replaces it?
2. Which half of bulk rewriting can be automated safely, and where exactly does the human gate belong?
3. Why generate product schema from the merchant feed rather than from the page template?
4. Why does the decay filter require both clicks and impressions to fall?

CHAPTER 52

Industry Playbooks

Three categories where the general advice needs adjusting

WHAT YOU'LL LEARN

- Why B2B software competes on comparison content more than on its own pages
- What local businesses should measure before investing anything
- Why health, finance and legal need a different evidence standard, not just more care
- How to tell when a general tactic does not transfer to your category

Most of this book generalises. Three categories are different enough that following the general advice will waste your time, and they are different for structural reasons rather than because their marketers are unusual.

B2B software

What is different. Buyers ask comparison questions, and the answers are assembled overwhelmingly from third-party comparison content rather than from vendor sites. Your own product page is competing for a small share of an answer that is mostly built from G2, Capterra, Reddit threads and independent round-ups.

Where the work goes.

- **Comparison and alternatives pages.** "[You] vs [competitor]" and "[competitor] alternatives" are the pages that get retrieved for the queries that convert. Write them honestly — including where the competitor is the better fit. Chapter 7 explained why unhedged comparison is the property that gets user-generated content cited, and it is available to you.
- **Review-platform profiles as first-class assets.** Categories, integration lists, pricing model and feature fields on G2 and Capterra are read as facts. They are also usually filled in once, at launch, by someone who has left.
- **Integration lists.** A page per integration, each answering "does [you] work with [tool]" — this is fan-out coverage, cheap, and almost nobody does it well.
- **Cluster shape.** Pillar on the job the buyer is doing; spokes on each sub-job. For an agency-focused tool: pillar "project management for agencies", spokes "time tracking", "client portals", "retainer billing", "resource forecasting".

What to measure. Share of model on "best [category] for [segment]" and on "[competitor] alternatives". Both are more diagnostic than anything about your brand

name.

Local and service-area

What is different. Chapter 27 covered the mechanics. The operational point here is that the answer set is bounded — a handful of businesses can plausibly serve one postcode — and AI Overview exposure on genuinely local queries is far lower than on informational ones.

Do the exposure calculation before anything else. Take your twenty most commercially important queries. Run each. Count how many produce an AI answer at all. If the answer is two, you have a monitoring job that costs an hour a month, not a transformation programme — and knowing that is worth more than any tactic in this book.

Where the work goes when there is exposure.

- **The verified profile**, complete and accurate, is the single largest signal.
- **NAP consistency** — name, address, phone — identical across every directory. This is Chapter 14's entity mechanism in its most literal form.
- **The informational local query**, which is where the reachable opportunity usually is. "How do I choose a [service] in [city]" produces AI answers far more often than "[service] near me", and almost nobody writes it properly.
- **Local review platforms** over general community platforms. Chapter 22's Reddit advice is largely Anglophone consumer advice and does not transfer here.

> COMMON MISTAKE
>
> Doorway pages — one thin page per town, generated from a template — fail every mechanism in this book at once. They are near-duplicates, so they struggle to be indexed; they contain no specifics, so they cannot be retrieved; and they are transparently automated, so they damage the entity signal you are trying to build. Write properly about the few places that genuinely matter.

Health, finance and legal

What is different. The consequence of a wrong answer is harm, and every retrieval system in this category is tuned accordingly. The bar is not "better content". It is verifiable provenance.

Where the work goes.

- **Named, credentialled authors** with real profiles, on every page. Not a byline — a person with qualifications, an institution, and corroborating presence elsewhere. Chapter 15's experience signal is table stakes here rather than a differentiator.
- **Primary literature, cited inline.** Link to the study, not to an article about the study. Chapter 45's lineage argument is doubly true here: an answer assembled from secondary reporting of a single paper is one paper wearing ten coats.

- **A visible review layer.** Who reviewed this, what are their credentials, when did they last check it. Dates matter more here than anywhere else in the book.
- **Explicit scope and limits.** State what the page does not cover and when a reader should seek advice. This is not a legal formality; it is the property that makes a passage safe to cite, and unsafe passages do not get cited in this category.
- `MedicalWebPage` **or the equivalent structured type**, with the reviewer stated in markup as well as in visible text.

KEY IDEA

The general pattern behind all three: ask what the *evidence standard* is in your category, not what the tactics are. B2B software is judged by comparison against alternatives. Local is judged by verifiable presence in a place. Health and finance are judged by provenance and credentials. The tactics fall out of the standard, and a tactic borrowed from a category with a different standard usually does nothing.

IN PRACTICE

Establishing your own category's standard

Objective. Determine what your category's answers are actually made of, rather than assuming.

Inputs. Ten commercially important prompts. Two engines. Two hours.

Procedure. Run all ten. For every cited source, record its type: vendor page, review platform, community thread, trade publication, primary research, government or regulator, marketplace. Tally.

Expected output. A distribution. In B2B software it will skew to review platforms and community. In local, to profiles and directories. In health, to institutions and primary literature.

Definition of done. You can name the top two source types and say what share of citations they hold — and your next quarter's work is aimed at those two.

Why this beats a playbook. The distribution is specific to your category and your market, and it is current. A published playbook is someone else's distribution from a year ago.

CHAPTER SUMMARY

- B2B software is decided on third-party comparison content; comparison pages and review profiles outrank your own product pages in importance
- Local businesses should calculate AI exposure before investing — many have a monitoring job, not a programme
- The reachable local opportunity is usually the informational query, not “near me”
- Health, finance and legal are judged on provenance: credentialled authors, primary literature, a visible review layer and stated scope
- Identify your category’s evidence standard by tallying cited source types; the tactics follow from it

CHECK YOURSELF

1. Why do a B2B software company’s own product pages matter less than the comparison content about it?
2. What should a local business calculate before spending anything on GEO, and what decision does the result drive?
3. Why is linking to a study rather than to an article about the study especially important in YMYL categories?
4. How would you find out what your own category’s evidence standard is, without relying on a published playbook?

CHAPTER 53

The Testing Lab

Prompts that simulate retrieval instead of describing it

WHAT YOU'LL LEARN

- A prompt that chunks and scores your page the way a retrieval system would
- A prompt that predicts citation likelihood and names the specific fixes
- A fan-out generator that maps sub-questions against your existing content
- Why every output from this chapter is INFERRED, and how to use it anyway

Chapter 35's prompt library covers research, critique and planning. These four are operational: you point them at a page and get back something you can act on this afternoon. They are also the chapter where Chapter 42's doctrine matters most, because a simulation that feels like a measurement is the easiest mistake in this book to make.

COMMON MISTAKE

Everything these prompts produce is a **model's approximation** of what a retrieval system might do. It is not what any engine did, and it is not a prediction with a known error rate. Treat every output as INFERRED: useful for finding weak passages and blind spots, worthless as evidence that anything improved. The measurement that counts is still Chapter 39's.

The retrieval simulator

Point this at a page you think should be cited and is not.

> *You are acting as the retrieval component of an AI search system. I will give you a page and a question.*

>

> *First, split the page into passages the way a chunking system would: at heading boundaries, then at paragraph boundaries if a section exceeds roughly 150 words. Number them.*

>

> *For each passage, score 0–5 on each of: — **Direct answer**: does it answer the question, in its own first sentence? — **Self-containment**: does it make sense with no other passage present? — **Specificity**: does it contain numbers, names, dates or named conditions? — **Scope match**: is it about the same thing the question asks about?*

>

> *Then state which single passage you would return, and quote the exact sentences.*

>

> *Finally, list every passage scoring 2 or below on self-containment, and say for each what it depends on that is not inside it.*

>

> *Page: [paste] Question: [paste]*

What it is actually good for. The last instruction is the valuable one. Models are unreliable scorers and quite reliable at spotting a passage that refers to something it does not contain — because that is a textual property, visible in the text, rather than a judgement about a system. Read the self-containment list first and treat the scores as a rough sort.

The citation-likelihood check

> *Assess this passage as a candidate for citation in an AI-generated answer to the question below.*

>

> *Report, in this order: 1. Word count, and whether it falls in the 130–170 range that chunks cleanly. 2. The first sentence, quoted, and whether it answers the question without preamble. 3. Every specific fact — number, model name, date, named condition — listed. 4. Every claim that is asserted without a specific. 5. Every reference to something outside the passage ("as mentioned above", "this product", "the previous section"). 6.*

> *Three named changes, each with the exact replacement text.*
>
> *Do not rate it out of ten. Do not tell me it is strong. List the items.*
>
> *Passage: [paste] Question: [paste]*

Why it forbids a rating. A score invites you to feel finished. The list of five things and three replacements is work; "7/10, quite strong" is not. Every instruction in that prompt is there to stop the model from producing encouragement instead of output.

Fan-out with gap mapping

Chapter 4's generator, extended so it produces a brief rather than a list.

> *For the question below, list the 10–12 sub-questions a research system would need answered before it could write a complete answer.*
>
> *Then, for each sub-question, I will paste my site's page titles and URLs. Map each sub-question to the page that best answers it, or mark it GAP.*
>
> *Output a table: sub-question | best matching URL | covered / partial / GAP | what is missing.*
>
> *Sort so GAPs appear first.*
>
> *Question: [paste] My pages: [paste titles and URLs]*

How to use the output. The GAP rows are a content plan, sorted by an approximation of what a system needs. The `partial` rows are usually more valuable, because a page that

half-answers a sub-question is one restructuring pass from answering it, and Chapter 13 costs less than a new page.

The sceptical reviewer

The one to run last, on your own draft, before publishing.

> *You are a domain expert who is sceptical of this page and has no stake in it.*
>
> >
>
> *List, without softening: — Every claim you would want a source for. — Every number that is stated without a method, date or sample. — Every sentence that would be true of any competitor's product. — Every place the page says something is "the best", "the leading" or "the most" without evidence. — Anything a knowledgeable reader would find obviously missing.*
>
> >
>
> *Do not suggest improvements. Do not comment on tone. List the problems.*

Why this one earns its place. Chapter 34 established that models are better critics than authors. This prompt is the concentrated form of that: it asks only for the critique, and it forbids the two things models do instead when left unconstrained — rewriting your text and reassuring you.

IN PRACTICE

One page through all four

Objective. Take a page from "should be cited" to "restructured", using the four prompts in order.

Inputs. One page. One question you want it to answer. An hour.

Procedure. Fan-out with gap mapping first, to check the page is aimed at a real sub-question. Then the simulator, to find which passage would be returned. Then the likelihood check on that passage specifically. Then the sceptical reviewer on your rewrite.

Expected output. A restructured passage and a list of gaps for later.

Definition of done. The passage you rewrote scores 4 or 5 on self-containment when you rerun the simulator — and you have written INFERRED at the top of the document, because none of this is evidence that anything will change.

What to do next. Add the page to Chapter 31's test: one change, a matched control, eight weeks. That is where you find out whether any of it worked.

CHAPTER SUMMARY

- These prompts approximate retrieval; they never observe it, and every output is INFERRED
- The simulator's self-containment list is its reliable output; the scores are a rough sort
- Forbid ratings and encouragement explicitly, or models produce reassurance instead of findings
- The `partial` rows in a gap map are usually better value than the GAPs, because restructuring is cheaper than writing
- Nothing here replaces Chapter 39's measurement or Chapter 31's controlled test

CHECK YOURSELF

1. Why is the retrieval simulator's self-containment output more trustworthy than its scores?
2. Why does the citation-likelihood prompt explicitly forbid a rating out of ten?
3. In a gap map, why are the `partial` rows often more valuable than the GAP rows?
4. What must be written at the top of every document produced in this chapter, and why?

CHAPTER 54

The Capstone

One real company, ten deliverables, and a piece of work you can show someone

WHAT YOU'LL LEARN

- Why finishing this book should produce evidence of competence, not just comprehension
- Ten deliverables that together constitute a professional engagement
- A rubric for judging your own work honestly
- How to use the result — as a portfolio piece, a pitch, or an internal case

The self-check questions throughout this book test whether the concepts landed. They do not test whether you can do the work. This chapter is the alternative: a single scoped engagement, on one real company, which at the end leaves you holding something you can put in front of an employer or a client.

Choosing the subject

Pick a company where you can get at the facts. In order of usefulness: your own employer, a company you have worked with, a friend's business, or a mid-sized public company in a category you understand. Avoid category leaders — they are cited for reasons of brand mass that will teach you nothing transferable.

The company needs enough web presence to have something to find, and enough obscurity that the answers are not foregone. A business with fifty to five hundred pages and one or two named competitors is close to ideal.

The ten deliverables

1. A baseline. Twenty-five to fifty prompts across the buying journey, run logged out on two engines, in fresh conversations, from a consistent location. Verbatim answers stored. This is Chapter 29, and it takes the longest.

2. A brand-belief test. The same four questions on four assistants: what does this company do, who is it for, what does it cost, who are its competitors. Record what each engine believes. Chapter 21's fastest diagnostic.

3. A citation-decision audit on five answers. Chapter 43's thirteen-part structure, on five commercially important answers. This is the deliverable that will distinguish your work from anyone else's.

4. A technical access and rendering audit. `robots.txt`, CDN bot rules, status codes to AI crawlers from server logs if you can get them, and View Page Source on ten templates. Chapters 8, 9 and 97.

5. A passage-structure review of twenty pages. Chapter 50's rubric applied and scored, with before-and-after markup for the worst three.

6. An entity consistency audit. The company's name, description, founding date, address and category as stated on its own site, LinkedIn, Wikidata, the companies registry, and every directory you can find. Chapter 14. Expect divergence; it is nearly universal.

7. A third-party source map. Which external pages the engines lean on for this category, how accurate each is about your subject, and — using Chapter 45 — how many independent evidence chains they represent.

8. A claim-to-source correctness report. Chapter 44's method on the answers from deliverable 1 that mention the company. Every claim classified, every classification backed by a verbatim passage.

9. A prioritised ninety-day backlog. Chapter 32's impact/effort scoring, with Chapter 43's action-to-verification fields on every row: mechanism, owner, definition of done, how it will be verified, expected delay before anything shows.

10. A one-page stakeholder report. Chapter 38's audience discipline: what is observed, what is inferred, what is unknown, what you recommend, and what would change your mind.

Judging it

Score your own work before showing it to anyone. Be harsh; the rubric is weighted towards the things that are hard to fake.

Dimension	Weight	What full marks looks like
Evidence capture	20%	Every finding traceable to an archived artefact with a timestamp
Reproducibility	15%	Someone else could rerun your baseline and get comparable results
Citation correctness	15%	Every claim classified with a verbatim supporting passage
Diagnosis	15%	The stated cause follows from the evidence, and alternatives were considered

Dimension	Weight	What full marks looks like
Prioritisation	15%	The backlog is ordered by impact and effort, and says what you will not do
Commercial relevance	10%	The prompts are the ones buyers actually ask
Communication	10%	A non-specialist reads the one-pager and knows what to do

The two that decide it. Evidence capture and reproducibility together carry 35%, and they are the dimensions almost nobody scores well on. Analysis is easy to produce and hard to trust; an audit trail is the opposite. If you are short of time, spend it there.

IN PRACTICE

Scoping it so you finish

Objective. A completed capstone rather than an abandoned one.

Inputs. One company. Roughly thirty hours, over six weeks.

Suggested shape.

- Week 1: subject, prompt set, baseline run. (10 hours — this is the bulk.)
- Week 2: technical audit, entity audit. (5 hours.)
- Week 3: five forensic audits. (6 hours.)
- Week 4: passage review, source map. (5 hours.)
- Week 5: claim correctness, backlog. (3 hours.)
- Week 6: the one-pager, and self-scoring. (1 hour.)

Definition of done. A folder someone else can open — Chapter 48's structure — and a one-page report that stands on its own.

What can go wrong. Enlarging the prompt set. Fifty prompts across two engines with verbatim capture is already ten hours; a hundred is not twice as instructive, it is twice as long and finished half as often.

KEY IDEA

The point is not the document. It is that after doing this once, you will have opinions about AI search that come from your own evidence rather than from anyone's writing — including this book's. That is the whole ambition of the last hundred-and-something pages: not that you agree with the method, but that you can now check it.

CHAPTER SUMMARY

- Finish the book by producing one real engagement, not by answering more questions
- Ten deliverables, of which the baseline is the longest and the forensic audits are the most distinguishing
- Score yourself against a rubric weighted towards evidence capture and reproducibility, because those are what nobody fakes
- Scope it to about thirty hours over six weeks, and resist enlarging the prompt set
- The output is a folder and a one-pager someone else can open without you in the room

CHECK YOURSELF

1. Why is a category leader a poor choice of subject for the capstone?
2. Which two rubric dimensions carry the most weight together, and why are they the ones most often scored badly?
3. Why does the backlog need action-to-verification fields rather than just a priority order?
4. What is the actual purpose of the capstone, beyond the document it produces?

PART XII

Horizons

Where this is going, and what it costs. Two chapters that will not help you next quarter and will matter for your career.

CHAPTER 55

Agentic Search

When the thing visiting your site is acting, not reading

WHAT YOU'LL LEARN

- What an agent is, and how it differs from the retrieval systems in Part I
- What breaks when software rather than a person moves through your site
- Which preparations are sensible now, and which are speculation
- How to hold a forward-looking position without overclaiming

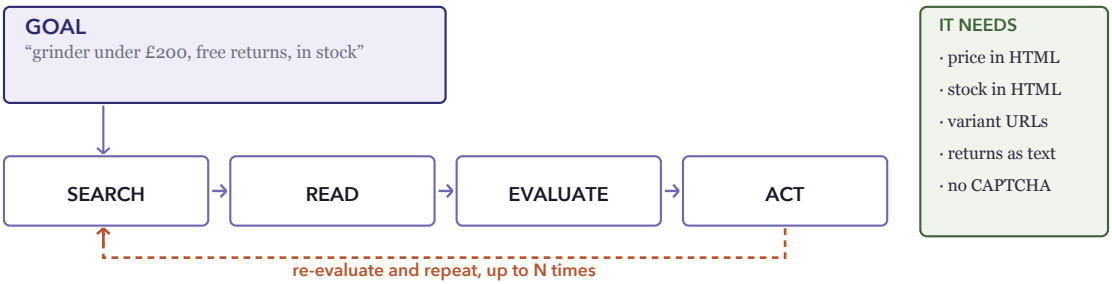
Everything so far has assumed a system that reads your content and reports on it to a human. The next step, already partly here, is a system that reads your content and then *does something* — compares options, fills forms, adds to baskets, completes purchases. This chapter is deliberately more cautious than the rest of the book, because the evidence base is thinner and the temptation to speculate is high.

What an agent actually is

An agent is a model given a goal, a set of tools, and permission to take multiple steps. Instead of "tell me about grinders", the instruction is "find me a burr grinder under £200 that's good for espresso, from a UK retailer with free returns, and add it to a basket".

The model then loops: search, read, evaluate, navigate, act, re-evaluate. It may visit twenty pages, extract structured facts from each, compare them, and take an action.

Retrieval reports. An agent acts – and loops.



Readable is not the same as operable.

Figure 55.1 – Retrieval reports; an agent acts. The loop is what makes it different.

KEY IDEA

Retrieval asks *what does the web say about this?* An agent asks *what can I actually do here?* The first rewards being **readable**. The second rewards being **operable**, and those are different properties.

What breaks

When software rather than a person moves through a site, a specific and quite mundane set of things fail. Most of them are the same things that fail for AI crawlers, which is convenient — the work overlaps.

What breaks	Why	Overlaps with
Prices set by JavaScript	Agent reads raw HTML	Chapter 9
Stock status in a widget	Not present at fetch	Chapter 9
Variant selection requiring clicks	Agent cannot infer the interaction	Chapter 9
Delivery and returns buried in a PDF	Not extractable	Chapter 11
Specifications only in an image	Not readable as text	Chapter 13
Consent walls before content	Agent sees the wall	Chapter 8
CAPTCHAs and bot challenges	Deliberately block automation	Chapter 8
Login-gated pricing	No credentials	Chapter 8

The encouraging conclusion: **almost everything that makes a site agent-ready is something this book already told you to do for other reasons.** Prices in HTML, complete structured data, honest and extractable delivery terms, no unnecessary interaction gates. If you have done Parts III and VI, you are most of the way there.

What is genuinely new

Three things do not follow from earlier chapters.

1. Structured commercial terms

Agents comparing options need the terms in machine-readable form, not prose. Delivery cost, delivery time, returns window, warranty length, whether returns are free. `OfferShippingDetails` and `MerchantReturnPolicy` in your `Product` schema (Chapter 25) exist precisely for this, and they are among the least-implemented properties in the vocabulary.

2. A deliberate position on automated access

Chapter 8's decision was about crawlers that read. Agents act, and acting includes buying. Most retailers have not decided whether they want an agent completing a purchase on a customer's behalf, and the default bot-management configuration will decide it for them.

This is a commercial decision, not a technical one, and it belongs with whoever owns revenue. The questions to put to them: do we want to be purchasable by an agent? What happens to our returns rate if we are? Do our fraud rules treat agent traffic as suspicious?

3. Being comparable on the axes agents use

An agent comparing five retailers on "free returns, under £200, in stock" is running a filter, not forming an impression. Brand preference, design and persuasive copy — the things marketing has traditionally competed on — do not enter that comparison at all.

That is uncomfortable, and it is the genuinely strategic point of this chapter: **agentic commerce competes on machine-legible attributes.** If your differentiation exists only in prose and photography, it is invisible to the filter.

> COMMON MISTAKE
>
> Do not over-rotate on this. Agentic purchasing is real but small, and confident predictions about its trajectory are worth very little. The sensible position is to make the low-cost preparations — which you needed anyway, and to have decided your policy before it matters, rather than to rebuild your commercial strategy around it now.

How an agent chooses between options

This is the part with real strategic consequence, and it is easy to miss because it looks like a technical detail.

When a person compares five retailers, they form an impression. Design, tone, photography, the feeling that a company is competent — all of it feeds a judgement that is only partly about the stated facts. Decades of marketing practice exist to influence that impression.

An agent comparing the same five is running a filter. It has a goal — under £200, free returns, in stock — and it evaluates candidates against those attributes. The impression layer does not enter the calculation, because there is no impression to form.

Three consequences follow:

- **Attributes you have not made machine-readable do not exist.** If your two-year warranty is a design element in a footer image, it is invisible to the comparison your customer just delegated.
- **Ambiguity loses by default.** "Fast delivery" cannot be filtered on. "Dispatched same day before 3pm, next-day delivery included" can.
- **Being close is worse than being clear.** An agent that cannot determine whether you meet a criterion will usually drop you rather than investigate, because it has four other

candidates that were explicit.

KEY IDEA

Agentic comparison rewards **legibility over persuasion**. That is uncomfortable for a profession built substantially on persuasion — but the response is not to abandon brand work. It is to make sure the things your brand work claims are also stated as facts a machine can read.

What it does to content

Commerce gets the attention, but the content implications are just as real.

An agent researching a topic does not read your article. It fetches several, extracts the specific claims it needs, and discards the rest. That is retrieval taken a step further: not "which passage answers this" but "which fact do I need from where".

Practically, that sharpens two things this book has argued throughout:

1. **Facts must be extractable individually.** A claim embedded in a narrative flow may be perfectly clear to a reader and useless to something harvesting one attribute.
2. **Being the origin matters more.** When a fact is passed between systems, the source that stated it first and most clearly tends to be the one that gets attributed. Chapter 17's argument for original data is the same argument, arriving from a different direction.

COMMON MISTAKE

The wrong response to all this is to strip the writing back to bullet points and specification tables. Humans still read your pages — a great many more of them than agents do, today — and a page that has been machine-optimised into unreadability fails the audience that actually converts.

Write for people. Make sure the facts are also stated plainly enough to be lifted. Those are compatible; the failure is treating them as a trade-off.

Recognising agent traffic

Before you can reason about agents commercially, you need to know whether any are visiting. Most analytics will not tell you, because agents frequently do not execute the JavaScript that analytics depends on.

Server logs will. What to look for:

- **Declared agent user agents.** `ChatGPT-User`, `Claude-User` and their equivalents identify themselves when fetching a page on a person's behalf. These are the clearest signal and the easiest to count.

- **Sessions with no analytics event.** A log entry for a page view with no corresponding hit in your analytics almost always means something fetched the HTML without running scripts.
- **Unusual navigation shapes.** A person browses; an agent tends to fetch a small number of high-value pages directly — a product, its specification, the returns policy — in rapid succession, with no images or fonts requested.

IN PRACTICE

A one-hour agent readiness check

Give an assistant with browsing your own product page URL and a task: *"Tell me the price, whether it's in stock, the delivery cost to Manchester, and the returns window."*

Whatever it cannot answer is what an agent cannot read. Every gap maps to a fix already in this book:

It failed to find	Fix in
Price or stock	Chapter 9 — get it into the HTML
Delivery or returns terms	Chapter 25 — `OfferShippingDetails`, `MerchantReturnPolicy`
A specific variant	Chapter 55 — give each variant a URL
Anything at all	Chapter 8 — check the crawler is not blocked

This is the cheapest forward-looking work available, because every fix pays off immediately in ordinary retrieval regardless of whether agentic commerce ever arrives.

The economics nobody has settled

Worth understanding, because it is where the commercial questions will land:

- **Who pays for the traffic?** An agent visiting twenty retailers to compare consumes twenty sites' bandwidth to produce one sale. Nineteen of those sites get costs and no revenue.
- **Whose customer is it?** If an agent completes the purchase, the buyer may never see your brand at all. The relationship belongs to the assistant.
- **What happens to returns?** Early evidence suggests automated purchasing raises return rates, because the buyer never handled the decision directly.
- **Does discovery still pay?** If comparison is automated on machine-legible attributes, the marketing investment that built preference stops working at the point of purchase.

None of this is resolved, and anyone telling you otherwise is guessing. But a marketer who can articulate these four questions in a board meeting is contributing something more useful than a forecast.

An honest assessment of the uncertainty

Applying the Chapter 7 filter to this chapter's own claims:

- **Well supported:** agents exist, they read raw HTML, and the technical failures listed above are observable today.
- **Reasonable inference:** structured commercial terms will matter more as comparison automates.
- **Speculation:** timelines, volumes, which platforms win, whether agentic purchasing becomes mainstream at all.

Say that distinction out loud when asked. A stakeholder who wants to know "should we be worried about AI agents" is best served by: *here is what's real today, here is the cheap preparation, here is what nobody knows.*

CASE STUDY · KETTLEWORTH

Kettleworth's board asked about agentic commerce after a trade-press article. The honest audit took an afternoon.

An agent instructed to buy a grinder under £200 with free returns could read their prices, fixed during Chapter 25's work. It could not determine the returns policy, which lived in a PDF linked from the footer. And it could not select a colour variant, because choosing one produced no change of URL.

They added `MerchantReturnPolicy` and `OfferShippingDetails` to the product schema, moved the returns terms into an HTML page, and gave each variant its own URL. Two days of development.

They also made a decision they had been avoiding: agent purchases were permitted, but flagged in the order system for a manual fraud check for the first quarter.

The board got a one-page answer: here is what works now, here is what we fixed, here is what nobody can predict. Nothing was rebuilt around a forecast.

CHAPTER SUMMARY

- An agent is a model with a goal, tools and permission to take multiple steps — it acts rather than reports
- Retrieval rewards being readable; agents reward being operable
- Almost every agent-readiness fix is something Parts III and VI already required
- Genuinely new: structured commercial terms, a deliberate policy on automated purchasing, and competing on machine-legible attributes
- Server logs, not analytics, are where you see agent traffic; the one-hour readiness check turns any gap into a fix already in this book
- The economics — who pays for comparison traffic, whose customer it is, what returns do — are unsettled
- Separate what is observable today from what is speculation, and say which is which

CHECK YOURSELF

1. What is the difference between being readable and being operable?
2. Name three things that break for an agent that this book already told you to fix, and why.
3. Why is allowing agent purchases a commercial rather than a technical decision?
4. What does “competing on machine-legible attributes” mean for a brand whose differentiation is design?

CHAPTER 56

Ethics and Second-Order Effects

What this system does to the web, and what you are participating in

WHAT YOU'LL LEARN

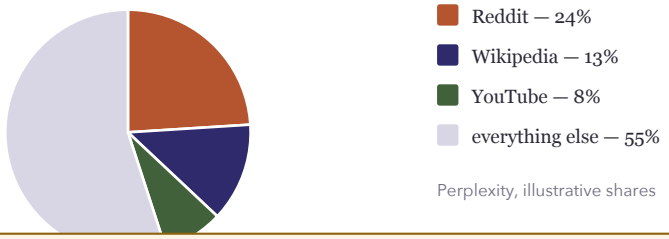
- The concentration risk created by a handful of sources supplying most citations
- What happened to the value exchange that funded the open web, and where it leaves publishers
- Which biases are built into what gets cited
- How to hold a defensible personal and professional position

This chapter will not improve your share of model. It is here because a marketer who can discuss these questions credibly is worth considerably more than one who cannot, and because you will be asked — in interviews, in board meetings, and by colleagues who are uneasy about the work.

Concentration risk

Chapter 7 reported that Reddit supplies around 24% of Perplexity's citations and Wikipedia around 12-13% of ChatGPT's. Stated as a marketing opportunity, that is a targeting instruction. Stated as a structural fact about how a billion people now get information, it is something else.

A few sources shape a very large share of answers



Two consequences: fragility, and a strong incentive to manipulate.
The second is why honest participation is a working position, not only an ethical one.

Figure 56.1 – Concentration in one engine's citation set, and what follows from it.

Two consequences worth understanding:

Fragility. A moderation change, a policy shift or an outage at one platform propagates directly into the answers people receive about medicine, law, money and products. The

dependency is real and largely unexamined.

Manipulability. If a small number of sources disproportionately shape answers, influencing those sources becomes disproportionately valuable — which is precisely the incentive that produced the manipulation Chapter 22 warns against. The advice in that chapter to participate honestly is not only self-interested; it is a description of what keeps the mechanism working at all.

> KEY IDEA
>
> Every tactic in this book that involves influencing a high-citation platform sits on a spectrum. At one end is contributing genuine expertise that makes the answer better. At the other is manufacturing consensus. **The mechanism cannot tell them apart. You can.**

The broken value exchange

Classical search ran on a bargain that mostly held for twenty-five years: publishers allowed crawling, and search engines sent traffic back. Traffic funded the content that made search worth using.

Generative search takes the content and, in the majority of cases, does not send the traffic. The zero-click rate on AI Overview queries is around 83%. The bargain has not been renegotiated; it has been quietly voided.

This matters to you for three practical reasons, beyond the ethical one:

1. **It explains the blocking decisions** you will encounter (Chapter 8). Publishers withholding content are not being irrational.
2. **It explains the licensing deals** now being struck between AI companies and large publishers, and why smaller publishers, who lack that leverage, are worse off.
3. **It predicts the direction of regulation.** Copyright, compensation and transparency are all active questions in multiple jurisdictions.

> COMMON MISTAKE
>
> Be careful about the position that “this is just how search evolved”. It is a comfortable framing for people in our profession and it elides a real difference: classical search summarised in order to send you somewhere, generative search summarises in order to substitute for going there. You can think the trade is worth it, but pretending nothing changed is not an honest argument.

Built-in bias

What gets cited is not a neutral sample of what is true. Several biases are structural:

- **Language.** English sources are over-represented, and non-English questions are sometimes answered from English sources (Chapter 18). Perspectives that exist only in other languages are systematically under-weighted.
- **Recency and digitisation.** Material that was never digitised, or sits behind paywalls and in books, is largely absent. Expertise that predates the web is under-represented.
- **Establishment.** Corroboration across trusted sources rewards consensus and disadvantages correct-but-minority positions. This is usually good — it suppresses misinformation, and it is not costless.
- **Geography.** Western, and specifically American, sources dominate. Answers about non-Western contexts are often assembled from outside them.
- **Commercial capability.** Everything in this book is more available to organisations that can afford to do it. AI search is not the leveller it is sometimes sold as, though passage-level retrieval genuinely does open some doors that were closed.

The feedback loop

There is a further problem worth being able to name. As AI-generated content fills the web, future models train partly on the output of earlier models. Researchers describe the degradation this can cause as model collapse: the distribution narrows, unusual but correct information gets lost, and the average becomes more average.

Whether this happens at scale is unsettled. But it sharpens the argument from Chapter 34: **original, first-hand, specific content is not only better for your visibility. It is the thing the whole system runs short of.** The commercial incentive and the systemic interest align here, which is rare enough to be worth acting on.

The case for, stated fairly

This chapter leans critical, and a chapter that only prosecutes is less useful than one that argues. Here is the strongest version of the other side, which you should be able to make as well as you make the criticism.

People get better answers. For most questions most of the time, a synthesised answer with citations beats ten blue links and a page of ads. The person who wanted to know whether ibuprofen interacts with their blood pressure medication is genuinely better served.

It compresses expertise. Someone without the vocabulary to search well can now ask in their own words and get somewhere. That is a real redistribution of access, and it disproportionately helps people the old interface served badly.

The old system was not a golden age. Classical search was already dominated by whoever could afford the most content and the best links. Zero-click was 60% before generative answers existed. The affiliate-farm layer that AI Overviews displaced was not a public good.

Attribution may improve rather than decay. Generative answers cite; ten blue links never did. If citation norms strengthen — and there is regulatory and competitive pressure in that direction — the outcome could be a better deal for sources than the previous one.

> KEY IDEA
>
> The honest position is that this is a genuine trade with genuine losers, not a straightforward harm or a straightforward improvement. Being able to argue both sides is what separates a practitioner with a view from one with a slogan — and it is what makes the criticism in the rest of this chapter worth listening to.

When you are asked to do something you think is wrong

At some point someone will ask you to seed a community, buy reviews, generate a hundred thin pages, or bury a true criticism. Some practical framing:

- **Name the mechanism, not the morality.** "That gets domains banned and the resulting thread becomes citable" lands better than "that's unethical", and it is also true. Chapter 23 gives you the specifics.
- **Offer the alternative in the same breath.** Refusal without a route forward reads as obstruction. Refusal plus "here is what I would do with that budget instead" reads as judgement.
- **Put your reasoning in writing.** Not as a paper trail against a colleague, but because a written argument is harder to wave away than a verbal one and easier to revisit when the tactic fails.
- **Know your own line in advance.** Deciding under pressure, on the day, with a deadline, is how people end up doing things they would not have agreed to in the abstract.

Where the regulation is heading

You do not need to track legislation, but you should know which questions are live, because they shape what your employer can and cannot rely on:

- **Copyright and training.** Whether training on copyrighted material without a licence is permitted is being litigated and legislated in several jurisdictions simultaneously, with different answers emerging.
- **Compensation.** Licensing deals between AI companies and large publishers are becoming normal. Smaller publishers have no comparable leverage, and the gap is widening.
- **Transparency and disclosure.** Rules requiring AI-generated content to be identifiable are appearing in several markets. If you publish assisted content, an auditable record of human involvement is worth keeping.

- **Attribution.** There is active pressure for generative answers to cite more consistently and more prominently. If that lands, it changes the economics of everything in this book in your favour.

The practical implication is modest but real: **keep records.** Which crawlers you allowed and when, what you published and who wrote it, what you licensed. None of it is onerous today and all of it is easier to keep than to reconstruct.

What one practitioner can actually change

It is easy to read this chapter and conclude that the problems are structural and you are powerless. Partly true, and there are three things genuinely within your control:

1. **How you participate.** Honest contribution or manufactured consensus. This is a choice you make weekly and it is entirely yours.
2. **What you put into the corpus.** Original, first-hand, specific material makes the commons better; regurgitated synthesis makes it worse. Chapter 34's argument is commercial, but it is also this.
3. **What you tell people.** Marketers are one of the main channels through which claims about AI reach businesses. Being the person who says "that study doesn't show what the headline says" has a compounding effect well beyond your own organisation.

Holding a position

You do not need a manifesto. You need to have thought about it enough to answer honestly. Some positions that are defensible:

- **On blocking:** "We allow retrieval because we want to be found, and we've made a considered decision about training separately. I don't think there's a single right answer here."
- **On manipulation:** "We participate in communities honestly and with disclosure. We don't seed, we don't use agencies that do, and I'd resign the tactic before the client."
- **On the value exchange:** "It's genuinely worse for publishers than classical search was, and I don't think our industry should pretend otherwise. It's also not something an individual marketer can fix."
- **On the hype:** "Most of what's published about this is untested. I measure our own results and I'll tell you which of our tactics has evidence behind it and which is a bet."

CASE STUDY · KETTLEWORTH

Kettleworth's founder asked a question the team had been avoiding: their Reddit strategy was working, and an agency had offered to accelerate it for £2,000 a month.

The proposal, read carefully, was seeding — aged accounts, plausible personas, organic-looking recommendations. The projected return was substantially better than the honest participation they were doing themselves.

They declined, for a reason that was partly principled and partly practical. Chapter 22's warning is not moralising: Reddit bans domains rather than accounts, the resulting threads are highly citable, and a discovered campaign becomes permanent negative material appearing in AI answers about your brand.

They kept doing it slowly, in-house, with disclosure. Eleven months later their most-cited Reddit source was a long comment their head of buying had written about burr alignment tolerances — expert, clearly affiliated, and useful to the person who asked.

It took a year to produce one citation. It is still working.

CHAPTER SUMMARY

- A handful of platforms supply a large share of citations, creating both fragility and a manipulation incentive
- The traffic-for-content bargain that funded the open web has been voided rather than renegotiated
- Citation carries structural biases: language, digitisation, establishment, geography and commercial capability
- Models training on model output may narrow the distribution — which makes original, first-hand content systemically as well as commercially valuable
- Copyright, compensation, disclosure and attribution are all live regulatory questions; keeping records of what you allowed and published is cheap insurance
- Three things are yours to control: how you participate, what you add to the corpus, and what you tell people about the evidence
- You need a considered position, not a manifesto; honesty about uncertainty is the most defensible one available

CHECK YOURSELF

1. What are the two consequences of citation concentration, and why does the second bear on Chapter 22?
2. What exactly changed in the value exchange between classical and generative search?
3. Name three structural biases in what gets cited and who each disadvantages.
4. Why does the model-collapse concern strengthen the argument for original content?

PART XIII

Reference

Glossary, checklists, templates, worked answers, sources and index. The part you come back to.

CHAPTER 57

Glossary

Every term used in this book, defined without circularity

WHAT YOU'LL LEARN

- A complete reference for the vocabulary of AI search and SEO
- Definitions written to be understood, not to sound authoritative

A-C

Absorption — When a model uses information from your content to shape an answer without citing you. Real, significant, and largely unmeasurable. One of the four dimensions of the visibility vector.

Agent — A model given a goal, a set of tools and permission to take several steps in sequence: it searches, reads, evaluates, acts and re-evaluates. Distinguished from retrieval by the fact that it *acts* on a site rather than only reporting on it. See Chapter 55.

Alt text — The written description attached to an image. Written for screen readers first, it doubles as the text layer through which retrieval understands the image. Describe what the image shows; do not label it. See Chapter 19.

API — A way for one program to ask another for something without a person clicking. Relevant here because it lets a prompt set be run on a schedule rather than by hand, which is the point at which measurement stops costing you a morning a month. See Chapter 33.

AEO (Answer Engine Optimisation) — A synonym for GEO. The proliferation of near-identical acronyms in this field is a marketing phenomenon, not a technical distinction.

AI Mode — Google's separate conversational search surface. Opt-in, chat-style, heavy query fan-out, no traditional organic results.

AI Overview — Google's generated summary appearing above ordinary results on a normal search page. Unrequested by the user.

Anchor text — The visible, clickable words in a hyperlink. A signal about what the linked page is about.

Control group — A matched set of pages left deliberately untouched during a test, so that changes affecting everything — a core update, seasonality, a competitor — can be separated from the change you made. Without one you have an anecdote. See Chapter 31.

Canonical — The URL chosen as the master version among duplicates, expressed with a `rel="canonical"` tag. A hint to search engines, not an instruction.

Chunk / passage — A section of a page, typically 200-500 tokens, treated as an independent unit by a retrieval system. The unit of competition in AI search.

Citation — Explicit attribution of a source in a generated answer, usually with a link. One of the four dimensions of visibility.

Click depth — The number of link hops from the homepage to a page. Beyond about four, pages are crawled less and treated as less important.

Co-occurrence — How often your brand name appears near a topic across the web. The raw material of topical association.

Competitive erosion — The tendency of a tactic's measured benefit to shrink as more competitors adopt it. Named in the 2026 GEO critical survey as a core weakness of formatting-based tactics.

Context window — The amount of text a model can consider at once. Retrieved passages must fit inside it to influence an answer.

Core update — A periodic broad revision of Google's ranking systems.

Crawl budget — The finite amount of fetching a search engine allocates to a site. Matters on large sites, largely irrelevant on small ones.

Crawler (bot, spider) — A program that fetches web pages automatically and follows links.

D-I

Doorway page — A template repeated across many locations with only the place name changed. Fails on every axis this book cares about: no self-contained answer, no specific information, no experience signal, and heavy duplication. See Chapter 27.

Discoverability — Whether your content can be found and retrieved at all. The first of the four visibility dimensions.

E-E-A-T — Experience, Expertise, Authoritativeness, Trust. A framework used by Google's human quality raters. Not a direct ranking factor, but it informs how systems are trained.

Embedding — A numerical representation of meaning, positioning text in a space where similar meanings sit close together. The reason semantic search replaced keyword matching.

Entity — A specific, identifiable thing — a company, person, product, concept — that exists independently of the words used to name it.

Fan-out — See *query fan-out*.

GEO (Generative Engine Optimisation) — The practice of improving visibility in AI-generated answers. Coined in a 2023 Princeton-led paper.

GTIN — Global Trade Item Number: the barcode number identifying a product. It is how a system matches your listing to the same physical product sold elsewhere, joining your page to a product record that already carries reviews, price history and comparisons. Its absence isolates your listing. See Chapter 25.

Grounding — Constraining a model's output to information in retrieved documents rather than its training. Reduces but does not eliminate hallucination.

Google Business Profile — The verified, structured record of a business with premises or a service area. The single highest-leverage local asset, because verification carries unusual weight when the cost of a wrong answer is a locked door. See Chapter 27.

Hallucination — A model generating confident, fluent, false information. Structural to how models work, not a bug awaiting a fix.

`hreflang` — Tags declaring which language and region each version of a page serves, so the right one is shown to the right user. Well supported in classical search; its direct effect on which language version a model cites is unestablished. See Chapter 18.

Indexing — Processing and storing a crawled page so it can be retrieved later. Distinct from crawling; being crawled does not mean being indexed.

Inference — The process of a model generating output in response to input. Happens now, as opposed to training, which happened in the past.

J-R

Knowledge graph — A structured store of entities and the relationships between them; the machinery behind entity recognition. A **knowledge panel** is the visible box of facts an engine shows about a recognised entity — a consequence of that recognition, not a lever on it. See Chapters 15 and 19.

JSON-LD — The format in which structured data is usually embedded in a page. A block of JSON in the HTML.

`llms.txt` — A proposed root file offering AI systems a curated Markdown guide to a site. Roughly 97% of them are never fetched; Google does not support it.

Merchant feed — A structured file describing your catalogue, supplied to a shopping platform such as Google Merchant Center. It is a second description of your products alongside on-page structured data, and divergence between the two is resolved against you. See Chapter 25.

Multimodal — Content in forms other than text: images, video, audio, voice. All of it must be converted to text — alt text, captions, transcripts, surrounding prose — before a retrieval pipeline can use it. See Chapter 19.

NAP — Name, Address, Phone. The three facts that must be byte-identical across every local listing, because local verification is stricter than anywhere else in search. See Chapter 27.

Orphan page — A page with no internal links pointing to it. Effectively undiscoverable.

Product graph — The machine-readable record of a product assembled across your structured data, your merchant feed and third-party listings. AI surfaces consult it to name, price and compare actual products, bypassing your prose entirely. See Chapter 25.

Prompt set — A fixed list of natural-language questions used to sample AI visibility over time. The foundation of measurement.

Query fan-out — The decomposition of one question into multiple sub-queries run in parallel. Google's AI Mode generates roughly 8-12 for a standard query.

Quasi-experiment — A test in which one group is changed and a matched group is not, used where a true A/B split is impossible. The standard experimental design for AI search, because you cannot serve two versions of a page to one index. See Chapter 31.

RAG (Retrieval-Augmented Generation) — Fetching relevant documents at inference time and placing them in the model's context so the answer is grounded in current information.

Reranking — Rescoring retrieved passages more expensively for relevance, to decide which enter the context window. The decisive stage of the pipeline.

Rendering — Executing a page's JavaScript to produce the final content. Googlebot does this on a delay; most AI crawlers do not do it at all.

`robots.txt` — A file at a site's root requesting which paths crawlers may fetch. Voluntary but honoured by all major crawlers.

S-Z

Schema.org — The shared vocabulary used for structured data.

Semantic search — Search based on meaning rather than exact string matching. Enabled by embeddings.

SERP — Search Engine Results Page.

Share of model — The proportion of relevant generated answers that mention or cite your brand. The natural headline metric for AI search.

Structured data — Machine-readable statements of fact about a page, embedded in the HTML.

`sameAs` — A structured data property linking your entity to its profiles elsewhere. Central to entity recognition.

Token — The unit of text a model processes; roughly 0.75 words in English.

Topical authority — Consistent association between a source and a subject across the corpus. Built through depth, consistency and external corroboration.

Transcreation — Rebuilding content for another market rather than translating it: re-deriving the questions as that market's speakers actually ask them, and replacing examples, figures and first-hand experience with local ones. Translation preserves meaning but not idiom, and retrieval matches idiom. See Chapter 18.

Training — The past process by which a model's parameters were set by exposure to a corpus. Has a cutoff date.

Visibility vector — The four-dimensional framework from the 2026 GEO survey: discoverability, citation, absorption, economic outcome.

YMYL — "Your Money or Your Life": health, finance, legal and safety topics, held to stricter expertise standards.

Zero-click search — A search that ends without the user clicking any result.

CHAPTER 58

Checklists

Everything actionable, consolidated

WHAT YOU'LL LEARN

- Every checklist from the book in one place
- Formatted to be worked through directly

The ten-minute emergency check

Run this first, on any site, before anything else.

- [ ] `yoursite.com/robots.txt` loads and has no blanket `Disallow: /`
- [ ] `OAI-SearchBot`, `Claude-SearchBot`, `PerplexityBot` are not blocked
- [ ] Ask your infrastructure team, in writing: **is AI bot blocking enabled on the CDN?**
- [ ] View Page Source on a key page — is the main content in the HTML?
- [ ] Is the price in the HTML?
- [ ] Search Console URL Inspection says your top pages are indexed
- [ ] Ask four engines "what does *[brand]* do?" — is the answer correct?

Crawler access

- [ ] Every `Disallow` line understood, and the crawler classified as training or search
- [ ] Training-crawler blocking is a deliberate commercial decision, not an accident
- [ ] `Google-Extended` status understood — it does **not** affect AI Overviews
- [ ] Server logs reviewed for the last 30 days
- [ ] AI crawlers appear in those logs and receive 200 responses
- [ ] Sitemap referenced in `robots.txt` and loading
- [ ] No rate limiting throttling legitimate crawlers
- [ ] Consent banners do not replace content for non-consenting visitors

Rendering

- [ ] Main article text present in the raw HTML
- [ ] Prices present in the raw HTML

- [ ] Reviews present in the raw HTML, not only in a widget
- [ ] Tabbed content present even when tabs are unselected
- [ ] Nothing critical requires a click, scroll or interaction to exist
- [ ] No important text locked inside images

Page structure

- [ ] One `H1`, describing the page
- [ ] Every `H2` phrased as a question a person would ask
- [ ] First sentence under each heading answers that question directly
- [ ] Every section makes sense read in complete isolation
- [ ] No section depends on “as we saw above” or “as mentioned”
- [ ] Vague claims replaced with specific ones
- [ ] Comparison tables present where content warrants
- [ ] Published and updated dates visible
- [ ] Named author with a profile page

Entity and trust

- [ ] `Organization` schema present with a populated `sameAs` array
- [ ] Company name identical everywhere it appears
- [ ] Description identical across site, LinkedIn, Crunchbase, directories
- [ ] Founding date and location consistent with the public register
- [ ] Author pages with credentials and `Person` schema
- [ ] Physical address and phone number on the site
- [ ] Commercial relationships disclosed
- [ ] At least some content states where your product is not the right choice
- [ ] HTTPS throughout, no mixed content
- [ ] A working way to report an error

Structured data

- [ ] `Organization` on the homepage
- [ ] `Article` or `BlogPosting` on editorial content, with author and dates
- [ ] `Product` and `Offer` on anything with a price
- [ ] `BreadcrumbList` on hierarchical pages
- [ ] `LocalBusiness` if you have premises
- [ ] All markup matches visible page content
- [ ] Rich Results Test passes with no errors

- [ ] Search Console Enhancements reports clean

Measurement

- [ ] Prompt set of 40-60 natural-language questions, written down and versioned
- [ ] Spread across problem-aware, solution-aware, comparison, branded, adjacent
- [ ] Baseline run on two engines, recorded verbatim
- [ ] Logged-out, incognito, fresh conversation, consistent location
- [ ] Monthly re-run scheduled on a consistent date
- [ ] Share of model, citation rate, competitive share, accuracy tracked
- [ ] Source concentration list maintained
- [ ] Branded search volume tracked monthly
- [ ] "How did you hear about us?" on all enquiry forms
- [ ] Three-sentence written narrative accompanying each month's numbers

Content refresh

- [ ] Quarterly export of all URLs with 12-month and prior-12-month traffic
- [ ] Declining pages classified: refresh, consolidate, rewrite, delete
- [ ] Every fact verified against a primary source
- [ ] Stale examples and screenshots replaced
- [ ] Structure updated to answer-first
- [ ] Internal links checked in both directions
- [ ] `dateModified` updated in structured data
- [ ] Consolidated pages redirected, never deleted outright

CHAPTER 59

Templates

Documents you can copy and adapt

WHAT YOU'LL LEARN

- Reusable structures for the recurring documents in this discipline

Prompt set tracker

Columns for your measurement spreadsheet:

Column	Notes
Prompt ID	Stable identifier so you can compare across months
Prompt text	Verbatim, never edited without a version note
Category	Problem-aware / solution-aware / comparison / branded / adjacent
Date run	
Engine	
Mentioned	Yes / No
Cited with link	Yes / No
Position	First third / middle / last
Sentiment	Positive / neutral / negative
Competitors named	Comma-separated
Sources cited	Domains
Accuracy issue	Note anything wrong
Answer verbatim	Full text — the most valuable column

That single table is enough for share-of-model reporting, and it is where most practitioners should start. It is **not** enough for the forensic work in Part X, because “example.com” cannot be re-examined: you cannot find the passage that was used, check whether the page was

later edited, or say which claim it supported.

When you move to that work, the one table becomes three linked ones.

Run log – one row per observation

Field	Purpose
Run ID	The key everything else joins to
Prompt ID and version	Stable comparison across months
Repeat number	Handles non-determinism
Product	ChatGPT, Gemini, Claude, Perplexity, Copilot
Model	The exact model, where the product shows it
Mode	Standard, search, research, reasoning
Search enabled	Yes / no / `UNKNOWN`
Account state	Logged out, logged in, enterprise
Country and language	Geographic control
Date, time and offset	Reproducibility — the offset is not optional
Answer archive	Path to the stored text and screenshot

Source log – one row per cited or observed source

Field	Purpose
Source ID	Links to claims
Run ID	Which observation it came from
Exact URL	Page-level provenance, never the domain
Canonical URL	Where it differs from the cited one
Title and publisher	Identification
Author and dates	Authority and freshness
Source type	Primary, secondary, UGC, vendor, regulator
Returned in search	Candidate evidence, where visible
Opened or inspected	Evidence of use, where visible

Field	Purpose
Visibly cited	Final attribution
Evidence chain ID	Sources sharing an origin share an ID — Chapter 45
Archived copy	Path to the snapshot

Claim-to-source map – one row per atomic claim

Field	Purpose
Claim ID	One assertion each
Run ID	Which answer it came from
Exact claim	Quoted from the answer
Source ID	The citation attached, or `none`
Supporting passage	Quoted from the source — the field that does the work
Support level	Direct / mostly / partial / indirect / none / contradicted
Unsupported qualifier	The specific overreach: number, date, scope, superlative, causal
Correction required	Yes / no, and to which page

> NOTE
>
> Three tables joined on Run ID and Source ID, not one wide sheet. The reason is that the relationships are one-to-many in both directions — one run has many sources, one source supports several claims, one claim may rest on several sources — and a single flat table forces you to repeat the run metadata on every row, where it silently drifts out of agreement with itself.

Citation incident log

For anything wrong that you are trying to get corrected. Chapter 47 is the procedure; these are the fields that let you close a row rather than lose track of it.

Field	Purpose
Issue ID	Something to refer to
Incorrect claim	Exact wording as generated
Engine, mode, prompt, date	Reproduction instructions

Field	Purpose
Reproduced	How many of three runs
Source URL	Where the wrong fact came from
Cause class	Stale, entity confusion, self-contradiction, loud negative, hallucination
Correct evidence	Your primary document
Owner contacted, date	Starts the ten-day clock
Page corrected	Date observed
Recrawl observed	Log line or cache date
Answer corrected	Date, and on which engines
Closed	Date, or escalation route

Intervention and verification tracker

Chapter 32's fields, as a sheet: finding, mechanism, action, owner, definition of done, immediate verification, measurement delay, success metric, failure response, rollback.

Monthly measurement note

Written for someone who was not in the room and will read it in ninety seconds. The order is deliberate: the number first, then how it was produced, then what you concluded, then what you will do. Stating the method before the interpretation is what stops a routine fluctuation being read as a result — and it is why the "Method" line is not optional even when nothing about it changed.

```
AI SEARCH — [MONTH]

Share of model: [X]% (previous [Y]%)
Method: [N] fixed prompts, [engines], logged out, run [date]

What changed:
[Three sentences on what moved and why you think so]

New in the citation set:
[Any third-party source now appearing that wasn't before]

Accuracy issues found:
[Anything an engine now says about us that is wrong]

Action this month:
[One or two specific things]
```

Quarterly stakeholder report

```
AI SEARCH VISIBILITY — Q[N]

1. COMMERCIAL
   Organic conversions: [X] ([+/−]% vs previous quarter)
   Branded search volume: [X] ([+/−]%)
   Direct traffic: [X] ([+/−]%)

2. WHAT WE DID
   - [Specific piece of work]
   - [Specific piece of work]
   - [Specific piece of work]

3. VISIBILITY
   Share of model: [X]% → [Y]%
   Method: [N] fixed prompts across [engines], sampled monthly

   Example — "[prompt]"
   Before: [verbatim, brand absent]
   Now: [verbatim, brand present]

4. WHAT WE CANNOT MEASURE
   Our content can inform answers without being credited, and most
   assistants don't pass referrer data. We therefore report sampled
   visibility and proxy measures rather than a direct traffic figure.
   [Keep this wording consistent every quarter.]

5. NEXT QUARTER
   - [Priority and the gap it addresses]
   - [Priority and the gap it addresses]
```

Developer request

Written to be forwarded without editing. It leads with the mechanism rather than the request, because a developer who understands *why* client-side rendering is invisible to a crawler will usually propose a better fix than the one you asked for. Keep the named crawlers in: “AI can’t see us” is arguable, “these three specific user agents receive an empty page” is testable, and a developer can verify it in a minute.

```
Subject: Content visibility to AI crawlers — [page templates]

Context: Several major crawlers (OAI-SearchBot, ClaudeBot,
PerplexityBot) do not execute JavaScript. Anything rendered
client-side is invisible to them, which means we can't be cited
in ChatGPT, Claude or Perplexity answers.

What I've found:
- [Template]: main content not present in initial HTML response
- [Template]: price not present in initial HTML response
```

```
- [Template]: reviews loaded via third-party widget

What I'm asking for:
Server-side rendering or static generation for these templates, so
the main content, prices and review text are present in the HTML
before JavaScript executes.

How to verify: View Page Source (not Inspect) and search for the
body text. If it isn't there, a crawler can't see it.

Happy to prioritise if all three isn't feasible — [template] is the
most commercially important.
```

Content brief

```
BRIEF: [Working title]

Target question: [The question this page must answer completely]

Fan-out — sub-questions this page must answer in self-contained
sections:
1. [sub-question]  → H2: [question-shaped heading]
2. [sub-question]  → H2: [question-shaped heading]
...

What only we can say:
- [Original data / client example / failure mode / practitioner detail]

Specifics required — no vague claims:
- [Figure or fact needed]  → source: [where it comes from]

Author: [Named person with relevant credentials]

Structural requirements:
- Every H2 phrased as a question
- First sentence under each H2 answers it directly
- Each section readable in isolation
- One comparison table
- Published and updated dates visible

Success measure: appears in answers for [specific prompts from the
tracked prompt set]
```

CHAPTER 60

Answer Key

Worked answers to all 172 self-check questions

WHAT YOU'LL LEARN

- Worked answers for all 172 self-check questions, chapter by chapter
- Written to be compared against your own, not memorised

Compare rather than copy. Where your answer differs but is defensible, yours may well be better — several of these questions have more than one good response. Where you cannot answer at all, the chapter reference tells you where to go back to.

Part I – Foundations

ANSWERS

Chapter 1 · What Actually Changed

1. The system does not choose whole pages; it pulls out sections and evaluates them separately. The fourteenth-ranked page may contain one section that answers this precise question cleanly and completely, while the second-ranked page discusses it in passing inside a section about something else. Being retrieved is about the quality of the specific passage, not the standing of the whole page.

2. First: which queries were sampled? The figure varies enormously by query intent, so an aggregate is meaningless without the mix. Second: which tracker, and over what period? Estimates from different providers ranged from 25% to 60% for AI Overview prevalence in 2026, with Google itself citing roughly 50%.

3. The generative layer retrieves its candidates from the classical index. If you are not crawled, indexed and ranking for something, you are not in the pool from which the answer is assembled. Technical SEO stopped being the prize and became the entry fee, but you still have to pay it.

4. It showed that among passages *already placed in a model's context*, those containing statistics, citations and quotations were quoted more often. It did not show that those edits help a passage get retrieved, that the effect transfers between engines, that it survives competitors doing the same, or that it produces traffic.

ANSWERS

Chapter 2 · How Search Engines Actually Work

1. Crawled, then indexed, then ranking — in that order. Check whether the crawler is permitted and able to fetch the URL; then whether the page is in the index (Search Console will say “Crawled — currently not indexed” if it was fetched and rejected); only then investigate ranking.

2. Crawl budget is the finite fetching allocation an engine gives your site. A 200-page site is fully crawled regardless. A 500,000-page site can burn its allocation on parameter URLs and near-duplicates, leaving real products crawled rarely and updates registering weeks late.

3. Modern retrieval matches meaning rather than character strings, using embeddings. The system already understands that “trainers”, “sneakers” and “running shoes” are near-identical in meaning, so repeating variants adds nothing and makes the copy worse for the humans whose engagement does still matter.

4. Duplicates split your signals. If the same content sits at four URLs, links, engagement and authority scatter across all four instead of concentrating on one, and the engine, not you, decides which becomes canonical.

ANSWERS

Chapter 3 · How Language Models Actually Work

1. The model predicts plausible continuations of text; it does not consult a store of facts. Fluency and accuracy are produced by different things, so nothing forces them to coincide. There is no internal signal for “I do not know”.

2. The system converts text into a position in a space where similar meanings sit close together. It already knows your synonyms are the same idea, so listing them adds no new signal — it only makes the sentence worse to read.

3. Training influence means your content was in the corpus used to set the model’s parameters; it is slow, indirect, has a cutoff, and cannot be measured. Retrieval influence means your page is fetched at the moment of the question; it is fast, direct, and the only one of the two you can measure and act on.

4. Either the price is not present as text in the HTML — it is rendered by JavaScript or sits inside an image, so no crawler can read it; or a stale third-party source states an old price and is being retrieved instead of you.

ANSWERS

Chapter 4 · How AI Search Systems Work

1. The question arrives conversationally; it fans out into sub-queries such as nutritional comparison, environmental impact, allergen considerations and cost; each retrieves candidates from the index; those documents are chunked into passages; passages are reranked and the survivors enter the context window; the model writes and attributes. A good page can fail at any stage — blocked crawler, not indexed, no passage matching a particular sub-query, a passage that only makes sense in context, or being used without attribution.

2. Because one question becomes eight to twelve retrieval events. A single page cannot be the best passage for all of them, but a cluster of pages covering the facets thoroughly can be the best passage for several.

3. Read the section in complete isolation, with only its heading. If a stranger would understand it and get an answer from it, it is retrievable. If it depends on "as we saw above", it is not.

4. Absorption is when a model uses information from your content to shape an answer without citing you. It makes reporting hard because the influence is real, potentially large, and produces no citation, no click and no measurable trace.

Part II – The Landscape

ANSWERS

Chapter 5 · Google: AI Overviews and AI Mode

1. AI Overviews are generated summaries placed above ordinary results on a normal search page, which the user did not request. AI Mode is a separate, opt-in conversational surface. AI Overviews affect your traffic far more today, because that is where the volume is.

2. Export your top queries by impressions, classify them by intent, sample a subset manually to see which trigger an AI Overview, then calculate the share of your *impressions* — not your query count — sitting on triggering queries.

3. They have opted out of having the content used to train Gemini. They have done nothing about AI Overviews, which retrieve from the ordinary Google index via `Googlebot`.

4. The summary satisfies the people whose question was simple. The ones who still click have a more specific or more serious need, so the remaining traffic is smaller but better qualified.

ANSWERS

Chapter 6 · The Assistants

1. Reddit supplies roughly 24% of Perplexity's citations and around 0.1% of Gemini's. The same tactic is therefore close to decisive on one engine and irrelevant on another.

2. Perplexity always retrieves and always cites, so it shows its sources transparently — ideal for diagnosis. It should not drive strategy because its user base is a fraction of ChatGPT's; measure where you can see, prioritise where your buyers are.

3. `GPTBot` collects training data; `OAI-SearchBot` fetches pages for ChatGPT's search retrieval. Blocking the first protects your content from training use. Blocking the second makes you invisible in ChatGPT answers — a very different and usually unintended outcome.

4. That no source dominates. There is no incumbent occupying the answer the way a single site can occupy position one, so a well-executed newcomer has a genuine opening.

ANSWERS

Chapter 7 · Who Gets Cited, and How to Judge a Claim

1. Genuine first-hand experience — add original testing, photographs and named failure modes. Question-and-answer structure — make your headings the questions. Dense specifics — replace vague claims with figures. Unhedged comparison — publish honest comparisons that concede where a competitor is better.

2. What was measured — citations, mentions, or traffic? On which engine, and was it tested on more than one? Did it affect retrieval or only usage once retrieved? Does it survive competitors publishing research too? Is there a mechanism — yes, unique data is the only source for a fact, which is a strong and plausible one. It survives the filter better than most claims, which is why the book recommends it.

3. Because the prompt set *is* the measurement instrument. Changing the questions changes the result for reasons unrelated to your work, making before-and-after comparison meaningless.

4. Because specific tactics have a half-life of months in this field, and the filter does not. It lets you evaluate claims that did not exist when this book was written — including claims that contradict it.

Part III – Technical Foundations

ANSWERS

Chapter 8 · Crawlability and the AI Crawlers

1. Do you mean the training crawlers or the search crawlers, because blocking the second removes us from the answers our customers read? And do we have a licensing position that blocking would strengthen, or are we giving up visibility for no commercial gain?

2. It stops your content being used to train Gemini. It does not remove you from AI Overviews, does not affect `Googlebot`, and does not reduce your exposure to generated summaries in Google Search at all.

3. Whether the CDN or WAF is blocking bots above the application layer; whether server logs show `PerplexityBot` actually arriving and what status code it received; and whether the pages are indexed and reachable at all, rather than orphaned or JavaScript-rendered.

4. `robots.txt` records your intent. Logs record what actually happened. A CDN rule, rate limit or IP block can prevent access without any change to `robots.txt`, and only the logs will show it.

ANSWERS

Chapter 9 · Rendering, JavaScript and What Bots Can Read

1. View Page Source shows the raw HTML as delivered by the server. Inspect Element shows the page *after* JavaScript has run. Most AI crawlers see only the first, so Inspect Element will reassure you about content a crawler cannot actually see.

2. Googlebot renders JavaScript in a second pass, so it eventually sees client-side content. Most AI crawlers do not render JavaScript at all, so the same content is permanently invisible to them.

3. Reviews — the most citable content you own, and commonly delivered by a third-party widget. Prices — decisive for commercial queries and frequently script-injected. Tabbed content, including specifications — often absent from the HTML unless the tab is selected.

4. *We need the main content of these page templates present in the initial HTML response, before JavaScript runs. Several important crawlers don't execute JavaScript, so anything rendered client-side is invisible to them, which keeps us out of AI answers. Can we use server-side rendering or static generation for these templates? You can verify with View Page Source — if the body text isn't there, a crawler can't see it.*

ANSWERS

Chapter 10 · Architecture and Internal Linking

1. Because architecture is how a machine works out what your passages are about and how much standing they carry. Internal links supply discovery, authority distribution and semantic context — all of which feed the retrieval candidacy of individual passages.

2. No passage on it answers anything, so nothing on it is retrievable. A hub that only routes readers elsewhere is invisible to a system that evaluates sections in isolation.

3. "What does *[product]* integrate with?"; "How do you actually work with clients?"; "What support do you get, and how fast?"

4. Because anchor text and the sentence around it are the semantic signal telling a system what the target page is about, and that association matters more when the system is matching meaning rather than serving a ranked list.

ANSWERS

Chapter 11 · Structured Data and Schema

1. Most confident: it enables rich results, which is directly observable. Next: it disambiguates entities, connecting your brand to known identifiers. Least, though still real: it assists machine understanding and verification, which both Google and Microsoft have confirmed helps their AI systems.

2. Because it is almost certainly a correlation with a large confound. Sites with complete, accurate structured data are overwhelmingly sites with competent technical teams, better content and stronger authority. The schema marks quality operations rather than necessarily causing the visibility.

3. It links your entity to its profiles elsewhere — LinkedIn, Wikipedia, Crunchbase. It matters because entity recognition works by corroboration across independent sources, and `sameAs` states those connections explicitly rather than leaving them to be inferred.

4. First, it violates the guidelines and can cost you rich results entirely. Second, and more importantly here, you are feeding a cross-checking verification system a contradiction between your markup and your visible page — which damages exactly the trust signals you want.

ANSWERS

Chapter 12 · llms.txt and the Emerging Standards

1. A curated Markdown guide at your site root, telling AI systems what matters and how to describe you. The idea is reasonable because it adds editorial curation on top of the access control `robots.txt` provides.

2. That almost nothing reads it. The most telling figure is that in a 90-day window covering over 500 million AI bot visits, only 408 requests targeted `llms.txt` — with roughly 97% of such files never fetched at all.

3. Displacement is spending effort on a novel, interesting-sounding file while the things that demonstrably matter — crawler access, rendering, passage structure — go unaddressed. It is the real risk because the cost is not the hour spent, it is the work not done.

4. *I've added it — it's an hour's work and costs nothing. But I want to be straight that the data doesn't support treating it as a priority: around 97% of these files are never fetched and Google has said it doesn't support them, so I've prioritised crawler access and page structure instead.*

Part IV – Content

ANSWERS

Chapter 13 · Writing for Retrieval

1. If this section were lifted out and shown to a stranger with nothing else from the page, would it make complete sense and answer something? If not, it cannot be cited.

2. *"What do you actually do for clients?"* followed by something specific: *"We run three services: operations audits, systems implementation and fractional ops leadership. Most clients start with a four-week audit at £6,500."*

3. Because it usually restates content already on the page in question form, giving the retrieval system two near-duplicate passages competing with each other rather than one strong one.

4. Because AI systems cross-check claims across sources. An invented figure that contradicts other sources damages your trust signals, whereas honest vagueness merely fails to help. Being wrong is worse than being unhelpful.

ANSWERS

Chapter 14 · Entities and Topical Authority

1. Because the system cannot resolve your brand name to a specific, identifiable thing. If it cannot confidently tell who you are, it will not confidently associate you with a topic, and it will not offer you as an answer, however good the page is.

2. Conflict-of-interest editing breaches Wikipedia policy, is detected routinely, and the article will likely be deleted while attracting lasting negative attention. The better route is to declare the conflict on the talk page and request changes with sources, and to use Wikidata, which has a much lower barrier and directly supports entity recognition.

3. Commit or cut. For each tail topic, either plan six to ten more pieces to build genuine depth, or accept you will not be an authority there and stop publishing into it. Owning two topics beats being adjacent to fourteen.

4. Entity systems connect organisations to people. A named author with a real profile page, stated credentials, `Person` schema and `sameAs` links to external profiles becomes a recognisable entity, and that recognition transfers to the organisation they write for.

ANSWERS

Chapter 15 · E-E-A-T, Trust and Why It Now Matters More

1. Because a wrongly cited source produces a confidently wrong answer with a name attached, which is more damaging than a bad link in a list, so retrieval systems weight source reliability heavily. And because systems cross-check retrieved passages against each other, so corroboration and consistency decide which source survives.

2. Because the other three take time you cannot compress: credentials, years of reputation, sustained consistency. Experience you can add immediately — original photographs, specific numbers from your own work, named failure modes, and process detail only a practitioner would know.

3. Your Money or Your Life: health, finance, legal and safety topics. Expertise requirements are much stricter, and content without demonstrable professional credentials struggles badly, so credentialed authorship becomes a precondition rather than an enhancement.

4. A price on your site differing from your pricing page or a marketplace listing; a headcount differing from LinkedIn; a founding date differing from the public company register. Any of these gives a cross-checking system an unresolvable contradiction.

ANSWERS

Chapter 16 · Citable Assets

1. Because it competes with nothing. Formatting tactics work only while few people use them and erode as adoption rises, whereas if you are the only source for a fact, anyone answering a question involving that fact must reference you.

2. Survey your own audience — 200 responses supports honest medians. Report your own operational data, anonymised and aggregated. Test products systematically against consistent criteria. A fourth option: start tracking one measurable thing quarterly, which compounds into the only longitudinal source in your field.

3. Because stated limitations make the data verifiable and let others use it responsibly. A survey with a declared sample size and recruitment method is safer to cite than one presenting a bare percentage, and safer to cite means more cited.

4. The asset's value comes entirely from being freely retrievable. A gated report earns no citations, no mentions and no authority — you trade the whole strategic benefit, which compounds, for a handful of email addresses, which do not.

ANSWERS

Chapter 17 · Content Decay and the Refresh Discipline

1. In classical search an outdated page slowly lost rankings. In AI search it can actively contradict current information from other sources, and unresolvable contradictions damage the trust signals for your whole domain. Decay became a liability rather than merely a lost opportunity.

2. Refresh — still relevant and getting traffic, but the facts have aged. Consolidate — overlaps with two or more other pages. Rewrite — the topic matters but the execution is weak. Delete — no traffic, no relevance, no links.

3. Because it starts from existing authority, links and history, whereas a new page starts from nothing. The same effort applied to an established URL compounds with what is already there.

4. Deleting the removed URLs instead of redirecting them. A 301 passes most of the accumulated value to the merged page; deletion discards every link and signal those pages had earned.

ANSWERS

Chapter 18 · International and Multilingual AI Search

1. Because thinner corpora mean fewer well-structured answers already exist. In English, being the best answer to a specific question is hard; in many other languages the question often has no good answer at all, so the bar to being retrieved is far lower.

2. Literal translation of headings, which breaks alignment with the sub-queries people actually ask; translated examples and figures, which leaves wrong currencies and jurisdictions; and translated first-hand experience — which is the one that destroys the experience signal, because an account of testing something in Manchester demonstrates nothing about the German market.

3. Because it measures your performance against questions nobody types. Translation preserves meaning but not idiom or query formulation, so your share of model is calculated against a set of phrasings that do not occur.

4. When your entity signals in the new market are weak and you need the parent domain's accumulated authority. A country-code domain signals local commitment but starts from zero standing — a serious cost when retrieval candidacy depends so heavily on the classical layer.

Part V – Authority Beyond Your Site

ANSWERS

Chapter 19 · Multimodal GEO

1. Because the data — the thing that makes it citable at all — is sealed inside a picture that the retrieval pipeline cannot read. The article describes the method and the conclusion, the figures exist only as pixels, and an assistant can therefore confirm the study exists but not say what it found. Always publish the numbers as text alongside the chart.

2. Something like: `alt="Organic sessions falling 34% year on year, with the steepest drop from January"`. It describes what the image shows and contains the specific claim, so it can answer a question. `alt="dashboard"` can answer nothing.

3. The first sentence must stand alone when read aloud — no dependence on a table, a list or a preceding sentence — and it must resolve the question in roughly forty words, because longer spoken answers get truncated or passed over.

4. You can influence the corroborated facts about you across independent sources, your Wikidata entry, your `Organization` and `Person` schema with `sameAs`, and your Google Business Profile. You cannot create a panel on demand or edit one directly. Panels are a consequence of entity recognition, not a lever on it.

ANSWERS

Chapter 20 · Links, Mentions and What Changed

1. A model learns associations from text. If many sources write that your brand specialises in a topic, the association is learned from the sentences themselves — a hyperlink is a machine-readable pointer that the language understanding does not require.

2. Co-occurrence is how often your brand name appears near a topic across the web. Measure it by searching your brand name in quotes, reading several pages of results, and tallying what topic your brand is associated with in each — then comparing that against the topic you want to own.

3. Because links still feed the classical index, the classical index supplies the retrieval candidates, and link-derived authority remains an input to reranking. The accurate statement is that links are necessary but no longer sufficient.

4. Podcast appearances; expert commentary in response to journalist requests; genuine participation in professional communities. Inclusion in third-party comparison round-ups is a fourth.

ANSWERS

Chapter 21 · Share of Model and Brand Presence

1. The proportion of relevant generated answers that mention or cite your brand. It beats rank because in AI search there is no list for the customer to scan — being part of the answer is the outcome that matters, and rank no longer describes it.

2. Retrieval-time presence is fastest — your pages, or pages mentioning you, being fetched and used. Training-data presence you cannot influence directly at all; it is the slow accumulation of having been written about before a cutoff you do not control.

3. Because they are the pages engines actually retrieve for commercial queries, so inclusion changes what answers say. They are also unusually actionable: many accept corrections, and fixing an outdated price on a heavily cited round-up is often a two-line email.

4. Because share of model is only meaningful relative to a stable instrument. Change the questions and the number moves for reasons unconnected to anything you did.

ANSWERS

Chapter 22 · The Platforms That Feed the Models

1. Because the engines weight sources very differently. Reddit content is specific, experiential, comparative and answer-shaped, which suits Perplexity's retrieval; Gemini leans on established web sources, Wikipedia and YouTube instead.

2. Declare your conflict of interest on the article's talk page and request the change with sources. Never edit the article about your own organisation directly.

3. Because the transcript is text, and text is what gets chunked, retrieved and cited. The video itself is not the retrievable asset; the accurate caption file is.

4. Because these communities detect and punish it. Reddit bans domains rather than just accounts, and the resulting critical threads are themselves highly citable, so a discovered campaign becomes permanent negative material appearing in AI answers about your brand.

ANSWERS

Chapter 23 · Digital PR for AI Visibility

1. Because the mechanism is corroboration — your brand associated with your topic across many independent sources. Twenty trade mentions repeat that association twenty times in the relevant topical context; one national feature states it once, to a larger audience but with less corroborative weight.

2. Original data; a specific quotable claim attributed to a named person; a number that becomes the reference figure for a question; a named phenomenon that sticks; and a clear extractable definition.

3. Because unlinked mentions now carry much of the weight links used to, so refusing coverage that will not link means refusing something genuinely valuable for no gain.

4. Because cross-checking systems find the discrepancy. A figure appearing as "30%", "over a third" and "nearly 40%" across three articles creates exactly the contradiction pattern that pushes a source down.

Part VI – Commerce and Conversion

ANSWERS

Chapter 24 · When the Answer Is Wrong About You

1. Stale source — fix the source page, not yours. Entity confusion — strengthen entity signals (Chapter 15). Your own contradiction — fix your site. A loud negative — build corroboration rather than suppressing. Hallucination — often unfixable; a clear canonical answer gives the system something better to reach for.

2. Check the source before touching anything. Ask the engine where it got the figure, open that page, and confirm the wrong price is stated there. In the worked case it was a stale marketplace listing, so rewriting your own pricing page would have achieved nothing at all — the answer was never coming from your site.

3. Because it reconciles the contradiction rather than adding a third competing claim. A retrieval system that finds £149 on one page and £189 on another has to choose; a page that says "it is £189 now, it was £149 during a 2024 promotion that has ended" explains both and supersedes the old one.

4. A model's training-data associations, genuinely true criticism, pure hallucinations with no traceable source, and another company's reputation when you share a name. Saying so early stops a team burning weeks on the unfixable and lets you spend the effort on rungs one and two, where corrections actually work.

ANSWERS

Chapter 25 · E-commerce and Product Discovery

1. Through content — a page of yours is retrieved and cited like any other passage. And through the product graph — structured product data from your schema, merchant feed or marketplace listings used to name, price and compare actual products, bypassing your prose entirely.

2. Because `gtin` is how a system matches your listing to the same physical product elsewhere. Without it your listing is isolated; with it, it joins a product record that already carries reviews, price history and comparisons. A description can be inferred from the page; identity cannot.

3. Because a perfect spread reads as unreliable to a cross-checking system, and because it gives that system nothing to answer "what are the downsides" with. Someone will answer that question; it is better that the source is you.

4. Because query fan-out generates use-case sub-queries. A page answering the narrower question has far less competition and matches a specific sub-query directly, whereas "best grinder" competes with every major publisher.

ANSWERS

Chapter 26 · After the Citation

1. Verifying an answer they already have, rather than seeking one. They are further down the funnel because the comparison happened inside the assistant, more sceptical because they are checking a claim, landed deep with no context about who you are, and impatient with repetition.

2. Because a page perfectly optimised for retrieval is by construction the page an assistant can summarise most completely, so the visitor arriving from that summary has already read its substance. Unless the page holds more than the answer, it has nothing left to offer.

3. It answers a question this visitor has already resolved. They have decided you are roughly right; the appropriate next step is concrete — see the price, check stock, book the call — not another stage of research.

4. Because withholding the answer removes you from the answer entirely: you lose the citation that brought the visitor and the visitor with it. The correct move is to be fully citable while holding more depth than a summary can carry.

Part VII – Measurement

ANSWERS

Chapter 27 · Local and Service-Area GEO

1. The answer set is bounded by geography, so the candidate pool is dozens rather than millions. AI Overviews trigger far less on local and transactional queries. And verification carries unusual weight, because sending someone to a business that has closed is a much worse failure than a slightly wrong statistic.

2. Because the cost of being wrong is higher and more immediate. A wrong statistic is an inaccuracy; wrong opening hours send a real person to a locked door. Systems compensate by leaning on verified, structured, corroborated data — which is why the business profile outweighs almost everything else.

3. They contain no self-contained answer, no specific information and no experience signal, and they create a substantial duplicate-content problem. Every quality mechanism in this book scores them badly. Write properly about the few places that matter and cover the rest with one well-structured service-area page.

4. Because “plumber near me” rarely triggers a generative answer at all — it returns a map. The informational version does trigger one, is a question you can genuinely answer better than competitors, and reaches people at the point where they are deciding how to choose rather than who to call.

ANSWERS

Chapter 28 · The Measurement Problem

1. Discoverability — partially observable via server logs. Citation — observable by sampling a fixed prompt set. Absorption — barely observable at all. Economic outcome — observable only indirectly, through proxies.

2. Absorption is your information shaping an answer without attribution. Raise it proactively because it is a genuine and permanent gap in your reporting; a stakeholder who discovers it themselves concludes your measurement is unreliable, whereas one you told in advance concludes you are rigorous.

3. Because it rises when more people encounter your name and later look you up — which is precisely the behaviour an AI citation produces. It is also free, reliable, and part of the feedback loop where citation drives branded search, which strengthens the entity signals that drive citation.

4. *Across a fixed set of 50 prompts run monthly on ChatGPT and Perplexity, logged out, we were mentioned in 34% of answers this month against 26% last month — seventeen answers against thirteen. That is a sample rather than a census, and four answers is within the run-to-run variation we have measured, so I am treating it as encouraging rather than as a trend. If it holds next month, it is a trend.* The named method, the raw counts and the refusal to promote a four-answer move are what make it defensible; Chapter 29's repeated-run protocol is what earns the right to say more.

ANSWERS

Chapter 29 · Building a Measurement System

1. Because that is how the systems are used, and they behave differently for each. Measuring keyword-shaped inputs produces a number describing behaviour that does not occur.

2. Because conversation memory and personalisation contaminate the result — the engine remembers you asked about your own brand and treats you differently. A fresh, logged-out session from a consistent location is the only way to get comparable measurements over time.

3. Source concentration. It tells you which third-party domains the engines actually rely on for your category — which becomes your target list for corrections, inclusion and participation, and it is information you cannot get any other way.

4. When you need to track hundreds of prompts, cover multiple markets or languages, or demonstrate rigour to a board, and when the analyst time saved exceeds the licence cost. Not when you are one person tracking one market with 50 prompts.

ANSWERS

Chapter 30 · Attribution and Proving Value

1. Someone asks an assistant and is told about you; they do not click, they read; days later they search your brand name directly; they arrive as direct or branded organic traffic; they convert, and the analytics platform credits direct or branded search.

2. Because a stakeholder given a precise figure that later collapses stops trusting everything else you report. Volunteering the limitation converts a vulnerability into evidence of rigour, and it removes the possibility of being caught out.

3. Because they make the abstract concrete. Reading an assistant recommend their own company, when it did not three months earlier, communicates the value immediately in a way no chart does.

4. *Referral data shows around 400 sessions, but that materially undercounts it — most assistants don't pass a referrer and most people never click at all. The better indicators are that branded search is up 22% and fourteen people told us on forms this quarter that they found us through an AI assistant. I'd rather give you two real numbers than one fabricated one.*

Part VIII — Practice

ANSWERS

Chapter 31 · Testing and Experiment Design

1. Because there is one index and one version of your page — you cannot serve variant B to half the retrieval systems. The replacement is a quasi-experiment: change one thing for a test group of pages, leave a matched control group untouched, and compare the two trajectories.

2. Core updates and engine changes that move everything at once; seasonality in category demand; and competitor activity. It also protects you from yourself, by revealing when a rise you were about to claim credit for happened to untouched pages as well.

3. Because it stops the result being reinterpreted after the fact, which is the commonest way marketing tests deceive the people running them. A recorded prediction and a stated falsifier turn a demonstration into a test.

4. That the direction is encouraging and the size is not measurable. At 50 prompts and a rate near 20%, the noise is roughly ±11 points, so a six-point move is inside it. Report it as suggestive, and either enlarge the prompt set or wait for the trend to accumulate.

ANSWERS

Chapter 32 · The Audit Playbook

1. Each stage is cheaper than the one after, and an early failure makes later stages irrelevant. Reordering means auditing content quality on a site whose crawlers are blocked — days of work producing findings that cannot matter until the blocker is fixed.

2. Asking your infrastructure team explicitly whether AI bot blocking is enabled on the CDN or WAF. It is now the leading cause of unexplained AI invisibility and is invisible in `robots.txt`.

3. High-effort, low-impact work: replatforming for SEO reasons, chasing every engine. Naming it explicitly is useful because it converts a vague "we're not doing that" into a defensible decision you can justify when someone proposes it.

4. Because it must summarise what the audit actually found, and because it is the only section many readers will read. Writing it first means writing it before you know the answer.

ANSWERS

Chapter 33 · The GEO Toolkit

1. Whether AI crawlers are being blocked, and what assistants actually say about you. Both cost nothing — the first from your server logs and `robots.txt`, the second from asking the assistants directly in a logged-out session.

2. That the engine fetched your page, evaluated it, and decided not to keep it. It is not a crawling problem and not an error; it is a quality or duplication judgement, and the fix is on the page rather than in your `robots.txt`.

3. Roughly above 150 tracked prompts, or when you cover two or more markets — the point at which the analyst hours saved exceed the licence. Below that, a spreadsheet gives you the same signal and a better understanding of it.

4. Because three months of reading answers yourself teaches you what your category actually sounds like, and means you can tell when a dashboard is wrong. Teams that start with a platform learn the numbers without learning the market and cannot sanity-check either.

ANSWERS

Chapter 34 · Using AI to Do the Work

1. Because a model produces the statistical average of what has been written on a topic, and citation goes to what is uncommon — specificity, first-hand experience, original data. Generating content with a model optimises directly against the property that earns visibility.

2. Choosing the topic and angle, because it requires knowing what you know that others do not. Filling the gaps only you can, because that is the original substance. Writing sections containing experience or judgement, for the same reason. Verifying every fact, because models confabulate confidently. Addressing the critique and adding the point of view, because models produce hedged consensus. Reading each section in isolation, because that judgement is the whole test.

3. Because you are asking the same class of system that performs the decomposition to show you how it would decompose. It is not a proxy for the mechanism; it is the mechanism.

4. That there is no penalty to detect or appeal. Google judges quality, not production method, so nothing flags your content. It simply never gets retrieved, because nothing about it is worth retrieving, and the outcome looks like ordinary underperformance.

ANSWERS

Chapter 35 · The Prompt Library

1. The fan-out generator, then the content-gap prompt, then the content brief. The order is the point: the fan-out tells you which sub-questions exist, the gap prompt tells you which of them you currently fail to answer in a self-contained passage, and only then do you know what the page actually has to contain. Reversing it means briefing a page before you know what it is competing for.

2. Because models default to affirming the work put in front of them. Left unconstrained, the sceptical-review prompt returns a list of strengths and two mild suggestions, which is useless. Explicitly forbidding compliments removes the politeness bias and produces criticism specific enough to act on.

3. *Review the product page copy below. List every claim that is vague, unquantified, or could be made by any competitor selling the same category. For each, show the phrase, the specific fact that would replace it — price, dimensions, compatibility, dispatch time, warranty, and where in our product data that fact would come from. Do not rewrite; just identify.*

4. Generic inputs produce generic outputs. With nothing specific to work from, the model returns the average of its training corpus, which is exactly the undifferentiated content that never gets retrieved (Chapter 34). The bracketed sections are where your own specificity enters the prompt; leaving them unfilled removes the only thing that could make the output worth publishing.

ANSWERS

Chapter 36 · Strategy and the First 90 Days

1. Because without a baseline you cannot demonstrate that anything you did worked, and you will be asked. It is also the only way to know which problem you actually have before spending a quarter solving a different one.

2. The baseline measurement, and the crawler access check. The first makes everything afterwards measurable; the second is where the single largest and most common failure hides.

3. `llms.txt` implementation — 97% of such files are never fetched. Tracking all five engines — two done properly beats five done badly. Buying a visibility tool in month one — build the manual system first or you will not understand the numbers.

4. Invert months two and three. With no existing content there is nothing to restructure, so build one small, deep topic cluster instead, and halve the prompt set to around 25 to keep the measurement sustainable for one person.

ANSWERS

Chapter 37 · The GEO Practitioner's Week

1. Because answers vary between runs even for the same prompt on the same engine. A daily reading measures that variance, not your performance, and it invites reacting to noise. Monthly sampling against a fixed set is the smallest interval that carries signal.

2. The thirty minutes of genuine platform participation. It compounds because standing in a community accumulates slowly and then gets cited for years — and it is dropped first because it produces nothing measurable in the week you do it.

3. The technical fixes are complete and their gains banked, while the content and authority work has not yet landed. Announcing it in advance turns a predicted feature of the work into evidence your model is right; discovering it unannounced looks like failure and is where programmes get cancelled.

4. Because tools change and every skill on the list transfers. Someone who can read a page as a machine does, write answer-first, and evaluate evidence will learn any new platform in a week. Someone who only knows a platform has to start again when it changes.

ANSWERS

Chapter 38 · Working with Stakeholders

1. Validate the concern, then quantify it specifically: calculate what share of your impressions sit on AI-Overview-triggering queries, present that figure, and give a plan for those queries and a separate one for the rest. Specificity converts panic into confidence.

2. That the published evidence for specific tactics is weak — a 2026 survey of 45 studies found almost nothing with a proven cross-platform effect — which is exactly why you measure your own results against a fixed prompt set rather than trusting industry claims. What is well supported is that crawler access, content structure and entity consistency affect whether you can be retrieved at all.

3. That you are sampling rather than measuring, with the method stated. And that some of your influence will be invisible because of absorption and referrer loss.

4. Because uniform confidence is indistinguishable from not having checked. Saying "this is well supported, this is plausible but untested" demonstrates that you have evaluated the claims, which makes your confident statements worth more.

Part IX – Instrumentation

ANSWERS

Chapter 39 · Measuring Citations in Practice

1. A mention means your name appeared in the prose; it does not establish that anything of yours was retrieved. The model may be reciting something absorbed during training, which can be months out of date and cannot be corrected by publishing. A citation means retrieval happened on that run, from a live document — which is a system you can influence this quarter.

2. Because prompts the engine answered generically, naming nobody at all, say nothing about your competitive position. Including them in the denominator drags every brand's number down and makes the metric jump around as engines change how often they name anyone. Dividing by prompts where naming occurred isolates the question you actually care about: when the engine picks names, does it pick yours?

3. Retrieval, not reputation. The model knows you well enough to name you from training data, but is not pulling a live document of yours into the context window. That points at crawlability, freshness or passage structure — Part III first, then Chapter 13 — rather than at authority work, which is what a low share of model would indicate.

4. Run the identical prompt set twice within one week, changing nothing in between, and compare share of model. Whatever it moves by is noise. It matters because on a 50-prompt set a single prompt is two percentage points, so movements of several points are routinely meaningless — and reporting them as results destroys your credibility the first time someone checks.

ANSWERS

Chapter 40 · Building a Citation Tracker

1. Because it is the only part that knows anything about a particular vendor — the endpoint, the request shape, and where citations sit in the response. Isolating it means switching engines, or surviving a vendor's change of response format, is an edit to one short function rather than to the whole tool. Everything else is yours and is durable.

2. A key pasted into a cell travels with the sheet. Anyone you share it with gets it, and shared sheets are copied, exported and inherited long after the person who made them has moved on. Script properties are not visible to viewers of the spreadsheet. If a key has already been exposed this way, revoke it rather than hoping.

3. Because a true/false column cannot be re-examined. If you later sharpen your definition of *named*, or want to know whether you were described as a specialist or as one of several options, the classification cannot answer it and the evidence is gone. The verbatim column lets you re-score old runs against a new definition without re-running them.

4. Any two of: it measures an API rather than the consumer product, so treat it as an index and not a replica — check periodically by hand. It cannot see Google AI Overviews, which have no supported API, so those stay a manual check on the Chapter 29 routine. It has no view on whether you were described well, so a human reads the verbatim answers monthly. And it will drift when vendors change models, so investigate the tool before believing a sharp movement.

ANSWERS

Chapter 41 · Watching the Crawlers

1. Summaries are aggregated according to the tool's idea of what matters, and AI crawlers are commonly bundled into a generic "other bots" category. That hides both which specific crawler arrived and what response it received — the two things you need. Raw logs also let you pivot by URL, which shows that the pages being fetched are often not the ones you assumed.

2. That your server refused the request. The crawler tried, and was turned away. This is invisible from every other vantage point: `robots.txt` will still state your intent to allow it, the page will still load fine in your browser, and nothing in your analytics will show it. It is gate one failing silently, and it is usually a CDN or firewall rule rather than anything you configured deliberately.

3. Take the IP address from the log line and look up its reverse DNS name; then look that name back up and confirm it resolves to the same IP. A genuine crawler resolves to a hostname on the operator's own domain, and that hostname resolves back to the original address. An impostor fails at the first step, having no reverse record on the operator's domain, or at the last, where the name resolves somewhere else entirely.

4. Part IV and Part V. A `200` proves access is working, so the failure is upstream of anything technical: either the passage is not structured to be liftable and self-contained (Part IV, and Chapter 13 in particular), or nothing corroborates you elsewhere so the system has no reason to prefer you (Part V). This is the *crawled but not cited* row, and it is routinely misdiagnosed as a crawlability problem because access is the easier thing to check.

Part X — Forensics

ANSWERS

Chapter 42 · Evidence Status: Actual, Inferred, Unknown

1. Because it is generated after the fact by a system with no privileged access to the reasoning that produced the earlier answer. It is not reading a log; it is producing the kind of text that tends to follow that question. An accurate account is a coincidence of a good guess with the truth, and you have no way to tell which of those you are holding.

2. ACTUAL, but only within the panel's documented scope. The panel is a record of what was *cited*. It is not a record of what was retrieved, what was opened, or what was considered and dropped — those remain unknown, and treating the panel as a complete retrieval record is the most common overreach in this area.

3. "Preferred" and "because" are both claims about an internal decision nobody can observe, stated in the past tense as though recorded. Rewrite as: "The competitor's guide was cited and ours was not. Theirs carries a 2026 update date and ours is undated — but the selection reason is not observable." That version keeps the evidence, drops the invention, and leaves the investigation open.

4. It tells you which conclusions are safe to revise. When new evidence contradicts something you labelled INFERRED, you change one line. When it contradicts something you reported as fact, you have to explain to the person who acted on it why they were told something untrue — and that cost is paid in credibility rather than in effort.

ANSWERS

Chapter 43 · Citation Decision Forensics

1. Because improvement only operates on pages that are in the pool. If the retrieval layer never surfaced your page as a candidate — because it is not indexed for that sub-question, or the crawler was blocked — then the quality of its writing is irrelevant to the outcome. You would be improving something that was never being evaluated.

2. OPENED, and only for your own pages. A server log records which crawler fetched which URL and what status it received. It says nothing about anyone else's pages, and nothing about what happened to yours after it was fetched.

3. Returned-but-not-opened is a promise problem: the title and snippet did not signal that an answer was inside, so the remedy is question-shaped headings and front-loaded specifics. Opened-but-not-used is a passage problem: the content was read and did not stand alone or did not answer directly, so the remedy is Chapter 13's restructuring. One is about being chosen; the other is about being usable once chosen.

4. Because several of the thirteen sections are genuinely unobservable — which sources were considered and rejected, the reranking weights, whether personalisation applied. A report with no unknowns has filled those in with plausible sentences, which means the confident parts and the invented parts are now indistinguishable to whoever reads it.

ANSWERS

Chapter 44 · Claim-Level Validation

1. It establishes that the source was involved somewhere in producing that region of text. It does not establish that the source states the claim, that it states all of the claim, or that it does not contradict it. Placement is an output of generation, not an assertion of support.

2. Because generated answers routinely add qualifiers the source did not contain — a rounded number, a tightened scope, a superlative. The source carries the underlying fact and not the addition, which is by definition partial support. Qualifiers are cheap to generate and each one is a separate claim requiring separate evidence.

3. It means your passage is being read as answering a question it does not answer, which will produce further misattributions and cannot be defended if challenged. The fix is on your side: make the section's actual scope explicit in the heading and the first sentence, so the boundary between what you claim and what you do not is visible in the text rather than left to inference.

4. Because the verbatim passage converts the request into a five-minute administrative fix. The publisher can see the discrepancy themselves, in their own words, without taking your word for anything or investigating. A description of the error asks them to do the work of finding it, which is the difference between a correction and an argument.

ANSWERS

Chapter 45 · Source Lineage and Independence

1. Because ten domains may carry one piece of evidence. Syndication, press-release propagation, a shared study, common ownership and copied errors all produce many URLs from a single origin. Corroboration only counts when sources could have disagreed and did not — and pages repeating the same release could not have disagreed.

2. Search a distinctive phrase of roughly twelve words from the cited page, in quotation marks. Syndicated copies surface immediately, because the text is identical. It is the fastest way to collapse a domain list into a chain list.

3. Because there is no independent evidence in the category, so original data from you creates a new chain rather than joining a crowded one. Publish a survey of your own audience, your own operational data, or systematic comparative testing — with the method stated, so it can be cited as an origin rather than as another opinion.

4. Because it overstates your position to the people making decisions on it. One release, one blog post and three pickups is a single chain with five URLs. Reporting it as five independent sources means a stakeholder believes your evidence base is five times more robust than it is, and the correction arrives at the worst possible moment.

ANSWERS

Chapter 46 · Counterfactual Testing

1. Because the citation is then a property of one particular wording rather than of your content. A user asking the same thing in any of the other four ways does not encounter you. Reporting the successful run as visibility describes a position you do not hold.

2. That you are winning retrieval but are absent from what the model knows without searching. Your content is being found and used when the system goes looking; it did not shape the model's unaided knowledge. That is a good position to be in — it is the half you can influence this quarter — but it means the standard-mode answer is not something content work will change.

3. Because within one conversation the model carries its earlier answers as context and converges on its own previous output. The paraphrases then agree because they are downstream of each other, not because your position is stable, and the run manufactures exactly the finding it was supposed to test.

4. Because the rerun produced a new answer under a new instruction. It shows what the evidence landscape looks like with that domain excluded. It says nothing about what the original, unconstrained system would have done — which is unobservable, and therefore UNKNOWN rather than inferable from the probe.

ANSWERS

Chapter 47 · The Correction Runbook

1. Because the answer may not reproduce. Generated answers are non-deterministic, products change, and the version you found may be gone within a day. An incident you cannot show is one you cannot escalate to a publisher, a platform or your own leadership, however clearly you remember it.

2. A single-engine error points at one source that engine favours, which narrows the investigation to that engine's citation set. An error present on all four points at something upstream that everyone reads — a widely syndicated article, a marketplace listing, a registry entry — and the correction there fixes all four at once.

3. Because if the wrong fact exists in a real source, this is an outreach problem with a specific target: get that page corrected. If it exists in no source, it is a hallucination, and outreach has nothing to aim at — your only remedy is publishing a canonical first-party statement and building enough corroboration that the true fact is easier to retrieve than the invented one.

4. Because the corrected page has to be recrawled before any engine can see the change, and because the error may persist on engines other than the one that surfaced it. Closure requires observing the recrawl — a log line or a cache date — and then rerunning the affected prompts across every engine, with the dated evidence of both states.

ANSWERS

Chapter 48 · The Evidence Archive

1. Any three of: answers are non-deterministic and may not reproduce; source pages are edited without a visible correction note; publication dates are silently re-stamped by a CMS; URLs redirect to something other than what was cited; source panels are redesigned or removed between product versions; models are retired; cited pages are assembled dynamically and no longer contain the passage.

2. Because a domain cannot be re-examined. You cannot find the passage that was used, check whether the page was later edited, or verify what it said at the time. The URL is the unit at which evidence exists; the domain is the unit at which reporting happens, and conflating them destroys the audit trail.

3. It lets you filter, count and compare across runs without opening anything. The per-run folders hold the evidence behind each row; the manifest is what makes a hundred runs a dataset rather than a hundred folders.

4. Because the gap between capture and decision is typically several days, and that is exactly the window in which a publisher quietly corrects a page. Archive at capture and you can show what it said. Archive at decision and you may be left asserting something you can no longer evidence.

Part XI – Playbooks

ANSWERS

Chapter 49 · Copy-Paste Implementation

1. Because a crawler obeys only the block that names it. Once it finds its own user-agent, it stops looking, so the `User-agent: *` group applies only to crawlers not named anywhere else. This is why a named crawler cannot inherit the general rules, and why forgetting to name one has a very different effect from denying it.

2. It lists other places the same entity is described — a LinkedIn company page, a Wikipedia article, a Wikidata item. That makes the corroboration mechanism from Chapter 14 explicit rather than leaving a system to infer that the “Kettleworth” on your site and the one on LinkedIn are the same organisation.

3. Because a tag manager injects the markup with JavaScript after the page loads, and most AI crawlers do not run JavaScript. The markup is present for a browser and absent for the crawlers whose behaviour you were trying to influence, so it looks deployed and does nothing.

4. You want to see `200`. A run of `403` means your server is refusing the crawler — access is failing at gate one, `robots.txt` notwithstanding, and it is almost always a CDN or firewall rule rather than anything in the file you published.

ANSWERS

Chapter 50 · The Page Teardown

1. Rewrite the headings as questions. It is the highest-leverage single change, because the heading is where chunk boundaries fall and where sub-query matching happens — and because writing the question first exposes which sections have no answer, which makes the rest of the edits obvious.

2. Because it refers to something the passage does not contain. Lifted out and shown alone, the sentence points at a product the reader has no information about, which is exactly what self-containment fails. The information is not wrong; it is simply unusable in isolation.

3. Self-containment. It asks whether the passage carries its own context, or depends on material that will not travel with it when a retrieval system extracts it. A section that fails this can be perfectly good prose and still be unusable as a passage.

4. Because the heading is the specification. Once it states the question the section must answer, most of the body edits follow mechanically — the answer goes first, the qualification after, and anything not bearing on that question becomes visibly out of place.

ANSWERS

Chapter 51 · GEO at Scale

1. Somewhere around two thousand pages, and above twenty thousand nothing requiring a person per page is possible at all. What replaces it is systematic work: template-level changes, filtered queues that surface only the pages worth touching, and sampled human review rather than exhaustive review.

2. Extraction and transformation — pulling headings out, reordering an existing sentence, drafting a question-shaped heading from supplied text. The human gate belongs wherever facts could enter: every model-drafted row is reviewed before publication, and every batch is spot-checked for numbers that were not in the input.

3. Because it gives the schema, the merchant feed and the product pages a single source, so they cannot drift apart. Chapter 25 established that contradictions between those three resolve against you; generating from one source makes the contradiction structurally impossible rather than something to be audited for.

4. Because either alone is mostly noise. Clicks fall when SERP layout changes or an AI Overview appears, without any loss of position. Impressions fall with seasonality and query volume. A page down on both has most likely lost retrieval or ranking, which is the population worth a person’s attention.

ANSWERS

Chapter 52 · Industry Playbooks

1. Because buyers in that category ask comparison questions, and the answers are assembled overwhelmingly from third-party comparison content — review platforms, community threads and independent round-ups. The vendor's own page competes for a small share of an answer mostly built from other people's writing about it.

2. Its AI exposure: how many of its twenty most commercially important queries produce an AI answer at all. If the number is very low, the correct response is a monitoring job costing an hour a month rather than a transformation programme — and knowing that saves more money than any tactic.

3. Because an article about a study is a secondary source, and in YMYL categories provenance is the standard being applied. It is also Chapter 45's argument at its sharpest: ten articles about one paper are one piece of evidence wearing ten coats, and a category judged on provenance will discount them accordingly.

4. Run ten commercially important prompts on two engines and tally the *type* of every cited source — vendor page, review platform, community thread, trade publication, primary research, regulator, marketplace. The resulting distribution is your category's standard, it is current, and it is specific to your market in a way no published playbook can be.

ANSWERS

Chapter 53 · The Testing Lab

1. Because self-containment is a textual property, visible in the text itself: does this passage refer to something it does not contain? That is the kind of question models answer reliably. The scores are judgements about how a system would behave, which is exactly what a model has no privileged access to.

2. Because a score invites you to feel finished. "7 out of 10, quite strong" ends the exercise; a list of five specific problems and three replacement sentences is work you can do. Every constraint in that prompt exists to stop the model producing encouragement in place of findings.

3. Because a page that half-answers a sub-question is one restructuring pass from answering it, and restructuring costs an afternoon. A GAP requires writing and publishing something new. The partial rows are the cheaper win, and they are usually more numerous than anyone expects.

4. The word INFERRED. Everything these prompts produce is a model's approximation of what a retrieval system might do — not an observation of what any engine did, and not a prediction with a known error rate. Chapter 42's doctrine applies most sharply here, because a simulation that feels like a measurement is the easiest mistake in the book to make.

ANSWERS

Chapter 54 · The Capstone

1. Because category leaders are cited largely for reasons of brand mass — they are named in training data, they are the subject of the comparison articles, they are the default answer. Nothing you learn from studying them transfers to a company that has to earn its position, which is every company you will actually work for.

2. Evidence capture and reproducibility, at 35% between them. They score badly because analysis is easy to produce and an audit trail is not: writing a confident diagnosis takes an afternoon, while archiving every artefact with timestamps and metadata is unglamorous work that nobody sees until someone challenges a finding.

3. Because a priority order tells you what to do next and nothing about whether it worked. The verification fields — mechanism, owner, definition of done, how it will be verified, expected delay — turn the backlog into a project-management document, and make it possible to close a row rather than merely doing it.

4. That afterwards you hold opinions about AI search that come from your own evidence rather than from anyone's writing, including this book's. The ambition is not agreement with the method; it is that you are now equipped to check it.

Part XII — Horizons

ANSWERS

Chapter 55 · Agentic Search

1. Readable means a system can extract and understand your content. Operable means a system can complete a task on your site — select a variant, determine returns terms, reach a price without an interaction it cannot perform.

2. Prices set by JavaScript, because the agent reads raw HTML (Chapter 9). Stock status in a widget, for the same reason. Consent walls served before content, because the agent indexes the wall rather than the page (Chapter 8). Each was already a problem for AI crawlers.

3. Because the question is not "can we technically permit it" but "do we want to be purchasable this way, what does it do to our returns rate, and how do our fraud rules treat it". Those belong to whoever owns revenue, not to whoever owns `robots.txt`.

4. That the differentiation is invisible to the comparison. An agent filtering on price, stock and returns terms never forms an impression, so brand, design and persuasive copy do not enter the decision. The response is to ensure your genuine advantages exist as machine-readable attributes, not only as prose and photography.

ANSWERS

Chapter 56 · Ethics and Second-Order Effects

1. Fragility — a moderation change or outage at one platform propagates directly into answers about medicine, law and money. And manipulability — if few sources shape many answers, influencing them becomes disproportionately valuable. The second bears on Chapter 22 because that incentive is precisely what produces the seeding and sockpuppeting the chapter warns against.

2. Classical search summarised in order to send you somewhere; generative search summarises in order to substitute for going there. Publishers allowed crawling in exchange for traffic, and the traffic side of that bargain has largely stopped arriving — around 83% of AI Overview queries end without a click.

3. Language — non-English perspectives are under-weighted and sometimes answered from English sources. Digitisation — expertise that predates the web or sits behind paywalls is largely absent. Establishment — corroboration rewards consensus and disadvantages correct-but-minority positions. Geography and commercial capability are two further valid answers.

4. Because if models increasingly train on model output, the distribution narrows and unusual but correct information is lost. Original, first-hand, specific content is then not just commercially advantageous to you but the input the whole system runs short of — a rare case where the incentive and the systemic interest align.

CHAPTER 61

Sources and Further Reading

Where the evidence in this book came from, and how to keep it current

WHAT YOU'LL LEARN

- The evidence base, with its dates
- How to keep this material current after the book dates

Every time-sensitive figure in this book carries a date, because the mechanics will outlast the numbers. This chapter lists what the claims rest on and, more importantly, how to re-check them yourself.

Primary evidence used

Generative Engine Optimisation, critical survey (arXiv:2607.14035, July 2026) — A review of 45 GEO studies from 2023-2026. Source for: the finding that only topical relevance and context position replicate reliably; the visibility vector; competitive erosion; quality degradation; the reassessment of the original GEO paper's claims. The most important single source behind Part I and Chapter 7.

GEO: Generative Engine Optimization (arXiv:2311.09735, November 2023) — The original Princeton / Georgia Tech / Allen Institute / IIT Delhi paper. Source of the widely quoted 40% figure, and of the conditions that figure depends on.

Ahrefs `llms.txt` study (137,000 sites) — search: *"We Analyzed 137K Sites: 97% of llms.txt Files Never Get Read"*. — Source for the finding that 97% of `llms.txt` files are never fetched.

SE Ranking `llms.txt` adoption study (300,000 domains) — search: *"State of llms.txt"*, SE Ranking, 2026. — Source for the ~10.1% adoption figure.

Semrush and Ahrefs AI Overview CTR analyses — Source for the 30-50% CTR reduction range on the top three positions.

Semrush most-cited domains study (325,000 prompts) and **Similarweb / Peec AI citation analyses** — Sources for citation share by platform, including the Reddit figures across Perplexity, ChatGPT and Gemini, and the ~5% single-domain ceiling.

Wellows AI Overview passage analysis (15,847 AI Overview results, December 2025) — Source for the 134-167 word citation band.

SparkToro zero-click study (first four months of 2026) — Source for the 68% US zero-click figure and the 60.45% 2024 baseline. Search: *"zero-click searches 2026 SparkToro"*.

Tinuiti Q1 2026 AI Citations Trends Report — Nine commercial categories across seven AI platforms over four months to January 2026. Source for Reddit's aggregate citation share and its rate of growth.

Conductor Q1 2026 benchmark and Xponent21 April 2026 measurement — The two ends of the AI Overview prevalence range, quoted to show the spread rather than to endorse either.

Rankability llms.txt adoption tracker (June 2026) — Source for the 8.7% adoption figure among the top 1,000 sites.

Google Search Central documentation and public statements — Source for snippet controls, `Google-Extended` behaviour, and Google's position on `llms.txt`.

How to keep this current

The mechanics in Parts I, III and IV will age slowly. The numbers in Parts II and VI will age quickly. To refresh:

1. **Re-run your own prompt set.** Your data about your market beats any published study, and you already have the method.
2. **Check crawler documentation quarterly.** OpenAI, Anthropic, Perplexity and Google publish their user agents and update them.
3. **Watch arXiv rather than blogs** for the underlying research. Search for "generative engine optimization" and sort by date.
4. **Apply the Chapter 7 filter** to everything new: what was measured, which engine, retrieval or usage, does it survive competition, is there a mechanism.

> KEY IDEA
>
> If you have been answering the self-check questions as you went, worked answers to all 172 are in Chapter 60 — compare rather than copy. The index in Chapter 62 lists every substantive term with the pages that discuss it, the main treatment in bold.
>
> The most durable thing in this book is not any tactic. It is the **method for evaluating claims** in Chapter 7 and the **measurement system** in Chapter 29. Tactics have a half-life of months. Those two have a half-life of years.

A closing note

This field rewards two qualities that are rarer than technical knowledge: the willingness to say "the evidence for that is weak", and the discipline to measure your own results rather than trusting someone else's.

Most of what is published about AI search in any given month will be forgotten within a year. The people still doing this well in five years will be the ones who built a baseline, changed one thing at a time, and were honest about what they could not prove.

That is available to you immediately, and it does not require permission, budget or a technical background.

CHAPTER SUMMARY

- Every time-sensitive figure in this book is dated; the mechanics will outlast the numbers
- The July 2026 critical survey is the most important single source, and it is largely a caution against overclaiming
- Refresh by re-running your own prompt set, checking crawler docs quarterly, and watching arXiv over blogs
- The durable assets are the claim-evaluation filter and the measurement system, not any individual tactic

CHAPTER 62

Index

Every substantive term, with the pages that discuss it

Pages are listed in order. **Bold** page numbers mark where a term is the subject of the discussion rather than mentioned in passing — start there. A trailing ellipsis means the term also appears on further pages not listed.

A

- **absorption** — **46**, 47, 48, **176**, 178, **179**, 189, 232, **367** …
- **agentic search** — 324, 325, **326**, **327**, 328, **329**, 381
- **AI Mode** — 43, 45, 47, **50**, **51**, **53**, 54, **56**, **355** …
- **AI Overviews** — 24, **25**, 26, 28, **50**, **51**, **53**, **69**, **355** …
- **alt text** — 127, **128**, 129, 130, **131**, 206, 339, 341, 362
- **anchor text** — 34, 79, 80, 81, 83, 84, 136, 339, 358
- **answer-first pattern** — **96**, 97, 101, 119, 124, 129, 130, 143, **229** …
- **API** — 91, **208**, 243, **245**, **247**, **248**, **249**, **250**, **373** …
- **attribution** — 46, 47, 164, 167, 177, **187**, **189**, 190, **238** …
- **authoritativeness** — 107, 109, 340

B

- **branded search** — 25, **140**, **178**, **179**, **187**, **188**, **189**, **190**, **368** …

C

- **canonicalisation** — 30, **32**, 33, 103, 153, **155**, **289**, **290**, **340** …
- **case study (Kettleworth)** — **40**, **66**, 77, **106**, **131**, **136**, **145**, **149**, **155**, **160**, **162**, **172**, **190**, **302** …
- **chunking** — 23, **42**, **44**, 45, 47, **82**, **99**, **302**, **314** …
- **citation rate** — 45, 63, 64, 180, 183, **194**, **195**, **240**, **242** …
- **ClaudeBot** — 56, 68, 69, 70, 74, 199, 253, 294, 298 …
- **co-occurrence** — 133, **134**, 135, **136**, 137, 148, 149, 340, 363
- **competitive erosion** — 46, 64, 112, 113, 116, 275, 340, 383
- **consolidation, content** — 117, 118, **119**, **120**, 223, 227, 308, 344, 346 …
- **content decay** — 82, **117**, 119, **120**, 200, 223, **227**, 361
- **context window** — 26, 27, 36, 37, **38**, 39, 40, 45, 46 …
- **control group** — **192**, **193**, **194**, **195**, **196**, 233, 339, 368

- **correction ladder** — 151, 152, 154, 155, 282
- **crawl budget** — 30, **31**, 35, 295, 340, 354
- **crawlers, AI** — **31**, **68**, **70**, **71**, **72**, **73**, **255**, **357**, **378** …

D

- **digital PR** — 134, **147**, **148**, 149, 150, 273, 274, 275, 276 …
- **discoverability** — 27, **46**, 65, **176**, 179, 185, 340, 343, 367
- **doorway pages** — 171, 172, 311, 340

E

- **e-commerce** — 24, 31, 35, 105, 144, **158**, **159**, 160, **161** …
- **E-E-A-T** — 33, 63, 85, **107**, 109, **111**, 126, 165, 224 …
- **embeddings** — 36, **37**, 38, 40, 41, 340, 342, 354
- **entities** — 32, 80, 86, 88, 89, **102**, **103**, 104, 105 …
- **experience (E-E-A-T)** — 63, 123, 142, 210, 211, 343, 356, 362, 370
- **experiment design** — **192**, 194, **196**, 260, 342, **368**

F

- **fan-out, query** — **42**, **43**, **47**, **80**, **81**, **162**, **214**, **215**, **340**, **370** …
- **freshness** — 33, 36, 46, 57, **117**, 120, 240, 257, **278** …

G

- **GEO (generative engine optimisation)** — **27**, **46**, **64**, **127**, **169**, **171**, **173**, **205**, **226**, **341**, **383** …
- **Google Business Profile** — 103, 130, 144, 170, 341, 362
- **Google-Extended** — **53**, 54, 68, **69**, 70, **73**, 199, 253, 294 …
- **GPTBot** — 55, 61, 68, 69, **70**, **71**, 74, 75, **253** …
- **gtin** — 159, **160**, 161, 162, **163**, 285, 296, 297, 341 …

H

- **hallucination** — 36, **39**, 40, 107, 151, 152, **155**, **341**, **365** …
- **hreflang** — 121, **124**, **126**, 207, 341

I

- **indexing** — **26**, **30**, **32**, **34**, 51, **206**, **209**, **341**, **354** …
- **internal linking** — 31, 34, **79**, 81, **83**, 84, 119, 194, 308 …

J

- **JavaScript rendering** — 30, 31, 34, **74**, **75**, 76, **77**, 78, **357** …

K

- **Kettleworth** — **40**, **66**, 77, **106**, **131**, **136**, **145**, **149**, **155**, **160**, **162**, **172**, **190**, **302** …
- **keyword density** — 32, 35
- **knowledge graph / panels** — **127**, **130**, **131**, **341**

L

- **links** — **133**, **135**, **136**, 207
- **llms.txt** — 63, **64**, **90**, **91**, **92**, **93**, 94, 100, **383** …
- **local SEO** — 33, 51, **169**, 170, **171**, **172**, 173, **311**, 313 …

M

- **measurement** — 57, 125, 126, **138**, 139, 140, **141**, **240**, **242** …
- **mentions, unlinked** — 109, 133, **134**, 136, 137, 148, 364
- **multilingual** — **121**, **123**, **124**, 125, **126**, 343, 362
- **multimodal** — **127**, 129, 131, 341, 362

O

- **OAI-SearchBot** — 55, 61, 68, 69, **70**, 71, 72, 73, **252** …
- **orphan pages** — **83**, 84, 128, 206, 342, 357

P

- **passage retrieval** — **23**, **28**, **42**, **44**, **45**, **47**, **79**, **96**, **158**, **239**, **240**, **272**, **315**, **318**, **355** …
- **Perplexity** — 26, 40, **55**, **56**, **59**, **60**, **61**, **69**, **73** …
- **practitioner cadence** — 108, 111, 174, 203, 217, 220, **226**, **229**, **233** …
- **prompt set** — **125**, **126**, **180**, **181**, **183**, **186**, **195**, **196**, **208**, **224**, **226**, 227, **233**, **241**, **244**, **321**, **372** …

R

- **RAG** — **39**, **342**
- **Reddit** — 55, **56**, **59**, 60, **62**, **66**, **142**, **144**, **145** …
- **rendering** — 33, 34, **74**, 75, **76**, 77, 93, **297**, **357** …
- **reranking** — 26, **42**, 44, 45, **47**, 51, 96, 112, **258** …
- **reviews** — 43, 66, 75, **76**, 77, 85, **160**, **161**, **162** …
- **robots.txt** — 30, 31, **70**, **71**, **72**, **73**, **294**, **299**, **369** …

S

- **sameAs** — 86, **87**, **88**, **89**, 102, 104, 106, 109, **296** …

- **schema markup** — 64, 65, 76, 81, 82, **85**, **86**, **87**, **89** …
- **Search Console** — 31, 32, 51, 52, 71, **207**, **209**, **226**, **229** …
- **self-containment** — 23, **24**, 44, 45, 46, 47, 97, 166, **229** …
- **server logs** — 68, 72, **73**, 92, 93, **176**, **251**, **258**, **263** …
- **share of model** — 57, 125, 126, **138**, 139, 140, **141**, **240**, **242** …
- **sitemap** — 30, **31**, 34, **70**, **71**, 79, 83, 199, **295** …
- **specificity** — **96**, **98**, 99, **100**, **101**, **119**, 199, **217**, **370** …
- **structured data** — 76, 81, **85**, **86**, 87, 88, **89**, 102, **342** …
- **surveys, original** — **27**, 38, **46**, 64, **113**, **115**, **116**, **274**, **361** …

T

- **tokens** — 36, **37**, 38, 40, 44, 53, 340, 342
- **tooling** — **25**, 64, 91, 144, **185**, **206**, **207**, **209**, **383**
- **topical authority** — **102**, 103, **104**, **105**, 106, 200, 343, 360
- **training vs inference** — **38**, **39**, 40, 41, 55, 58, 68, 89, **260** …
- **transcripts** — 127, 128, **129**, 131, 136, **143**, 145, 146, 205 …
- **trust** — 99, 107, 110, 118, 119, 123, 149, 162, 165 …

V

- **visibility vector** — 46, 176, 179, 339, 343, 383
- **voice search** — 106, **127**, **129**, 131, 138, 176, 341

W

- **Wikidata** — 103, 104, 106, 130, 131, 143, 145, 296, 320 …
- **Wikipedia** — 55, **56**, **59**, **103**, **104**, **122**, **143**, **249**, **331** …

Y

- **YMYL** — 109, 111, 224, 313, 343, 380
- **YouTube** — **56**, 59, 62, 66, 122, 142, 143, **144**, 145 …

Z

- **zero-click** — 22, **25**, 332, 333, 343, **384**

About the author



Luke McLaughlin has spent ten years running organic growth and demand generation for technical B2B products — SEO, content, paid media and lifecycle marketing, owned end to end and run across eight international markets.

His route into the field was not the usual one. An MSc in molecular biology came before a decade in search, and the working knowledge of Python, SQL and applied AI he picked up along the way is what he uses to build the kind of measurement systems this book argues for. The work has spanned editorial and SEO leadership, content programmes for B2B SaaS, and regulated medical writing — technical subjects where being wrong is expensive.

He writes at Digital Marketing Laboratory and builds local-first AI tools at MultiAgentAI. The cover artwork is his.

lukekevinmclaughlin.com · linkedin.com/in/luke-kevin-mclaughlin · github.com/lukekevinmclaughlin-oss

THE AI SEARCH HANDBOOK

Search stopped handing people a list. It started writing them an answer.

This handbook explains the machinery underneath that change — how retrieval, chunking and citation actually work — and turns it into practical craft you can apply without a technical background. It is deliberately unexcited: where the evidence is strong it says so, and where a widely repeated tactic has been contradicted by data, it says that too, and shows you how to tell the difference yourself.

INSIDE

- How language models and AI retrieval actually work
- Crawlers, rendering and the blocks you did not know about
- Writing passages a machine can lift, and why structure beats length
- Entities, E-E-A-T and the authority you earn off your own site
- International, e-commerce and what to do after the citation
- A measurement system you can run yourself, and a 90-day plan

62 chapters · 62 diagrams · 224 worked answers · quick-reference card
Cover artwork by the author
First edition · Revision 3 · August 2026
Luke McLaughlin · lukekevinmclaughlin.com