NM

PRIMARY RESEARCH · AUGUST 2026

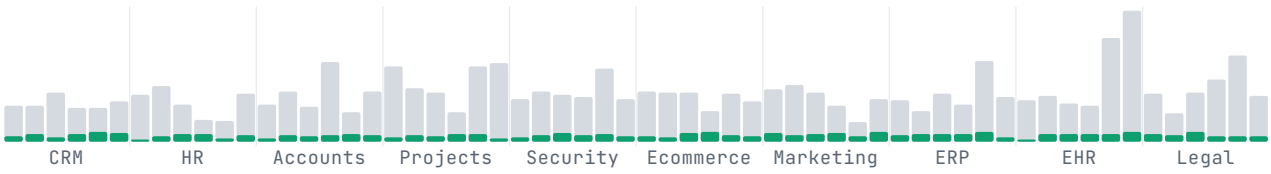
# How ChatGPT Shortlists Software Brands

An audit across 10 categories and 60 buying questions. I recorded what ChatGPT reads, throws away and links to when a buyer asks it which software to buy – and what that decides.


NM

**Nathan Mzumara**

AI Search & Digital Growth Strategist · nathanmzumara.com


Each column is one of the 60 questions. Tall grey bar = results ChatGPT read. Short green bar = links it showed the reader.

| 60 | 10 | 2,680 | 367 | 150 |
|---|---|---|---|---|
| QUESTIONS ASKED | SOFTWARE MARKETS | RESULTS READ | LINKS SHOWN | BRANDS NAMED |

Captured 26 August 2026 · ChatGPT gpt-5-6-mini · every figure is measured, none are estimates

KEY TAKEAWAYS

# Six findings that should change a plan

If you read nothing else, read this page. Every number below comes from the 60 captured answers, not from a survey or an estimate.

**65.9%**

### Your own website is the biggest source of evidence

Two in every three links ChatGPT showed pointed at a software company's own site. In 26 of the 60 answers, every single link was a vendor page. Pricing, plans and feature pages are doing more work than anything else you publish.

**69.3%**

### Most recommendations are not backed by anything

248 of 358 brand recommendations had no link to that brand anywhere in the answer. The links sit next to checkable facts like prices. The advice itself – who is best – usually arrives with nothing attached.

**4.1%**

### G2 and its rivals are barely in the room

All the review sites combined produced 15 of 367 links. Reddit, LinkedIn and YouTube produced none at all. Meanwhile 43 small comparison sites nobody tracks produced 14.4%.

**11.1%**

### Almost everything it reads is thrown away

Of the 2,327 different pages ChatGPT pulled into these 60 answers, only 258 became a visible link. Of 564 websites it consulted, 446 were dropped without trace – so citation-only tracking misses 79.1% of what it actually looked at.

**86.3%**

### Old pages barely get a look

Of the pages that carried a publication date, 86.3% were dated this year and 90.4% within twelve months. Anything meant to be compared against rivals needs a visible date and a real refresh cycle.

**0.34**

### The shortlist changes with the wording

Six phrasings of the same buying question returned brand lists overlapping by a Jaccard score of only 0.338. In HR and payroll, all six questions produced a different winner. Tracking one prompt tells you almost nothing.

THE ONE-LINE VERSION

ChatGPT is not reading review sites to decide which software to recommend. It is reading vendors' own pages to check facts, and pulling the recommendation itself out of what it already believed.

CONTENTS

# What is in this report

Twenty-one sections in four parts. The eight findings each lead with the number they are about.



### If you have five minutes

- Read the key takeaways on page 2 – every headline number with the sample it rests on.
- Then section 15 for the six things I would actually do about it.

### If you want to check the work

- Section 03 covers the capture method, 04 covers the model, 16 lists what this audit cannot tell you.
- Section 19 holds every number behind every chart, and section 20 lists ten further reads.

A NOTE ON THE NUMBERS

Nothing here was typed by hand. Charts and prose read from one computed dataset, and every headline figure was independently recomputed from the raw captures before publication.

01

# Why I audited ChatGPT's software picks

Your buyers have started asking ChatGPT which software to buy. Nobody can tell you how it picks.

This is not an accident of usage. OpenAI has been building the commercial machinery for it – see why hiring a chief revenue officer turns ChatGPT into a B2B buying channel.

I have worked in search for over twelve years. For most of that time the job had a shared map. You could see the results page, you could see who ranked, and you could work out roughly why. That map does not cover what is happening now. A buyer types a question, gets a shortlist of five vendors, and there is no results page to inspect. The shortlist simply appears.

Most of what gets written about this is guesswork. The common story goes: AI reads G2 and Reddit, so collect reviews and get talked about on Reddit. I wanted to know whether that was true, so I stopped guessing and measured it.

I ran 60 software buying questions through ChatGPT and recorded everything it did. Not just the answer, but the searches it ran, the pages it read, the pages it threw away, and the handful it finally showed. Then I sorted every one of those websites by hand into what kind of site it actually was.

This report is what came back. Some of it confirmed what I expected. Several findings did not, and one of them should change how most software companies spend their next quarter.

### Who this is for

If you run marketing, growth or a software company, the practical question is simple: when a buyer asks ChatGPT about your category, are you on the list, and what evidence is it showing them? This report answers the second half of that question with measured data, and gives you a way to test the first half yourself.

> **WHAT MAKES THIS DIFFERENT**
>
> Every AI visibility tool on the market shows you citations – the links that appear in an answer. None of them show you what the model read and discarded. I captured both, and the gap between them turns out to be where most of the story is.

02

# What I tested: 10 markets, 60 questions

Ten software markets. Six ways of asking. Sixty questions in total, each in its own fresh chat.

Ten categories covering most of the B2B software market. They are deliberately unlike each other: some publish pricing openly, some hide it; some have two obvious leaders, some have thirty vendors fighting.

| CATEGORY | WHAT A BUYER IS LOOKING FOR | PRICING |
|---|---|---|
| **CRM** | Sales software for a growing team | Public |
| **HR & payroll** | People and pay systems for 50–500 staff | Mixed |
| **Accounting** | Bookkeeping and finance software | Public |
| **Project management** | Work and task tracking for teams | Public |
| **Cybersecurity** | Endpoint protection for a company network | Mixed |
| **Ecommerce platforms** | The system an online shop runs on | Public |
| **Marketing automation** | Email, campaigns and lead management | Mixed |
| **ERP & manufacturing** | The core system a factory runs on | Gated |
| **Healthcare EHR** | Patient record systems for clinics | Gated |
| **Legal practice management** | Case and matter software for law firms | Mixed |

## Six ways of asking the same thing

This turned out to matter more than anything else. Real buyers do not all ask the same way. Some type four words, some name two vendors, some state a budget. Each category was asked all six ways.

| | SHAPE OF THE QUESTION | THE CRM VERSION I USED |
|---|---|---|
| **A** | Plain best-of | What is the best CRM software? |
| **B** | Best-of with a company size | What is the best CRM software for a 200-employee UK company? |
| **C** | Head-to-head | Compare Salesforce vs HubSpot for a mid-market B2B sales team. |
| **D** | Alternatives to a brand | What are the best alternatives to Salesforce? |
| **E** | Budget and requirements | I need a CRM under $50 per user per month that integrates with Xero and Outlook. |
| **F** | Category and vendor landscape | What is CRM software and who are the leading vendors in 2026? |

The full 60-question matrix, every answer and every link is in the data appendix.

03

# How I recorded what ChatGPT was doing

I did not copy answers off the screen. I recorded the live data stream sitting behind them.

When ChatGPT answers you, it sends a continuous stream of data to your browser. That stream carries far more than the words you see. It carries the name of the model doing the work, every search result the model received, and the exact links it chose to show.

I added a small piece of code to the page that copies that stream as it arrives. The browser displays one copy as normal; my copy is saved and analysed. Nothing about the question or the answer was changed.

### What was recorded for all 60 questions

- **Which model answered.** Read from the stream itself, not from the interface selector.
- **Every search result received** – web address, page title and publication date. This is *what it read*.
- **Every link shown to the user** – the source buttons in the answer. This is *what it cited*.
- **The full answer text**, and which brands it named, in the order it named them.
- **How long each answer took**, to the millisecond.

The gap between the second and third of those is the heart of this audit.

### How the websites were classified

The 200 most-read websites, plus every website that earned at least one link – 225 in total – were sorted by hand into ten types, and those ten names are used everywhere in this report: vendor's own site, comparison site, review site, news and trade press, analyst, agency or partner, testing lab, government, Reddit, LinkedIn and YouTube, and Wikipedia. Where a site's nature was not obvious, I opened it and read it before deciding. Domains were normalised, so `help.asana.com` and `www.asana.com` count as one website.

| 60 | 2,680 | 367 | 358 |
|---|---|---|---|
| QUESTIONS RUN | RESULTS CAPTURED | LINKS CAPTURED | BRAND PICKS LOGGED |

04

# Why the free model, and why it matters

Every question in this audit ran on the free version of ChatGPT. That was a deliberate choice, and it is the version most of your buyers are using.

### Most people asking are not paying

OpenAI has reported roughly **50 million paying subscribers** across Plus, Pro, Team and Enterprise, against a user base that passed **one billion weekly users** in 2026. Depending on which figures you line up, published estimates put the free tier at somewhere between **81% and 94%** of active users.

The exact number is arguable. The direction is not. The overwhelming majority of people typing "what is the best CRM" into ChatGPT are on the free tier. If you want to know what your buyers are being told, that is the version to measure.

**~1bn**
WEEKLY CHATGPT USERS, MID-2026

**~50m**
PAYING SUBSCRIBERS, ALL TIERS

**81–94%**
ESTIMATED SHARE ON THE FREE TIER

### Which model actually answered

The free tier does not let you pick a model – a router chooses one for you. So rather than trust the interface, I read the model name out of the response stream for every question. All 60 resolved to the same model, `gpt-5-6-mini` – not one was handled by a larger reasoning model, and not one was answered without searching the web first: retrieval fired on **100% of questions**.

### How each question was run

- Each question ran in its **own fresh conversation**, so no answer could influence the next.
- 54 of 60 ran in **temporary chat**, excluded from history and personalisation. The other six are declared in the limitations.
- No custom instructions, no uploaded files, no prior context – just the question, on 26 August 2026 from a single UK connection.

**WHAT THIS DOES AND DOES NOT COVER**

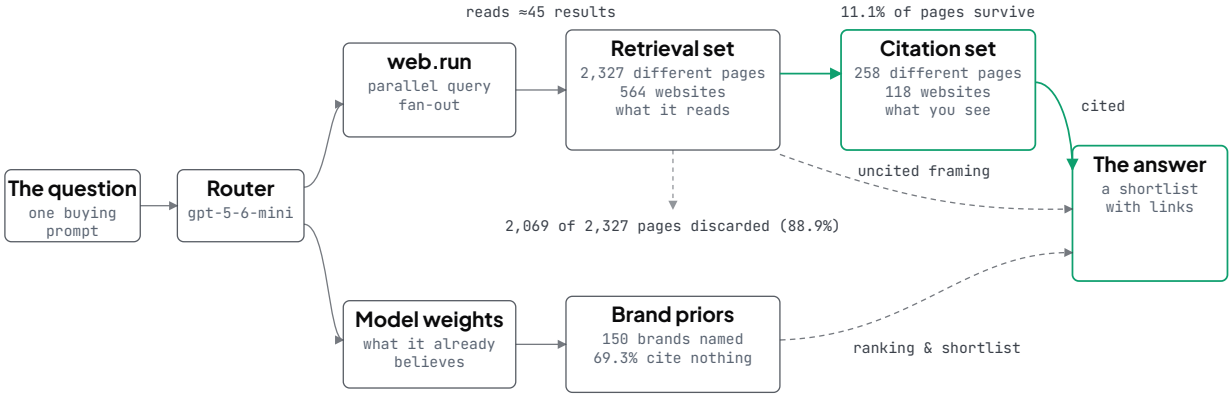
This audit describes the free tier, which is what most buyers see. It does not describe the paid tier. A Plus or Pro account on a larger reasoning model may search more widely and cite more independently – the biggest open question left by this work.

Usage sources: getairefs.com, August 2026 (citing BNN Bloomberg, 31 July 2026 and OpenAI, April 2026); backlinko.com, February 2026 (citing Bloomberg, Reuters, CNBC). They disagree on the free-tier share because they compare weekly active users with total subscribers.

05

# How a shortlist is actually assembled

A recommendation is not looked up. It is assembled – and most of the assembly happens where nobody can see it.




SOLID = THE READER CAN SEE IT · DASHED = IT SHAPES THE ANSWER INVISIBLY

Three things reach the answer. Only one – the 367 links – is visible to the reader. The other two, 2,069 discarded pages and what the model already believed about the category, do most of the work on the shortlist and leave no trace on the page.

Three separate things feed the answer, and only one of them is visible to the buyer.

| WHAT REACHES THE ANSWER | HOW MUCH | CAN THE BUYER SEE IT? |
|---|---|---|
| **The links shown as sources** | ~6.1 per answer | Yes – this is the whole evidence trail |
| **Pages read but never linked** | ~39 per answer | No |
| **What the model already believed** | Most of the ranking | No |

That is the single most important idea in this report. When a buyer sees five sources under an answer, they read it as “this is where the recommendation came from”. It usually is not. It is where a few checkable facts came from. The recommendation itself often arrived with nothing attached.

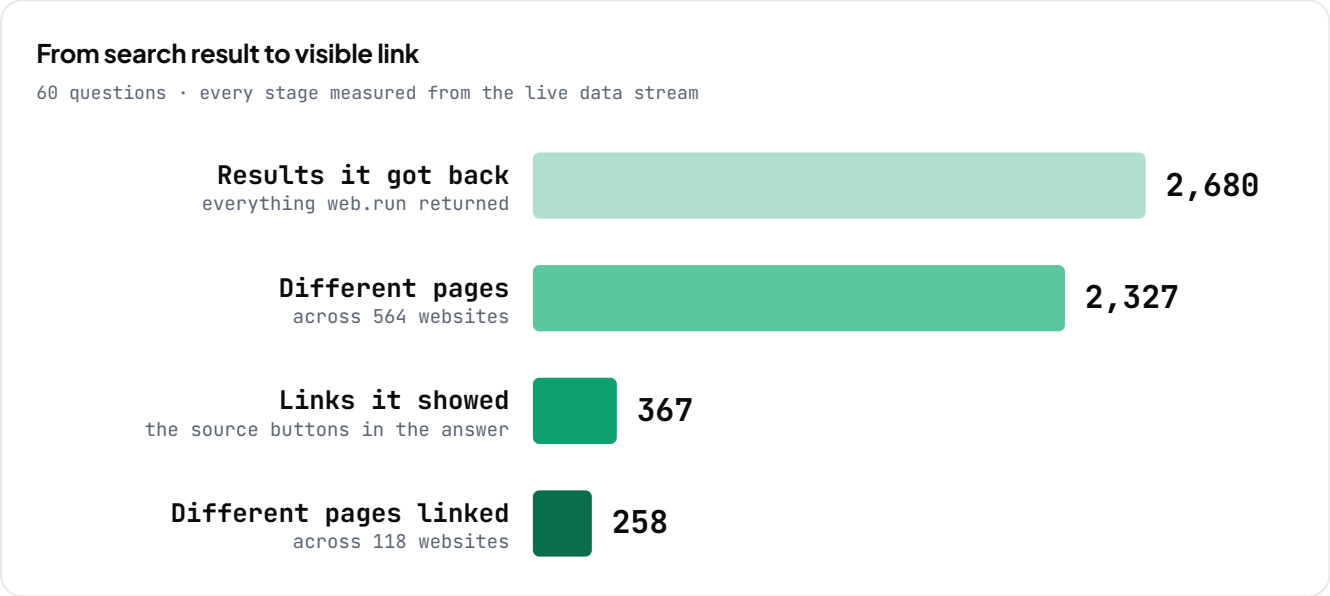
THE SHORT VERSION

ChatGPT reads about 45 search results to write about 6 links. The other 39 still shape the answer. They just leave no trace anywhere you can measure with a citation tracker. I have written separately on how the shortlist gets written before the search even runs.

06

# 88.9% of what ChatGPT reads never shows

Almost everything ChatGPT reads is thrown away before the answer reaches the buyer. That gap is where visibility is decided.




**11.1%**
OF PAGES READ BECAME A VISIBLE LINK

**20.9%**
OF WEBSITES READ WERE NAMED ANYWHERE

**446**
WEBSITES READ THEN DROPPED ENTIRELY

Two survival rates matter and they say different things. **11.1% of pages survive**: if ChatGPT reads your page, there is roughly a one in nine chance it links to it. **20.9% of websites survive**: if it reads anything at all from your site, there is roughly a one in five chance your site gets named.

The second number should change how you measure. ChatGPT read pages from 564 different websites across these 60 questions and linked to 118. If you only track citations, you are blind to 79.1% of the sites it actually consulted – and some of those are shaping how the category gets described.

One more pattern: the reading scales, the linking does not. The heaviest question pulled 120 results, the lightest 18. But every answer showed between 2 and 9 links. Contested categories generate enormous invisible research and the same thin visible evidence.

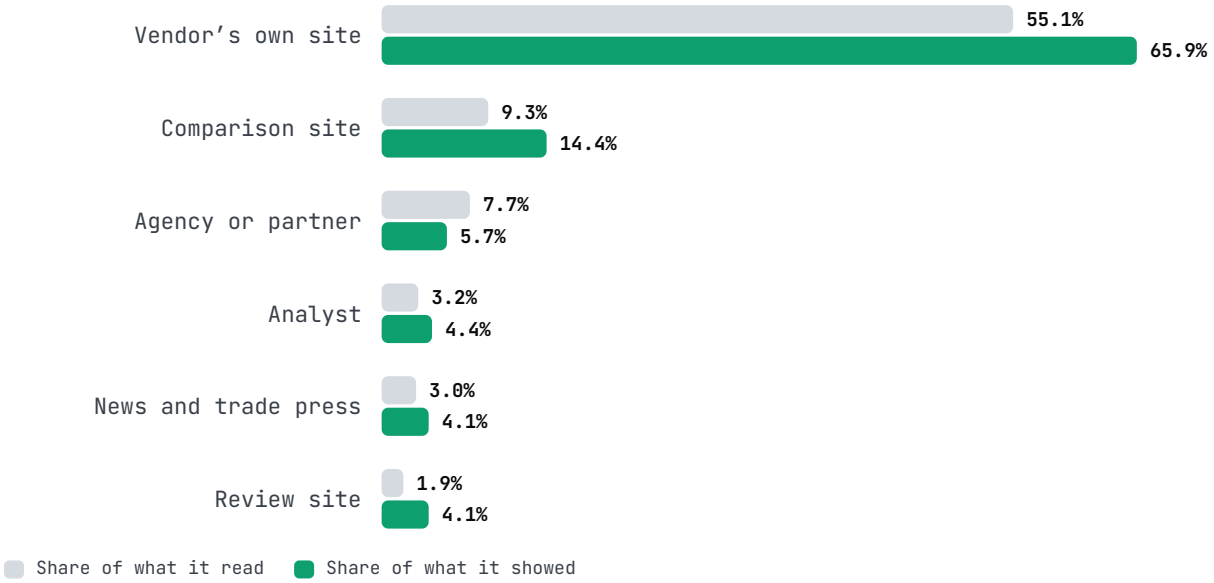
07

# 65.9% of links point to vendors' own sites

Not G2. Not Reddit. Two in every three links pointed at a software company's own website. This was the biggest surprise in the audit.

### What ChatGPT reads compared with what it shows

The six kinds of website that produced 1% or more of the 367 links. Full table on page 30




**65.9%**
OF ALL LINKS ARE VENDOR PAGES

**26/60**
ANSWERS WHERE EVERY LINK WAS A VENDOR

**4/60**
ANSWERS WITH NO VENDOR LINK AT ALL

**278**
PRICING PAGES READ

This is not ChatGPT trusting marketing copy. It is ChatGPT looking for facts it can check, and your pricing page is genuinely the best source for what your product costs.

The practical read is simple. If your pricing, plans, features and integrations are not written down clearly on your own site, you have handed that job to someone else – and someone else will guess.

There is a second effect worth sitting with. In 26 of 60 answers, every single source was a vendor site. The buyer sees a confident, well-sourced recommendation in which the only evidence comes from the companies selling the thing. It looks independent. It is not.

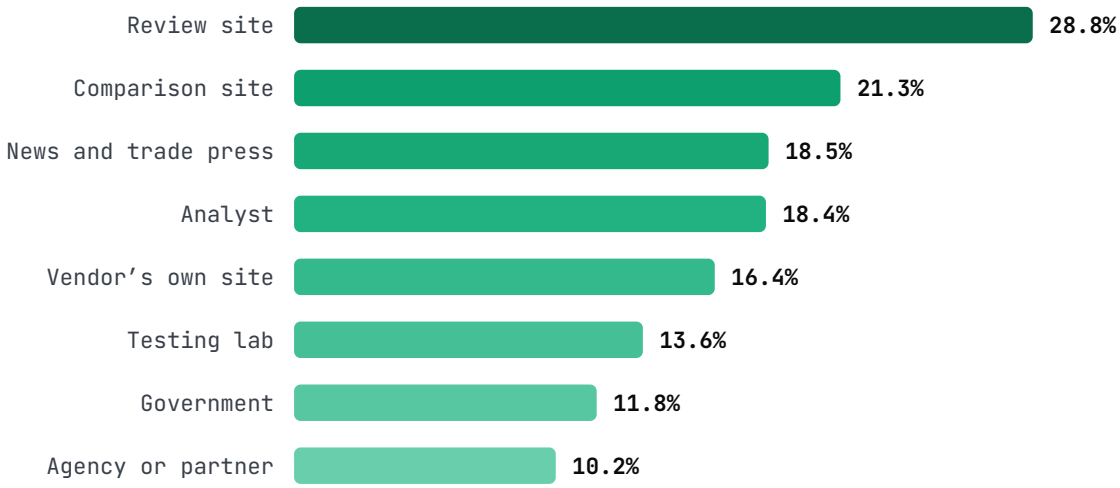
08

# 4.1% of links come from G2 and its rivals

The review sites are trusted. They are also almost never read. That combination is far worse for them than it sounds.

**How likely each kind of site is to be linked, once read**

Of the pages read from that kind of site, the share that became a visible link




**4.1%**
OF LINKS FROM ALL REVIEW SITES COMBINED

**28.8%**
OF REVIEW PAGES READ BECAME A LINK – THE BEST RATE OF ANY SOURCE

**0**
LINKS FROM REDDIT, LINKEDIN OR YOUTUBE

Read those together and the mechanism is clear. When ChatGPT does find a G2 or Capterra page it links to it more often than anything else – 28.8% of the time. It almost never finds one: across 60 questions and 2,680 results, review sites came up 52 times. They are not losing on trust. They are losing at the search step, before trust comes into it.

The community story does not survive contact with the data either. Reddit was read six times across the whole audit and linked zero times. LinkedIn, six times and zero links. YouTube, nineteen times and zero links. If your AI visibility plan rests on seeding Reddit threads, this audit gives you no evidence it works for software buying questions.

WORTH SITTING WITH

A lot of software marketing budget currently points at review sites on the assumption that AI reads them. Here they produced 4.1% of what buyers saw. Vendor websites produced 65.9%. That is a sixteen-fold difference, and the cheaper of the two is the one that wins.

09

# 14.4% of links come from sites nobody tracks

The second biggest source of links was a group of small comparison sites that almost no marketing team is monitoring.

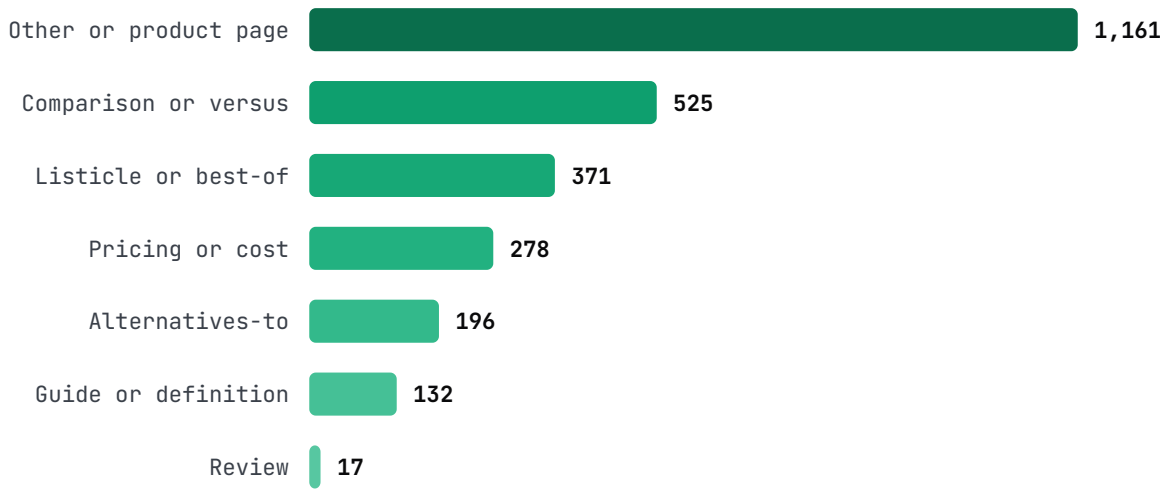
43 of them earned 53 links between them – 14.4% of everything shown, and 3.5 times what G2, Capterra, Software Advice, TrustRadius and TechnologyAdvice managed together. I opened four of them to find out what they actually are.

| SITE | LINKS | WHAT IT ACTUALLY IS |
|---|---|---|
| **erpresearch.com** | 11 | Publishes ERP rankings and comparisons. Says no vendor pays for placement. Does not say who owns it, who writes it, or how it makes money. |
| **ciopages.com** | 10 | Calls itself vendor-independent with no sponsorships and no affiliate revenue, while selling directory listings at $199 a year. No named writers. |
| **toolradar.com** | 2 | Has a named author, a monthly review cycle and a published method, and says it does not sell placement. The most transparent of the four. |
| **stackscored.com** | 1 | Template-built comparison pages with an affiliate disclosure: it earns commission on purchases made through its links. |

These sites are not winning on authority. They are winning on *shape*. They publish exactly the kind of page the search step is hunting for: a dated, titled table comparing named products with prices. And they publish it fast.

**What kind of pages ChatGPT actually read**

Classified from page titles · 2,680 results

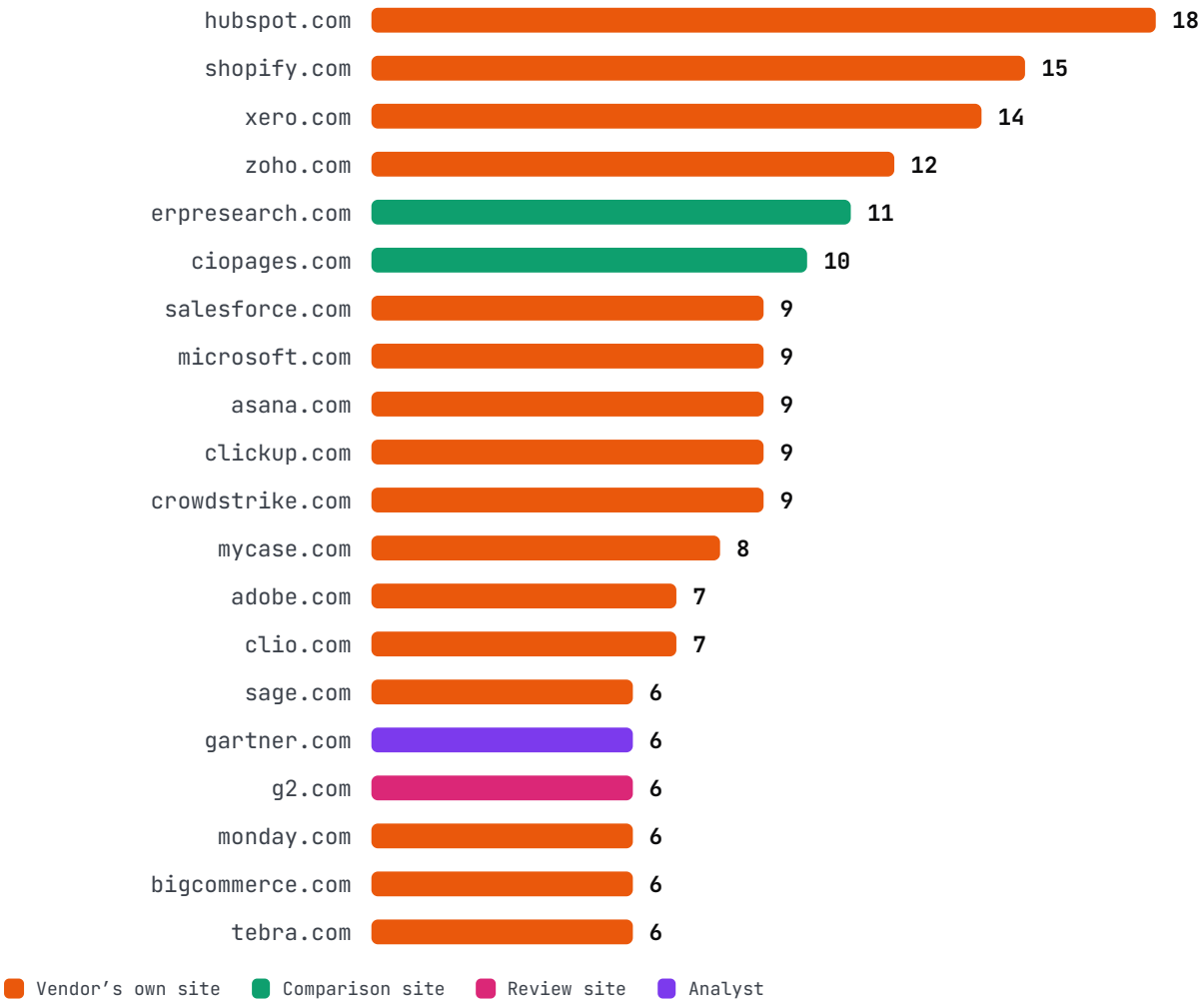



The top row is a catch-all, mostly ordinary product and pricing pages. Of the pages that describe themselves as something specific, comparison and “versus” pages are the biggest group by a clear margin.

09 · 14.4% OF LINKS COME FROM SITES NOBODY TRACKS · CONTINUED

### The 20 most-linked websites

Colour shows the kind of site · 118 websites were linked in total




Notice how flat that is. The most-linked website in the entire audit holds 4.9% of all links; the top ten hold 31.6%; and 46.6% of linked sites appeared exactly once.

**118**
DIFFERENT WEBSITES LINKED

**31.6%**
HELD BY THE TOP TEN

**46.6%**
LINKED EXACTLY ONCE

WHAT TO DO WITH THIS

Find the small comparison sites ranking for "best [your category] software" and "[competitor] alternatives" in your market, and make sure your product is listed accurately. In this audit that long tail out-produced every review site combined.

10

# 69.3% of recommendations cite no source

This is the finding I keep coming back to. Seven out of ten brand recommendations in this audit had nothing behind them at all.

**358**
BRAND PICKS RECORDED

**110**
HAD A LINK TO THAT BRAND

**69.3%**
HAD NO SUPPORTING LINK

**28.3%**
OF TOP PICKS UNSOURCED

I made this test as easy to pass as I could. A recommendation counts as "sourced" if *any* link anywhere in that answer points at the recommended brand's own website. It does not have to support the specific claim. It just has to be there. Even so, 248 of 358 failed.

It improves slightly at the top. The single brand ChatGPT names first – the "I'd go with" brand – has no source 28.3% of the time. So the headline pick is more likely than average to be evidenced. Roughly one in four still is not.

### Two different problems, two different budgets

| | GETTING CITED | GETTING RECOMMENDED |
|---|---|---|
| **What it is** | A search and content problem: can the fan-out find a page of yours that answers the question in a checkable way? | A reputation problem: does the model already associate your brand with this category? |
| **How fast it moves** | Weeks. Publish a page, it gets indexed, it starts getting picked up. | Years. It comes from how much is written about you across the whole web over time. |
| **How measurable** | Easy. The link either appears or it does not. | Hard, and hard to shortcut. This is where 69.3% of the outcome is decided. |
| **Who owns it** | SEO and content | Brand, PR and product marketing |

Almost all the AI visibility advice being sold right now addresses the first column. The data says the second column is where the decision is actually made. The mechanism behind that is worth reading on its own: brands named in ChatGPT's own search query are mentioned 68.9% of the time, against 2.1% for brands it only meets in the results.

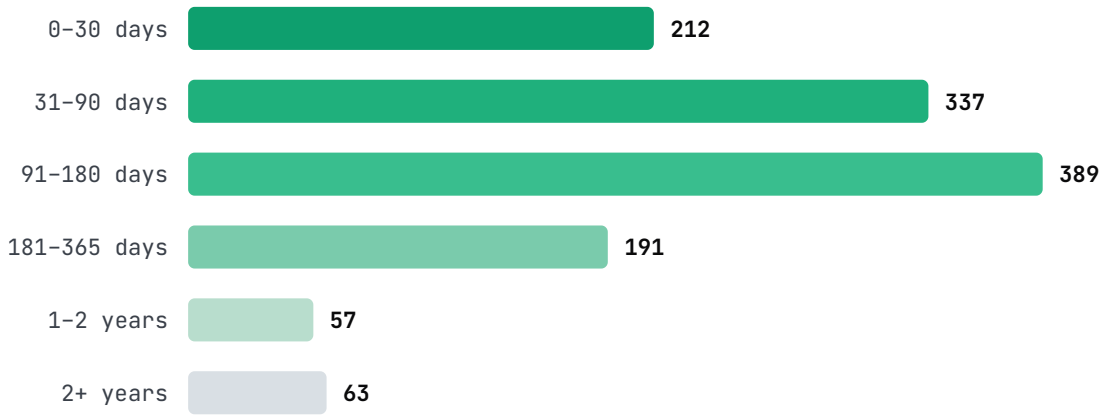
11

# 86.3% of dated pages were published in 2026

Where a page carried a publication date, that date was almost always recent. Dated content older than two years barely entered the picture.

**How old the pages were that ChatGPT read**

1,249 of the 2,680 results carried a readable publication date (46.6%)




**90.4%**
OF DATED PAGES PUBLISHED IN THE LAST YEAR

**86.3%**
CARRIED A 2026 DATE

**5.0%**
WERE MORE THAN TWO YEARS OLD

Two caveats keep this useful. First, 53.4% of the pages read carried no date at all – mostly vendor product and pricing pages, which are clearly not punished for it. Second, a "2026" date is a claim the publisher makes, and small comparison sites have an obvious incentive to restamp monthly.

So the accurate version is not "fresh content wins". It is: **the search step strongly prefers pages that say they are current.** That preference is hardening as answers get faster – at 750 tokens a second, feed freshness starts behaving like a ranking signal.

**WHAT TO DO WITH THIS**

An undated evergreen page is fine when you are the vendor and the page is the definitive source for a fact about your own product. Any page whose job is to be compared against rival pages needs a visible date and a genuine refresh cycle, or it loses retrieval to a site that restamps monthly.

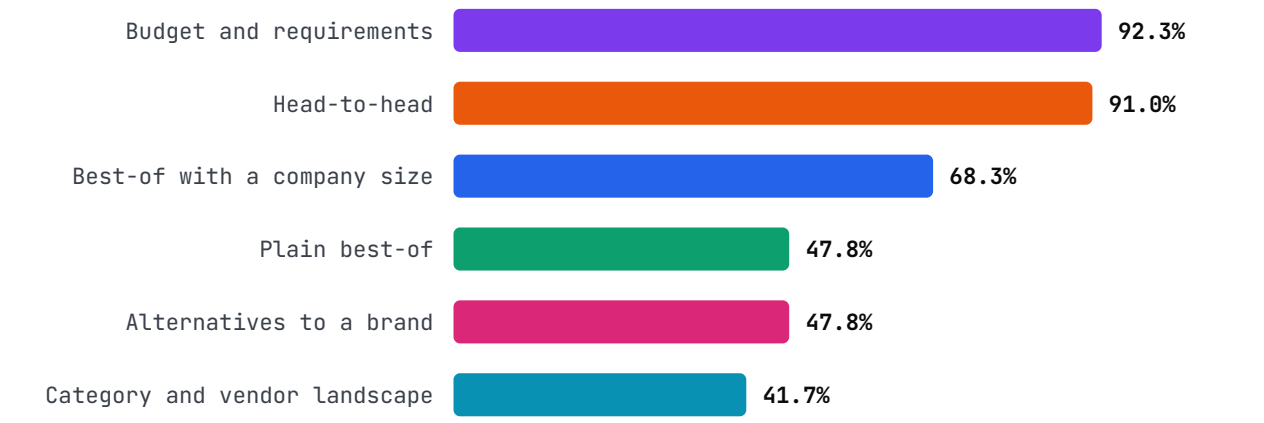
12

# Only 34% of a shortlist survives rewording

Same buyer, same category, six different phrasings – and the answer changes completely. This was the largest single source of variation in the audit.

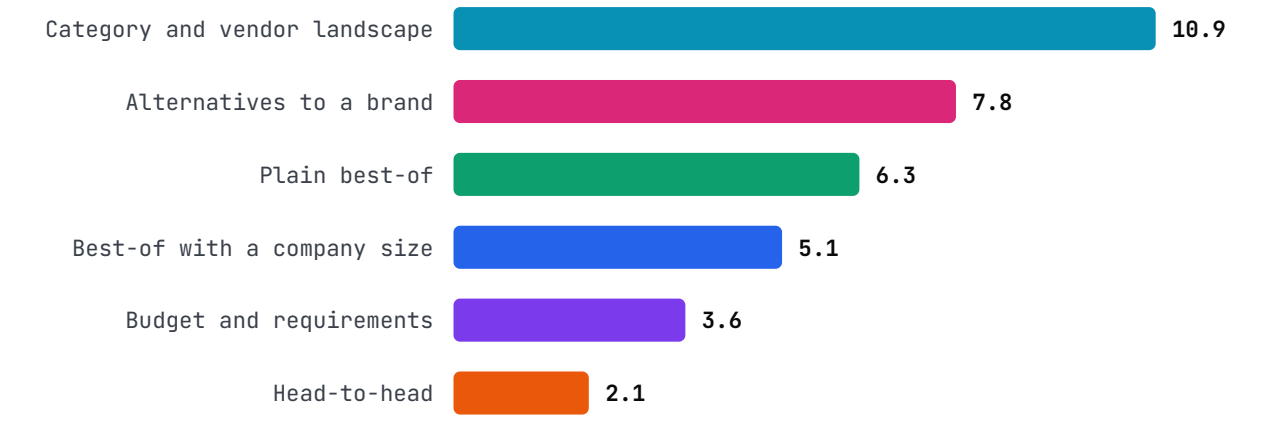
**How much each kind of question leans on vendor websites**

Higher means the answer is built mostly from the vendors' own pages




**How many brands each kind of question puts in front of a buyer**

Average number of different products named




The two charts move in opposite directions, and the reason is clean. When a buyer names two vendors, the shortlist is already decided. ChatGPT is no longer shopping; it is fact-checking. So it names just 2.1 brands and takes 91.0% of its sources from those two companies' own websites. A budget question does the same for a different reason – 92.3% vendor sources, because only the vendor publishes the price.

At the other end, a broad category question names 10.9 brands and takes only 41.7% from vendors. That is the one question shape where analysts, review sites and comparison sites actually get a hearing.

12 · ONLY 34% OF A SHORTLIST SURVIVES REWORDING · CONTINUED

| | SHAPE OF QUESTION | READ | LINKS | BRANDS | WORDS | FROM VENDORS |
|---|---|---|---|---|---|---|
| A | Plain best-of | 43.2 | 4.6 | 6.3 | 404 | 47.8% |
| B | Best-of with a company size | 41.8 | 6.0 | 5.1 | 827 | 68.3% |
| C | Head-to-head | 41.3 | 6.7 | 2.1 | 1,034 | 91.0% |
| D | Alternatives to a brand | 37.7 | 6.9 | 7.8 | 741 | 47.8% |
| E | Budget and requirements | 52.3 | 6.5 | 3.6 | 460 | 92.3% |
| F | Category and vendor landscape | 51.7 | 6.0 | 10.9 | 929 | 41.7% |

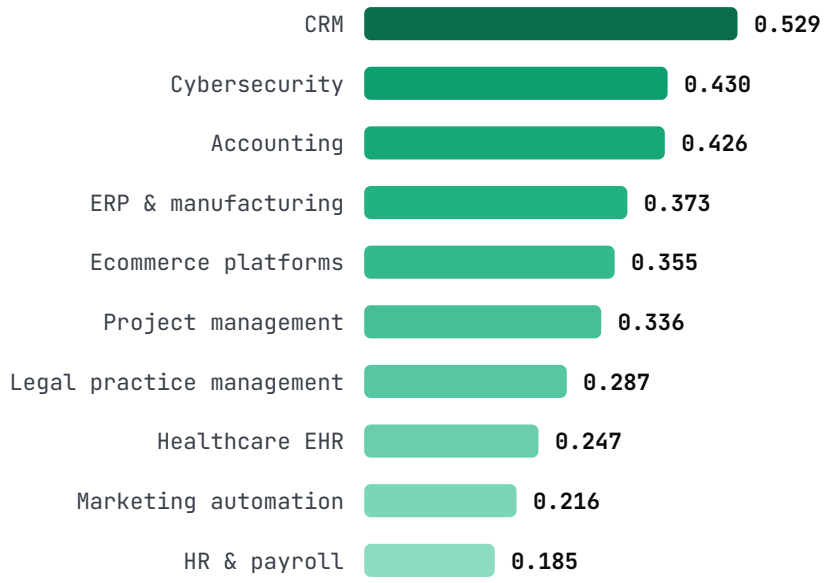
## How unstable is the shortlist?

For each category I compared the six answers against each other and measured how much the brand lists overlapped. A score of 1.00 would mean the phrasing makes no difference at all. A score of 0.00 would mean no brand ever appears twice.

The average across all ten categories was **0.338**. Only about a third of the brands survive between two ways of asking the same question.

**How stable each category's shortlist is**

1.00 would mean phrasing makes no difference. Higher is more settled.



12 · ONLY 34% OF A SHORTLIST SURVIVES REWORDING · CONTINUED

| CATEGORY | DIFFERENT BRANDS RANKED FIRST, OUT OF 6 | MOST FREQUENT LEADER | HOW OFTEN IT LED |
|---|---|---|---|
| HR & payroll | 6 | Rippling | 17% |
| ERP & manufacturing | 5 | SAP S/4HANA | 33% |
| Marketing automation | 5 | HubSpot | 33% |
| Cybersecurity | 4 | CrowdStrike | 50% |
| Ecommerce platforms | 4 | Shopify | 50% |
| Accounting | 3 | QuickBooks | 50% |
| Healthcare EHR | 3 | Athenahealth | 50% |
| Legal practice management | 3 | Clio | 67% |
| Project management | 3 | Asana | 67% |
| CRM | 2 | HubSpot | 67% |

CRM is the most settled at 0.529. HR and payroll is close to random at 0.185. In HR the brand ranked first was different in all six questions: Rippling, Paylocity, Workday, ADP, BrightHR and Oracle HCM each led exactly once. There is no such thing as “what ChatGPT thinks the best HR software is”. There is only what it says to one particular question.

### The same category, six ways of asking

| SHAPE OF THE QUESTION | RANKED FIRST IN CRM | RANKED FIRST IN HR & PAYROLL |
|---|---|---|
| Plain best-of | HubSpot | Rippling |
| Best-of with a company size | HubSpot | Paylocity |
| Head-to-head | HubSpot | Workday |
| Alternatives to a brand | Salesforce | ADP |
| Budget and requirements | HubSpot | BrightHR |
| Category and vendor landscape | Salesforce | Oracle HCM |

Two brands cover all six CRM answers. Six cover all six HR answers. If you sell HR software and your dashboard tracks one prompt, it is reporting a coin toss as a trend.

IF YOU ARE TRACKING AI VISIBILITY TODAY

Most tools track a fixed list of prompts and report a single score. With a shortlist overlap of 0.338 and an average of 3.8 different category leaders across six phrasings, one tracked prompt is close to noise. Track a spread of question shapes, and report the variation, not just the average.

13

# 48.7% to 90.9%: vendor share by category

The ten markets behaved so differently that a single strategy would be wrong in most of them. The bar in each card splits that category's links: colour is the vendors' own sites, grey is everyone else.

### CRM

210 results read · 40 links · 9 brands named

The most settled market I tested. Five brands turn up in nearly every answer and HubSpot led four of the six questions. If you are not on that list already, you are arguing with a consensus the model has formed.

60.0% vendor's own site 40.0% third party

**2/6** brands ranked first

**56.7%** picks with no source

**16** sites linked

**0.529** shortlist overlap

**Always there:** HubSpot, Salesforce, Zoho CRM, Pipedrive +1 more

**Most-linked site:** hubspot.com

### Accounting

258 results read · 33 links · 10 brands named

Effectively a two-horse race read off two pricing pages. Xero and QuickBooks appear in every answer and nine in ten links point at a vendor's own site.

90.9% vendor's own site 9.1% third party

**3/6** brands ranked first

**46.7%** picks with no source

**9** sites linked

**0.426** shortlist overlap

**Always there:** QuickBooks, Xero

**Most-linked site:** xero.com

### Project management

331 results read · 32 links · 14 brands named

Lots of reading, very little showing. ChatGPT pulled 55 results per answer and cited just eight sites, almost all vendor help centres and pricing pages.

87.5% vendor's own site 12.5% third party

**3/6** brands ranked first

**62.5%** picks with no source

**8** sites linked

**0.336** shortlist overlap

**Always there:** Asana, ClickUp, Monday.com, Jira

**Most-linked site:** asana.com

### Ecommerce platforms

245 results read · 37 links · 14 brands named

Shopify is the centre of gravity. Named in all six answers and linked 15 times. Everyone else competes for the space left over.

83.8% vendor's own site 16.2% third party

**4/6** brands ranked first

**65.6%** picks with no source

**11** sites linked

**0.355** shortlist overlap

**Always there:** Shopify, BigCommerce

**Most-linked site:** shopify.com

13 · 48.7% TO 90.9%: VENDOR SHARE BY CATEGORY · CONTINUED

### Cybersecurity

275 results read · 36 links · 12 brands named

The only category where independent test labs get cited. AV-TEST and AV-Comparatives both made it in, so the buying culture leaks into the model.

66.7% vendor's own site 33.3% third party

**4/6**
brands ranked first

**68.8%**
picks with no source

**10**
sites linked

**0.430**
shortlist overlap

**Always there:** CrowdStrike, SentinelOne, Microsoft Defender, Sophos
**Most-linked site:** crowdstrike.com

### HR & payroll

211 results read · 30 links · 19 brands named

The least stable category here. Six questions produced six different winners and four in five picks had no source. There is no incumbent in the model's head.

63.3% vendor's own site 36.7% third party

**6/6**
brands ranked first

**80.0%**
picks with no source

**13**
sites linked

**0.185**
shortlist overlap

**Always there:** Rippling, ADP, Workday
**Most-linked site:** gusto.com

### Marketing automation

235 results read · 43 links · 22 brands named

The widest field and weakest agreement. Twenty-two brands across six answers, sixteen appearing exactly once. The category is still being defined.

60.5% vendor's own site 39.5% third party

**5/6**
brands ranked first

**76.5%**
picks with no source

**20**
sites linked

**0.216**
shortlist overlap

**Always there:** HubSpot, ActiveCampaign
**Most-linked site:** hubspot.com

### How to read these cards

The same six questions were asked in every category

- **The bar** splits that category's links: colour is the vendors' own websites, grey is everyone else.
- **Brands ranked first** counts how many different brands took the top spot across the six questions. 6/6 means a different winner every time.
- **Picks with no source** is the share of brand recommendations with no link to that brand's own site anywhere in the answer.
- **Shortlist overlap** scores how much the six answers agreed, from 0.00 to 1.00.
- **Always there** lists the brands appearing in at least four of the six answers.

13 · 48.7% TO 90.9%: VENDOR SHARE BY CATEGORY · CONTINUED

### ERP & manufacturing

259 results read · 40 links · 17 brands named

The clearest case of one small publisher steering a market. erpresearch.com was cited more often than SAP or Microsoft in this category.

52.5% vendor's own site 47.5% third party

**5/6**
brands ranked first

**72.7%**
picks with no source

**13**
sites linked

**0.373**
shortlist overlap

**Always there:** SAP S/4HANA, Epicor, NetSuite, Acumatica +3 more
**Most-linked site:** erpresearch.com

### Healthcare EHR

363 results read · 39 links · 19 brands named

Heaviest reading, lowest vendor share of anything tested. Because pricing is hidden, the model falls back on Becker's, KLAS and specialist sites.

48.7% vendor's own site 51.3% third party

**3/6**
brands ranked first

**76.9%**
picks with no source

**19**
sites linked

**0.247**
shortlist overlap

**Always there:** eClinicalWorks, Epic, Tebra
**Most-linked site:** tebra.com

### Legal practice management

293 results read · 37 links · 17 brands named

Clio is close to a default answer, but the evidence behind it is mostly small legal-tech blogs. The softest citation base of any category with a clear leader.

54.1% vendor's own site 45.9% third party

**3/6**
brands ranked first

**76.9%**
picks with no source

**16**
sites linked

**0.287**
shortlist overlap

**Always there:** Clio, MyCase, PracticePanther, Smokeball
**Most-linked site:** mycase.com

### The rule these ten cards share

**Where a category publishes its prices, the vendors own the evidence. Where pricing is gated behind a sales call, third-party publishers own it instead.**

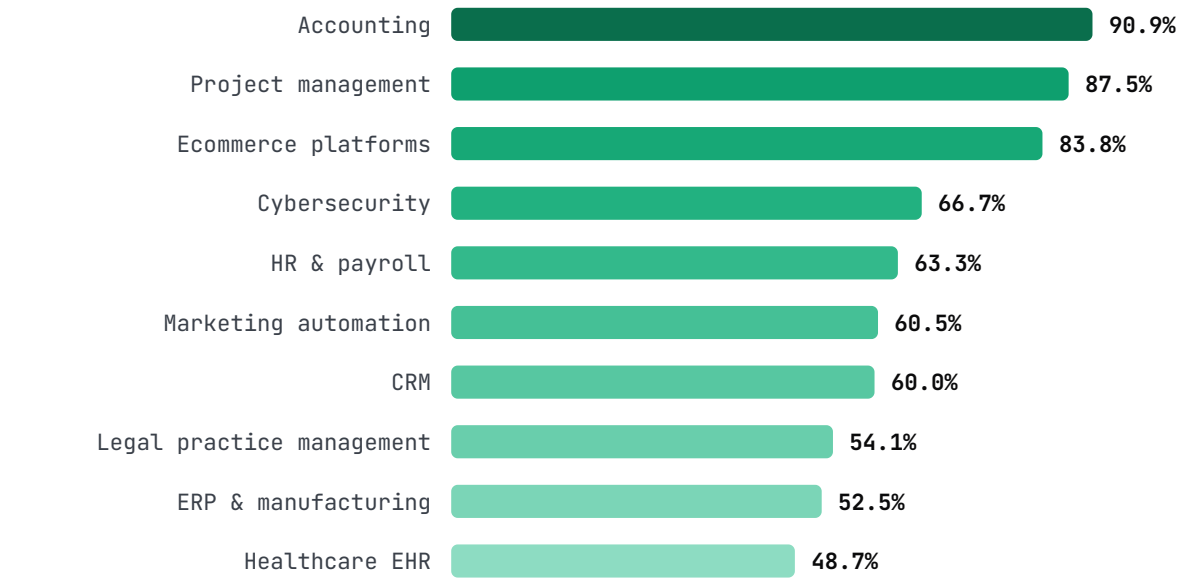
That is the single strongest predictor in this section – stronger than category size, stronger than how many vendors compete, and stronger than how mature the market is.

Accounting, the most open about price, is the most vendor-dominated market here at 90.9%. Healthcare EHR, the most gated, is the least at 48.7%. Everything else falls between them in roughly the order you would guess from how easy it is to find a price.

13 · 48.7% TO 90.9%: VENDOR SHARE BY CATEGORY · CONTINUED

**How much each category leans on vendor websites**

Six questions per category




The three categories where pricing is most consistently published – accounting, project management and ecommerce – sit at 90.9%, 87.5% and 83.8%. The two where pricing is gated behind a sales call – ERP and healthcare EHR – sit at 52.5% and 48.7%. The difference goes to third parties.

If you sell enterprise software with "contact sales" pricing, you have handed roughly half your AI evidence base to publishers you do not control. That is now a marketing decision as well as a sales one.

One more pattern: CRM produced 9 brands across six answers, five appearing almost every time. Marketing automation produced 22, sixteen appearing exactly once. Across the audit, 47.1% of brands showed up in only one of six answers for their category.

**IF YOUR CATEGORY GATES ITS PRICING**

You will not win the evidence back by publishing a price you cannot honour. Publish the *structure* instead: what you charge for, what drives the number up or down, what a typical deployment includes and what a realistic range looks like. That gives ChatGPT a checkable fact with your name on it, which is the only kind of page it reliably links to. The alternative is that erpresearch.com or a healthcare trade title supplies the number for you.

14

# Four real answers, taken apart

Averages hide how this feels in practice. These four are real captures: what the answer said, which brands it named, what it read and every link it showed. Grey marks are pages read; green marks are pages linked.

P17 · ACCOUNTING · BUDGET AND REQUIREMENTS

### "I need accounting software under $40 a month that handles multi-currency invoicing and US sales tax. What should I buy?"

WHAT CHATGPT SAID

> *"If $40/month is a hard ceiling, I'd buy Zoho Books Professional – with one important caveat: its $40 price is the annual-billing rate; month-to-month is $50."*

BRANDS NAMED, IN ORDER – 1 OF 3 SOURCED

| | | |
|---|---|---|
| 1 | Zoho Books | sourced |
| 2 | Xero | no source |
| 3 | QuickBooks | no source |

WHAT IT READ VS WHAT IT SHOWED

27 results from 5 sites → 7 links from 1 site

EVERY SOURCE IT LINKED TO

| | |
|---|---|
| **zoho.com**<br>Vendor's own site | 7× |

**This is what a monoculture looks like.** A buyer with a real budget got a confident, specific, well-sourced answer – in which every single source was the website of the company being recommended. Nothing independent was shown at all.

P27 · CYBERSECURITY · HEAD-TO-HEAD

### "Compare CrowdStrike vs SentinelOne for endpoint detection and response."

WHAT CHATGPT SAID

> *"CrowdStrike: stronger overall ecosystem, threat intelligence, threat hunting and mature SOC workflows. SentinelOne: autonomous response, operational simplicity and rollback."*

BRANDS NAMED, IN ORDER – 2 OF 2 SOURCED

| | | |
|---|---|---|
| 1 | CrowdStrike | sourced |
| 2 | SentinelOne | sourced |

WHAT IT READ VS WHAT IT SHOWED

43 results from 3 sites → 8 links from 2 sites

EVERY SOURCE IT LINKED TO

| | |
|---|---|
| **crowdstrike.com**<br>Vendor's own site | 4× |
| **sentinelone.com**<br>Vendor's own site | 4× |

**Naming two vendors hands them the evidence.** ChatGPT read 43 results but from only three websites, and all eight links went to the two competitors' own pages. Once the buyer supplies the shortlist, ChatGPT stops shopping and starts fact-checking.

14 · FOUR REAL ANSWERS, TAKEN APART · CONTINUED

P46 · ERP & MANUFACTURING · ALTERNATIVES TO A BRAND

### "What are the best alternatives to NetSuite?"

WHAT CHATGPT SAID

> *"If you're evaluating alternatives to Oracle NetSuite, there isn't one universal winner… these are the alternatives I'd shortlist based on the current 2026 ERP landscape."*

BRANDS NAMED, IN ORDER – 0 OF 12 SOURCED

| | | |
|---|---|---|
| 1 | Oracle | no source |
| 2 | NetSuite | no source |
| 3 | Microsoft Dynamics 365 Business Central | no source |
| 4 | Acumatica | no source |
| 5 | SAP Business One | no source |
| 6 | Odoo | no source |

and 6 more, 6 of them with no source

WHAT IT READ VS WHAT IT SHOWED

34 results from 9 sites → 7 links from 2 sites

EVERY SOURCE IT LINKED TO

**erpresearch.com** 5×
Comparison site

**cudio.com** 2×
Comparison site

**One small site defined a whole shortlist.** erpresearch.com supplied five of the seven links. Twelve ERP brands were named and not one had a source behind it. The entire recommendation rests on one third-party page plus the model's memory.

P07 · HR & PAYROLL · PLAIN BEST-OF

### "What is the best HR and payroll software?"

WHAT CHATGPT SAID

> *"There isn't one universal 'best' HR/payroll platform. Rippling – best overall if you want HR + payroll + benefits. Gusto – best for simplicity and value. ADP – best for larger organisations."*

BRANDS NAMED, IN ORDER – 2 OF 5 SOURCED

| | | |
|---|---|---|
| 1 | Rippling | sourced |
| 2 | Gusto | sourced |
| 3 | ADP | no source |
| 4 | Paychex | no source |
| 5 | Deel | no source |

WHAT IT READ VS WHAT IT SHOWED

43 results from 17 sites → 2 links from 2 sites

EVERY SOURCE IT LINKED TO

**rippling.com** 1×
Vendor's own site

**gusto.com** 1×
Vendor's own site

**The links follow the facts, not the advice.** Two links were shown, both pricing pages, both for the two brands ChatGPT happened to quote a price for. ADP, Paychex and Deel were each given a market position with no source at all.

The same machine shows up in all four. Links cluster around checkable facts – prices, plans, feature availability. The advice itself usually arrives with nothing attached.

15

# Six things I would do about this

Six moves that follow directly from the data, roughly in order of how much they are worth.

### 1. Treat your own website as your main AI channel

65.9% of links went to vendor websites, and in 26 of 60 answers they were the only source shown. The pages that earn those links are the dull ones: pricing, plan comparison, feature availability, integration lists, supported platforms. Anything stating a checkable fact in a clear, scannable form. This is the cheapest work in the report and almost nobody is prioritising it.

### 2. Publish your prices

The three categories that publish pricing most consistently sat at 83.8–90.9% vendor sources. The two that gate pricing behind a sales call sat at 48.7–52.5%. Hiding your price does not keep you out of the answer. It just means somebody else's guess at your price is what gets shown. If you cannot publish a number, publish the structure: what you charge for, what drives it up, what a typical deployment includes.

### 3. Stop measuring AI visibility on review sites alone

All the review sites combined produced 15 of 367 links. 43 small comparison sites produced 53, and 46.6% of linked sites appeared only once. You need monitoring built for a long tail, and it needs to watch what gets *read*, not only what gets *linked*.

### 4. Publish the page shape the machine is hunting for

Setting ordinary product pages aside, comparison and "versus" pages were the most-read shape at 525 results, ahead of best-of lists at 371. Dated. Titled. Tabular. Products named. The small sites winning links are not writing better than you. They are formatting better than you.

### 5. Separate two very different goals

Getting cited and getting recommended are not the same job. With 69.3% of recommendations carrying no source, most of the outcome is decided before any search runs. Budget for both, and do not let the fast, measurable one crowd out the slow one.

### 6. Test the range of questions, not one question

With a shortlist overlap of 0.338 and an average of 3.8 different category leaders across six phrasings, tracking one prompt tells you very little. Track a spread and report the spread.

15 · SIX THINGS I WOULD DO ABOUT THIS · CONTINUED

## Where to start, depending on your job

### CEO or founder

- Put one question in the next board pack: when a buyer asks ChatGPT about our category, are we on the list? Get it tested across six phrasings, not one.
- Revisit the pricing decision. Gated pricing cost roughly half the evidence base here. That is now a marketing cost, not just a sales preference.
- Know which game you are in. A settled category means fighting an established default for years. An unsettled one means the door is open now.

### CMO

- Move budget towards your own factual pages – pricing, comparison, integrations, specs. That is where 65.9% of the visible evidence came from.
- Re-examine review site spend. It produced 4.1% of what buyers saw.
- Build a dated comparison page for every serious rival. It is the most-read page shape in the audit.
- Change the reporting: ask for a distribution across question shapes with variance, not one score, and remember that citation share is not traffic.

### Growth or SEO lead

- Audit which of your pages state checkable facts in a form a machine can lift. Most product pages do not.
- Find the long tail. Identify the small comparison sites ranking in your category and get listed accurately.
- Put visible dates on anything comparative and refresh on a real schedule.
- Measure retrieval, not just citation – citation-only tracking misses most of the picture. Start with the GA4 Source Group dimension and hostname filters.

THE UNCOMFORTABLE VERSION

In this audit the most reliable route into a ChatGPT software recommendation was to be a vendor with a clear public pricing page, or a six-month-old comparison site with a dated table and a plausible domain name. Deep editorial authority was not the mechanism. Structure, recency and checkability were.

**Weeks**
TO CHANGE WHAT GETS CITED

**Years**
TO CHANGE WHAT GETS RECOMMENDED

**Both**
NEED A BUDGET LINE

## If you only get one quarter

Spend the first month on your own pricing and comparison pages, because that is the fastest lever in this data and the one you fully control. Spend the second getting listed accurately on the small comparison sites already ranking in your category, since 43 of them produced 14.4% of everything buyers saw. Spend the third building a measurement habit: the same six question shapes, run monthly, recorded with what the model read as well as what it linked. That last one is what turns this from a report you read into a number you manage.

16

# What this audit does not tell you

A study about citation quality should be honest about its own. Here is everything I would want to know before quoting these numbers.

- **Sixty questions is a probe, not a census.** The headline figures rest on 60 observations. The per-category figures rest on six each. Treat category numbers as directional, not precise.
- **One model, one account tier.** All 60 answers were handled by `gpt-5-6-mini` on a free account. A paid account with a bigger reasoning model may search and cite differently. This is the biggest open question here.
- **Geography was varied with words, not location.** I changed the wording ("UK company", "US medical group") but every question came from one UK connection. Real location effects are not measured.
- **Six of the sixty leaked into chat history.** The temporary-chat setting did not hold for six runs. ChatGPT's memory feature could in principle have nudged later answers in the same account.
- **Link counts are a floor.** Where ChatGPT groups several sources behind a "+1" button, I counted the main one only. The true source count is somewhat higher than 367.
- **Brand detection used a fixed list.** Answers were matched against 350 product names spanning the ten categories, scoped so each answer was only checked against its own category. A product outside that list would not have been counted.
- **A publication date is a claim.** The recency finding measures what pages say about themselves, not when they were really written.
- **One point in time.** Everything was captured on 26 August 2026. These systems change constantly. Treat this as a snapshot and re-run it if you need current numbers.
- **Organic answers only.** No paid placement appeared in any of the 60 captures, but ChatGPT Ads are now live across 31 European markets, and that will change what sits beside an answer.
- **The search box is not the only surface.** Personalisation from ChatGPT's Computer History feature and agentic AI browsers acting on the user's behalf both route around the question shape entirely. Neither is measurable with this method.

HOW THE NUMBERS WERE CHECKED

Every headline figure in this report was recomputed independently from the raw captures by a second script using a different code path, and cross-checked against the first. All 31 checks reconcile. No figure in this document was typed by hand – charts and prose read from the same computed dataset.

17

# Four follow-up tests worth running

Four experiments that would each be worth more than adding more questions to this one.

### 1. Paid tier versus free tier

Run the identical 60 questions on a Plus or Pro account with a larger reasoning model. If the bigger model reads more widely and cites more independently, then everything in this report describes the tier most buyers use but not the tier your enterprise buyer uses. That is a material difference and nobody has published it.

### 2. Real geography

Run the same questions from US, UK and EU connections. My expectation is that the comparison-site long tail is heavily local, which would change the shopping list for anyone selling across borders.

### 3. Does it actually move?

Publish a properly structured, dated comparison page for one category, then re-run its six questions monthly. This is the only way to find out how long the loop takes from publishing to being cited – and that number does not currently exist anywhere.

### 4. ChatGPT versus everyone else

The same 60 questions through Gemini, Claude, Perplexity and Google AI Mode. My expectation is that the vendor-site dominance is a ChatGPT trait rather than a universal one, but that is a hypothesis, not a finding.

IF YOU RUN ONE OF THESE

The capture method is reusable and the analysis is scripted. If you run any of these tests I would genuinely like to see the results – and I will publish comparisons with attribution.

**60**

QUESTIONS IN THIS AUDIT – THE METHOD SCALES TO ANY NUMBER

**~30 min**

TO RUN A FULL 60-QUESTION SWEEP ONCE SET UP

18

# Glossary of terms used in this audit

Plain definitions for every term in this report.

| | |
|---|---|
| **Retrieval** | Everything ChatGPT searched for and received – the pages it read. Here, 2,680 results from 564 websites. Almost none of it is visible to the reader. |
| **Citation** | The clickable source buttons shown in an answer. What the reader actually sees. Here, 367 links from 118 websites. |
| **Query fan-out** | When ChatGPT turns one question from you into several searches run at once. It is why one prompt pulls back 40 or 50 pages. |
| **web.run** | The name of ChatGPT's web search tool, visible in the live data stream. When you see "searching the web", this is it. |
| **Conversion rate** | Of the pages read from a kind of website, the share that became a visible link. It measures how much ChatGPT trusts a source once it has found it. |
| **Survival rate** | The share of pages, or websites, making it from read through to linked. 11.1% of pages and 20.9% of websites survived here. |
| **Unsourced pick** | A brand recommended in an answer where no link points at that brand's own website. 69.3% of picks here. |
| **Model memory** | What the model knows from training, before it searches anything. Where most brand rankings come from. |
| **Question shape** | A type of question rather than a specific one. This audit used six, from plain best-of to head-to-head. |
| **Shortlist overlap** | A 0 to 1 score for how much two brand lists share. 1.00 means identical. The average here was 0.338. |
| **Resident brands ("always there")** | My term for brands appearing in at least four of the six answers in a category – effectively the category default. |
| **Concentration (HHI)** | A standard 0–10,000 measure of how concentrated a market is. Lower means more fragmented. Links here scored 177. |
| **Vendor's own site** | A software company's own website, including documentation, help centre, support and investor pages. |
| **Review site** | G2, Capterra, Software Advice, TrustRadius, TechnologyAdvice and equivalents. |
| **Comparison site** | A third-party site whose main output is software comparison, ranking or "alternatives to" pages, and which is not an established review site. |
| **GEO / AEO** | Generative Engine Optimisation and Answer Engine Optimisation – the emerging labels for making a brand more likely to appear in AI answers. |

19

# The full numbers behind every chart

## Every kind of website, read and linked

| KIND OF WEBSITE | SITES | READ | SHARE READ | LINKS | SHARE OF LINKS | LINK RATE |
|---|---|---|---|---|---|---|
| Vendor's own site | 79 | 1,476 | 55.1% | 242 | 65.9% | 16.4% |
| Comparison site | 43 | 249 | 9.3% | 53 | 14.4% | 21.3% |
| Agency or partner | 58 | 206 | 7.7% | 21 | 5.7% | 10.2% |
| Analyst | 10 | 87 | 3.2% | 16 | 4.4% | 18.4% |
| News and trade press | 20 | 81 | 3.0% | 15 | 4.1% | 18.5% |
| Review site | 5 | 52 | 1.9% | 15 | 4.1% | 28.8% |
| Testing lab | 2 | 22 | 0.8% | 3 | 0.8% | 13.6% |
| Government | 4 | 17 | 0.6% | 2 | 0.5% | 11.8% |
| Reddit, LinkedIn, YouTube | 3 | 31 | 1.2% | 0 | 0.0% | 0.0% |
| Wikipedia | 1 | 6 | 0.2% | 0 | 0.0% | 0.0% |
| Long tail, never linked | 339 | 453 | 16.9% | 0 | 0.0% | 0.0% |

## Every category, side by side

| CATEGORY | READ | LINKS | BRANDS | SITES | VENDOR | NO SOURCE | LEADERS | OVERLAP |
|---|---|---|---|---|---|---|---|---|
| Accounting | 43.0 | 5.5 | 10 | 9 | 90.9% | 47% | 3/6 | 0.43 |
| Project management | 55.2 | 5.3 | 14 | 8 | 87.5% | 62% | 3/6 | 0.34 |
| Ecommerce platforms | 40.8 | 6.2 | 14 | 11 | 83.8% | 66% | 4/6 | 0.35 |
| Cybersecurity | 45.8 | 6.0 | 12 | 10 | 66.7% | 69% | 4/6 | 0.43 |
| HR & payroll | 35.2 | 5.0 | 19 | 13 | 63.3% | 80% | 6/6 | 0.18 |
| Marketing automation | 39.2 | 7.2 | 22 | 20 | 60.5% | 76% | 5/6 | 0.22 |
| CRM | 35.0 | 6.7 | 9 | 16 | 60.0% | 57% | 2/6 | 0.53 |
| Legal practice management | 48.8 | 6.2 | 17 | 16 | 54.1% | 77% | 3/6 | 0.29 |
| ERP & manufacturing | 43.2 | 6.7 | 17 | 13 | 52.5% | 73% | 5/6 | 0.37 |
| Healthcare EHR | 60.5 | 6.5 | 19 | 19 | 48.7% | 77% | 3/6 | 0.25 |

19 · THE FULL NUMBERS BEHIND EVERY CHART · CONTINUED

## The 20 most-linked websites

| WEBSITE | KIND OF SITE | LINKS | ANSWERS |
|---|---|---|---|
| **hubspot.com** | Vendor's own site | 18 | 7 |
| **shopify.com** | Vendor's own site | 15 | 5 |
| **xero.com** | Vendor's own site | 14 | 5 |
| **zoho.com** | Vendor's own site | 12 | 5 |
| **erpresearch.com** | Comparison site | 11 | 4 |
| **ciopages.com** | Comparison site | 10 | 4 |
| **salesforce.com** | Vendor's own site | 9 | 6 |
| **microsoft.com** | Vendor's own site | 9 | 4 |
| **asana.com** | Vendor's own site | 9 | 4 |
| **clickup.com** | Vendor's own site | 9 | 3 |
| **crowdstrike.com** | Vendor's own site | 9 | 4 |
| **mycase.com** | Vendor's own site | 8 | 2 |
| **adobe.com** | Vendor's own site | 7 | 2 |
| **clio.com** | Vendor's own site | 7 | 4 |
| **sage.com** | Vendor's own site | 6 | 4 |
| **gartner.com** | Analyst | 6 | 4 |
| **g2.com** | Review site | 6 | 3 |
| **monday.com** | Vendor's own site | 6 | 3 |
| **bigcommerce.com** | Vendor's own site | 6 | 3 |
| **tebra.com** | Vendor's own site | 6 | 3 |

19 · THE FULL NUMBERS BEHIND EVERY CHART · CONTINUED

## The 16 most-recommended brands

| BRAND | ANSWERS NAMING IT | TIMES RANKED FIRST | AVERAGE POSITION | TOTAL MENTIONS |
|---|---|---|---|---|
| HubSpot | 11 | 6 | 1.91 | 109 |
| Microsoft Dynamics 365 | 9 | 0 | 4.44 | 27 |
| QuickBooks | 6 | 3 | 1.83 | 32 |
| Xero | 6 | 2 | 1.67 | 58 |
| CrowdStrike | 6 | 3 | 1.83 | 51 |
| SentinelOne | 6 | 0 | 3.17 | 38 |
| Shopify | 6 | 3 | 1.50 | 80 |
| BigCommerce | 6 | 1 | 2.33 | 52 |
| Salesforce | 5 | 2 | 1.80 | 57 |
| Zoho CRM | 5 | 0 | 3.60 | 29 |
| Pipedrive | 5 | 0 | 4.80 | 26 |
| NetSuite | 5 | 0 | 5.20 | 24 |
| Asana | 5 | 4 | 1.60 | 44 |
| ClickUp | 5 | 1 | 2.80 | 28 |
| Monday.com | 5 | 0 | 2.80 | 29 |
| SAP S/4HANA | 5 | 2 | 4.40 | 15 |

## How concentrated the links were

| MEASURE | VALUE |
|---|---|
| Different websites linked | 118 |
| Concentration score (HHI) | 177 |
| Share held by the biggest site | 4.9% |
| Share held by the top 10 | 31.6% |
| Websites linked exactly once | 55 of 118 (46.6%) |

20

# Ten more things worth reading next

Ten pieces from my insights archive that pick up where this audit stops. Each one is listed with what it covers and which section of this report it speaks to.

01 **ChatGPT Picks Winners Before It Searches**

Brands named in ChatGPT's own self-generated search query are mentioned in the answer 68.9% of the time. Brands it only meets in the fetched results: 2.1%.

THE MECHANISM BEHIND SECTION 10

02 **Citation Share Is Not Traffic**

Three different AI search architectures, three different things worth optimising for. Perplexity cites the most and sends the fewest clicks.

WHAT SECTIONS 07–09 DO NOT TELL YOU ABOUT TRAFFIC

03 **The AI Search Glossary**

54 working definitions across metrics, platforms, crawlers, tactics and structured data, kept current.

THE LONGER VERSION OF SECTION 18

04 **GA4 Source Group Dimension: Measure AI Traffic Now**

How to separate AI referral traffic from everything else in Google Analytics, with the hostname filters that make it clean.

THE PRACTICAL HALF OF RECOMMENDATION 6

05 **14x Speed Makes Latency a Ranking Signal**

When answers are generated at 750 tokens a second, feed freshness and server speed start deciding which sources make the cut.

WHY SECTION 11'S RECENCY BIAS IS GETTING STRONGER

06 **OpenAI Hires First CRO: ChatGPT Becomes a B2B Channel**

The commercial signal that software buying inside ChatGPT is being built deliberately, not happening by accident.

THE CONTEXT FOR SECTION 01

20 · TEN MORE THINGS WORTH READING NEXT · CONTINUED

07 **ChatGPT Computer History: Search Just Ended for Brands**

Once the model can see what you actually use, visibility stops being about the search box and starts being about the workflow.

A LIMIT THIS AUDIT COULD NOT MEASURE – SECTION 16

08 **AI Browsers Are Rewriting the Rules of Customer Acquisition**

Agentic browsers read and act rather than browse. What that does to a funnel built for human attention.

WHERE THE RECOMMENDATIONS IN SECTION 15 ARE HEADING

09 **ChatGPT Ads Expand to 31 EU Markets**

Paid placement inside the same answers this audit measured organically, and why "ads sit inside a goal" rather than beside a keyword.

THE SURFACE THIS AUDIT DELIBERATELY EXCLUDED

10 **The Discovery Digest**

A weekly read of what Google, OpenAI, Anthropic, Perplexity, xAI, Meta and Microsoft actually shipped, every Friday.

HOW TO KEEP THIS REPORT CURRENT

## Running this audit on your own category

The method in this report is reusable and I have deliberately written it up in enough detail to be copied. Section 03 covers the capture – what to record and, more importantly, the difference between what the model reads and what it shows. Section 02 has the six question shapes; use those rather than one prompt, because the shortlist is largely written before the search runs and a single phrasing will not show you that. Section 16 lists what this method cannot see, which is the part worth reading before you quote any of it back at a board.

KEEPING THIS CURRENT

Everything here is a snapshot of one week in August 2026, and the systems it measures change monthly. The Discovery Digest tracks what Google, OpenAI, Anthropic, Perplexity, xAI, Meta and Microsoft actually ship, every Friday – subscribe free if you want the running commentary rather than the annual audit. There is more on the work and how to get hold of me at nathanmzumara.com/about.

21

# How to cite this audit and its data

Everything here is original data I collected. Please use it – I would just ask that you point back so people can check the method.

**Mzumara, N. (2026).** *How ChatGPT Shortlists Software Brands: an audit across 10 categories and 60 buying questions.* Primary research, captured 26 August 2026, ChatGPT gpt-5-6-mini. nathanmzumara.com

If you quote a single number, please carry the sample with it. "65.9% of links were to vendor sites" is only meaningful as "65.9% of 367 links across 60 questions in 10 software categories, ChatGPT gpt-5-6-mini, August 2026".

## What ships with this report

| FILE | WHAT IS IN IT |
|---|---|
| **Data appendix (.xlsx)** | All 60 answers, 367 links with their web addresses, 225 classified websites and 358 brand recommendations. |
| **Summary deck (.pptx)** | Fifteen slides covering the same findings, with speaker notes. |
| **Editable report (.docx)** | The full report in Word, for internal circulation. |

## Corrections

If you find an error in this report I want to know. The full dataset is included precisely so the numbers can be checked, and corrections get the same prominence as the original claim.

ABOUT THE AUTHOR

**Nathan Mzumara** is an AI Search and Digital Growth Strategist with over twelve years in search and growth marketing. He publishes the Discovery Digest, a weekly briefing tracking ten AI search platforms, at nathanmzumara.com/newsletter. More about him and the work at nathanmzumara.com/about.