GEO IS MONEYBALL

Why small businesses get cited by AI search, and how to get on base.

By Anthony Barranco, Protony · 2026-11-03 · 12 minute read · Download the PDF

EXECUTIVE SUMMARY

AI search engines cite pages, not brands. That one fact changes who can win. In classic Google, twenty years of backlinks and brand authority decide who ranks. In ChatGPT, Perplexity, Gemini and Google AI Overviews, the engine reads a handful of pages and quotes the ones it can actually use. A small business with a clear, structured, attributed page gets on base. A famous brand with a messy page does not.

This is the Moneyball opportunity. In 2002 the Oakland A's won 103 games on a third of the Yankees' payroll by buying the undervalued skill that actually won games, on-base percentage, while everyone else paid for stars. AI search is mispricing the same way right now. The skills that earn citations are cheap, measurable and mostly ignored by the companies with the biggest budgets.

The evidence is young but it points one direction. The study that coined GEO found that adding citations, statistics and quotes to a page lifted its visibility in AI answers by up to 40 percent, and that the gain was far larger for lower-ranked sites than for the site already in first place. Ahrefs' 2026 analysis of 4 million AI Overview citations found that 62 percent of cited pages do not rank in Google's top 10 for the query, and 31 percent do not rank in the top 100. Backlinks, the currency of the old game, show the weakest correlation with AI visibility of any factor measured.

This paper is for owners and marketers at small and mid-sized businesses who assume AI search is a big-brand game. It lays out what the engines actually score, what the research shows, what our own AI Visibility Index shows, and a 30-day plan to get on base.

THE MONEYBALL ARGUMENT

The A's didn't buy stars. They bought base runners. In 2002 Oakland had one of the lowest payrolls in baseball and won 103 games, the same as the Yankees, who spent roughly three times as much. Billy Beane did it by noticing that the market paid for the

visible stuff, home run power, big names, the way a player looked, while the scoreboard paid for something boring and measurable: whether a hitter got on base. Walks counted. Singles counted. Oakland bought the skill and let New York buy the brand.

AI search is being scored the same way, and almost nobody has noticed yet.

Traditional search rewards brand. Twenty years of Google have trained every marketer to believe the big site wins, because in a ranked list of ten links, domain authority and backlinks really do decide the order. That belief is the Yankees theory, and for a ranked list it is mostly true.

An AI engine is not producing a ranked list. It has to write one answer and defend it with sources. To do that it retrieves a handful of pages, reads them, pulls the sentences that answer the question, and cites the pages those sentences came from. The question it asks of your page is not "how famous is this company." It is "can I quote this." Is the answer stated plainly? Is there a number or a name I can use? Can I tell who wrote it and when? Did my crawler even get in?

That is on-base percentage. It is scored page by page. And it is exactly the kind of boring, cheap, measurable skill that a market built on brand consistently misprices.

The big brands are the Yankees here. Their pages are written by committees, wrapped in marketing language, loaded with scripts crawlers can't read, and nobody has touched the metadata since the last redesign. They are paying for home run power in a league that has started rewarding walks. A two-person company that answers the question in the first sentence, names its author and lets the crawlers in can get cited ahead of them. Not sometimes. Reliably, and the research in the next two sections shows why.

HOW AI SEARCH ENGINES PICK THEIR SOURCES

Every major AI answer engine works in three steps: retrieve, read, cite. Understanding those steps is the whole game, because each one is a place a small site can win or lose.

Retrieve. When someone asks ChatGPT "who should I hire to redo my website" or "best accounting software for a five-person firm," the model does not answer from memory. It turns the question into several search queries (Google calls this query fan-out), runs them against a live search index, and pulls back a few dozen candidate pages. If your page is blocked to the crawler, or buried in JavaScript it can't render, you lose here and nothing else matters. Ahrefs' April 2026 analysis of 1.4 million ChatGPT prompts found roughly 33 candidate URLs retrieved per prompt, of which about half end up cited.

Read. The engine then reads the candidates looking for sentences that directly answer the question. The same Ahrefs study found the strongest single predictor of citation was plain relevance: how closely the page's title and content matched the question. Pages with natural-language URL slugs were cited 89.8 percent of the time versus 81.1 percent for pages without. This is where answer-first writing pays. A page that opens with the answer gives the model a sentence it can lift. A page that opens with three paragraphs of brand story does not.

Cite. Finally the engine picks which pages to attribute. This is where trust signals matter: a named author, a date, structured data that says what the page is, specifics the model can verify. It is also where the old game stops helping. Ahrefs' May 2025 study of 75,000 brands found the number of backlinks had a correlation of just 0.22 with AI Overview visibility, the weakest of every factor measured, while plain web mentions of the brand scored 0.66.

None of those three steps asks how big the company is. They ask whether the page can be read, whether it answers the question, and whether someone stands behind it. That is why the Moneyball analogy holds: the scoreboard is counting something different from what the market is paying for.

WHAT THE RESEARCH SAYS

Four studies, from two independent sources, support the same conclusion: page quality moves AI citations, brand authority barely does, and the gains go disproportionately to sites that are not already on top.

The paper that named GEO found the underdog effect first. Researchers at Princeton and IIT Delhi coined the term Generative Engine Optimization in a paper presented at KDD 2024 (Aggarwal et al., arXiv 2311.09735). Testing nine content changes across 10,000 queries, they found that adding citations, quotations and statistics to a page raised its visibility in AI answers by 30 to 40 percent. Keyword stuffing, the oldest big-site trick, did nothing and on Perplexity performed about 10 percent worse than doing nothing. The finding that matters most for this paper: the gains were not evenly spread. Adding citations raised visibility by 115 percent for the site ranked fifth in search results, while the top-ranked site's visibility fell by 30 percent. The engines were redistributing attention away from the incumbent and toward whoever gave them the better page.

Most AI Overview citations don't come from top-ranking pages. Ahrefs' March 2026 analysis of 863,000 keywords and 4 million AI Overview citations (reported by Search Engine Journal) found that only 38 percent of cited pages ranked in Google's top 10 for

the query. 31 percent ranked between 11 and 100. The remaining 31 percent did not rank in the top 100 at all. A separate BrightEdge analysis in February 2026 put top-10 overlap even lower, at about 17 percent. Whatever the exact number, the pages Google's own AI chooses to cite are mostly not the pages Google's own ranking puts first.

Backlinks are the weakest signal measured. Ahrefs' May 2025 study of 75,000 brands (ahrefs.com) measured which factors correlate with being mentioned in AI Overviews. Web mentions of the brand came first at 0.66. Branded search volume was 0.39. Domain Rating was 0.33. Referring domains were 0.30. Raw backlink count was last at 0.22. The authority metrics that defined SEO for two decades sit at the bottom of the list.

Relevance beats everything else page by page. Ahrefs' April 2026 study of 1.4 million ChatGPT prompts (ahrefs.com) found that the closeness between a page's title and the user's question was the clearest separator between cited and uncited pages. Cited pages were also not especially fresh: their median age was about 500 days. The engine was not rewarding the newest page or the biggest site. It was rewarding the page that most plainly answered the question.

FINDING	FIGURE	SOURCE, DATE
Visibility lift from adding citations, stats, quotes	+30 to 40%	Aggarwal et al., KDD 2024
Lift for site ranked 5th vs change for site ranked 1st	+115% vs -30%	Aggarwal et al., KDD 2024
Effect of keyword stuffing on Perplexity	about -10%	Aggarwal et al., KDD 2024
AI Overview citations ranking outside Google top 10	62%	Ahrefs, March 2026
AI Overview citations ranking outside top 100	31%	Ahrefs, March 2026
Correlation of backlinks with AI Overview visibility	0.22 (weakest)	Ahrefs, May 2025

Correlation of brand web mentions	0.66 (strongest)	Ahrefs, May 2025
Citation rate, natural-language URL vs not	89.8% vs 81.1%	Ahrefs, April 2026

Taken together: the engines reward pages that answer clearly, cite specifics and can be read, and they reward them more when the site is not already the incumbent. That is a market paying for on-base percentage. Most of the league is still paying for home runs.

THE SECOND WAY ON BASE: BE MENTIONED WHERE THE ENGINES ALREADY READ

A hitter can get on base two ways: a hit or a walk. Fixing your own pages is the hit. The walk is being mentioned on the pages the engine already trusts. Both count the same on the scoreboard, and most companies only practice one of them.

The research is clear on how much the walk is worth. In the Ahrefs brand study, web mentions of a brand were the strongest single correlate of AI Overview visibility at 0.66, three times the correlation of backlinks. The engines are not counting links to you. They are reading what the pages they trust say about you, and the trust goes to a concentrated set of sources. A handful of publications, communities and reference sites get cited again and again for a given category, and most of them are not the outlets with the biggest name recognition. They are the ones that answer the question plainly, which is the same test the engine applies to your own site.

That changes what a PR list is for. The traditional media list is built on reach and recognition: the outlets executives know, the ones that produce volume. The list that matters for AI citations is different and smaller. It is the set of sources an engine actually pulls from when someone asks the question you want to be the answer to, and you can only find it by asking the engines and recording who gets cited. A trade publication nobody in the boardroom reads can be worth more than a national paper if the engine keeps citing it.

Three practical moves follow. First, run the twenty prompts that matter to your business through ChatGPT, Perplexity and Google AI Overviews once a month and list every source cited. That list is your real media list. Second, split it in two: sources where you are

already mentioned and should stay current, and sources where competitors are mentioned and you are absent. The second column is the opportunity. Third, check the date on your existing mentions. The engines favor pages that are current and specific, so a two-year-old mention on a trusted source is a base runner who has wandered off the bag. A fresh mention, a quote, a data point or a product update on that same source puts you back on.

None of this needs a budget either. It needs a spreadsheet, a monthly hour and the willingness to pitch an outlet because the engine reads it rather than because the CEO does.

WHAT OUR OWN DATA SHOWS

Company size does not predict how readable a site is to AI. The Protony AI Visibility Index scores real websites every month on the six things AI engines need in order to read, trust and cite a page: crawler access, structured data, content structure, discoverability, authority signals and technical health. The November 2026 edition covers 26 companies, from household names to small SaaS tools, scored 2026-11-03. Scores run from 0 to 100.

The headline is not that small sites score higher than big ones. It is that size doesn't predict the score at all. The spread among big brands is wider than the gap between big and small.

COMPANY	SCORE	GRADE	BIGGEST GAP
Zendesk	96	A	No author attribution
Stripe	93	A	No publication date
Wix	92	A	No author attribution
Salesforce	91	A	No author attribution
Zoom	82	B	No question headings

Nike	78	C	No question headings
Slack	77	C	No JSON-LD schema
Dropbox	75	C	No social metadata
Hello Bonsai (small freelancer tool)	74	C	No question headings
Shopify	71	C	No JSON-LD schema
Jobber (small business software)	70	C	No JSON-LD schema
monday.com	66	D	AI crawlers blocked
Gumroad	64	D	No JSON-LD schema
Carrd (one-person website builder)	57	F	No JSON-LD schema

Read across that table and the Moneyball pattern is right there. Nike, one of the most recognizable brands on earth, scores a 78. Hello Bonsai, a tool for freelancers that most people have never heard of, scores a 74. Slack and Dropbox, with engineering teams in the thousands, sit in the mid-70s next to a two-person SaaS. monday.com, a public company, grades a D because its robots.txt blocks the AI crawlers outright, so nothing else on the page can count. Zendesk scores 96 not because it is big but because its pages carry full schema, clean structure and open crawler access. Carrd scores 57 for the opposite reason, and a Carrd competitor that added schema and fixed its structure would pass it in a week.

The lesson for a small business is the Oakland lesson. You are not competing against a brand's reputation. You are competing against a specific page, and most of those pages have gaps you can close for the cost of an afternoon.

The Index is re-scored on the first of every month and published at protony.ai/ai-visibility-index. Any site can get the same score in 12 seconds with the free GEO Grader.

THE ON-BASE STATS

If on-base percentage is the number that wins baseball games, these are the numbers that win citations. They are listed in the order the evidence supports, strongest first. None requires a budget. All of them are things a small team can do in a week.

1. **Let the crawlers in.** Check robots.txt for rules that block GPTBot, ClaudeBot, PerplexityBot or Google-Extended, and make sure your important pages render as plain HTML rather than only through JavaScript. A blocked page is a strikeout before you reach the plate. This is the single most common fatal error we see, and it is a one-line fix.
2. **Answer the question in the first sentence.** Every important page should open with the plain answer to the question that page exists to answer. Not a warm-up paragraph. The answer. The Princeton paper found this kind of clarity was worth 15 to 30 percent in visibility on its own, and Ahrefs found relevance between the question and the page was the clearest separator between cited and uncited pages. This is the sentence that gets quoted.
3. **Put specifics on the page.** Numbers, named examples, definitions, a quote from a real person. Adding citations, statistics and quotations was the single strongest change in the Princeton study, worth up to 40 percent. Vague pages are not quoted because there is nothing to quote.
4. **Put a name and a date on it.** An author with a bio, an about page, a visible published and updated date. The engine is deciding whether a human stands behind the claim, and anonymous pages lose that call.
5. **Add structured data.** Organization, Article, Person and FAQ schema tell the engine who you are and what the page is without making it guess. It is roughly thirty lines of code a developer can paste into the page head.
6. **Add an llms.txt file.** A short markdown file at /llms.txt that names your site and links your key pages. We list it last on purpose. It costs ten minutes and it cannot hurt, but the evidence that engines read it today is thin (see the limits section below). Treat it as cheap insurance, not a lever.

Do the first four on your ten most important pages and you have a lineup. The Protony GEO Grader checks all six and ranks the fixes by points available.

A 30-DAY PLAN FOR A SMALL BUSINESS

You do not need a rebuild. You need to get on base on the pages that matter, in this order.

Week 1: find out where you stand. Run your site through the GEO Grader and write down the score and the top three fixes. Open ChatGPT and Perplexity and ask the five questions a customer would ask before hiring you. Note whether you appear, and who does instead. That list of competitors is your scouting report: they are the pages you are actually up against.

Week 2: open the gates. Fix robots.txt so the AI crawlers can get in. Confirm your key pages render without JavaScript. Add Organization and Person schema and an about page with a real name and face. Add llms.txt. This is plumbing and it takes one afternoon with a developer.

Week 3: rewrite the openings. Pick your ten most important pages: home, services, top blog posts, the pages that answer "how much" and "how long" and "is this right for me." Rewrite the first paragraph of each so the first sentence is the answer. Add one number, one named example or one short quote to each. Put an author and a date on every one.

Week 4: re-score and re-ask. Run the grader again. Ask ChatGPT and Perplexity the same five questions. Technical fixes are typically picked up within days to a few weeks as crawlers revisit; content changes take a month or two to show in answers. You will not win every question in 30 days. You will see the score move and usually one or two questions flip.

Then keep going, ten pages a month. That is the whole swing. Singles and walks. The home runs can wait.

WHERE THE EVIDENCE IS STILL EARLY

GEO is three years old and the engines change monthly. A paper that overstated its case would be the one thing that undercut it, so here is what we do not yet know.

llms.txt is widely recommended and rarely read. Originality.ai's tracking study found adoption grew 8.8 times between June 2025 and May 2026, to about 36,000 sites. But Ahrefs' server-log analysis of 137,000 domains in May 2026 found that 97 percent of llms.txt files received zero requests, and AI retrieval bots made up about 1 percent of the requests that did arrive (PPC Land, July 2026). Google has said the file does not affect

Search visibility either way. We still recommend it because it is free and the engines may adopt it, but we score it at 4 points out of 100 and you should treat anyone who sells it as a primary lever with suspicion.

Correlation is not causation. The Ahrefs brand study measured which factors travel with AI visibility, not which ones cause it, and its sample was limited to domains with a Domain Rating above 40, so it says little about very small sites directly. The Princeton paper is experimental and does show cause, but it was run on the engines of 2023 and 2024.

The engines concentrate citations. Several analyses have found that a small set of large domains, Wikipedia and Reddit among them, take a disproportionate share of ChatGPT citations overall. The Moneyball argument is not that small sites win most citations in aggregate. It is that for the specific, commercial questions a business cares about, the page that answers best gets cited regardless of who owns it, and the research above shows the incumbent advantage is far weaker than it is in classic search.

Methodology is moving. Ahrefs' own top-10 overlap figure fell from 76 percent to 38 percent between July 2025 and March 2026, and the company notes part of that drop reflects better parsing rather than a change in Google's behavior. Every figure in this paper carries a date for that reason. We will update it as the data does.

FREQUENTLY ASKED QUESTIONS

What is Generative Engine Optimization?
GEO is the practice of making a website easy for AI answer engines like ChatGPT, Perplexity, Gemini and Google AI Overviews to read, trust and cite. It focuses on page structure, structured data, crawler access and authority signals rather than backlinks and domain authority.

Do small businesses have a real chance of being cited by ChatGPT?
Yes. AI engines cite individual pages, not brands. A clear, well-structured, attributed page on a small site can be cited ahead of a larger brand with a weaker page. The Princeton GEO study found that adding citations raised visibility 115 percent for the site ranked fifth while the top-ranked site lost 30 percent.

Is GEO the same as SEO?

No. SEO competes for a spot on a ranked list of links, where backlinks and domain authority decide the order. GEO competes to be the sentence inside a generated answer, where clarity, specifics and attribution decide. Many GEO fixes help SEO too, but the priorities are different.

Is llms.txt worth adding?
It is free and takes ten minutes, so yes, but it is not a lever. Ahrefs server-log data from May 2026 found 97 percent of llms.txt files received zero crawler requests, and Google says it does not use the file. Fix crawler access, answer-first structure and author attribution first.

How long does GEO take to work?
Technical fixes like crawler access and schema are picked up within days to a few weeks as AI crawlers revisit the site. Content changes take longer, usually one to two months, because they depend on recrawl cycles and the engine's confidence in the page.

What does the Protony AI Visibility Index measure?
It scores real websites each month on six things AI engines need to read, trust and cite a page: crawler access, structured data, content structure, discoverability, authority signals and technical health, out of 100. The same scoring runs the free GEO Grader at protony.ai.

SOURCES AND METHODOLOGY

All external figures were taken from the source pages below, opened and read on 2026-11-03. Protony Index figures were scored 2026-11-03 with the current grader weights, which value open crawler access over llms.txt.

Changelog. 2026-11-03: first published. 2026-11-03, later the same day: added the section on earned mentions ("The second way on base"), drawing on the Ahrefs brand study figures already cited.

SOURCE	WHAT IT IS	USED FOR

Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan, Deshpande. GEO: Generative Engine Optimization. KDD 2024, arXiv 2311.09735v3	Peer-reviewed experimental study, 10,000 queries, nine content methods	Visibility lifts, keyword stuffing result, rank-5 vs rank-1 effect
Ahrefs, Why ChatGPT Cites One Page Over Another, April 15, 2026	Analysis of 1.4M ChatGPT prompts	Retrieval counts, relevance finding, URL slug citation rates, page age
Ahrefs AI Overview citation study, March 2026, via Search Engine Journal	863,000 keywords, 4M AI Overview citations	Share of citations outside top 10 and top 100
Ahrefs, An Analysis of AI Overview Brand Visibility Factors, May 26, 2025	75,000 brands, Spearman correlation	Backlinks vs mentions correlations
PPC Land, llms.txt adoption rises 8.8x but 97% of files get zero AI requests, July 2, 2026	Originality.ai adoption tracking and Ahrefs server-log data	llms.txt adoption and crawler request rates
eWeek, ChatGPT nears 1 billion weekly users, February 27, 2026	OpenAI announcement as reported	Scale of AI search adoption (900M weekly users)

How the Index scores. Each site is fetched and checked against six categories with fixed point values: AI crawler access (20), structured data (20), content structure (20), discoverability (15), authority signals (15) and technical health (10). Checks are automated and identical for every site. The same scoring runs the free GEO Grader at protony.ai. The Index is re-run on the first of each month.

About the author. Anthony Barranco runs Protony, a GEO agency for small businesses, startups and B2B brands, and leads AI search visibility work for a Fortune 100 technology brand. He has run paid media and digital marketing for Johnson & Johnson, Zoetis, Pepsi, Smile Doctors and Cardless.

Cite this paper: Barranco, A. (2026). GEO Is Moneyball: Why small businesses get cited by AI search. Protony. https://www.protony.ai/research/geo-is-moneyball/