FIELD REPORT NO. 03 · REDLINE ADVISORS

The sound of *sameness*

What twenty-nine AI companies told us about the price of sounding like everyone else.

AUTHOR

Mick Hollison
Founder & CEO

PUBLISHED

July 2026
Park City, Utah

DATASET

29 homepages
Signal Scan, July 2026

SERIES

No. 01 Defying the Algorithm
No. 02 Speed of Thought

Two field reports ago I made a claim I couldn't prove.

I said B2B technology had drifted into a Sea of Sameness. That AI had democratized bland. That every company in every category had started to sound like every other company in that category, and that buyers had quietly lost the ability to tell them apart.

I believed it. I had thirty years of pattern recognition behind it. IBM. Microsoft. Citrix. Cloudera. What I didn't have was a number.

Now I do. It is seventy.

In July 2026 we batch-scanned the public homepages of 29 venture-backed AI companies, seed stage through growth, plus one large incumbent for ballast. Same scoring engine as the free Signal Scan anyone can run at redlinecs.ai/signal-scan. Same extraction, same scorer, no manual adjustment, no thumb on the scale. We fed it the URL, and we read what came back.

I ran it out of curiosity, not strategy. I wanted to know how the best of the best were doing. Not the stragglers and not the me-too crowd. The front of the pack in the defining technology shift of our lifetimes, the companies with the most money, the best engineers, and the sharpest people in the building.

Then Anthropic scored a 50.

I name them here as a deliberate exception to my own rule. Every other company in this index stays anonymous because they are live prospects and naming them burns the relationship. Anthropic is different. I depend on them, I'm transparent about that dependency throughout this report, and the irony only works if you know who I'm talking about.

I want to be precise about why that one landed the way it did. I run my business on Anthropic's tools. I build products with them. I used them to help construct the report

you're holding. And the company I lean on that hard couldn't clear a 50 on its own front door.

I didn't expect that. I expected them to do better.

They'd fallen into the trap that has swallowed this industry for forty years. They described their what. Endlessly, capably, in enormous detail.

None of which means you can't build an extraordinary company that way. Anthropic may be the fastest-growing revenue company in history. But excitement about a what is ephemeral. It's real and it's powerful and it can evaporate in a quarter, the moment human buyers stop believing in the why, or never learn what it was.

Eleven of the 29 landed on exactly the same score.

Not near it. On it.

This is what that means.

01

The scoreboard

Before the argument, the arithmetic.

29 COMPANIES SCANNED, JULY 2026	65 AVERAGE SCORE OUT OF 100	38% LANDED ON EXACTLY 70
7% REACHED STRONG SIGNAL, 80 OR ABOVE	31% HIT THE FLOOR ON CONCRETE	4 LOWEST SINGLE-DIMENSION SCORE IN THE SAMPLE

The Signal / Whisper Ratio™ scores copy from 0 to 100 across four published dimensions: Buzzword-free, Authentic, Concrete, and Punchy. The bands are Whisper from 0 to 35, Faint signal from 36 to 55, Signal forming from 56 to 79, and Strong signal from 80 to 100.

Average across the 29: 65. The dead middle of signal forming.

That is the whole problem in one number. A 65 isn't broken. A 65 is competent. A 65 is a headline that went through three rounds of review, survived a brand workshop, got a thumbs-up from the CEO, and shipped. 65 reads fine.

A 65 also doesn't get repeated.

There is a particular kind of noise that carries no information. It's not silence. Silence at least gets noticed. It's the sound of a room where everyone is speaking, fluently and politely and at length, and nothing anyone says would change if you swapped the speakers. That is the sound this index recorded. 29 companies, competent to a fault, and almost none of them saying anything the others couldn't have said.

Nobody's going to fire a marketing team over a 65. That is exactly why the 65 is still up there.

A 65 is a session musician. Perfect pitch, perfect timing, shows up early, nails the take in one pass. And nobody's ever bought a record for the session musician.

02

The ceiling at seventy

Here is the finding that reorganized how I think about my own work.

Eleven of 29 companies, 38 percent of the sample, scored exactly 70. Not 69. Not 71. Seventy.
That's not a coincidence and it isn't a rounding artifact. It's a gate.

One rule inside the scoring logic matters more than the rest: copy can't climb above 70 unless it contains something only that company could say. A uniqueness claim. A named, owned mechanism. A category the company is willing to plant a flag in. Absent that, the ceiling holds no matter how good everything else is.

And everything else was good. That is the part worth analyzing further.

These 11 companies wrote clean copy. Buzzword-light. Human. Readable. Free of the substance-free, buzzword bingo that made B2B a punchline for two decades. Somebody on each of those teams did real work, and by every conventional measure of quality they did it well.

They aren't failing on polish. They're failing on soul.

I learned this from Mark Templeton, who was CEO at Citrix, who learned it from his son. It went from a young man's recommendation to the operating system of a multibillion-dollar company in about a year, which should tell you something on its own.

Everyone can tell you what they do. Some can tell you how they do it. Almost nobody can tell you why they exist, and **the ones who can are the ones people follow**.

I had never seen it measured. Now I have.

The distinctiveness gate and the why gate turn out to be the same gate.

Those 11 companies articulated a what and a how, cleanly and competently and with real craft, and then stopped at the edge of the only question that would have separated them.
Seventy is the score you get for a well-executed what.

I've watched this from four seats and the pattern holds at every one.

IBM had a clear why and it was honestly about the customer, meeting the need and then beating it. That made us competent. It didn't always make us compelling.

Microsoft, during my stretch there, was in a hard few years. We were tremendous iterators and not always great innovators. The products had turned inward about the same way the culture had. The why was about how things got done rather than what actually got accomplished for anyone in the market.

Citrix was the interesting one. Templeton put the why at the center of everything, marketing most of all, and people believed it. We thought we could deliver a great computing experience, secure by design, at a fraction of what any other model cost. We went from roughly $600 million in revenue to north of $3 billion while I was there. There was a lot of love in that building. We were right for a long time, right up until we weren't.

Cloudera was where I got to run the play myself, with a C-suite of Berkeley and Stanford people who all had strong opinions about why we existed. It took almost three months to land on 38 words covering the why, the how, and the what. Ten years later that why is still largely intact, and I'm prouder of it than almost anything else on my resume.

Which is why I run those sessions with the C-suite and nobody below it. My favorite moments are the ones where it gets loud. Executives calling bullshit on each other, constructively, blood pressure up, because they're arguing about the why and they know what it's worth. That argument is the DNA. Get it right and every decision downstream is easier and clearer and better. Miss it and things go sideways in a hurry.

Over a third of a sample of well-funded AI companies wrote a homepage headline that any of their competitors could paste onto their own site without anyone noticing. Not because the writers were lazy. Because nobody upstream of the writers ever decided what the company was actually claiming, and a writer can't manufacture a differentiator that the business hasn't chosen.

The copy is a symptom. The unmade decision is the disease.

Two companies of 29 cleared 80. Both did the same thing to get there: they carried a named, owned idea in the headline itself. One went so far as to claim a world-first category outright. You can argue with a claim like that. Reasonable people will. But you can't confuse it with anyone else, and in a market where 38 percent of your peers are interchangeable, being arguable beats being invisible every day of the week.

7 percent. That is the strong-signal rate among 29 companies who collectively raised enough capital to run a small nation.

03

The AI companies gave the AI nothing to repeat

Nine of 29, 31 percent, scored the absolute floor on Concrete. A 12.

Floor on Concrete means the line contained no numbers. No named customers. No verifiable facts. Nothing a reader could check and nothing a machine could carry.

Read that back with the sample in mind. These are AI companies. Their entire pitch is that machines can extract meaning from information. And they built a front door with no extractable information in it.

When a buyer asks ChatGPT to compare vendors in that category, the model reaches for whatever it can hold onto. Numbers. Named deployments. Specific mechanisms. Dated proof. A company whose homepage offers none of those isn't misrepresented in the answer. It is absent from the answer. There was nothing to pick up.

This is the mechanism underneath the Sea of Sameness, and it is more banal than anyone wants it to be. Companies aren't losing the AI-mediated shortlist because of some exotic algorithmic bias. They're losing it because they wrote nothing down that a machine could carry across the room.

One more from the sample, because it earned its place. A single homepage lead sentence ran 47 words and scored a 4 out of 100 on Punchy. The lowest single-dimension read we recorded.

47 words. One sentence. A machine will compress that to nine words and choose which nine. If you didn't choose, it will choose for you, and it won't call to ask.

04

The second surface

Everything above concerns the words. There is a second failure mode that has nothing to do with the words, and I only understood it clearly a few weeks ago.

The two surfaces take two different disciplines, and confusing them is most of the problem.

The human surface is crafted. Artisanal work, iterative and judgment-driven, built to be felt rather than skimmed. It runs on belief, on story, on why.

The machine surface is engineered. Cold, structural, unforgiving work, built to be extracted rather than admired. It runs on markup, on verifiable fact, on what.

Which puts a wrinkle in the most quoted idea in modern marketing. Sinek's line is that people don't buy what you do, they buy why you do it. That is still true of the humans. It's not true of the machines. A model doesn't buy anything and it can't be moved. It has no capacity for belief and no interest in your origin story. It needs the what, in plain text, ranked in the markup, and checkable against something.

So the job has doubled. Same page, two audiences, opposite requirements. Craft the why for the human. Engineer the what for the machine. Get either one wrong and the other one can't save you.

Craft isn't polish. Polish is what you get when a talented team applies real skill to a question nobody upstream ever answered, which is a fair description of 11 companies sitting at exactly 70. And engineering isn't optional simply because nobody on the marketing team owns it.

A great dish is crafted. The walk-in cooler is engineered. Nobody's ever photographed the walk-in, and the restaurant closes without it.

Most companies polish one surface and inherit the other.

On July 27 we pulled apart the homepage of ServiceNow. Not a company from the index, and not a company with a copy problem. Their hero line scored a 72, the highest single read in anything we have published in this series:

> *AI without workflows is just expensive advice. Most companies have AI. Few have the data to ground it, the workflows to deploy it, and the security to govern it.*

That is good writing. Buzzword-free scored 100. Punchy scored 100. A sharp, opinionated, load-bearing sentence.

Then we asked four AI models the same plain question. What is the current headline on servicenow.com?

We got four different answers.

Gemini returned the indexed metadata layer and an entirely separate line about being an AI control tower for business reinvention. ChatGPT returned the hero pitch above. Perplexity returned an event recap for a user conference. Claude returned all three slides of a rotating carousel, correctly enumerated, and volunteered that it couldn't determine which one loads first.

Four extractors. One question. Four answers, none of them wrong.

Then we counted the markup. 149 headings on the page. Zero h1 elements. Zero h2 elements. 147 of the 149 were navigation menu labels. Exactly one heading was visible on load. And that excellent 72-scoring pitch line? It didn't appear in the heading list at all. It's not marked up as a heading of any kind.

Here is the precision that matters, and I got it wrong myself on the first pass.

The failure isn't that the machines couldn't read the page. Claude read the whole carousel perfectly, in order. Reading was never the constraint.

The failure is that nothing on the page declared which claim was the claim. 149 candidate headings, no ranking among them, and four extractors each obliged to guess.

If complexity were the problem, a Wikipedia article would be unreadable. It is structurally far more complex than servicenow.com. Deeper nesting, vastly more text, orders of magnitude more links. And every model on earth agrees what a Wikipedia article is about, because its title, its opening heading, and its first paragraph all say the same thing.

Complexity isn't the problem. Unranked priority is.

Now the caveat, because I'd rather state it than have it used against me. The missing h1 isn't the cause of anything. Google has said so directly and repeatedly, for years: a page ranks perfectly fine with no h1 tags or with five, and their systems work with the HTML as they find it, including pages with no semantic structure at all. Anyone who tells you the h1 tag is a lever is selling you 2011.

The h1 count is a proxy, and a good one, which is a different job. It's the cheapest observable stand-in for the thing that actually matters, which is whether anything on the page agrees with anything else about which claim is primary. On servicenow.com the title said one thing, three rotating slides said three more, and no element was marked as dominant. You can argue about heading semantics all day. You can't argue with four extractors returning four answers.

So the requirement isn't a tag. The requirement is that one unambiguous, prominent, server-rendered sentence states what you are, and that your title, your opening body copy, and your heading structure all agree on it. The mechanism is redundancy and agreement. The h1 is just where you can count it in thirty seconds.

So there are two failure modes, and they take different repairs:

Readability failure. The claim exists only after JavaScript executes, or it is baked into an image. Some machines physically can't retrieve it. The fix is server-side rendering and live text.

Salience failure. The claim is perfectly readable and completely unranked. Every machine retrieves it. None of them knows it matters. The fix is heading hierarchy and title agreement.

ServiceNow is the second kind, which is the more interesting kind, because it is invisible to every tool that checks whether a site works. The site works. The site is fast, accessible, well-built, and professionally

maintained. It simply never says which sentence is the important one.

A three-slide carousel is three coequal claims with identical markup weight. That's not a design decision. It is an unmade positioning decision, shipped to production. The rotation is the symptom. The tie is the disease.

We saw the same shape at SAP two weeks earlier. A 65 composite, a 32 on Concrete, and at least three different homepage headlines being served simultaneously depending on who was looking. My own browser saw one. Two models saw a second. A third model returned the title and meta layer. Every model matched a real layer of a real site. There simply was no canonical headline to match.

I want to be careful here, because this is where commentary usually outruns evidence. We never proved why SAP's variants split. Rendering depth, metadata staleness, and differently-timed cached copies are all live candidates and we didn't rule any of them out. What we can say is narrow and sufficient: four observers, one question, no agreement.

Humans tolerate ambiguity. Machines resolve it. And they don't confer.
One more piece of intellectual honesty, and it turns out to be the most useful paragraph in this report.

We asked four models what the current headline on servicenow.com is. That is a question about a page. It proves the page has no canonical claim. It doesn't prove the models are confused about what ServiceNow does, and if you go ask them that instead, you will get four fairly consistent answers.

The reason is worth understanding, because it is the whole game. A model's sense of a company comes from two places. There is what it absorbed in training, which is the entire published corpus: news coverage, analyst notes, review sites, forum arguments, conference transcripts, filings, Wikipedia, and every article a journalist wrote by paraphrasing your positioning. And there is what it retrieves at the moment somebody asks. Your homepage is one document competing against all of that.

By raw volume your homepage is a rounding error. Its importance isn't volume. It's that the homepage is the seed. It's the canonical self-description that everybody downstream paraphrases. Reporters lift it. Analysts quote it. Partners paste it into listings. Your own reps internalize it and repeat it in calls. Wikipedia editors cite it. Over two or three years a distinctive claim propagates into hundreds of derivative documents, and that propagation is what the model absorbs. A generic claim propagates too. It propagates as category noise, indistinguishable from what your competitors seeded.

Which produces the finding that should worry the people reading this most.

Homepage influence is inversely proportional to corpus density. ServiceNow survives its own markup because two decades of coverage give the models somewhere else to look. Their homepage might be a thousandth of what a model knows about them. A three-year-old AI company with forty press mentions has no such cushion. For them the homepage isn't a rounding error. It is most of the file.

Every company in our index of 29 is in the second category. So is almost every company that will read this.

05

What the research actually says

I didn't want this report to rest on three homepages and a strong opinion. So I went and read the literature, and the literature turns out to be considerably more interesting than the marketing built on top of it.

In July 2026 a critical survey of Generative Engine Optimization research was published on arXiv, reviewing 45 studies from November 2023 through July 2026. It's the most honest document I've read on this subject and almost none of it has made it into the trade press, presumably because it doesn't sell anything.

Five findings from that body of work deserve to be in every marketing leader's head.

Fixed tactics don't generalize. The C-SEO Bench study tested optimization methods across six domains and roughly 1,900 queries. Of 54 method-and-domain combinations, three came back significantly positive in the main experiment. None were positive in question answering. Several transformations actively lowered rank.

Gains erode as adoption rises. The same work found that improvements decline as more competitors adopt the same technique, trending toward a zero-sum contest. Which is precisely what a tactic is. When everyone runs the play, the play stops working.

Optimizing for citation can cost you retrieval. The SAGEO Arena study reinstated the full pipeline across 171,000 documents and found that optimizing a page body alone reduced its presence in the top twenty results by roughly 9 percent, reduced top-ten presence after reranking by 16 percent, and reduced final citation by 6 percent. Winning the quote and losing the room.

Engines don't share sources. One 2026 audit reported URL-level overlap between organic Google, AI Overviews, and Gemini in the range of 0.11 to 0.18. Another found that 53 percent of domains cited by Google's AI Overviews don't appear in the organic top 10, and 27 percent don't appear in the top 100. There's no single ranking to win. There are several different weather systems and they aren't correlated.

Being known isn't the same as being surfaced. A 2026 study of 112 startups found that ChatGPT recognized 99 percent of the products when they were named, but surfaced them in roughly 3 percent of organic discovery queries. Perplexity ran 94 percent recognition against 8 percent discovery. That single preprint uses two models and should be read as directional rather than settled, but the distinction it draws is the whole ballgame. The machine knows who you are. It doesn't think of you.

The survey's own summary conclusion is worth stating plainly, because it is the opposite of what the category sells: across all 45 studies reviewed, no technique demonstrated a stable, longitudinal, cross-platform causal effect on organic discoverability or on downstream clicks and conversions. Content already

retrieved can be made to matter more. Getting retrieved in the first place remains largely unsolved by tactics.

Scott's own research puts a floor under the urgency, and it is worth stating in his numbers rather than mine. Drawing on the G2 2026 AI Search Insight Report, he reports that a slim majority of B2B software buyers now begin research with an AI assistant more often than with a traditional search engine, that roughly seven in ten use one somewhere in the buying process, and that close to a third never leave the conversation to visit a website at all. His broader estimate is the one that should keep marketing leaders up at night: of the more than six million corporations operating in the United States, he puts the share effectively invisible in category-first AI answers, unless a buyer names them outright, at 95 to 97 percent. He calls the resulting winner-take-most dynamic the Great Marketplace Compression, and the population it strands the corporate graveyard of AI invisibility.

His diagnosis of the tooling market matches what the academic literature found. G2 now lists over 400 AEO vendor solutions. Nearly all of them report what is happening. Very few explain why it is happening, which is the difference between a dashboard and a diagnosis, and it is the gap his index was built to close.

And here is my favorite artifact of the tactical era. For 18 months the industry has been telling everyone to publish an llms.txt file. Reported adoption sits at roughly one in ten domains. Google has said on the record that it doesn't use the file. Crawler-traffic analyses suggest the answer engines overwhelmingly skip it and read the HTML directly. Those adoption and traffic figures come from vendor research rather than peer review and should be treated as directional. The direction is clear enough.

An entire discipline spent a year and a half on a file that the engines largely don't read, while 38 percent of a sample of AI companies sat pinned at 70 because nobody had decided what the company was.

06

Three layers, and where everything sits

The confusion in this market is that AEO, GEO, SEO, and positioning all get discussed as if they were competing for the same job. They're not. They operate at different layers, and the layers are sequential.

Layer one is substance. Do you have a claim that only you could make? This is the Signal Core™, the irreducible differentiator, expressed as a Brand Checksum™ built to survive compression. Nothing in the GEO or AEO tool landscape works at this layer. There's no platform that measures it, because it isn't a measurement problem until somebody has made a decision. For what it is worth, this is the specific area where RedLine excels.

Layer two is structure. Is that claim readable, ranked, and unambiguous on the surfaces you control? One prominent sentence that says it, with the title tag, the opening body copy, and the heading structure all agreeing. This overlaps technical SEO at the edges, but the question is different. SEO asks whether the page can be indexed. This asks whether the page declares a priority.

Layer three is distribution. Do the engines retrieve, cite, and repeat it? This is GEO and AEO proper. This is where Scott Hebner's AEO Advantage Index sits, and it is a leading approach to the question of external, buyer-facing visibility. It is where the monitoring platform category sits too, and that category is doing real work: tracking mention frequency, share of voice, sentiment, and citation sources across engines over time. Those tools measure something genuine that we don't measure.

The relationship is straightforward once the layers are separated.

GEO asks whether the machines can find you. The Signal / Whisper Ratio™, or SWR, asks whether what they find is worth repeating.

Sequential, not competing. And the sequence only runs one direction. You can't optimize the distribution of a claim that doesn't differentiate. Do that successfully and all you have done is distribute the sameness more efficiently. 11 companies in our sample could hire the best answer-engine firm in the world tomorrow and the result would be 11 companies that are more findable and still interchangeable.

Which brings me back to the erosion finding, because it is the strategic center of this whole report.

Layer three is a treadmill, and it speeds up as more people climb on. That's not a criticism of the discipline. It's a structural property the research has now documented: individual gains decay as adoption rises.

Layer one doesn't decay. Distinctiveness is the only lever in the entire stack that doesn't erode when your competitors adopt it, for the plain reason that it is definitionally the thing they can't adopt.

Your competitors can copy your schema markup by Friday. They can't copy the customer who chose you over the incumbent and can explain exactly why.

The RedLine advantage

Here is how our own instruments map onto that stack, stated plainly so nobody has to guess.

Signal Scan is a free, single-surface read at layer one. It scores the most load-bearing sentence you own and it takes about a minute. It's a directional read of the front door, not a full-site audit.

The Signal / Whisper Ratio™ spans layers one and two, through two lenses. Signal Fidelity looks inward, at how much of the intended narrative survives LLM compression. Signal Discovery looks outward, at whether buyers searching the category find you, hear a distinct story, and receive the whole thing rather than a

fragment. That outward lens is where our work touches layer three and where the AEO and monitoring tools are genuinely complementary rather than redundant.

The Signal Core™ and the Brand Checksum™ live at layer one exclusively. The Signal Brief™ governs layer one and, in the revision this research has forced on us, layer two as well. Signal Watch™ monitors drift across all three over time, because every one of these is a moving target and none of them holds still after launch.

Layer three is where Scott Hebner's work comes in. Scott is principal analyst for AI and CMO practice leader at theCUBE Research, he joined RedLine as an operating partner this month, and the AEO Advantage Index is his instrument, not ours. It measures the outside-in half: whether the engines discover you, cite you, shortlist you, and recommend you across the AI-mediated buyer journey, and why they do or do not.

I'm not going to pretend these are one methodology. They're two instruments pointed at different layers, now under one roof, and we're still working out how they fit together. But the division is clean enough to state. We measure whether the claim is worth repeating. His index measures whether the machines are repeating it. You want both numbers, and until recently almost nobody had either.

And one asymmetry between the layers that took me too long to see clearly.

Markup is addressing. It routes attention on pages you own. The checksum is the payload. It determines whether the claim survives being compressed to 15 words and restated by somebody who isn't you.

Compression is lossy and it's selective. Specific, unusual, checkable things survive it. Generic things dissolve, because a model can rebuild them from the category without you.

Scouts don't write "he's a good athlete." They write six foot four, 240 pounds, runs a 4.5. The first one dissolves the moment anybody repeats it. The second one gets you drafted.

That is the mechanism under our Concrete floor of 12. Nothing survived, because nothing was unreconstructable.

But here is the part that settles the priority question. Markup only works on surfaces you control. The checksum works everywhere. Your h1 can't travel into a G2 profile, an analyst note, a podcast transcript, or an argument on Reddit. Your checksum can, because human beings repeat it. Layer two is local and it stays local. Layer one compounds for years.

If you have budget for exactly one of them, this isn't a close call.

That last point isn't a preference. The literature demands it. Reported daily source-level overlap between repeated identical queries runs in the range of a third to two-fifths. One audit found that even with the models locked into their most predictable, least random setting, asking the identical question again changed between nine and 28 percent of the answers. Another found that a majority of ChatGPT repetitions in its configuration didn't activate web search at all.

Researchers in this field now recommend repeating each prompt seven to eight times before believing the result, and varying across run, paraphrase, date, and engine.

Which means a single AI visibility check isn't a measurement. It is an anecdote with a number attached.

Anyone selling you a one-time audit of how the machines see you is selling you one frame of a film.

07

The honest part

I'd rather disclose the limits than have someone else find them.

Sample. 29 companies. One vertical, AI and software. July 2026. One scan per company. Single-line homepage reads. No follow-up, no time series, no control group. 29 is a sounding, not a survey. It is enough to establish that the ceiling at 70 exists and enough to show its shape. It's not enough to generalize to all of B2B, and I'm not going to.

Method. Every company in the index was scanned the same way anyone can scan their own line, through the public tool. The ServiceNow copy discussed in section 04 was an exception. It was pasted as text rather than scanned by URL, because the scanner timed out on their site. SAP and Palantir were URL scans. I state that because the alternative is defending it later.

What I'm not claiming. I'm not claiming the scanner timeout says anything about ServiceNow's performance. It says something about our timeout threshold. I'm not claiming to know why SAP served three headlines. I'm not claiming that any of these companies will lose deals over this, because I've no attribution data and neither does anyone else. The research is explicit that the link from citation metrics to revenue is the weakest evidence in the entire field.

Third-party figures. The Forrester numbers below come from Forrester's own published materials. The GEO research findings come from a July 2026 critical survey and the studies it reviews, several of which are preprints and are flagged as such by the survey authors. The llms.txt adoption and crawler figures are vendor research and should be treated as directional. Where I couldn't get to a primary source, I left the number out rather than round it into place.

One more disclosure, offered in the spirit of the rest. Not a single company has come back to dispute a score. That's not evidence the scoring is right. It might only mean nobody enjoys arguing in public about their own homepage. I mention it because I'd report the pushback if there were any, and there hasn't been any yet.

We're expanding the index. The next cut will be larger and cross-vertical, and it will include a structural pass on heading markup so that section 04 stops being three specimens and starts being a correlation. Confidence with disclosure beats false precision every time.

08

What I'd do Monday morning

Not a methodology. Four things any team can do this week without hiring anybody.

Run your own line. Paste your homepage headline and supporting sentence into the free scan at redlinecs.ai/signal-scan. It takes a minute. If you come back at exactly 70, you now know precisely what is wrong, and it isn't the writing.

Check whether your page agrees with itself. Open your homepage, view source, and compare three things: the title tag, the first heading, and the first sentence of body copy. If they say three different things, or if the sentence you actually care about isn't among them, you have a salience problem and your copywriter can't fix it. Counting h1 tags is the fast version of this check, not because the tag is a lever, but because it is the cheapest place to see whether anyone ever decided.

Ask four models the same question. Open clean sessions in ChatGPT, Gemini, Claude, and Perplexity. Ask each one what your company does and who it is for. Count the distinct answers. One is a good day. Four is the finding.

Then do the hard part. Find the sentence only you can say. Not the sentence you wish were true, and not the sentence the category says. The one your best customer already uses when they explain you to a colleague. That sentence is sitting in a call recording somewhere in your organization right now.

Anyone can count your h1 tags. Anyone can run the four-model test. What none of them can tell you is which sentence belongs in the h1. That is the work, and it has never been a machine's job.

The finding I wasn't looking for

I built Signal Scan expecting founders and marketers. Positioning people. The audience I've sold to my whole career.

Then a former colleague mentioned he was running his cold and warm outreach through it. Sales emails. BDR copy. He works at ServiceNow, which, given section 04, I find delightful. He wasn't scoring a homepage. He was scoring the first line of a note to a stranger.

I never saw that coming, and he isn't alone. Sellers have turned out to be some of the heaviest users of the tool, which means I now have to teach the scorer to tell sales copy from marketing copy, because they aren't the same job and shouldn't be judged on the same scale.

That's a whole other report. But it tells you something that the people who understood this first weren't the brand teams. They were the ones who find out in about nine seconds whether a sentence worked.

The close

Positioning used to end at the words. Now it ends at the markup, and then it ends again at the retrieval, and every one of those layers is a place where a decision you never made gets made on your behalf.

Field Report No. 01 claimed the ocean existed. This one measures it. 65 out of 100 on average, with a hard shelf at 70 that 38 percent of a well-capitalized sample is sitting on right now, writing perfectly good sentences that could belong to anybody.

The companies that break through won't be the ones with the best schema markup. The research is now fairly clear that the tactical edge decays on contact with competition. They will be the ones willing to say something arguable, put it in one sentence, mark it up so a machine knows it is the sentence, and then check every month whether the machines still repeat it correctly.

That's not optimization. That is deciding who you are and refusing to let a summarizer do it for you.

The window is open and it is closing. The models are forming their version of your company right now, from whatever you left lying around on the front door.

Go read your own homepage the way a machine reads it. Then go fix the sentence.

Defy the algorithm.

Sources and notes

1. **RedLine Signal Scan Index, July 2026.** 29 venture-backed AI and software companies, public homepage headline and supporting line, scored 0 to 100 on the Signal / Whisper Ratio™ across Buzzword-free, Authentic, Concrete, and Punchy. One scan per company. Company names withheld. RedLine Advisors proprietary data.

2. **Martinez, O. (2026).** *Optimizing Visibility in Generative Engines: A Critical Survey of Generative Engine Optimization (2023–2026).* arXiv:2607.14035, July 15, 2026. Critical scoping review of 45 studies. Source for the C-SEO Bench, SAGEO Arena, cross-engine overlap, run-to-run stability, recognition-versus-discovery, and repetition-protocol findings cited in sections 05 and 06. Several underlying studies are preprints or forthcoming at the time of that review and are flagged as such by its author.

3. **Forrester Research.** *The State Of Business Buying, 2026*, published January 21, 2026, drawing on Forrester's Buyers' Journey Survey, 2025. 94 percent of business buyers report using AI in their buying process, up from 89 percent the prior year. Twice as many buyers named generative AI or conversational search as a more meaningful information source than any other source, ahead of vendor websites, product experts, and sales. Typical buying decision now involves 13 internal stakeholders and nine external influencers.

4. **ServiceNow homepage structural analysis.** Console DOM capture, July 27, 2026. 149 headings, zero h1, zero h2, 147 navigation labels. Four-model comparison run the same day in clean sessions. RedLine Advisors primary observation.

5. **Google's stated position on heading tags.** John Mueller, Google Search Advocate, in Ask Google Webmasters and Office Hours sessions and on record since 2019 and reaffirmed since: a page ranks fine with no h1 or with five, and Google's systems work with the HTML as found, including pages using no semantic heading structure. Cited in section 04 specifically to establish that the h1 is a diagnostic proxy and not a ranking mechanism.

6. **SAP homepage variant observation.** July 13, 2026. Four model sessions plus direct browser observation, five screenshots archived. Cause of the variant split was not determined and is not claimed.

7. **llms.txt adoption and crawler behavior.** Vendor-published research, 2026. Directional only, not peer reviewed. Google's stated position that it doesn't use llms.txt as a search signal is on the record from Google search relations staff, July 2025.

8. **Hebner, S. (2026).** *The Corporate Graveyard of AI Invisibility: How to Survive the Great AI Marketplace Compression.* theCUBE Research, June 26, 2026. https://thecuberesearch.com/corporate-graveyard-ai-invisibility/ Source for the invisibility estimate, the Great Marketplace Compression, the AEO vendor count, and the measure-versus-diagnose distinction cited in section 05. Buyer behavior figures in that brief are attributed by its author to the G2 2026 AI Search Insight Report and are reproduced here on that attribution rather than from the G2 original.

9. **AEO Advantage Index.** theCUBE Research. https://thecuberesearch.com/ai-engine-optimization-advantage-index/ Referenced in section 06 as the layer-three instrument.

10. **Sinek, S. (2009).** *Start With Why: How Great Leaders Inspire Everyone to Take Action.* Portfolio. The why, how, what rubric referenced in sections 02 and 04. Used here as a diagnostic lens on the index findings, not as a description of RedLine methodology.

11. **30 percent average signal clarity improvement post-engagement** is RedLine client outcome data and is separate from all index measurements above.

Disclosure. Scott Hebner joined RedLine Advisors as an operating partner in July 2026. The AEO Advantage Index is his work, published through theCUBE Research, and RedLine has a commercial interest in his engagements. His research is cited here on its merits and readers should weigh the relationship accordingly.

RedLine Advisors · Park City, Utah · contact@redlinecs.ai · redlinecs.ai