RANKJUDGE

AI SEO Fact Check 2026

Eleven claims about AI search, robots.txt, llms.txt and AI Overviews, checked against the primary sources



Eleven AI SEO claims, checked against their sources. One survived intact.

FALSE	4	claims
MISREAD	2	claims
UNPROVEN	2	claims
NARROWER	2	claims
CONFIRMED	1	claim

Every claim below was checked by opening the document it rests on. Where we could not confirm something, the entry says so instead of guessing. Nothing here was tested by us: RankJudge runs no benchmarks and never claims to.

First edition · 25 October 2026 · Ron Kloth, founder of RankJudge

rankjudge.com

Free. No email address required. Quote it with attribution.

THE FINDINGS

Seven findings about AI crawlers, AI citations and SEO evidence

The short version, in the order they appear. Each one is set out in full later in this paper, with the primary source and the date somebody opened it.

1. Robots.txt governs what a crawler may request, not what is done with the page afterwards. Google supports four fields in it and silently discards every other line, so an invented no-AI-training rule does nothing.

2. Blocking Google-Extended does not keep you out of AI Overviews. Those are built from the ordinary search index and follow Googlebot.

3. llms.txt has been measured on its own three times, across three publishers, and reported anything from a 19.7% fall to an 81% rise. The values cross zero and there is no median.

4. Citation scatters where ranking concentrates. Across 1,050 citations, 85.3% of the domains cited appeared exactly once, and the ten most cited took 11.5% between them.

5. "AI cites Reddit" is a statement about one engine. Every Reddit and YouTube citation in that dataset came from Gemini. ChatGPT cited neither once in 216 searches.

6. Roughly 63% of every controlled SEO test in a 9,249-document review comes from a single company, and 51% of the tactics published in that review rest on a source that changed more than one thing at once.

7. AI Overviews cut clicks and hand them to whoever is cited: a 58% drop for position one when uncited, a 120% rise when cited, and 99% of what gets cited was already in the top ten.

SUMMARY

What this fact check covers, and who should read it

Eleven things the AI SEO industry repeats as fact. We opened the document behind each one and read it. Four turned out to be false. Two misread the source they cite. Two have been measured, and the measurements show nothing. Two are true of far less than they claim. One holds up, and even that one usually gets quoted with the wrong half of the finding.

This is not a tactics list. It is the part that saves you money: what you can stop doing, stop repeating, and stop worrying about, so the hours and the budget go somewhere they work. Every entry names the source, the date we opened it, and what we still could not confirm. Nothing in it was tested by us, because we do not test anything and never say we do.

Read this before you

- **Edit robots.txt to keep AI out of your content**
  Three of the four false claims here are about that one file. It does less than almost every post about it says.
- **Spend an afternoon adding llms.txt**
  It has been measured on its own three times. The results run from a 19.7% fall to an 81% rise, and cross zero.
- **Shortlist a tool to track your brand in AI answers**
  The one specification that separates a tracker which can see what is happening from one sampling too thinly to notice.
- **Quote an SEO case study in a pitch, a post or a board paper**
  The base rate for a published SEO result, and the single question that disqualifies half of them.
- **Compare software prices, for a client stack or one purchase**
  What a vendor pricing page does with the billing toggle, and the 13 corrections it cost us to find out.

A NOTE FROM RON KLOTH

> Taking eleven claims apart could read as taking people down, so let me be clear that it is not meant that way. The work quoted and checked on these pages is among the best this community produces. Krešimir Ćorluka gave away a summer of his own work for nothing. SearchPilot publishes the tests that failed, which almost nobody does. David Quaid published the raw data behind his study and not just the conclusions drawn from it. Those are not the habits of people trying to sell you something.
>
> Reading them properly did change how I think about one thing, and it is the reason this site exists. Every tool you pay for is optimising for something, and that something was chosen by a person with a view about how ranking works. If you do not know what your tool is trying to rank for, you cannot tell whether it is working. You can only tell whether the dashboard is green. That question sits underneath all eleven of these claims.
>
> *Ron Kloth, founder of RankJudge*

The prices behind claim 11, and current figures for all 17 tools we track, are at rankjudge.com, each one read off the vendor’s own page on a stated date.



THE METHOD

How to check an SEO or AI search claim in four questions

None of this is sophisticated. It is four questions, and almost nothing in circulation survives all four. That is the finding, not the method.

1

Name the organisation, not the article that mentioned it

Most SEO figures circulate as a number attached to a blog post that cites another blog post. Open the document the number came from. If you cannot reach it, the figure does not go on the page.

2

Date it, because a true figure goes stale quietly

A price, a crawler name and an AI engine's behaviour all move without announcement. A figure with no reading date is a claim about an unspecified past.

3

Ask what else changed

Half the published evidence in this field changed more than one thing at once. When it did, the direction is worth something and the size is worth nothing.

4

Ask for the route to the metric

When two credible tests disagree, do not take the bigger number, the newer test or the better-known publisher. Ask whether anybody can name the mechanism by which the change could reach the thing measured. A measured effect with no known mechanism is an open question, not a false claim, and that is how it should be written up.

WHAT EACH VERDICT MEANS

FALSE	The primary source contradicts the claim.
MISREAD	The source is real and says something different from what is quoted from it.
UNPROVEN	Measured, and the results do not establish an effect.
NARROWER	True of less than the claim covers.
CONFIRMED	The source says what is claimed, and the frame matters.



THE CLAIMS

Eleven claims about AI search and SEO, opened at the source

Each entry gives the claim as it circulates, what the primary source actually says, and what we could not confirm. The last line matters most: it is the part nobody else publishes.

01 Blocking AI crawlers in robots.txt stops AI companies training on your content.

FALSE

WHAT THE SOURCE SAYS

Gary Illyes co-wrote the standard that robots.txt runs on. He calls it an access preference protocol, and says it “doesn’t dictate what a system does with the data after it has already crawled it.” The file decides what a crawler may ask for. It has nothing to say about what happens next.

WHAT IS TRUE INSTEAD

Block GPTBot today and OpenAI’s crawler stops fetching your pages tomorrow. That is the whole of what you have done. It does not touch a page already collected, a copy of yours sitting on someone else’s site, or a dataset bought from a third party. You have fitted a filter, not issued a recall.

> **WHAT WE COULD NOT CONFIRM**
> *Whether any crawler drops what it already holds once a site starts blocking it. No operator publishes that, so nobody outside those companies knows.*

Source: Google Search Central video, published 21 October 2026

02 Add a “no AI training” line to robots.txt to opt out.

FALSE

WHAT THE SOURCE SAYS

Google reads four things in a robots.txt file: user-agent, allow, disallow and sitemap. Its own specification says invalid lines are ignored, and that anything other than allow, disallow and user-agent never reaches the parser.

WHAT IS TRUE INSTEAD

So the rule you invented, or copied from a post recommending one, is read and thrown away. You have not opted out of anything. You have written a comment. Something real is coming: draft-ietf-aipref-attach-05, by Gary Illyes of Google and Martin Thomson of Mozilla, dated 19 August 2026, would add a Content-Usage rule and a matching HTTP header. It is a draft rather than a standard, and it expires on 20 February 2027.

> **WHAT WE COULD NOT CONFIRM**
> *Whether crawlers outside Google treat an unsupported line as meaningful. The standard lets each one decide for itself, and none of them publish the answer.*

Source: Google robots.txt specification, read 24 October 2026; IETF Datatracker, read 24 October 2026

03 Blocking Google-Extended keeps you out of AI Overviews.

FALSE

WHAT THE SOURCE SAYS

Google's own crawler documentation is blunt about this. Google-Extended has no user agent string of its own, and the token "is used in a control capacity". It decides whether crawled content can train Gemini and ground answers in Gemini Apps and Vertex AI. The same page says it does not affect your inclusion in Google Search and is not a ranking signal.

WHAT IS TRUE INSTEAD

AI Overviews are built from the ordinary search index, which follows Googlebot. Google-Extended is the wrong lever. It is also the one place where a usage preference in robots.txt genuinely works, and it works because Google chose to honour it, not because the file can force anything.

> **WHAT WE COULD NOT CONFIRM**
> *Nothing. This one is stated plainly on Google's own documentation, which is rarer than it should be.*

Source: Google list of common crawlers, last updated 15 October 2026, read 24 October 2026

04 Adding llms.txt improves your visibility in AI answers.

UNPROVEN

WHAT THE SOURCE SAYS

The largest review of published SEO evidence found three independent tests, from three different publishers, of adding llms.txt on its own. They reported everything from a 19.7% fall to an 81% rise. The review gives no median, because the values cross zero.

WHAT IS TRUE INSTEAD

A fourth entry does show a clear gain. It also changed four things at once, one of which was unblocking AI crawlers, so it cannot tell you which one did the work. The review's own verdict is the useful part: no search engine consumes the file as a ranking input, and no major model provider has said it reads one. There is no route from the change to the number.

> **WHAT WE COULD NOT CONFIRM**
> *Whether any provider reads it quietly. No public commitment is not the same thing as no reading.*

Source: Krešimir Ćorluka, The SEO Meta Study, 10 September 2026

05 Information gain means adding something the top five results do not have.

MISREAD

WHAT THE SOURCE SAYS

The patent everyone cites says an information gain score “is indicative of additional information that is included in the given document beyond information contained in other documents that were already presented to the user.” Beyond what that reader has already been shown. Per person, per session.

WHAT IS TRUE INSTEAD

Not beyond the top five results. That reading is an interpretation somebody wrote down once and everybody repeated, and the cheat sheets claiming you “score zero” for covering the same ground as the top five are quoting a number the patent does not contain. Worth remembering too that a granted patent proves Google thought about something. It does not prove Google ships it.

> **WHAT WE COULD NOT CONFIRM**
> *Whether anything resembling this runs in Google Search today. Google has not said.*

Source: US 11,354,342 B2, “Contextual estimation of link information gain”, Google LLC, inventors Victor Carbune and Pedro Gonnet Anders, filed 18 October 2018, granted 7 June 2022

06 AI assistants love Reddit, so get mentioned there.

NARROWER

WHAT THE SOURCE SAYS

We counted the citations in a published dataset of 100 prompts put to ChatGPT, Claude and Gemini. reddit.com came back 20 times and youtube.com 22. Every single one of those citations came from Gemini.

WHAT IS TRUE INSTEAD

ChatGPT cited neither domain once, across 216 searches. So “AI cites Reddit” is a claim about one engine wearing the clothes of a claim about all of them. Which engine your buyers actually use decides whether any of this applies to you.

> **WHAT WE COULD NOT CONFIRM**
> *Whether it holds beyond these 100 prompts. A hundred prompts is evidence. It is not proof.*

Source: Dataset published by David Quaid, GEO Search Console, 7 October 2026. Counts are our own arithmetic over his file.

07 Here are the fifty websites AI cites most. Get on the list.

MISREAD

WHAT THE SOURCE SAYS

In the same dataset, 1,050 citations went to 784 different domains. 669 of those, which is 85.3% of them, were cited exactly once. The ten most cited took 11.5% between them, and the single biggest took 2.1%.

WHAT IS TRUE INSTEAD

Ranking concentrates. Citation scatters. A leaderboard describes 11.5% of what happened and reads as though it describes the lot. Across ten software prompts, G2 was cited once, Capterra and TrustRadius and Software Advice were not cited at all, and 109 domains nobody has heard of were.

> **WHAT WE COULD NOT CONFIRM**
> *Whether a citation brings a visit. Nothing in the dataset measures what happened next.*

Source: Dataset published by David Quaid, GEO Search Console, 7 October 2026. Counts are our own arithmetic over his file.

08 We changed one thing and traffic rose 40%.

UNPROVEN

WHAT THE SOURCE SAYS

A review of 9,249 collected documents, 1,391 of which passed screening as primary sources that measured something, found 602 of 769 scored tactics with no controlled test behind them at all. 138 rested on exactly one. Six rested on three or more. And of the 182 tactics it published, 92, which is 51%, rest on a source that changed something else at the same time.

WHAT IS TRUE INSTEAD

So half the published evidence in this field cannot tell you which change produced the number. The base rate is the other half of the problem: 81.9% of published outcomes in that corpus are positive, against 3.8% negative. The author's explanation is the honest one. "Agencies publish the studies that went well. Nobody writes up the quarter where the migration tanked."

> **WHAT WE COULD NOT CONFIRM**
> *The real failure rate of SEO work. The author says plainly that 3.8% is not it.*

Source: Krešimir Ćorluka, The SEO Meta Study, 10 September 2026

09 Controlled testing proves this tactic works.

NARROWER

WHAT THE SOURCE SAYS

We counted the evidence appendix of that review ourselves. 77 of the 123 source lines marked “Controlled A/B test” belong to one company, SearchPilot. That is roughly 63% of every controlled test in a 9,249-document review. The next biggest contributor has eight.

WHAT IS TRUE INSTEAD

They have earned the trust. 11 of their 67 entries carrying a figure are negative, which is 16% against 3.8% for the corpus, so they are over four times likelier than the field to publish something that failed. But their method splits pages rather than users, and that needs thousands of near-identical pages, which is why their named clients are enterprises. A result measured across tens of thousands of enterprise product pages is evidence about tens of thousands of enterprise product pages.

> WHAT WE COULD NOT CONFIRM
> *Whether any of it carries over to a site with dozens of pages. Nobody has tested that, because the test needs the pages.*

Source: Our count over the evidence appendix of The SEO Meta Study. SearchPilot methodology and client list from searchpilot.com, read 25 October 2026.

10 AI Overviews have destroyed organic clicks.

CONFIRMED

WHAT THE SOURCE SAYS

Ahrefs measured a 58% drop in click-through rate for position-one content when an AI Overview is present. Omnibound found AI Overviews on 48% of tracked queries, rising to 88% in healthcare. seoClarity found that 99% of Google AI Overviews source from the top ten web results.

WHAT IS TRUE INSTEAD

Here is the half that gets dropped. Seer Interactive measured a 120% rise in organic clicks when a page is cited in an AI Overview. The same feature that takes clicks from the uncited hands them to the cited, and 99% of what gets cited was already ranking. That is not an argument that ranking stopped mattering. It is an argument that it matters differently.

> WHAT WE COULD NOT CONFIRM
> *How these interact on any one site. Each figure comes from a different study with a different sample.*

Source: Ahrefs, Omnibound, seoClarity and Seer Interactive, as recorded on our benchmarks page, checked 5 September 2026

11 The price on the vendor’s pricing page is the price you will pay.

FALSE

WHAT THE SOURCE SAYS

Vendors lead with the annual rate and let you read it as monthly. In September 2026 we ran a routine re-check, opening our own records against the live pages. It produced 13 corrections to figures already published here, across 11 of the 16 tools we tracked at the time.

WHAT IS TRUE INSTEAD

Read the billing toggle both ways before you quote a number, and write down which one you read. A price with no billing basis attached is not a price. And a comparison that mixes the two will quietly reverse its own verdict.

> WHAT WE COULD NOT CONFIRM
> *Pricing for plans sold only through a sales call. We do not have it, and neither does anyone else publishing a comparison of it.*

Source: RankJudge price history, corrections published 30 September 2026

SCOPE AND LIMITS

What this fact check cannot tell you

Stated here rather than buried, because a paper that lists other people’s weaknesses and not its own has not earned the reader’s attention.

- We run no tests. RankJudge has no lab, no panel of sites and no benchmark suite, and nothing in this document was measured by us. What we do is open the source somebody else cited and read it.
- Four of the eleven entries rest in part on one review, and that review rests substantially on one testing company. We say so in the entries rather than hiding the concentration behind a source count.
- Two entries rest on our own arithmetic over a dataset of 100 prompts. One hundred prompts is evidence about one hundred prompts.
- This document is about which claims survive checking. It is not a tactics list and it will not tell you what to do on Tuesday.
- Everything here was true on the date printed beside it. Re-read anything you intend to rely on.

Credit: the SEO Meta Study, and the gap it leaves

Four entries here would not exist without Krešimir Ćorluka's The SEO Meta Study, published 10 September 2026: 9,249 documents collected, 1,391 screened as primary sources, 182 tactics published, and no email address required to download it. He paid for the pipeline himself and gave the result away, and the most useful parts of it are the limits he states about his own work rather than any tactic in it. This document exists because of one of those limits. His review maps what the field has published. It does not, and does not claim to, check whether the claims the field repeats survive contact with their sources. That gap is the whole of what is above.

> **Download his study, free and with no email required: ovisy.hr/blog/seo-meta-study**



SOURCES

Every source behind these eleven claims, named and dated

Thirteen entries. Open any of them. If one of these does not say what we report it saying, write to ron@rankjudge.com and the correction gets published with the date we made it.

robots.txt as an access preference protocol	Gary Illyes, Google Search Central · 21 October 2026
Supported robots.txt fields; invalid lines ignored	Google robots.txt specification · read 24 October 2026
Google-Extended is a control token with no user agent	Google list of common crawlers · updated 15 October 2026
Content-Usage proposal, draft-ietf-aipref-attach-05	Gary Illyes (Google) and Martin Thomson (Mozilla), IETF · dated 19 August 2026
Information gain, measured against documents already presented to the user	US 11,354,342 B2, Google LLC, Carbune and Gonnet Anders · granted 7 June 2022
llms.txt tests; tactic evidence gates; publication bias; bundling	Krešimir Ćorluka, The SEO Meta Study · 10 September 2026
Controlled A/B testing methodology, client list and pricing	SearchPilot · read 25 October 2026
Citation counts across ChatGPT, Claude and Gemini over 100 prompts	Dataset by David Quaid, GEO Search Console; counts our own · 7 October 2026
Position-one CTR drop when an AI Overview is present	Ahrefs · checked 5 September 2026
Organic click increase when cited in an AI Overview	Seer Interactive · checked 5 September 2026
Share of AI Overviews sourcing from the top ten results	seoClarity · checked 5 September 2026
Prevalence of AI Overviews across industries	Omnibound · checked 5 September 2026
Thirteen corrections across eleven tools	RankJudge price history · 30 September 2026



ABOUT THE PUBLISHER

Who publishes RankJudge, and how it is paid for

RankJudge is an independent comparison site for AI SEO and AEO software, founded in 2026. It tracks 17 tools and publishes 12 head-to-head comparisons. Every price on it was read off the vendor's own pricing page and is published with the date it was read and whether it is the monthly or the annual rate.

It runs no benchmarks and no product testing of any kind, and it never claims to. What it does is open the vendor's page, open the research somebody cited, and write down what is actually there. That is a smaller claim than most sites in this category make, and it is one we can stand behind on every page.

It is funded by affiliate links on its comparison pages. Those links never move a verdict, and several tools we rate well pay us nothing at all. The disclosure sits under the buttons and at rankjudge.com/affiliate-disclosure rather than being buried.

It collects nothing from the people who read it. No newsletter, no contact form, no email gate, including on this document. And it corrects itself in public: one routine re-check in September 2026 produced 13 corrections to figures already published on the site, across 11 of the 17 tools tracked. Those are listed with dates at rankjudge.com/price-history.

Founded	**2026**
Tools tracked	**17**
Comparisons published	**12**
Tests run by us	**None, ever**
Self-corrections published	**13 in September 2026**



ABOUT THE AUTHOR OF THIS FACT CHECK

Ron Kloth

Founder and sole author, RankJudge

Twenty years in financial services, seventeen running a local business, and two running Digital Papercuts, an SEO agency of his own. One person, writing under his own name.

The byline is on every page of RankJudge for a specific reason: the verdicts name losers, and a recommendation nobody signs is one nobody can be held to. The same applies to this document. Every judgement in it is his, every figure in it was checked before it was printed, and any correction will be published under the same name with the date it was made.

ron@rankjudge.com · linkedin.com/in/ron-kloth-970a1613 · rankjudge.com/about