GoGoChimp

GOGOCHIMP RESEARCH · REGIONAL STUDY · OCTOBER 2026

How citable are the North's top 50 digital agencies to AI search?

An October 2026 study of Prolific North's Top 50 Digital Agencies, scored for how likely AI engines are to quote them.

42
agencies fully crawled and scored

65.5
median Rubric score

48
median Trusted score, the weakest pillar

12 of 42
scored 70 or above

Chris McCarron
Founder, GoGoChimp

October 2026 · gogochimp.com

CONTENTS

What's in this paper

01	The median northern digital agency scores 65.5 out of 100
02	Trusted is the weakest pillar, with a median of 48
03	Five checks fail at more than half of the agencies
04	No agency blocks AI crawlers in robots.txt
05	Homepages outscore the rest of the site at 37 of 42 agencies
06	One site serves schema that crawlers without JavaScript can't see
07	Fifteen of the 42 agencies don't declare any sameAs links
08	Northern agencies score 4.5 points below a 415-agency benchmark
09	Who scored highest and lowest, and why?
10	Why weren't five of the agencies scored?
11	How did we crawl and score the sites?
12	How can you test your own site against the same checks?
·	References

How to read this paper. Agency figures come from GoGoChimp's Rubric crawl of 6 October 2026. Any other source is named where it's used. Rubric scores pages from 0 to 100, and 70 is the line at which a page counts as quotable. Rubric predicts citability, it doesn't count citations. Top performers are named, positively. No other agency is named. This study is independent. It isn't affiliated with or endorsed by Prolific North or Mustard Research, which compiled the list.

Forty-two of Prolific North's Top 50 digital agencies scored a median 65.5 out of 100 for AI citability. Trusted was the weakest pillar, at 48.

Crawled: 6 October 2026 · **Sample:** Prolific North Top 50 Digital Agencies 2025 · **Tool:** Rubric, scoring version 2026-10-01 · **By:** Chris McCarron, GoGoChimp

Only 12 of the 42 agencies we could fully crawl score 70 or above. If you run any of the 42, the technical basics aren't your problem. None of them failed Rubric's HTTP, indexability, canonical or robots.txt checks.

What's missing is markup and attribution. At the median agency, 94% of pages fail Rubric's content-type schema check and 79% fail its named-author check.

MAIN FINDINGS

01 The median score is 65.5 and the mean 66.6. Twenty-seven of the 42 agencies sit between 60 and 69. The top score is 80, shared by two agencies.

02 Trusted is the weakest pillar, with a median of 48. Twenty-two agencies score under 50 on it and only 2 reach 70.

03 Five checks fail on most pages at 25 or more of the 42 agencies. Content-type schema leads (32 agencies), followed by question-shaped headings (29).

04 The homepage beats the median inner page at 37 of 42 agencies, by a median of 12.5 points.

05 None of the 42 blocks an AI crawler in robots.txt. Three of the 47 sites we tried served our crawler a bot-challenge page on most URLs.

06 Against all 415 agency sites in the same crawl, the North scores 4.5 points lower. Its median is 65.5 against 70, and its Trusted median 48 against 52.

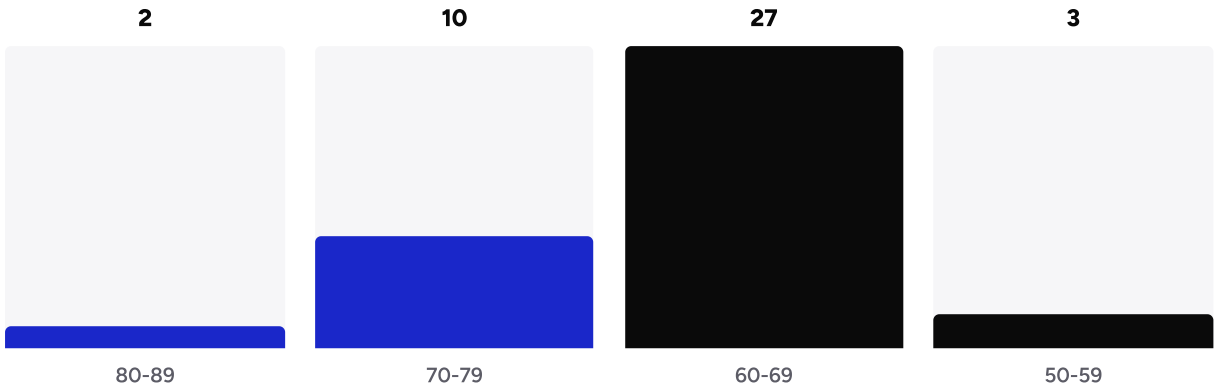
We name the top performers only. Any agency on the list can ask us for its own page-level results.

This study is independent. It isn't affiliated with or endorsed by Prolific North or Mustard Research, which compiled the list.

01 The median northern digital agency scores 65.5 out of 100

Across the 42 agencies we could fully crawl, the median score is 65.5 and the mean 66.6. The middle half sits between 64 and 70, so there isn't much spread. We crawled up to 40 pages on each agency's site and scored every page against Rubric's 44 citability checks.

Northern agencies by score band (n=42)



12 of the 42 fully crawled agencies scored 70 or above. Source: GoGoChimp Rubric crawl of 6 October 2026, scoring version 2026-10-01, up to 40 pages per site.

MEASURE	SCORE
Median	65.5
Mean	66.6
Middle half (interquartile range)	64 to 70
Highest	80
Lowest (fully crawled)	54

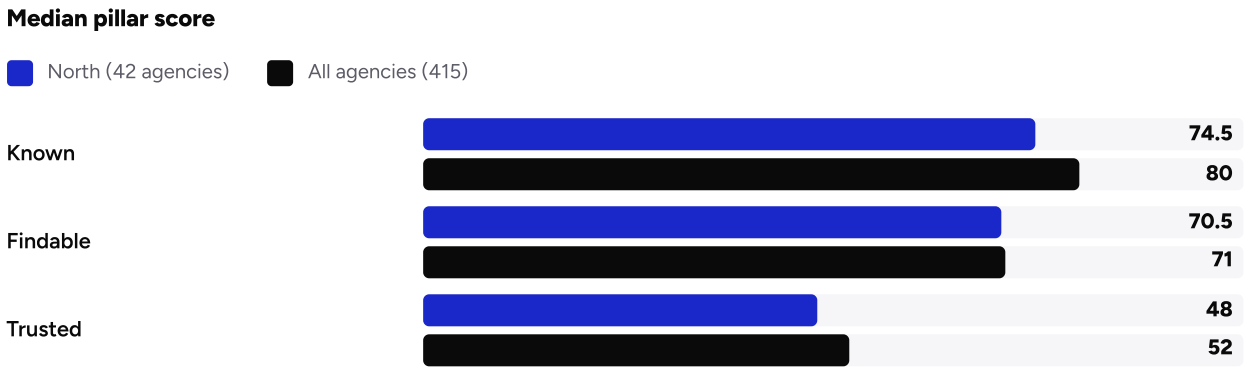
SCORE BAND	AGENCIES
80 to 89	2
70 to 79	10
60 to 69	27
50 to 59	3

All 42 clear the technical bar, then most of them down tools.

> *Twelve of the 42 agencies score 70 or above, and only two reach 80: The SEO Works and Embryo.*

02 Trusted is the weakest pillar, with a median of 48

Trusted is the lowest of Rubric's three pillars, at a median of 48. Known asks whether an engine can tell what a page is and who it belongs to. Findable asks whether bots can reach a page and lift an answer from it. Trusted asks whether an engine has a reason to believe the page.



Only 2 of the 42 northern agencies scored 70 or above on Trusted. Source: GoGoChimp Rubric crawl of 6 October 2026, scoring version 2026-10-01, up to 40 pages per site.

PILLAR	MEDIAN	AGENCIES SCORING 70+
Known	74.5	28 of 42
Findable	70.5	26 of 42
Trusted	48	2 of 42

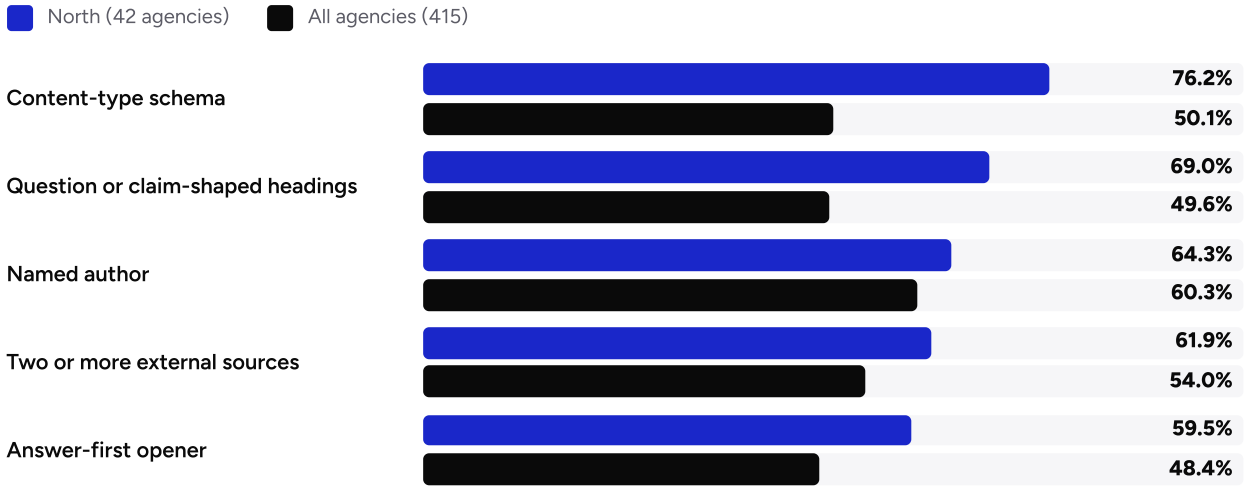
Twenty-two of the 42 score under 50 on Trusted. Only Rise at Seven (76) and Embryo (72) clear 70. The two Trusted checks that fail most widely are named authors (27 of the 42 agencies) and external source links (26).

> *Trusted is the pillar where the North's top digital agencies score lowest: a median of 48, against 74.5 for Known and 70.5 for Findable.*

03 Five checks fail at more than half of the agencies

Five checks fail at 25 or more of the 42 agencies. We count a check as failed at an agency when it fails on more than half of the pages where it applies. The next two aren't close: entity clarity and the H1 snippet-window check fail at 13 agencies each.

Share of agencies failing the check on most pages



North counts: 32, 29, 27, 26 and 25 of 42. Source: GoGoChimp Rubric crawl of 6 October 2026, scoring version 2026-10-01, up to 40 pages per site.

CHECK	AGENCIES FAILING ON MOST PAGES	MEDIAN SHARE OF PAGES FAILING
Content-type schema (Article, Service and so on)	32 of 42	94%
Question- or claim-shaped H2 and H3 headings	29 of 42	62%
Named author on the page	27 of 42	79%
Two or more external source links	26 of 42	61%
Answer-first opener (40-70 words)	25 of 42	56%

Content-type schema is the biggest single gap. Without it, an engine has to work out for itself whether it's reading a service page, a case study or a blog post.

Schema isn't a proven citation lever, though. Ahrefs compared 1,885 pages that added JSON-LD with 4,000 control pages. ChatGPT citations moved +2.2% and AI Mode +2.4%, both "statistically indistinguishable from zero". AI Overview citations fell 4.6% (Ahrefs, 2026).

Schema still tells an engine what a page is, and it's cheap to add.

Named authors fail on 79% of pages at the median agency, and external source links on 61%. A page that doesn't link to a single source gives an engine nothing to check its claims against.

Headings and openers are the content-shape problems, failing on 62% and 56% of pages at the median agency. A page that fails the opener check doesn't start with a direct answer of 40-70 words. One that fails the heading check labels its sections instead of stating a claim or asking the question a buyer would type.

If I ran one of these sites, I'd fix authors before schema. A byline is a template change, and it clears one of the two Trusted checks that fail most widely in this sample.

> *At the median agency, 94% of pages carry no content-type schema and 79% name no author. Neither fix needs a redesign.*

04 No agency blocks AI crawlers in robots.txt

None of the 42 fully crawled sites disallows an AI crawler in robots.txt. Rubric checks the file against 15 crawler user agents, including GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot and Google-Extended.

Eleven of the 42 publish an llms.txt file. We count a site if the file exists at all, so the 11 include two files with no markdown links. Rubric records llms.txt but doesn't score it. Google says a site doesn't need one to appear in AI Overviews (Google, 2025).

05 Homepages outscore the rest of the site at 37 of 42 agencies

The homepage scored higher than the median inner page at 37 of the 42 agencies. The median homepage scored 79.5 and the median inner page 65.5. Worked out site by site, the median gap is 12.5 points.

A homepage-only check is out by that much, because it sees one page of a 40-page crawl. The other 39 are where an agency's service and case-study pages live.

> *At 37 of 42 agencies, the homepage beat the median inner page, by a median of 12.5 points. A homepage-only check flatters the site.*

06 One site serves schema that crawlers without JavaScript can't see

One of the 42 agencies had Organization schema that appeared only after JavaScript ran, on 27 of its 40 crawled pages. Crawlers that read raw HTML miss it on most of the site.

In Vercel's 2024 analysis, OpenAI's, Anthropic's and Perplexity's crawlers didn't run JavaScript (Vercel, 2024). Google renders JavaScript, so Google does see script-added schema. The schema is there for any visitor who brings a browser, and missing for the AI crawlers that don't.

Rubric catches JavaScript-only schema by comparing each page's raw HTML with a Chrome-rendered copy. Across all 415 agencies in the crawl, 44 (10.6%) had at least one page like this. At 1 of 42 (2.4%), the North is cleaner than the wider benchmark here. If you're not sure where your own markup sits, check whether your schema needs JavaScript before assuming the AI crawlers see it.

07 Fifteen of the 42 agencies don't declare any sameAs links

Fifteen of the 42 sites don't declare any sameAs links, and the median site declares 2. None of the 42 links to a Wikidata entry.

The sameAs property ties a brand to its profiles on LinkedIn, Companies House, Wikipedia or review sites. Those links let an engine confirm the agency is the one it's read about elsewhere.

08 Northern agencies score 4.5 points below a 415-agency benchmark

The North's top digital agencies score a median of 65.5, 4.5 points below the 415-agency median of 70. Their Trusted median is 48, against the benchmark's 52. Every site ran in the same 6 October crawl, on the same Rubric engine and 40-page cap, so the scores compare directly.

We attempted 484 agencies from five lists and scored 415. Three of the lists date from August: a seven-country list on 12-13 August, then English-speaking and international lists on 23 August. The other two are 10 UK SEO agencies and this Prolific North list. The 42 northern agencies are part of the 415, and leaving them out wouldn't shrink the 4.5-point gap.

MEASURE	NORTH (42 AGENCIES)	ALL AGENCIES (415)
Median score	65.5	70
Scoring 70 or above	12 (28.6%)	222 (53.5%)
Known median	74.5	80
Findable median	70.5	71
Trusted median	48	52
Trusted under 50	22 (52.4%)	168 (40.5%)
Trusted 70 or above	2 (4.8%)	61 (14.7%)

The North has the lowest median of the five lists. It trails furthest on Known, by 5.5 points, then on Trusted by 4, and it's half a point behind on Findable. Entity links point the same way: 15 of the 42 northern agencies (35.7%) declare no sameAs links, against 92 of the 415 (22.2%).

For wider context, we crawled 244 other websites the next day with the same engine and page cap: online shops, blogs and publishers, SaaS companies and other businesses. Their median was 66 and their Trusted median 49. So the North's top digital agencies score about the same as the average website in that set, 65.5 against 66, while agencies as a whole sit four points above it.

Both groups share the same five most common failures, in a different order.

CHECK FAILING ON MOST PAGES	NORTH	ALL AGENCIES
Content-type schema	32 of 42 (76.2%)	208 of 415 (50.1%)
Question- or claim-shaped headings	29 of 42 (69.0%)	205 of 413 (49.6%)
Named author	27 of 42 (64.3%)	249 of 413 (60.3%)
Two or more external source links	26 of 42 (61.9%)	224 of 415 (54.0%)
Answer-first opener	25 of 42 (59.5%)	200 of 413 (48.4%)

Content-type schema shows the widest gap, at 26 percentage points. Named authors, the most common failure across the 415, are within 4 points.

> *In the same 6 October crawl, 415 agency sites have a median of 70 out of 100. The North's top digital agencies score 65.5, and their Trusted median is 48 against 52.*

The North's other patterns hold across all 415 agencies

Trusted is the weakest pillar on all five lists in the crawl. The North's Trusted median is the second lowest of the five, above the 10 UK SEO agencies at 46.5.

The homepage beats the median inner page at 37 of the 42 northern agencies (88.1%), by a median of 12.5 points. Across the 415, it's 373 sites (89.9%) and 12 points.

Bot challenges run at a similar rate: 3 of the 47 northern sites (6.4%), against 25 of the 484 agencies (5.2%).

09 Who scored highest and lowest, and why?

The SEO Works and Embryo lead on 80, while Rise at Seven has the best Trusted score. The three lowest scorers lose points for the same reasons the median agency does. The scores reflect page markup and structure on the crawl date, and they don't judge client work.

The SEO Works (80) scores 91 on Known. It names an author on 23 of 30 inner pages where the check applies, and its statistics carry a source on 24 of 30. The median agency fails the author check on 79% of pages.

Embryo (80) has the second-highest Trusted score in the sample, at 72. External sources and sourced statistics both passed on all 18 of its inner pages where those checks applied.

Rise at Seven (73) has the highest Trusted score in the sample, at 76. It has external sources on all 38 inner pages where the check applied.

The three lowest scorers sit at 54, 55 and 59, with Trusted scores of 38, 42 and 48. We haven't named them, and we're sending each agency its own results privately.

One of the three describes its work as software development, so some content-led checks fit it less naturally than they'd fit an SEO agency. For all three, the gaps are markup and content fixes that don't need a rebuild.

10 Why weren't five of the agencies scored?

Three of the five served our crawler a bot-challenge page on most URLs, and the other two exposed only their homepage. We didn't rank these five, because a score built on challenge pages or a single homepage would misrepresent the site. A failed crawl is still a data point, so we report them rather than drop them.

SITE	WHAT HAPPENED
1	A Cloudflare "Just a moment" challenge page on 39 of 40 URLs
2	A Cloudflare "Attention Required" page on all 40 URLs
3	A NinjaFirewall "403 Forbidden" page on all 40 URLs
4	Only the homepage was crawled, although it carries 38 internal links. The report doesn't record why
5	Only the homepage was reachable. Its main content runs to 35 words, the page links to nothing else on the site and Rubric found no XML sitemap, so the crawler had nothing to follow

Our crawler sends a standard Chrome user agent with browser headers. From outside, we can't tell whether verified AI crawlers get the same challenge.

If you run one of the three challenged sites, check your firewall's bot rules. A challenge page served to an AI crawler is a page that can't be cited.

We excluded three more agencies before crawling, because their listed domains now redirect to different brands.

11 How did we crawl and score the sites?

- **Sample.** The Prolific North Top 50 Digital Agencies 2025, published 16 June 2025. Mustard Research compiled it from turnover, headcount and pre-tax profit (Prolific North, 2025). It ranks agencies headquartered in the North of England, so the sample doesn't stand for the UK as a whole. It's the latest list Prolific North had published when we crawled.
- **Domains.** We confirmed each agency's website by fetching it and checking the page title. Where the obvious domain was parked or redirected, we used the live site.
- **Crawl.** We ran Rubric's audit engine locally rather than through the hosted Rubric service. It crawled up to 40 pages per site on 6 October 2026, with the broken-link check off, as part of the 484-agency benchmark crawl. We attempted 47 sites and fully crawled 42.
- **Scoring.** Rubric scoring version 2026-10-01: 44 checks across three pillars (Known, Findable, Trusted), rolled into a 0-100 score. Rubric also reports per-engine fits for ChatGPT, Perplexity, AI Overviews, Gemini, Copilot and Claude. They're modelled from Rubric's weightings rather than observed from any engine.
- **Version.** Scores from version 2026-10-01 aren't comparable with our August crawls, which ran on an older engine. The current engine judges named authors and content-type schema more strictly. So the benchmark comparison above uses the 6 October crawl of the August lists rather than their August scores.

- **Stability.** We first crawled these 47 sites on 5-6 October, on the same engine and page cap. The 6 October crawl, a day later, gave the same median (65.5) and Trusted median (48), and the median site moved 0 points.
- **What the score is not.** Rubric predicts citability, it doesn't count citations. Nor does it judge the quality of an agency's client work.

12 How can you test your own site against the same checks?

To test your own site, run a Rubric audit and start with the five checks in the failures table above. Compare your failing checks with ours rather than your headline score, because a different page count or scoring version can shift it.

References

- GoGoChimp. (2026, 6 October). *Per-site Rubric scores for the Prolific North Top 50 Digital Agencies 2025* (first-party crawl data, 6 October 2026). Not published: each listed agency can request its own results.
- GoGoChimp. (2026, 6 October). *Agency AI-citability benchmark* (first-party Rubric crawl of 484 agency sites, including this list, on scoring version 2026-10-01, up to 40 pages per site, 6 October 2026, 415 scored). Summary figures only. Per-site results aren't published.
- Google. (2025, updated 10 December). *AI features and your website* (Google Search Central). https://developers.google.com/search/docs/appearance/ai-features
- Linehan, L., with data by Guan, X. (2026, 11 May). Ahrefs study of JSON-LD schema and AI citations (1,885 pages that added schema, 4,000 control pages, difference-in-differences). Ahrefs. https://ahrefs.com/blog/schema-ai-citations/
- Prolific North. (2025, 16 June). *The Prolific North Top 50 Digital Agencies*, compiled by Mustard Research. https://www.prolificnorth.co.uk/insight/rankings/the-prolific-north-top-50-digital-agencies/
- Zecchini, G., Moore, A. A., Ubl, M., & Siddle, R. (2024, 17 December). Vercel and MERJ study of AI crawler behaviour (server logs on nextjs.org and Vercel's network). Vercel. https://vercel.com/blog/the-rise-of-the-ai-crawler

GoGoChimp

ABOUT THE AUTHOR

Chris McCarron founded GoGoChimp in Glasgow in 2013 and has 13 years of conversion rate optimisation experience. He was previously Head of Growth at Roadtrippers and FOMO, and he wrote *CITED*, a book on getting your business recommended by AI search.

Chris McCarron on LinkedIn

ABOUT RUBRIC

Rubric is GoGoChimp's AI-citability auditor. It crawls a site the way AI crawlers read it and runs 44 checks across three pillars, Known, Findable and Trusted, weighted for six AI engines: ChatGPT, Perplexity, Google AI Overviews, Gemini, Microsoft Copilot and Claude.

Rubric predicts citability, it doesn't count citations. 70 is the quotable line.

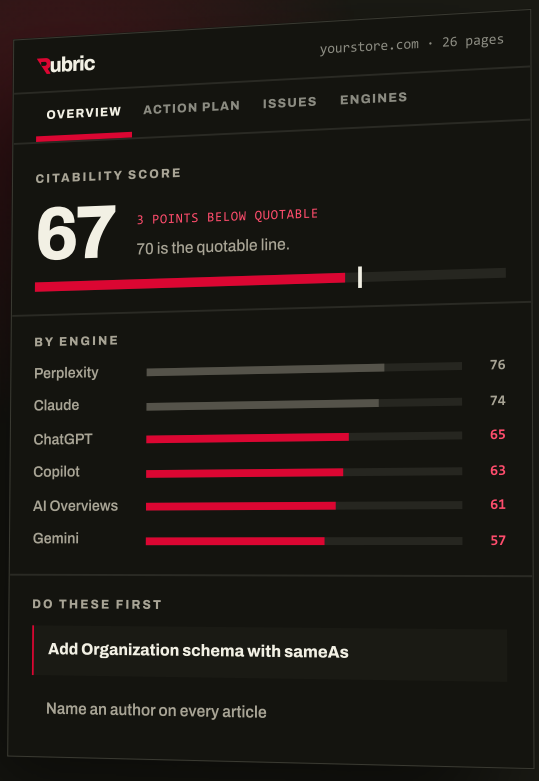
Rubric

FREE AI SEO AUDIT

See how your own site scores.

The Rubric audit is free and needs no sign-up.

RUN A FREE RUBRIC AUDIT



© 2026 GoGoChimp · Glasgow, Scotland