

vercatus . com

ANALYSED	PAGES EXAMINED	TEMPLATES	MACHINE IDENTITIES
16 September 2026	40 of 40 listed	5	9
INTENTS SIMULATED			
314			

What machines see

Googlebot sees a small, fast, fully indexable static site: 40 URLs in one sitemap, clean canonicals, TTFB around 35 to 48 ms, lab LCP comfortably inside the good band on every template except the homepage (1,296 ms, still good) and no render-blocking beyond one Google Fonts stylesheet. What it does not see is differentiation between pages: the link graph is navigation-flat, the same hubs receive inbound links from nearly every page, and no page accumulates the internal signal that would let Google decide which URL owns "pricing", "AI crawler access" or the brand name itself. Structured data stops at the marketing shell: the homepage carries a rich graph (Organization, WebSite, SoftwareApplication, Offer, FAQPage) but 28 docs pages emit nothing, and no @id links the graph together, so entity understanding does not propagate past the front door. The non-rendering AI crawlers (GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot) are allowed and get almost everything, with one caveat: 10 to 15 percent of the text on the homepage and the commercial templates appears only after JavaScript, which on the site's single most-cited page is a self-inflicted version of the exact problem the product diagnoses. Retrievers chunk this site into many short, topically overlapping passages: the docs guides template averages 238 words, two pages fall below 120 words, and the same claims about plans, crawler behaviour and retention are restated on four or more URLs, which is why the simulation shows near-total cannibalisation and why rerankers keep surfacing a legal or contact page for queries that should land on a definition page. Answer engines that do reach a passage find high-specificity numbers (the 78 percent JavaScript-added gallery figure, the worked pricing examples, the self-hosting cost benchmarks) but usually without a date, an author or a source, so a grader has no provenance to weigh; live-retrieval engines such as Perplexity and ChatGPT search cannot judge freshness on 34 of 40 pages at all. Agent-class clients find llms.txt and

nothing else: no feed, no api-catalog, no OpenAPI descriptor, no security.txt, which is a gap for a product whose buyers include engineers integrating a REST API.

AI-crawler requests served

100%

175 of 175 probes

Intents fully answered

71%

223 of 314

Entities with a defining page

8%

158 entities found

Findings

40

0 critical · 11 high · 30 rejected in review

For the board the same findings, for a reader who will not open the rest of this report

AI assistants can read every page they asked for , yet the site answers only 60 percent of buying questions .

Access is not the problem. Across 40 pages sampled from 5 page types, AI assistants such as ChatGPT, Claude, Perplexity and Google's AI read 100 percent of the pages they requested (175 of 175). The gap is in what the pages say. Of 314 questions readers ask in this field, the site fully answers 71 percent, and of 62 questions from people ready to buy, it fully answers 60 percent. Only 33 percent of the site's factual statements carry a source, figure or quote (93 of 285), which limits how confidently anyone, human or machine, can repeat them.

This is the first measurement, so there is no earlier run to compare against. It sets the baseline for future runs.

1. Approve one official price and buying page, and sign off the numbers once. · CEO with Head of Marketing

Plan prices and what each plan includes currently disagree across four pages, and the way to start a free report, book a demo or sign up is never stated. This is the shortest path between an AI answer and revenue.

2. Decide to publish definitive pages on our own subject, including honest competitor comparisons. · Head of Content

The terms and identities our product exists to measure have no defining page, and every question naming a competitor returns nothing. These are the questions answer engines are asked most often in this category.

3. Decide how much of our own research and evidence we publish, and under whose name.

Head of Editorial with CTO

Only a third of our statements carry a source, no page names an author and almost none show a date. Our strongest original measurement was published once and left incomplete.

These figures measure how much of the site machines can reach, understand and cite, not visitor numbers or sales, until analytics are connected.

Summary

Vercatus is machine-readable: all 40 sampled pages return 200, all are indexable, robots.txt disallows no AI crawler, and the five templates serve their substance in raw HTML. Access is therefore not the constraint, which is unusual and means every remaining point of leverage sits in content, entity definition and commercial clarity. The largest losses are commercial: plan facts disagree across /, /pricing/, /docs/reference/plans/ and /terms/ (Portfolio is both \$2,500/month and \$2,500 plus per-site fees; Enterprise includes ten sites in one place and is priced on request in another), no page carries Offer markup, and every transactional intent (free report, demo, signup) resolves only partially because the signup path, trial terms and booking link are never stated. The second loss is topical ownership: 174 of 194 informational intents are cannibalised, even the brand query "Vercatus" retrieves /terms/ and /contact/ before any page that defines the company, and the crawler identities the whole product is about (GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Google-Extended) have no defining page anywhere on the site. The third is a demand gap: category vocabulary that buyers actually type (GEO, query fan-out, llms.txt, AI Overviews, schema for AI) is unanswered, and every competitor-named comparison intent returns nothing, so the site is absent from exactly the "best tool" and "X vs Y" questions that answer engines serve most often for software buyers. Provenance is thin for a site whose product is evidence grading: zero of 40 pages name an author and only 6 show a visible date, on a subject that changes quarterly. The brand is still a string rather than a thing: the Organization node has no sameAs and no Wikidata item, and 28 docs pages emit no JSON-LD at all. The fastest wins, in order, are the transactional path and canonical pricing (days of work, directly revenue-linked), then defining passages for the crawler tokens, then comparison and category-vocabulary content built from the 138-entry knowledge base that is currently described but not published.

Machine access

Every sampled page was requested as each identity below and the responses compared. A cell shows pages served over pages requested; anything short of the full count means that identity was refused, challenged or served an error. Identity probes run on a per-template sample, so counts differ by column.

TEMPLATE	PAGES	PLAIN HTTP	BROWSER	BROWSER, NO JS	GOOGLEBOT	GPTBOT
/	1	1/1	1/1	1/1	1/1	1/1
/blog/ai-crawlers-do-not-run-javascript	2	2/2	2/2	2/2	2/2	2/2
/docs/guides/{slug}	14	14/14	14/14	14/14	11/11	11/11
/docs/concepts/{slug}	14	14/14	14/14	14/14	12/12	12/12
/about	9	9/9	9/9	9/9	9/9	9/9

■ served
 ■ partly refused
 ■ refused on every page
 robots.txt disallows no AI crawler

Performance by template

Lab figures only; no real-user data was available for this origin. Lab numbers compare templates with each other; they are not what a visitor sees.

TEMPLATE	PAGES	LCP	CLS	BLOCKING	TTFB	WEIGHT	RENDER-BLOCKING	REAL USERS · LCP
/docs/concepts/{slug}	14	712 ms	0.0089	61 ms	37 ms	56 KB	1 · 1 KB	-
/docs/guides/{slug}	14	592 ms	0.4246	34 ms	36 ms	55 KB	1 · 1 KB	-
/about	9	512 ms	0.0173	181 ms	36 ms	9 KB	1 · 1 KB	-
/blog/ai-crawlers-do-not-run-javascript	2	388 ms	0.0002	161 ms	28 ms	10 KB	1 · 1 KB	-
/	1	1,296 ms	0.0097	1,058 ms	35 ms	306 KB	1 · 1 KB	-

Chromium, desktop viewport, 4× CPU slowdown, unthrottled network, media requests aborted. Colours are Google's Core Web Vitals bands. AI crawlers read the HTML without rendering, so only first byte and weight reach them.

Answer coverage

Questions this site should be able to answer were run through a retrieval engine built over the site's own content, the same shape of pipeline an answer engine uses. Coverage

is judged from the retrieved passages alone.

INTENT CLASS	INTENTS	COVERAGE	ANSWERED
informational · en	194		77%
troubleshooting · en	39		64%
commercial · en	29		62%
transactional · en	22		73%
navigational · en	19		63%
comparison · en	11		27%

answered: a passage on the site fully answers it

partial: a passage exists but is incomplete

unanswered: no passage on the site answers it

Highest-value questions the site cannot fully answer

QUESTION	COVERAGE	WHAT WAS MISSING
best AI visibility tracking tools · commercial	partial	No comparison of multiple AI visibility tracking tools or vendors · No pricing, feature comparison, or ranking criteria
can AI crawlers read my website · informational	partial	No data specific to the querying user's own website; all evidence is from Vercatus's example sites (a news site and a sampled site with a Learning Center) · No confirmation via WAF/bot-management logs to rule out spoofed user-agent probes versus real blocking
why is my site not showing up in ChatGPT search · troubleshooting	partial	No specific diagnostic steps for the querying user's own site (e.g., how to test their own domain against GPTBot/OAI-SearchBot) · No information on other causes like lack of citable content, low authority, or indexing/sitemap issues
start a free AI visibility report for my site · transactional	partial	Explicit pricing/free-tier terms are not stated (only 'design partners' language, no confirmation this is literally free) · Domain verification process details are not described
best tools to check if AI crawlers can access your site · commercial	partial	No mention of alternative third-party tools (e.g., Screaming Frog, Cloudflare Radar, robots.txt testers) for comparison · No pricing or setup details for Vercatus itself
how to make website content citable by	partial	No concrete guidance on content formatting (structured data, schema markup, headings) beyond raw HTML availability · No

QUESTION	COVERAGE	WHAT WAS MISSING
ChatGPT and Perplexity · informational		specifics on how to phrase/structure answerable claims for extraction (e.g., direct answer paragraphs, FAQ format)
Cloudflare is blocking GPTBot, how do I allow it · troubleshooting	partial	Exact Cloudflare UI steps to enable GPTBot under Verified Bots or AI Crawl Control · Whether GPTBot itself carries a verifiable signature Cloudflare recognizes, versus only user-agent string
Perplexity is not citing my website, how do I fix it · troubleshooting	partial	No guidance on content/answer formatting or structured data that improves likelihood of being cited once a page is fetchable · No specifics on how to verify whether PerplexityBot specifically is being blocked on this exact site (only a general example from a different site)

Findings

Each finding names the evidence it rests on and the knowledge-base hypotheses that support it, with the strength of that evidence graded A (official documentation or standard) to D (practitioner consensus). 30 candidate findings were discarded in adversarial review before this list.

HIGH · 11

HIGH Category vocabulary (GEO, fan-out, llms.txt, AI Overviews) is unanswered while the 138-entry knowledge base stays privat

site · 10 pages

The head definitional intents of the category the product sells into are unanswered: "what is generative engine optimisation (GEO)" (0.62), "how does query fan-out work in Google AI Mode" (0.66), "should I use schema.org structured data for AI Overviews" (0.68), "how to add llms.txt to a website" (0.60), "how to check if your content appears in Google AI Overviews" (0.72). The site serves /llms.txt yet has no page explaining llms.txt, and four pages claim a "138 graded hypotheses" knowledge base that is not published, so the strongest citable asset on the site is invisible to retrieval.

WHAT TO DO

Publish the knowledge base as a public /learn/{concept} library, one page per concept, answer-first in the opening sentence, with the grade, the effect and the primary source link inline. Start with the five unanswered queries above plus entity, chunk, citability and machine identity. Each page becomes the entity home for a term the docs already reference, which also feeds fan-out branches for AI Mode.

Metric: unanswered informational intents (currently 12 of 194) and citations landing on /learn/ pages.

Measure: Unanswered informational intents drop from 12/194 to $\leq 4/194$ and at least 8 new /learn/ URLs appear in the most-cited pages list in the next full run.

Grounded in entity.entity-home C geo.query-fan-out A retrieval.query-decomposition A
content.original-data-citability C geo.citations-quotes-statistics B

HIGH No comparison or alternatives content : every competitor-named intent is unanswered

site · 6 pages

The comparison category is the weakest in the simulation (11 intents: 3 answered, 4 partial, 4 unanswered) and the specific vendor queries "Vercatus vs Semrush AI visibility", "Vercatus vs Ahrefs Brand Radar" and "Vercatus alternatives" return nothing, with no expected page at all. The profile also lists competitors as an empty set, so no page on the site names another tool. Buyers in a new category shortlist by comparison, and answer engines serve those intents from whoever publishes them, which today is competitors and review sites.

WHAT TO DO

Publish a /compare/ hub plus one page per named alternative (Semrush AI visibility/Brand Radar, Ahrefs Brand Radar, Profinder, Otterly, Peec AI, Screaming Frog for the crawler contrast), each with a real HTML comparison table (data source, crawler identities probed, retrieval simulation yes/no, price entry point, self-hosting) and an honest "when the other tool is the better fit" paragraph. Also add a /compare/alternatives/ page listing them. Target: comparison-category answered rate from 3/11 to $> 7/11$ and first citations on vendor-name-plus-vs queries.

Measure: comparison-category answered rate in the next run (cohort: 11 comparison intents) and citations on "vs"/"alternatives" queries

Grounded in commercial.comparison-intents C retrieval.markdown-and-tables C geo.query-fan-out A

HIGH No defining passage for the crawler tokens the whole product is about (GPTBot , OAI-SearchBot , ClaudeBot , PerplexityBot ,

site · 6 pages

GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Googlebot and Google-Extended are each mentioned 4 to 7 times but have no entity home, and the related intents stall: "what is OAI-SearchBot and how is it different from GPTBot" is partial with the gap "no explicit statement of OAI-SearchBot's purpose", "what is Google-Extended and should I block it" is partial, and "what does robots.txt actually control for AI crawlers" is partial.

The mentions are scattered inside /docs/concepts/machine-identities/ (cited 48 times) as a list of identities probed rather than as per-token definitions, so no passage answers a single-token question on its own.

WHAT TO DO

Create /docs/reference/crawler-tokens/ with one H2 per token, each section opening with a self-contained sentence of the shape "OAI-SearchBot is OpenAI's search crawler; it builds the index ChatGPT search cites, and blocking it removes the site from ChatGPT search while GPTBot only affects training." Include purpose, vendor, whether it honours robots.txt, the published IP or key location, and the correct robots.txt line. Link every scattered mention to that section with the token as anchor text. Target: the three partial token intents answered, and the new page appearing in cited pages for token queries.

Measure: coverage of crawler-token informational intents (partial to answered) and citations to /docs/reference/crawler-tokens/, cohort: informational intents naming a bot token

Grounded in entity.entity-home C retrieval.entity-explicitness C
access.ai-crawler-token-catalogue A geo.search-vs-training-bots A

HIGH Transactional path intents (free report , demo , signup) are all partial and /contact/ is thin

site · 3 pages

Every high-value transactional intent stalls on missing mechanics: "start a free AI visibility report for my site" (0.86) has no stated trial terms beyond "design partners" language, "request a Vercatus demo" (0.74) finds no booking link, and "sign up for Vercatus Single site plan" (0.66) has no checkout steps. The expected page for two of them, /contact/, is one of the two pages flagged as thin main content (deterministic e9ae380485f44a39), so there is no self-contained passage for an engine or an agent to quote.

WHAT TO DO

Create a /get-started/ page that states in plain text what a first run includes, whether it is free and under what terms, the domain verification method (DNS TXT or file, with the exact token format), typical run duration, and direct links to signup and to a booking calendar. Expand /contact/ above 250 words with named routes (sales, security review, support) and ContactPoint markup so the navigational and transactional intents resolve to one page each. Metric: transactional partial count.

Measure: Transactional intents partial count falls from 6/22 to <=2/22 and /get-started/ becomes the cited page for the free-report and signup queries in the next full run.

HIGH Plan entities have no entity home and their facts diverge across four pages

site · 4 pages

Single site plan, Portfolio plan and Enterprise plan are all homePageId null while plan facts are spread across /, /pricing/, /docs/reference/plans/ and /terms/, and "Vercatus pricing" is cannibalised across those three pages. The claims table records genuine divergence: enterprise plan includes ["10 sites, 10,000 pages per run"] vs ["SSO, residency, own model provider"] vs ["ten sites, \$250 each beyond..."], and portfolio plan costs ["\$2,500/month"] vs ["\$2,500/month + per-site fees"]. Offer markup exists only on the homepage, not on /pricing/ (Organization,WebSite only).

WHAT TO DO

Make /pricing/ the single entity home for all three plans: one HTML table with a row per plan (name, platform fee, included sites, pages per run, billing period, currency) mirrored in Offer nodes with stable @id per plan, and reduce /docs/reference/plans/ and the homepage to worked examples that link to it. State the Portfolio base fee and the per-site ladder in the same sentence so the qualifier cannot be dropped. Target: "how much does the Vercatus Enterprise plan cost" and the three plan intents move to answered with /pricing/ as the cited page.

Measure: coverage of the 29 commercial + 22 transactional intents mentioning plans; /pricing/ as cited page for "Vercatus pricing"

HIGH Troubleshooting intents stop one step short of the fix (no per-CDN unblock runbooks)

site · 5 pages

14 of 39 troubleshooting intents are only partial, and the recorded gaps are consistently the operational last mile: "Cloudflare is blocking GPTBot, how do I allow it" (0.76) fails against /docs/guides/allow-list/ for lack of the exact Cloudflare UI steps, and "my WAF returns 403 to AI crawlers but robots.txt allows them" (0.70) lacks the distinction between a default-on CDN AI block and a deliberate rule. "why is my site not showing up in ChatGPT search" (0.86) is spread thinly across three pages instead of one diagnostic.

WHAT TO DO

Ship three per-CDN runbooks (Cloudflare AI Crawl Control, Akamai, Fastly) with the actual menu paths and the verified-bot toggles, each opening with the token list to allow, plus one diagnostic page "my pages are not cited in ChatGPT: how to test" containing the curl commands per identity and the log-verification step that a user-agent probe cannot replace. Link them from /crawler/ and /docs/guides/allow-list/ with descriptive anchors. Metric: partial troubleshooting intents.

Measure: Partial troubleshooting intents fall from 14/39 to $\leq 6/39$ and the CDN runbooks appear as the expected page for the WAF/robots intents in the next full run.

Grounded in [access.cloudflare-ai-crawl-control](#) A [geo.waf-ai-blocking-default](#) A
[access.googlebot-verification](#) A [geo.crawler-access](#) A

HIGH The crawler identities the product is built around have no entity home

site · 6 pages

GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Googlebot and Google-Extended are each mentioned 4 to 7 times across the site but every one has homePageId null; they are described only in passing inside /docs/concepts/machine-identities/ and /crawler/. The simulation shows the cost directly: "what is OAI-SearchBot and how is it different from GPTBot" and "what is Google-Extended and should I block it" resolve partial, with gaps listed as the missing purpose-level definition of each bot, and "how to verify a bot is really Googlebot and not spoofed" is partial too.

WHAT TO DO

Publish one short definitional page per crawler token under /docs/crawlers/{token}/ (purpose: training vs search vs user-initiated, vendor, published IP/verification method, robots.txt token, what allowing or blocking it changes), each answer-first with the token named in the first sentence, cross-linked from /docs/concepts/machine-identities/ as a hub. This is the site's core topical territory and each page is a fan-out branch. Target: the 5 bot-definition intents move from partial to answered.

Measure: answered share of the 6 crawler-definition informational intents (currently partial/unanswered)

Grounded in [entity.entity-home](#) C [geo.query-fan-out](#) A [content.self-contained-passages](#) B
[access.ai-crawler-token-catalogue](#) A

HIGH Pricing facts are duplicated across four URLs with no canonical source , producing contradictions and split retrieval

site · 4 pages

The same plan facts appear on `/pricing/`, `/docs/reference/plans/` and `/terms/`, and the claim extractor already finds conflicts: portfolio plan costs "\$2,500/month" versus "\$2,500/month + per-site fees", and the Enterprise plan "includes" three different bundles across `/pricing/` and `/docs/reference/plans/`. The simulator confirms the retrieval cost: "Vercatus pricing" is contested by `/terms/`, `/pricing/` and `/docs/reference/plans/`, and "Vercatus Single site vs Portfolio vs Enterprise" by four own pages. When a reranker must choose between four near-duplicate passages with different numbers, the cited one is often not the page the business wants quoted.

WHAT TO DO

Make `/pricing/` the single source of plan facts (per-tier price, billing period, currency, included runs and pages, per-site increments) as one HTML table with per-row plan names, mirror it in Offer markup, and reduce `/docs/reference/plans/` and `/` to a one-line summary plus a link to `/pricing/` with the anchor naming the plan. Remove pricing numbers from `/terms/`. Target: pricing contradictions drop to zero and `/pricing/` becomes the sole cited page for pricing intents.

Measure: number of detected pricing contradictions (baseline 3 subject/predicate pairs) and the cited page for "Vercatus pricing" in the next full run

Grounded in retrieval.duplicate-passages-split-signal C entity.consistency-contradictions C
commerce.pricing-page-citability C retrieval.lost-in-the-middle B

HIGH Plan prices are stated without their qualifiers and disagree across pricing surfaces

site · 3 pages

The claims table records "A portfolio plan is \$2,500 a month" (`/`) alongside "Portfolio plan costs \$2,500/month + per-site fees" (`/pricing/`) and worked totals of \$3,400 (5 domains), \$6,250 (12) and \$10,150 (25) on `/docs/reference/plans/`; the Enterprise node carries three different "includes" objects across `/pricing/` and `/docs/reference/plans/`. The arithmetic is internally consistent, but the bare \$2,500 sentence is the quotable one, so an engine can state a price that is wrong for almost every real buyer. Omission of qualifiers is the dominant AI-answer fidelity failure mode.

WHAT TO DO

Rewrite every price sentence so the figure and its conditions sit in the same sentence ("Portfolio starts at \$2,500/month and includes 3 sites; additional sites are \$450 each up to ten and \$300 beyond, so 12 sites is \$6,250/month"). Designate `/pricing/` as the single source, make `/docs/reference/plans/` and the homepage block render from the same data, and give Enterprise one canonical inclusions list. Target: zero pricing contradictions in the next run's claims table and correct plan totals in sampled engine answers for "Vercatus pricing".

Measure: count of pricing contradictions in the claims table (cohort: /, /pricing/, /docs/reference/plans/) and price accuracy in sampled ChatGPT/AI Mode answers

Grounded in entity.consistency-contradictions C geo.ai-overview-claim-fidelity B
commerce.pricing-page-citability C

HIGH No third-party corroboration for "best tool" intents : no reviews , no benchmarks , no sameAs

site · 40 pages

Every "best X" commercial intent resolves partially with the same gap pattern: "best AI visibility tracking tools" (0.9), "best tools to check if AI crawlers can access your site" (0.82) and "best GEO tools for enterprise websites" (0.7) all report no independent review, benchmark or competitor naming. The homepage Organization node has zero sameAs links (deterministic 59d161ac7caf4a87), /product/ carries no testimonials or customer logos, and 0 of 40 pages have an author. Earned and third-party sources supply the large majority of citations in answers about a category, so a self-contained brand site cannot win these.

WHAT TO DO

Two tracks. Off-site: claim G2/Capterra listings, create a Wikidata item, and add sameAs (LinkedIn, GitHub, Wikidata, G2) to the Organization node so the brand resolves as a thing. On-site: publish a dated, methodology-first benchmark of machine-access and citability across a public sample of sites (the existing news-site render-delta data is the seed) as first-party data third parties can cite, and add named-customer or design-partner evidence with permission. Target: brand named in sampled engine answers for "best AI visibility tools" and non-zero sameAs.

Measure: brand mention share in repeated engine samples for category "best tool" queries; sameAs count on the Organization node

Grounded in geo.ai-citations-earned-media C entity.sameas-reconciliation B
geo.brand-mentions-corroboration C content.original-data-citability C

HIGH Link graph is navigation-flat , so no page owns a topic : 174 of 194 informational intents cannibalised

site · 40 pages

Every top hub has an identical inbound count (/docs/, /, /docs/concepts/knowledge-base/, /docs/trust/security/, /docs/reference/api/ all at 39; a second band at 27-28), which is the signature of global nav and footer links rather than topical hub-and-spoke linking. With no differentiated internal linking, the same crawler-access story is repeated on /crawler/, /docs/concepts/machine-identities/, /docs/guides/allow-list/ and /blog/ai-crawlers-do-not-run-javascript/, and the simulator shows 174/194

informational, 37/39 troubleshooting and 29/29 commercial intents competing between the site's own pages (for example "Cloudflare is blocking GPTBot, how do I allow it" splits across /crawler/, /docs/guides/allow-list/ and /). The /docs/ index also fails its own navigational intent ("Vercatus docs" partial: no section index), so it is a link sink, not a hub.

WHAT TO DO

Designate one owner page per concept cluster (machine identities and bot tokens, allow-listing, retrieval simulation, plans and pricing), and rewrite the other pages to summarise in one sentence and link out with a descriptive anchor naming the concept instead of repeating the explanation. Turn /docs/ into a real hub that lists every section with descriptive anchors, and add a bot-token glossary hub that links to per-bot sections. Target: cannibalised intent count falls from 174/194 informational and the owner page becomes the top-ranked passage for its cluster in the next full run.

Measure: cannibalised intent share (baseline 174/194 informational) and citation concentration on designated owner pages, measured on the full 40-page sample in the next full run

Grounded in content.internal-linking-topic-clusters A retrieval.duplicate-passages-split-signal C
retrieval.lost-in-the-middle B content.anchor-text-signal A entity.entity-home C

MEDIUM · 17

MEDIUM Transactional path is undocumented: no trial terms, no signup steps, no booking link

site · 2 pages

All three buying intents resolve only partially: "start a free AI visibility report for my site" (0.86) lacks free-tier terms and signup steps, "request a Vercatus demo" (0.74) finds no scheduling link or demo description, and "sign up for Vercatus Single site plan" (0.66) finds no checkout path. The expected page for two of them, /contact/, is flagged thin (<120 words, deterministic e9ae380485f44a39) and mixes sales, security and login into a page with no ContactPoint markup.

WHAT TO DO

Rebuild /contact/ (or a new /get-started/) with three explicitly labelled routes as visible prose plus ContactPoint markup: (1) what the "design partner" report includes, whether it costs anything, and the exact steps; (2) a demo booking URL with duration and agenda; (3) self-serve signup for Single site with payment method and billing terms. State domain verification as a numbered sequence, linking /docs/start/verify-domain/. Target: the three transactional intents move to answered and /contact/ above 120 words of main content.

Measure: answered rate on the 22 transactional intents and word count of /contact/
main content

Grounded in geo.self-contained-passages **B** content.self-contained-passages **B**
geo.information-agents-and-agentic-booking **A**

MEDIUM The site diagnoses machine visibility but publishes no remediation playbooks

site · 6 pages

Mid-funnel "how do I fix it" demand is uncovered: "how to write content chunks that AI answer engines can quote" (0.64) and "how to track AI referral traffic from ChatGPT and Perplexity in analytics" (0.70) are unanswered, "how to make website content citable by ChatGPT and Perplexity" (0.80) is partial and split across four pages, and "how do publishers stop losing traffic to AI Overviews" (0.68) lacks any structural guidance. This is exactly the demand the product creates, so the gap hands the informational entry point to competitors and to generic SEO blogs.

WHAT TO DO

Publish a remediation playbook cluster: answer-first passage patterns with before/after examples, one sub-question per H2 with the answer in the first sentence, entity naming instead of "our platform", and an attribution recipe naming `utm_source=chatgpt.com` and the referrer hosts to filter in analytics. Reuse the site's own run data as worked examples (the 78% JavaScript-added gallery figure already does this well). Metric: these four intents plus their sub-questions.

Measure: The four remediation intents move to answered and playbook URLs enter the top 12 most-cited pages, measured on the informational cohort in the next full run.

Grounded in content.self-contained-passages **B** retrieval.chunking-strategy **B**
geo.chatgpt-search-and-fetchers **A** geo.citations-quotes-statistics **B**

MEDIUM Pricing and plans pages carry no Offer markup and sit on the generic /about template

page · /pricing/ · 3 pages

/pricing/ and /docs/reference/plans/ hold the plan tables and the worked examples but are rendered by the /about template (fingerprint 04be7c224ed8afa8, JSON-LD: Organization, WebSite only) and the docs templates emit no JSON-LD at all (deterministic bb06e54649b2833, 28 pages). Offer and SoftwareApplication exist only on the homepage, whose price sentences are the unqualified ones. The pricing FAQ on /pricing/ is also unmarked, while the homepage FAQ is.

WHAT TO DO

Emit SoftwareApplication with one Offer per plan (name, price, priceCurrency USD, billingIncrement/unitText month, eligibleQuantity for included sites, url with fragment) on /pricing/, referencing the Organization @id, and mirror the same numbers in the visible HTML table with plan names in every row so each row chunks with its plan attribution. Add FAQPage to the existing /pricing/ Q&A. Target: price and plan facts extracted correctly for "how much does X cost" intents and "how much does the Vercatus Enterprise plan cost" moving from partial to answered.

Measure: Offer nodes present and price parity with visible text on /pricing/ and /docs/reference/plans/; answered rate on the 29 commercial intents

Grounded in commerce.pricing-page-citability C commercial.product-facts-structured A
tech.structured-data-types-that-matter A entity.structured-data-disambiguation A

MEDIUM First-party measurement is the site's strongest citable asset and it is published once , incompletely

site · 7 pages

Only 7 of 40 pages carry original data, yet the two most specific claims on the site are first-party measurements: "78% of the words on the gallery template and 51% on the article template appear only after JavaScript runs" and the per-template word-count table on /blog/ai-crawlers-do-not-run-javascript/, whose page-quality note records incomplete rows (n/a for article and home HTML word counts). That single post is already cited 54 times in the simulation, the same order as /product/ (54) and /pricing/ (51), on the strength of one dataset from one news site. Restated third-party facts dominate the rest of the corpus.

WHAT TO DO

Complete the existing table (fill the n/a cells, add the measurement date and the crawl profile) and then publish a recurring aggregate from the crawler's own corpus: render-delta by template type across N sites, share of sites where at least one AI token gets a non-200, and TTFB distribution by CDN, each as an HTML table with a stated sample size and date, under /research/. Give each statistic its own H3 so it can be quoted alone. Target: pages with original data rising from 7 of 40 and new /research/ URLs entering the most-cited list.

Measure: count of pages with original data and citation count of /research/ pages,
cohort: informational and commercial intents about AI crawler behaviour

Grounded in content.original-data-citability C geo.citations-quotes-statistics B
geo.ai-citations-earned-media C retrieval.numbers-dates-specificity B

MEDIUM The agency segment's only page is thin and priceless while the agency intent is live

Agencies are a named audience, and "white label SEO and AI visibility reports for agencies" (importance 0.7) resolves partially against /docs/guides/white-label/, which is flagged as thin main content (<120 words, deterministic ca17b36d4fa6bed0) and buries plan gating inside a settings table. The simulation records the exact missing facts: agency-specific reporting content, Portfolio pricing numbers and example report contents.

WHAT TO DO

Create a commercial /solutions/agencies/ page (not a docs guide) that states what is white-labelled (logo, domain, report cover, PDF), the Portfolio economics with the \$521-per-site worked example, client-seat and role behaviour, and a sample report; link the docs guide from it for implementation steps. Keep the docs page but expand it past 120 words with the setting table plus prose. Target: the agency intent moves from partial to answered and /solutions/agencies/ becomes the cited page.

Measure: coverage verdict for the agency white-label intent and main-content word count on the white-label pages

Grounded in commercial.comparison-intents C content.self-contained-passages B
commerce.pricing-page-citability C

MEDIUM The product entity is declared only on the homepage : /product/ carries no SoftwareApplication node

SoftwareApplication, Offer and FAQPage nodes exist only on / (template 9d8b9a2a2e236e6d); /product/ sits on the /about template with Organization,WebSite only, despite being the third most-cited page (54 citations) and the expected page for "best AI visibility tracking tools" and "best tools to check if AI crawlers can access your site". The product as a thing is therefore defined on a page whose passages also carry pricing and brand claims, which is why those commercial intents are cannibalised between / and /product/.

WHAT TO DO

Move the canonical SoftwareApplication node (stable @id, name, applicationCategory, featureList, offers referencing the plan Offer @ids on /pricing/) to /product/ and have the homepage reference that @id instead of redeclaring it; keep the 12-step pipeline and 138-hypothesis structure as labelled sections so each is a distinct retrievable passage. Target: /product/ becomes the cited page for the product and tool-selection commercial intents, removing the / vs /product/ split.

Measure: cited page for the 29 commercial intents; presence of a single SoftwareApplication @id referenced site-wide

MEDIUM Four audience segments share one page , so no segment intent gets a citable passage

page · /solutions/ · 6 pages

/solutions/ carries publishers, e-commerce, SaaS/B2B and agencies in one document, and every segment intent lands partially on it: "AI visibility monitoring for ecommerce product pages" (0.66) lacks product-schema checks and a worked example, "SaaS documentation AI search visibility tool" (0.56) lacks a documentation-specific case, "best GEO tools for enterprise websites" (0.70) lacks criteria. The agency intent "white label SEO and AI visibility reports for agencies" (0.70) falls to /docs/guides/white-label/, which is flagged thin (deterministic ca17b36d4fa6bed0).

WHAT TO DO

Split /solutions/ into four pages, each opening with the segment named in the first sentence and each carrying one worked example from a real run (template table with page counts, render delta, cited-page list) rather than generic framing. Add a marketing-side agency page with an example white-label report, distinct from the docs settings page, and keep /solutions/ as a hub linking out with descriptive anchors. Metric: segment commercial intents answered.

Measure: Commercial intents partial count falls from 11/29 to ≤5/29, with segment pages replacing /solutions/ as the expected page for the four segment queries.

MEDIUM Structured-data architecture stops at the marketing templates : 28 docs pages emit no JSON-LD and no @id graph links them

site · 28 pages

Rich JSON-LD exists only on / (SoftwareApplication, Offer, FAQPage, Organization, WebSite), /blog (Article) and the /about template (Organization, WebSite); both docs templates (6d2b07d7f2f35d3b, 58613178545d6698 = 28 pages, deterministic finding bbf06e54649b2833) emit none, despite being definitional pages with tables of facts (plans, env vars, API endpoints, roles). /pricing/ and /docs/reference/plans/ carry the actual worked prices (\$3,400 for 5 domains, \$6,250 for 12, \$10,150 for 25) as HTML tables with no Offer markup, while the homepage Offer states different numbers. There is also no evidence of a shared @graph: Organization is restated per template rather than referenced by a stable @id, and the homepage Organization has zero sameAs (59d161ac7caf4a87).

WHAT TO DO

Publish one Organization node with a stable @id (for example `https://vercatus.com/#organization`) plus sameAs to LinkedIn, Crunchbase and a claimed Wikidata item, and have every other page reference that @id as publisher. Add TechArticle (or Article) with headline, dateModified and about/DefinedTerm to both docs templates, Product plus per-plan Offer on /pricing/ and /docs/reference/plans/ mirroring the visible table rows, and Blog/ItemList on /blog/. Target: pages without JSON-LD falls from 28 to 0 and plan-price intents resolve as answered rather than partial.

Measure: count of sampled pages emitting JSON-LD (baseline 12/40) and valid Offer/TechArticle nodes per template in the next full run

Grounded in [entity.structured-data-disambiguation](#) A [tech.structured-data-types-that-matter](#) A
[entity.organization-home](#) A [entity.sameas-reconciliation](#) B [commerce.pricing-page-citability](#) C

MEDIUM

Organization node has no sameAs and no Wikidata item: the brand is a string, not a thing

site · 12 pages

The Organization JSON-LD emitted on / and the /about template carries 0 sameAs links (deterministic 59d161ac7caf4a87), and Vercatus itself is absent from the Wikidata-linked entity set while the third parties it writes about are resolved (Googlebot Q1425771, GPTBot Q121364002, Cloudflare Q4778915, Anthropic Q116758847). For a young B2B SaaS with no competitor set and no earned-media corpus, this leaves answer engines with no external record to reconcile the brand against, which is why brand and comparison intents ("Vercatus alternatives", "Vercatus vs Ahrefs Brand Radar") return unanswered.

WHAT TO DO

Add sameAs to the single Organization node (LinkedIn, GitHub, Crunchbase, X, and the Wikidata item once created), create a Wikidata item for Vercatus with the official website and founding facts, and keep those facts identical to /about/. Then pursue third-party corroboration (directories, press, developer registries) since a brand's own domain is a minority of the evidence engines use. Target: sameAs count ≥ 4 , a resolving Wikidata QID, and named-brand presence on the 19 navigational and 11 comparison intents.

Measure: Organization sameAs count ≥ 4 plus a live Wikidata QID; brand mention rate on navigational/comparison intent cohort (30 intents)

Grounded in [entity.sameas-reconciliation](#) B [entity.organization-home](#) A
[trust.consistency-across-web](#) C [geo.ai-citations-earned-media](#) C

MEDIUM**No author on any of 40 pages and only 6 with a visible date , on a topic that changes quarterly**

site · 34 pages

Page-quality extraction found author present on 0 of 40 pages and a visible date on 6, while the corpus makes time-bound claims (Cloudflare defaults, RFC 9421 draft status, "do Core Web Vitals affect Google rankings in 2026" resolves only partially). /privacy/ is the only page with an explicit version and date. For a vendor whose argument is that machines must be able to verify facts, anonymous and undated technical claims are the weakest configuration for both Google's provenance signals and for users who cross-check AI answers.

WHAT TO DO

Add a named author or reviewer with Person markup and sameAs (LinkedIn, GitHub) plus visible published and updated dates to the blog template (81c8cba51d0605f9) and to every /docs/concepts/ and /docs/guides/ page, with dateModified in JSON-LD matching the visible date. Put the date inside the passage for volatile claims (for example "as of March 2026, Cloudflare blocks known AI crawlers by default for domains onboarded after 1 July 2025"). Target: authorPresent and dateVisible on the docs and blog templates, and freshness-sensitive intents moving off partial.

Measure: authorPresent and dateVisible share on docs and blog templates, cohort: pages on 58613178545d6698, 6d2b07d7f2f35d3b, 81c8cba51d0605f9

Grounded in trust.provenance-signals **B** content.author-experience **B** content.freshness-signals **A**
geo.freshness-answer-engines **C**

MEDIUM**No page wins the brand's own name : "Vercatus" retrieves /terms/, /contact/, /crawler/ and the CLI reference**

site · 4 pages

The navigational intent "Vercatus" is cannibalised across /contact/, /docs/reference/cli/, /crawler/ and /terms/, and /about/ does not appear in the competing set at all. /about/ is written in a manifesto style rather than as a definitional entity home, so the strongest passages naming the company are legal and technical boilerplate. An entity with no single defining passage is resolved by engines from whichever copy scores marginally better, which here is the terms page.

WHAT TO DO

Rewrite the opening of /about/ as an entity-home definition: one answer-first paragraph naming Vercatus, what it is (machine-visibility intelligence SaaS), founding/operating facts, location and contact, followed by the manifesto. Point every internal brand mention and the Organization @id at /about/, and demote self-descriptive prose in /terms/ and /crawler/ to a single sentence that links to

/about/. Target: /about/ becomes the top-cited page for the brand navigational intents and the /terms/ competition disappears.

Measure: cited page for the 19 en navigational intents; /about/ citation count vs /terms/ citation count

Grounded in entity.entity-home C entity.organization-home A
retrieval.duplicate-passages-split-signal C retrieval.lost-in-the-middle B

MEDIUM The 28 most-cited docs pages emit no JSON-LD, so retrieved passages carry no entity anchor

template · 58613178545d6698 · 28 pages

Both docs templates (58613178545d6698 /docs/concepts/{slug} and 6d2b07d7f2f35d3b /docs/guides/{slug}, 28 pages) emit zero JSON-LD (deterministic bb06e54649b2833), yet these are the pages the simulator cites most after the homepage: /docs/concepts/machine-identities/ (48 citations), /docs/guides/allow-list/ (41), /docs/concepts/performance/ (39). Structured data exists only on the three marketing templates, so the pages that actually get retrieved make no machine-readable statement about who published them, when, or what term they define.

WHAT TO DO

Add a shared JSON-LD graph to the docs layout: a TechArticle (or DefinedTerm for /docs/concepts/) node with headline, description, datePublished/dateModified, plus publisher referencing the single Organization @id defined on /about/, and a BreadcrumbList. Use one stable @id per URL so entity and publisher relationships link across templates rather than being re-declared. Target: JSON-LD present on 40/40 sampled pages and publisher @id resolvable from every cited docs page.

Measure: share of sampled pages with valid JSON-LD referencing the Organization @id (cohort: 28 docs pages)

Grounded in entity.structured-data-disambiguation A entity.organization-home A
tech.structured-data-types-that-matter A

MEDIUM Near-total intent cannibalisation, including the brand query resolving to /terms/ and /contact/

site · 40 pages

286 of 314 simulated intents are cannibalised (174 of 194 informational), which is a structural pattern, not noise: docs, marketing pages and blog posts each restate the same definitions. The brand query "Vercatus" competes across /contact/, /docs/reference/cli/, /crawler/ and /terms/, with neither / nor /about/ winning, and "do AI crawlers execute JavaScript" splits between the blog post and

/docs/concepts/machine-identities/. When the best passage ranks behind the site's own weaker copies, grounding sets pick the wrong page.

WHAT TO DO

Assign one canonical passage per concept: keep the full definition on the entity home (e.g. JavaScript rendering on the blog post, machine identity on /docs/concepts/machine-identities/) and reduce every other mention to a single-clause pointer with a descriptive anchor. Make /about/ the brand entity home with the full "what Vercatus is" paragraph and add sameAs there (see deterministic 59d161ac7caf4a87). Consider data-nosnippet on the footer-like nav blocks that let /terms/ and /contact/ win topical queries. Metric: cannibalised intent share.

Measure: Cannibalised intent share falls below 60% of simulated intents, and the brand query "Vercatus" resolves to / or /about/ as the top-cited page in the next full run.

Grounded in retrieval.duplicate-passages-split-signal C retrieval.lost-in-the-middle B
entity.entity-home C content.anchor-text-signal A

MEDIUM

The /docs/guides template averages 238 words, too thin to yield a self-contained retrievable passage

template · 6d2b07d7f2f35d3b · 14 pages

Template 6d2b07d7f2f35d3b (14 pages) averages 238 words against 495 for the concepts template, and /docs/guides/white-label/ falls below 120 words (deterministic ca17b36d4fa6bed0); /contact/ has the same problem on the /about template (e9ae380485f44a39). Several of these pages are the expected answer page for real intents that come back partial: "white label SEO and AI visibility reports for agencies" (expected /docs/guides/white-label/) is missing report contents and plan pricing, and "Cloudflare is blocking GPTBot, how do I allow it" (expected /docs/guides/allow-list/) is missing the actual Cloudflare UI steps. After chrome removal these pages produce one or two short chunks that look like template boilerplate.

WHAT TO DO

Expand each guide to answer its named task end to end in ~400-600 words: one H2 per sub-question with the answer in the first sentence, the product or bot entity named explicitly in each section, and concrete steps or values (exact Cloudflare AI Crawl Control path, exact header names, what a white-label report contains). Where a guide cannot justify a page, merge it into its parent and redirect. Target: avg words per page on this template above 400 and the listed intents moving from partial to answered.

Measure: average main-content words on template 6d2b07d7f2f35d3b (baseline 238) and partial-to-answered conversion for the guide-owned intents

MEDIUM

Homepage and commercial template ship 10-15% of their text only after JavaScript

template · 9d8b9a2a2e236e6d · 10 pages

The homepage template (9d8b9a2a2e236e6d) has a JS-added text ratio of 0.15 over 1260 words, so roughly 190 words exist only after rendering, and the /about template (04be7c224ed8afa8, 9 pages including /pricing/, /product/, /solutions/) sits at 0.10, about 42 words per page. The page-level observation notes the homepage's live Cloudflare crawl-access table is the site's strongest first-party proof, and that interactive block is exactly the kind of content that is injected client-side. GPTBot, ClaudeBot, PerplexityBot and OAI-SearchBot fetch HTML and do not execute JavaScript, so the site's best evidence and possibly parts of its price grid are absent from the passages those engines can cite.

WHAT TO DO

Diff raw versus rendered HTML for <https://vercatus.com/>, /pricing/, /product/ and /solutions/ and identify which blocks are JS-only. Server-render (Astro SSG) the demo table, pricing figures and any FAQ text so the full word count is present in the initial HTML, leaving only interactivity to hydration. Target a JS-added text ratio below 0.03 on both templates, the level the blog template already achieves (0.03).

Measure: JS-added text ratio per template (targets: 9d8b9a2a2e236e6d and 04be7c224ed8afa8 below 0.03) and raw-HTML word count on /, /pricing/, /product/

MEDIUM

Docs claims are unsourced : 152 of 208 claims across the two docs templates carry no evidence

template · 58613178545d6698 · 28 pages

The docs templates hold the bulk of the site's claims but the weakest evidence density: /docs/concepts/{slug} has 111 claims with 91 lacking evidence (avg specificity 0.67) and /docs/guides/{slug} has 97 claims with 61 lacking evidence (0.66), against 0.80 on the blog template and 0.84 on the homepage where evidence is attached to almost every claim. Page-quality notes confirm the pattern: the numeric claims on /docs/concepts/sampling/ (50,000 cap, 20% cap, 1.3M URL example) and the cost benchmarks on /docs/operators/self-hosting/ are undated and unsourced. Unsourced numbers are quotable but not verifiable, which is what makes the difference for engines whose users cross-check.

WHAT TO DO

For each numeric or behavioural claim in the two docs templates, put the qualifier in the same sentence (measurement date, sample, model, hardware) and link the governing hypothesis id, which the blog template already does successfully (it cites `tech.raw-html-for-ai-crawlers` and `perf.field-over-lab` inline). Concretely: date the self-hosting cost table, state the run profile and page count behind each figure on `/docs/operators/cost/`, and add the source standard link for RFC 9421 and CrUX claims on `/docs/reference/crawler/` and `/docs/concepts/performance/`. Target: claims-without-evidence share on the docs templates falling from ~73% toward the blog's ~7%.

Measure: claims without evidence per docs template and avg claim specificity, cohort: pages on 58613178545d6698 and 6d2b07d7f2f35d3b

Grounded in `geo.citations-quotes-statistics` B `retrieval.numbers-dates-specificity` B
`geo.ai-overview-claim-fidelity` B

MEDIUM 34 of 40 pages carry no visible date , so live-retrieval engines cannot judge freshness

site · 34 pages

Only 6 of 40 sampled pages show a visible date and no author is present anywhere; the 28 docs pages emit no JSON-LD at all, so there is no `dateModified` either. The subject matter (crawler tokens, Cloudflare defaults, AI Mode behaviour) is time-sensitive, and Perplexity and ChatGPT search favour clearly dated pages for exactly these queries, which the simulation shows as partial: 'what is Google-Extended and should I block it' (0.64), 'do Core Web Vitals affect Google rankings in 2026' (0.60).

WHAT TO DO

Add a visible 'Last reviewed' date to every docs and reference page, generated from the content file's git timestamp, and mirror it as `dateModified` in per-page JSON-LD. Do not touch the date without a substantive edit. Track whether dated docs pages start winning the time-sensitive informational intents currently resolving as partial.

Measure: Share of pages with a visible date and matching `dateModified` (target above 90%); coverage on time-sensitive informational intents

Grounded in `geo.freshness-answer-engines` C `content.freshness-signals` A `trust.provenance-signals` B

LOW · 11

LOW No JSON-LD structured data

site · 28 pages

28 page(s) emit no schema.org JSON-LD. Structured data is the cheapest way to state entity identity, authorship, dates and relationships unambiguously; answer engines and knowledge graphs consume it directly.

WHAT TO DO

Add template-level JSON-LD: Organization/WebSite on the home page, Article/Product/FAQPage/HowTo per template, with @id links between them and sameAs to authoritative profiles.

Grounded in entity.structured-data-disambiguation A

LOW 15% of the homepage's text exists only after JavaScript , on the site's most-cited page

page · / · 1 page

The homepage has the highest JS-added text ratio on the site (0.15 versus 0.08 on docs and 0.03 on the blog template) while being the most retrieved page in the simulation (106 citations) and the carrier of the headline pricing claims (\$1,000/month single site; \$2,500 portfolio). AI crawlers and retrieval fetchers do not execute JavaScript, so roughly one word in seven of this page is unavailable to GPTBot, ClaudeBot, PerplexityBot and OAI-SearchBot, and the page is also the one whose live crawl-access table is described as a first-party demo. The same page carries 306 KB and 1,058 ms lab TBT, indicating a client-side hydration step that is doing the extra rendering.

WHAT TO DO

Diff the raw HTML against the rendered DOM for / and server-render (Astro static output) whatever text is currently client-injected, prioritising the pricing figures, the live access table and the FAQ answers already mirrored in FAQPage markup. Keep JavaScript for interactivity only. Target: JS-added text ratio on / below 0.03, matching the blog template, verified by the raw-versus-rendered word delta in the next run.

Measure: raw-vs-rendered JS-added text ratio on / (baseline 0.15) in the next full run

Grounded in tech.raw-html-for-ai-crawlers C perf.ai-crawlers-fetch-html-only B
tech.rendering-budget A geo.self-contained-passages B

LOW Trust and security facts about Vercatus contradict themselves across /privacy/ , /docs/trust/security/ and /docs/operator

site · 4 pages

The claims table shows five different objects for vercatus/retains ("page bodies for 30 days", "derived data...", "server logs, max 30 days", "page bodies for 30 days, derived

data for workspace lifetime") across `/`, `/privacy/` and `/docs/trust/security/`, and two different objects for `vercatus/lacks` ("SOC 2 attestation and regional data residency" vs "status page and generated API reference"). Each statement may be individually true, but retrieval returns one chunk at a time, so a security reviewer or an answer engine sees partial retention policies without their qualifiers.

WHAT TO DO

Designate `/docs/trust/security/` as the entity home for data-handling facts and state each one as a complete sentence carrying subject, object and period ("Vercatus retains fetched page bodies for 30 days; derived findings persist for the workspace lifetime; server logs are kept at most 30 days"). Have `/privacy/` and the homepage quote that wording verbatim and link to it, and separate "not yet available" roadmap gaps from compliance gaps so the two lacks claims stop colliding. Target: contradiction count for subject `vercatus` drops to zero in the next run.

Measure: number of same-subject/same-predicate contradictions for subject "vercatus" (currently 5 predicates affected)

Grounded in `entity.consistency-contradictions` C `geo.ai-overview-claim-fidelity` B
`trust.about-contact-policies` A

LOW

All 40 URLs are published through a news sitemap , and 34 of 40 pages carry no visible date at all

site · 40 pages

Discovery reports the single sitemap as a news sitemap carrying all 40 URLs, which is the wrong type for evergreen docs, marketing and legal pages: news sitemaps are scoped to recently published news articles and give Google no usable lastmod discipline for the rest. Only 6 of 40 analysed pages show a visible date, and neither docs template emits JSON-LD, so there is no `dateModified` anywhere on the 28 documentation pages that describe a product whose plans, bot tokens and env vars change. `/rss.xml`, `/feed.xml` and `/feed` all return 404, so the two-post blog has no push or subscription path either.

WHAT TO DO

Replace the news sitemap with a standard `urlset` sitemap listing the 40 canonical URLs with accurate lastmod, and keep a news sitemap only if genuine news is published. Add a visible "last updated" line plus `dateModified` in the new `TechArticle` markup on both docs templates, and add an RSS/Atom feed for `/blog/` plus `IndexNow` pings on publish for Bing and Copilot surfaces. Target: dated pages rising from 6/40 and sitemap type matching content type.

Measure: share of sampled pages with a visible date and machine-readable dateModified (baseline 6/40), and sitemap type classification in the next discovery pass

Grounded in tech.sitemap-hygiene A content.freshness-signals A geo.freshness-answer-engines C
tech.indexnow A

LOW Layout shifts poor (lab p75 CLS 0.4502)

site · 9 pages

9 measured page(s) shift after first paint. Element responsible for the largest shift: `div.right-sidebar.astro-okomqxn`. Images without declared dimensions: 0 per page. Shifts happen to people, not to crawlers; on ad-funded templates the usual cause is a slot injected above content.

WHAT TO DO

Reserve the space: width/height or aspect-ratio on images and embeds, fixed-height slots for ads and consent banners, size-adjust for web fonts. Measure: lab p75 CLS on this template.

Grounded in perf.cvv-thresholds A perf.cls-image-dimensions A

LOW Docs template shifts layout badly in lab (p75 CLS 0.4246 from div . right-sidebar) and the homepage blocks the main thread

template · 6d2b07d7f2f35d3b · 15 pages

Lab p75 CLS on template 6d2b07d7f2f35d3b is 0.4246, in the poor band (>0.25), with the largest shift attributed to div.right-sidebar.astro-okomqxn and zero images missing dimensions, so the cause is an unreserved sidebar column rather than media (deterministic 99867d/83090ff). The homepage records 1,058 ms lab TBT and 306 KB, a real INP risk, and every template loads one render-blocking stylesheet from fonts.googleapis.com. None of this reaches AI crawlers (TTFB is 28-48 ms and HTML is 9-306 KB, both healthy), and there is no CrUX data, so this is a human-experience and tie-breaker issue only.

WHAT TO DO

Reserve the right-sidebar column with a fixed width or grid template so it cannot reflow after hydration, self-host the font with font-display: swap and size-adjust to remove the third-party render-blocking request, and split or defer the homepage bundle causing the 1,058 ms TBT. Confirm with field data once CrUX has coverage. Target: lab p75 CLS on this template under 0.1 and homepage TBT under 300 ms.

Measure: lab p75 CLS on template 6d2b07d7f2f35d3b (baseline 0.4246) and homepage TBT (baseline 1,058 ms), plus CrUX LCP/CLS/INP once available

LOW Enterprise procurement facts (price , retention , attestations) are opaque or stated three ways

site · 4 pages

Enterprise is the highest-value plan yet its platform fee is "on request" ("how much does the Vercatus Enterprise plan cost" resolves partially), and the facts an enterprise reviewer checks are inconsistent across pages: retention appears as "page bodies for 30 days", "derived data for workspace lifetime", "server logs, max 30 days" across /, /privacy/ and /docs/trust/security/, and the "lacks" claims differ between /docs/operators/deploying/ and /docs/trust/security/. Agents comparing vendors read these policy statements directly.

WHAT TO DO

Publish a single machine-readable trust summary block (one table, repeated verbatim nowhere else) covering retention by data class, subprocessors, no-training commitment, SSO/OIDC, residency and attestation status with dates, and link every mention to it. Give Enterprise a published starting price or a stated price band with what changes it; "on request" cannot be cited. Target: zero retention or inclusions contradictions in the claims table and the Enterprise cost intent answered.

Measure: contradiction count for subject "vercatus" in the claims table (cohort: /, /privacy/, /docs/trust/security/, /docs/operators/deploying/)

LOW Agent and security surface is empty apart from llms.txt , and the crawler key directory is unprobed

site · 40 pages

/.well-known/security.txt, /.well-known/agents.json, /.well-known/tdmrep.json and /.well-known/ai-plugin.json all return 404, and there is no RSS or API catalog link; only llms.txt is served (llms-full.txt is 404, deterministic finding e5177e34b439c0ae). The audience explicitly includes security and infrastructure reviewers who must allow-list a crawler, and the site's own claim is that requests are signed with Ed25519 keys published at /.well-known/http-message-signatures-directory, an endpoint this run did not probe, so the central differentiator is currently unverified from the outside.

WHAT TO DO

Publish security.txt (contact, policy, expires) and confirm the http-message-signatures-directory resolves with the advertised JWKS. Expose the REST API

through an RFC 9727 api-catalog link or a linked OpenAPI document so agent clients can act on the product rather than only read about it. llms-full.txt is optional and low value; do it only because it is cheap.

Measure: Count of resolving well-known endpoints (security.txt, signature directory, api-catalog) and successful external verification of the Ed25519 JWKS

Grounded in geo.agent-surfaces D geo.web-bot-auth A trust.about-contact-policies A geo.llms-txt D

LOW An API-first product exposes no machine-actionable surfaces : no api-catalog , OpenAPI , security .txt or crawlable signup p

site · 40 pages

The product ships a REST API, CLI and self-hosting path, yet /.well-known/agents.json, /.well-known/security.txt, /.well-known/ai-plugin.json and /.well-known/tdmrep.json all return 404, /docs/reference/api/ is a hand-maintained endpoint table with no JSON-LD, and the site's own docs admit a missing "generated API reference".

Transactional intents reflect this: "sign up for Vercatus Single site plan" and "start a free AI visibility report for my site" both resolve partial with the exact signup steps and entry URL missing, meaning no crawlable page describes the conversion path an agent or a reader would follow. Security and infrastructure reviewers are a named audience, and security.txt is the file they look for.

WHAT TO DO

Publish an OpenAPI document and advertise it via RFC 9727 (/well-known/api-catalog and Link rel="service-desc"), add /.well-known/security.txt pointing at the security contact already listed on /contact/, and create a crawlable /start or /signup page in the sitemap that states the steps, what is required, whether a free or design-partner tier exists, and the domain-verification flow, with SoftwareApplication/Offer markup. Target: the two transactional intents move to answered and the well-known probes return 200.

Measure: transactional intent coverage (baseline 16 answered / 6 partial of 22) and 200 responses for /.well-known/api-catalog and security.txt

Grounded in geo.agent-surfaces D access.nlweb-and-schema-endpoints A trust.about-contact-policies A geo.information-agents-and-agentic-booking A

LOW Declared ai-train=no is contradicted by a robots .txt that allows every training crawler

site · 40 pages

robots.txt returns 200 and disallows no AI crawler at all, while the same file carries Content-Signal: search=yes, ai-input=yes, ai-train=no (deterministic finding 9ee4d1b370797744). GPTBot, ClaudeBot and Google-Extended are training tokens

and are currently fully permitted, so the declared preference has no matching rule and no TDMRep reservation exists (/well-known/tdmrep.json returns 404). For a vendor whose product explains per-token crawler policy, the mismatch on its own domain is both a policy gap and a credibility problem.

WHAT TO DO

Decide the intent and make robots.txt match it: either keep the blanket allow and drop ai-train=no, or add explicit Disallow blocks for GPTBot, Google-Extended and CCBot while keeping OAI-SearchBot, Claude-SearchBot, PerplexityBot and ChatGPT-User/Claude-User allowed, and publish /well-known/tdmrep.json if the operating entity is in the EU. Then use the resulting file as the worked example in /crawler/ and /docs/guides/allow-list/.

Measure: robots.txt per-token policy consistent with the Content-Signal line, verified in the next run's access matrix

Grounded in access.content-signals-policy A access.ai-crawler-token-catalogue A
geo.search-vs-training-bots A access.tdm-reservation-eu A

LOW

No feed and no IndexNow , so Bing and Copilot pick up changes only on their own schedule

site · 40 pages

/feed, /rss.xml and /feed.xml all return 404 and there is no IndexNow key in evidence, while the site ships a blog and 28 docs pages whose accuracy depends on fast-moving vendor behaviour. Copilot grounds on the Bing index, so publish-to-answer latency on Microsoft surfaces is entirely passive today. This is low leverage at 40 URLs but nearly free to fix.

WHAT TO DO

Add an IndexNow key and ping Bing on every deploy from the same build step that regenerates the sitemap, and publish an RSS or JSON feed for /blog/ so aggregators and agent tooling can follow updates. Verify submissions in Bing Webmaster Tools.

Measure: Median hours from deploy to Bing re-crawl of changed URLs, measured in Bing Webmaster Tools

Grounded in tech.indexnow A geo.copilot-uses-bing-index A

INFO · 1

INFO

Machine access is clean across all five templates , so leverage sits in content , not access

site · 40 pages

Probes with Googlebot, GPTBot, ClaudeBot, PerplexityBot, OAI-SearchBot and browser identities returned zero AI-blocked pages on every template, robots.txt is 200 with no AI disallows, all 40 sampled URLs are 200 and indexable, and TTFB is 28-48 ms. No snippet-suppressing directive (nosnippet, max-snippet:0) was observed. The single largest GEO failure mode, a refused fetch, does not apply here, which means the remaining upside is in what the passages say, not in whether they can be reached.

WHAT TO DO

Treat access as solved and reallocate effort to passage quality and comparison coverage. Confirm the probe result against edge or WAF logs once, since user-agent probes cannot prove a real crawler is served the same response, and re-run the access matrix after any CDN or bot-management change. Do not add access-side work to the roadmap on the strength of this run alone.

Measure: AI-blocked page count stays at 0 across all templates in the next full run, corroborated by one sample of edge logs

Grounded in `geo.crawler-access` A `access.googlebot-verification` A `geo.waf-ai-blocking-default` A

What to do , in order

Ranked by impact against effort, with access blockers first, because nothing else matters while machines cannot read the page. Each item names the metric and the pages to measure it on, so the next run can confirm whether it worked.

1 Make the transactional path answerable end to end S 3 weeks

Every transactional intent resolves only partially: "start a free AI visibility report", "request a Vercatus demo" and "sign up for Single site plan" all fail on the same missing facts (what the design-partner offer actually costs, the signup steps, a booking link, what domain verification requires). /contact/ is below 120 words and carries no ContactPoint markup. This is the shortest path between an AI answer and revenue, and the fix is copy plus one schema node.

EXPECTED	Moves 6 transactional and 2 commercial intents from partial to answered; gives assistants an explicit next action to quote when a user asks how to start.
METRIC	Simulated coverage for the 22 transactional intents (answered share), plus assisted signups and demo requests attributed to referrers carrying <code>utm_source=chatgpt.com</code> or <code>perplexity.ai</code>
MEASURED ON	Treated: /contact/, /pricing/, /docs/guides/white-label/ and a new /start/ page. Control: /docs/reference/* pages, unchanged in the same period.

2 Give every crawler identity a defining page M 6 weeks

GPTBot, OAI-SearchBot, ClaudeBot, PerplexityBot, Googlebot and Google-Extended are mentioned 4 to 7 times each and have no entity home, while intents such as "what is OAI-SearchBot and how is it different from GPTBot" and "what is Google-Extended and should I block it" resolve partially against a concepts page that lists tokens without defining them. These are the highest-volume informational queries in the site's own category and the site is the natural authority on them.

EXPECTED	Creates 6 to 8 self-contained, citable definition passages naming each token, its vendor, its purpose (training versus search versus user-initiated) and the robots directive that governs it; each becomes an entity home that internal anchors can point to.
METRIC	Answered share on the crawler-token intent subset (currently partial), count of entities with a non-null homePageld, and citation count per new page in the retrieval simulation
MEASURED ON	Treated: new /docs/reference/crawlers/{token} pages plus /docs/concepts/machine-identities/ as the hub. Control: /docs/concepts/{slug} pages not touched.

3

One canonical pricing surface , with qualifiers in the same sentence and Offer markup

S

3 weeks

Portfolio pricing is stated as \$2,500/month on one page and \$2,500/month plus per-site fees on another; Enterprise is described three ways across /pricing/, /docs/reference/plans/ and /docs/guides/allowances/; plan names have no entity home; and none of the pricing surfaces carry Product/Offer structured data. Contradictory facts about the same entity lower confidence in every statement about the brand, and AI Overviews misquote most often by dropping a qualifier that sits in a different sentence or page.

EXPECTED	Removes the four detected pricing contradictions, makes /pricing/ the single retrieval target for price intents instead of /terms/ and /docs/reference/plans/ competing, and makes prices machine-extractable for shopping and comparison surfaces.
METRIC	Contradiction count on subject "plan" in the claims table (target zero), number of distinct URLs cited for the query "Vercatus pricing" (target one primary), presence of valid Offer nodes on the pricing template
MEASURED ON	Treated: /pricing/, /docs/reference/plans/ (reduced to a pointer plus allowance mechanics), plan sections on /. Control: /solutions/ and /product/, unchanged.

4

Publish honest comparison and alternatives content

M

8 weeks

"Vercatus vs Semrush AI visibility", "Vercatus vs Ahrefs Brand Radar", "Vercatus alternatives", "best GEO tools for enterprise websites" and "AI crawler monitoring

tool vs Screaming Frog" are all unanswered or partial, and 4 of 11 comparison intents return nothing at all. Comparison and recommendation intents are among the most frequent commercial questions put to answer engines, and brand-only pages never capture them.

EXPECTED	Opens an intent class the site currently scores zero on; real HTML comparison tables with header rows are among the shapes answer engines quote most for X vs Y questions.
METRIC	Answered share of the 11 comparison intents and of the commercial "best tool" subset; count of comparison pages cited in the simulation
MEASURED ON	Treated: new /compare/{competitor} pages and a /compare/ hub, plus per-segment pages split out of /solutions/. Control: /product/ and /about/, unchanged.

5 Publish the category vocabulary and open part of the knowledge base L 12 weeks

GEO, query fan-out, llms.txt, AI Overviews measurement, schema for AI answers, chunk writing and AI referral tracking are all unanswered, while the site claims 138 graded hypotheses that exist only as a private asset described in prose. This is the largest addressable demand gap in the sample and the site already holds the raw material.

EXPECTED	Converts an internal asset into an entity-rich, citable public corpus with one sub-question per H2, which is what query fan-out rewards; also gives third parties something concrete to reference off-site.
METRIC	Answered share across the 194 informational intents and the number of distinct pages cited (breadth, not just volume); inbound references to knowledge-base URLs
MEASURED ON	Treated: new /kb/{hypothesis-id} pages and glossary entries. Control: existing /docs/concepts/{slug} pages, left as-is during the first measurement window.

6 Turn troubleshooting pages into per-CDN unblock runbooks S 5 weeks

"Cloudflare is blocking GPTBot, how do I allow it", "my WAF returns 403 to AI crawlers but robots.txt allows them", "why is my site not showing up in ChatGPT search" and "Perplexity is not citing my website" all stop one step short of the fix: the pages explain the diagnosis but never give the exact Cloudflare AI Crawl Control, Akamai or Fastly steps, or the reverse-DNS verification procedure. Troubleshooting queries are where an answer engine most needs a step list it can quote.

EXPECTED	Moves 14 partial troubleshooting intents toward answered and makes the site the quotable source for a problem its buyers search at the moment of pain.
----------	--

METRIC	Answered share of the 39 troubleshooting intents; citation count for /docs/guides/allow-list/ and the new runbook pages
MEASURED ON	Treated: /docs/guides/allow-list/ split into per-CDN runbooks. Control: /docs/guides/{slug} pages not in the runbook set.

7 Complete the entity layer : sameAs , Wikidata , product node , docs JSON-LD M 6 weeks

The Organization node has zero sameAs links and there is no Wikidata item, so knowledge graphs cannot reconcile the brand; /product/ carries no SoftwareApplication node even though the homepage does; 28 docs pages emit no JSON-LD and nothing links the graph by @id; and no page wins the query "Vercatus", which currently retrieves /terms/, /contact/, /crawler/ and the CLI reference.

EXPECTED	Anchors every retrieved docs passage to a named organisation and product, and gives external knowledge graphs a reconciliation target. Also resolves the trust-fact contradictions (retention stated five ways, attestations stated two ways) that currently undercut enterprise answers.
METRIC	Pages with JSON-LD (target 40/40), sameAs count on the Organization node, existence of a Wikidata item, and which URL is cited first for the brand query in the simulation
MEASURED ON	Treated: all /docs/* templates plus /about/ and /product/. Control: /blog/* pages, which already carry Article markup.

8 Add provenance : named authors , visible dates , dateModified , and fix the sitemap type S 4 weeks

Zero of 40 pages name an author and only 6 show a date, on a topic where crawler behaviour and engine features change quarterly. All 40 URLs are published through a news sitemap while 34 carry no date at all, which is a contradictory freshness signal. Provenance is the durable differentiator against machine-generated volume and the thing an AI-referred visitor checks before trusting the answer.

EXPECTED	Makes the site legible to live-retrieval engines for time-sensitive queries and adds the author and source signals that raters and LLM graders weigh; also lets the 152 unsourced docs claims be tied to the standards they derive from (RFC 9309, RFC 9421, Google documentation).
METRIC	Share of pages with a visible date and a Person author (target above 90 percent), share of docs claims carrying evidence, sitemap type corrected to a standard sitemap with accurate lastmod
MEASURED ON	Treated: /docs/concepts/{slug} and /docs/guides/{slug} (28 pages). Control: /about template pages, dated in a later phase.

9 Rebuild the link graph around topic hubs so pages stop competing

L

10 weeks

174 of 194 informational intents are cannibalised. The cause is structural: navigation links dominate, every page links to the same hubs, and no page accumulates the anchor text that would make it the owner of a topic. Duplicate and near-duplicate passages across URLs split citation credit and let rerankers pick the wrong copy, which is exactly why /terms/ surfaces for pricing and the brand name.

EXPECTED	Fewer, stronger passages per topic; the intended page wins its intent instead of the fourth-best copy. Also consolidates the 238-word guides template into passages long enough to stand alone.
METRIC	Cannibalised intent count (baseline 174 of 194 informational), number of distinct URLs competing per query, and inbound descriptive-anchor count per hub
MEASURED ON	Treated: the crawler-access cluster and the pricing cluster (hub plus spokes, descriptive anchors, deduplicated blurbs). Control: the operators cluster (/docs/operators/*), restructured later.

10 Build off-site corroboration and repeat the first-party measurements

L

16 weeks

Most of the evidence answer engines use about a brand comes from third parties, not the brand's own site, and the site currently has no reviews, no benchmarks, no directory presence and no sameAs. Meanwhile its strongest citable asset (the JavaScript render-delta measurement, the self-hosting cost benchmarks) was published once, on one site, with incomplete table rows and no date. Grade C and D evidence here, so treat as a believed rather than known lever, but the direction is well supported.

EXPECTED	Gives "best AI visibility tool" style answers a third-party source to name the brand, and gives journalists and other sites a reason to cite original numbers rather than restate vendor copy.
METRIC	Count of independent domains mentioning the brand, presence on G2 or similar with reviews, and brand mention rate in a fixed 30-query sample run against ChatGPT, Perplexity and Google AI Mode
MEASURED ON	Treated: the brand entity itself, measured by repeated engine sampling. Control: the same query set measured before the campaign starts.

11 Close the machine-surface and hygiene gaps

S

4 weeks

Lower-value but cheap items that a machine-visibility vendor should not be caught failing: 10 to 15 percent of the homepage text arrives only after JavaScript on the site's most-cited page; the docs template shows lab p75 CLS of 0.42 from the right sidebar and the homepage blocks the main thread for about 1,058 ms TBT; there is

no feed and no IndexNow, so Bing and Copilot learn about changes late; the agent surface is empty apart from llms.txt (no api-catalog, no OpenAPI descriptor, no security.txt) despite an API-first product; and the declared Content-Signal ai-train=no is contradicted by a robots.txt that allows every training crawler.

EXPECTED	Restores full homepage text to non-rendering fetchers, removes a visible credibility risk on the product's own subject matter, shortens publish-to-Copilot latency, and makes the declared content policy match the enforced one.
METRIC	JS-added text ratio per template (target below 0.03), lab p75 CLS on /docs/guides (target below 0.1), homepage TBT, and presence of feed, IndexNow key, api-catalog and security.txt
MEASURED ON	Treated: homepage and /docs/guides template. Control: /blog template, already at 0.03 JS ratio and 0.0002 CLS.

Entity graph

Entities resolved across pages and linked to Wikidata by exact label match. An entity with no defining page has no canonical home on this site for engines to resolve it to.

ENTITY	TYPE	WIKIDATA	MENTIONS	DEFINING PAGE
★ vercatus CLI	Technology	unlinked	42	/docs/reference/cli/
Googlebot	Technology	Q1425771	7	none
ClaudeBot	Technology	unlinked	6	none
PerplexityBot	Technology	unlinked	6	none
GPTBot	Technology	Q121364002	6	none
OAI-SearchBot	Technology	unlinked	6	none
Enterprise plan	Product	unlinked	5	none
Vercatus crawler	Technology	unlinked	5	/docs/reference/crawler/
Google-Extended	Technology	unlinked	4	none
RFC 9421 HTTP Message Signatures	Technology	unlinked	4	none
Portfolio plan	Product	unlinked	4	none

ENTITY	TYPE	WIKIDATA	MENTIONS	DEFINING PAGE
Ed25519	Technology	Q56099578	4	none
Cloudflare	Organization	Q4778915	4	none
Wikidata	Technology	Q2013	3	none

Most citable claims on the site

Answer engines quote specific, checkable statements. These scored highest for specificity: concrete, quantified and dated beats "industry-leading".

CLAIM	EVIDENCE	SPECIFICITY
This is Version 1 of the privacy policy, dated 16 September 2026. /privacy/	none	100
A Portfolio plan with 5 domains costs \$3,400 per month. /docs/reference/plans/	statistic	95
Measured on a full-profile run: about \$17 for 500 pages, \$80 for 2,500 pages, \$310 for 10,000 pages, or half that through the Batches API. /docs/operators/self-hosting/	statistic	95
A worked example shows 12 domains costing \$6,250/mo, or \$521 per site. /pricing/	statistic	95
USER_AGENT defaults to 'Vercatus/1.0 (+https://vercatus.com/crawler)' and is sent on every request. /docs/operators/configuration/	none	95
Requests are signed using RFC 9421 HTTP Message Signature with tag web-bot-auth and Ed25519, with keys at /.well-known/http-message-signatures-directory on vercatus.com /docs/reference/crawler/	first party data	95
Gallery template: 48 words in HTML, 205 words after JavaScript, 78% added by JavaScript. /blog/ai-crawlers-do-not-run-javascript/	statistic	95

CLAIM	EVIDENCE	SPECIFICITY
A Portfolio plan with 12 domains costs \$6,250 per month. /docs/reference/plans/	statistic	95

Method and limits

SAMPLING

40 pages were drawn from 40 sitemap URLs, stratified so every URL pattern and language is represented. Findings marked *template* or *site* generalise to pages outside the sample; page-level findings do not.

ANSWER SIMULATION

Retrieval combines lexical and multilingual dense search over the site's own passages, and an answer is judged from those passages alone. It models how an answer engine uses a site; it is not a measurement of any live engine.

ACCESS PROBES

Requests carried each crawler's published user-agent from this system's own network. A refusal proves the site refuses *that identity from this origin*; a site verifying crawlers by IP or reverse DNS may treat genuine traffic differently. Confirm against WAF and bot-management logs before acting.

PERFORMANCE

Timings come from single headless-browser loads, not field data. Treat them as indicative and confirm against real-user measurements before prioritising.