

RESEARCH REPORT · BALTIC AI VISIBILITY INDEX 2026

How AI assistants recommend companies in the Baltic states

A four-engine, two-wave measurement of visibility, choice-set size, stability, sourcing and accuracy across 358 of the largest companies in Latvia, Lithuania and Estonia

WebGlazer Research, Riga, Latvia

Correspondence: admin@webglazer.com

Data collected 5 to 14 August 2026 · Analysis closed 3 September 2026 · This version 23 September 2026

Abstract

Generative AI assistants increasingly answer commercial questions ("which bank should a small company in Latvia use?") with a short list of named companies rather than a page of links. No public measurement of this behaviour existed for the Baltic market. We constructed a frame of 358 companies from the official business registers of Latvia, Lithuania and Estonia (FY2025 revenue), classified each into one of 49 shared industry categories, and posed 2,530 unbranded buying questions in English, Latvian, Lithuanian and Estonian to four consumer-facing engines: ChatGPT, Google AI answers, Gemini and Perplexity. Every buying question was repeated one week later, and every company received four direct identity questions. The corpus comprises 25,932 answers, 66,716 extracted company mentions, 230,138 cited sources and 41,663 captured background search queries. A validated name matcher (96.0% precision, 95.2% recall on a held-out sample) scored 57,880 company-answer pairs.

The median buying answer named three companies (mean 3.05); Google AI answers named 1.9 and ChatGPT 4.4. Across all answers the engines named 11,136 distinct companies, half of them exactly once, and the largest firms in each category held 27.5% of the name slots. Within the tracked frame, 15.8% of companies (95% CI 11.8 to 19.4) were never named in their own category, and only 32 of 355 were named in more than 80% of answers. No country held a visibility advantage (all pairwise adjusted $p = 0.88$). English prompts surfaced local companies 4.4 percentage points more often than local-language prompts (paired $p < 0.001$). Company-level mention rates were highly reproducible across waves ($r = 0.992$), yet individual answers churned: a mention recurred one week later 82.6% of the time overall, but 90.2% for the most visible quartile and 38.7% for the least. Cited sources were long-tailed (19,615 hosts; company-owned domains 8.8% of citations in buying answers; Reddit the largest non-search host). Human-anchored accuracy of AI descriptions was 72.2% fully correct, 5.3% containing a major error. A no-search probe showed Estonian companies are better represented in model memory than Latvian ones in both engines tested (adjusted $p \leq 0.015$), the only country effect in the study. Joining external SEO metrics showed organic search traffic to be the strongest single predictor of being named (partial $\rho = 0.52$ controlling for revenue), and domain authority to predict visibility only when the domain belongs to the company itself: companies on a parent group's domain had the higher median domain rank yet a 21.9% mention rate against 34.7%.

Keywords: generative engine optimisation; AI visibility; brand mention; large language models; Baltic states; recommendation systems; measurement study

Key findings at a glance

 3 names	The median answer to a buying question names three companies. Google AI answers name 1.9 on average, ChatGPT 4.4. In 15.3% of answers no company is named at all.
 11,136	Distinct companies named across 20,200 buying answers. 5,433 were named exactly once. The ten most-named brands hold 6.4% of all name slots.
 32 / 355	Tracked companies named in more than 80% of their own category's answers. 56 (15.8%) were never named at all.
 82.6%	Probability that a company named in week 1 is named again in week 2 for the same question and engine. 90.2% in the top visibility quartile, 38.7% in the bottom one.
 8.8%	Share of citations in buying answers that point at a company's own website. The rest is spread over 19,615 third-party hosts led by Reddit and the national business registries.
 +4.4 pts	Higher mention rate for local companies when the question is asked in English rather than the local language, consistent across all three countries.
 0 of 53	Companies in the top domain-authority quartile of their own national domain that were never named by an assistant. On an international group domain, higher authority did not help.
 0.992	Wave-to-wave correlation of company-level mention rates. The instrument is stable; the churn is in individual answers, not in the measurement.

1. Introduction

Search engines present a ranked list and leave the choice to the user. Conversational AI assistants make part of that choice themselves: asked for a supplier, they answer with a small number of named companies and a rationale. For a company, appearing in that answer is now a precondition for being considered at all, and the mechanism that decides who appears is not observable from ordinary web analytics, because no click occurs until after the shortlist has been formed.

Commercial "AI visibility" tools report on this for individual brands, but there has been no systematic, public measurement of how the assistants behave across an entire economy: how many options they offer, how concentrated those options are, how stable the answers are from week to week, what sources they draw on, and how accurately they describe the companies involved. This study provides that measurement for the three Baltic states, whose combined market is small enough to census at the top of the revenue distribution and diverse enough (three languages, three registers, 49 industry categories) to expose language and country effects.

We had five questions. (1) When a buyer asks an assistant for a recommendation, how many companies are named and which ones? (2) Do the largest companies in a category dominate those answers? (3) Does the country, the language of the question or the choice of engine change the answer? (4) How stable are the answers over one week, and how much do the engines agree with each other? (5) Where do the assistants get their information, and how accurate is what they say?

A design principle shaped everything that follows: a company is only ever compared with other companies in its own industry category. A bank is measured against banks, a courier against couriers. Cross-industry averages appear only where the measurement is genuinely industry-neutral, such as whether the assistant can describe a company correctly.

2. Materials and methods

2.1 Company frame

The frame was built from the official business registers of Latvia (Uzņēmumu reģistrs, via data.gov.lv annual report data), Lithuania (Registru centras, via rekvizitai.vz.lt) and Estonia (Äriregister), using FY2025 revenue. Banks, which report income rather than revenue, were merged in by income. The frame began as the 100 largest companies per country and was extended with 58 smaller companies so that every industry category contained at least five members in at least one country. Eight of the largest firms had no website at all; they were replaced by the next-largest company and are reported as a finding rather than silently dropped. The final frame holds 358 active companies: Lithuania 128, Estonia 121, Latvia 117 (Supplement S17).

2.2 Category classification

Each company was classified once, at setup, into one of 49 plain-language categories shared across the three countries (for example banks, grocery retail, parcel and postal, road freight logistics, pharma and medical manufacturing) and into a market type: domestic consumer, domestic B2B or export. Every assignment was reviewed by a human. One category, unclassifiable wholesale traders, has no meaningful buyer question and was excluded from the recommendation analysis, so recommendation denominators refer to 355 companies rather than 358. The full company-to-category table is Supplement S17.

2.3 Prompt instrument

Two instruments were used. Buyer prompts are unbranded questions of the kind a real customer asks in a category: "Which bank is best for a small company in Latvia?", "Which parcel service in Lithuania is most reliable for e-commerce deliveries?". Ten were written per category (480 in total), each instantiated per country and translated into the local language by native-speaking reviewers, giving 2,530 distinct buyer prompts: 1,270 in English, 460 in Lithuanian, 400 in Latvian and 400 in Estonian. Export-oriented categories received English buyer-perspective prompts ("Nordic timber suppliers") rather than domestic consumer ones, because that is the conversation that costs those companies business. Identity prompts are four direct questions per company: what it does (English), what it does (local language), whether it is trustworthy, and how it compares with its main rival. Every identity prompt names the country, because pilot testing showed that a bare brand name such as "LMT" resolves to unrelated global entities. The buyer prompt book is Supplement S16.

Table 1. Example buyer prompts (English instantiation and Latvian translation)

Category	English	Latvian
grocery retail	What is the best supermarket chain in Latvia?	Kura ir labākā lielveikalu ķēde Latvijā?
grocery retail	Which grocery store in Latvia is best for a family doing a big weekly shop?	Kurš pārtikas veikals Latvijā ir vislabākais ģimenei, kas veic lielus iknedēļas iepirkumus?

Category	English	Latvian
parcel postal	Which parcel delivery service in Latvia is the best?	Kurš sūtījumu piegādes pakalpojums Latvijā ir labākais?
parcel postal	Which courier in Latvia is best for a small online shop sending daily parcels?	Kurš kurjers Latvijā ir vislabākais mazam interneta veikalam, kas sūta sūtījumus katru dienu?
banks	What is the best business bank in Latvia?	Kura ir labākā biznesa banka Latvijā?
banks	Which bank is best for a small company in Latvia?	Kura banka ir vislabākā mazam uzņēmumam Latvijā?

2.4 Engines and collection

Four consumer-facing engines were measured. Gemini was queried through the Vertex AI API with Google Search grounding enabled (model gemini-3.1-pro-preview, EU-hosted); a pilot hand-checked a sample of grounded API answers against the consumer Gemini application to confirm they track. ChatGPT, Google AI answers (the AI answer block on Google search results) and Perplexity were collected from the real consumer products via a scraping provider rather than through their APIs, because an earlier validation on a single brand showed that API answers without live search were materially thinner than what a user sees and recalled roughly half as many competitors. Claude was planned as a fifth engine but its web-search tool was blocked by an organisation policy that could not be changed in time; the study therefore covers four engines, not five.

Wave 1 ran from 5 to 7 August 2026 and comprised all 2,530 buyer prompts and all identity prompts on all four engines. Wave 2 repeated the 2,530 buyer prompts on all four engines one week later, with the instrument frozen. Of 25,932 collected answers, 25,928 succeeded. Direct API and scraping cost for the entire study was USD 139.60; the engines made 35,620 background search calls, of which 41,663 individual queries were captured (Google AI answers expose none, which is a collection gap rather than a finding).

Table 2. Corpus composition

Component	Count	Notes
Companies in frame	358	LT 128, EE 121, LV 117; 355 in recommendation denominators
Industry categories	49	48 with buyer prompts
Distinct buyer prompts	2,530	EN 1,270; LT 460; LV 400; ET 400
Buyer answers (two waves, four engines)	20,200	$2,530 \times 4 \times 2$, minus 40 failures
Identity answers	5,728	$4 \text{ prompts} \times 358 \text{ companies} \times 4 \text{ engines}$
Company mentions extracted	66,716	61,547 inside buying answers
Company-answer mention checks	57,880	validated matcher, both waves
Cited sources	230,138	176,673 inside buying answers; 19,615 distinct hosts
Background search queries captured	41,663	ChatGPT, Gemini, Perplexity
Identity answers graded for accuracy	2,864	100-row human anchor
No-search memory probes	716	$358 \text{ companies} \times 2 \text{ engines}$

2.5 Mention detection and its validation

Whether an answer names a given company was decided by an alias matcher over the full answer text and the extracted competitor list. The alias table (800 aliases, 35 restricted to competitor-list matching) was built from legal names, trading names and human-reviewed overrides. Two estimand rules were fixed in advance and applied throughout: a parent or umbrella brand counts as a mention when the entity trades under that brand (LMT for Lmt Retail & Logistics, Coop for the county cooperatives), and a product brand a company merely imports does not count (the "Peugeot" alias for Auto-Bon Baltic was removed).

The matcher was validated in two stages. A human labelled a stratified 200-row sample; after two corrections to the sample itself (12 apparent false positives were an artefact of the sample showing only the first 1,500 characters of each answer) the matcher scored 99.0% precision and 84.9% recall, and the seven genuine misses were traced to seven fixable alias classes. Because those fixes were tuned on that sample, an independent check was run on a fresh 400-row sample judged by a Gemini Flash model that had first been calibrated against the human labels (96.0% agreement; the model only ever missed mentions, never invented them). On that held-out sample the matcher scored 96.0% precision and 95.2% recall.

2.6 Accuracy grading

Ground truth for each company was condensed from its homepage (registry description for 48 companies whose sites were uninformative). All 2,864 identity answers were graded by a Gemini Flash judge (temperature 0, JSON schema, 98.6% self-agreement on a 286-answer re-run) under two prompts: a strict one that counts any claim absent from the ground truth as an error, and a contradiction-only one. The two disagreed materially (15.3% versus 4.4% major errors), and the gap was traced to ground-truth thinness rather than model behaviour. A 100-row handcheck, weighted toward disagreements, was therefore graded by a human with live web search and reweighted to the population strata; the published accuracy shares come from that anchor. Per-engine and per-country comparisons use the per-row judge grades and are reported as ceilings, because a control slice showed the automated judges miss entity-boundary errors.

2.7 Parametric memory probe

To separate what a model knows from what it retrieves, every company received a bare "What is {brand}?" with web search disabled, once to gemini-2.5-flash-lite on Vertex AI and once to the consumer ChatGPT interface with search off. Answers were graded into knows, partial, wrong entity, or does not know by a language-model grader against the ground truth; grades are single-pass and not human-calibrated, and the two engines differ in both model tier and access path, so this layer compares what a consumer meets rather than models under laboratory control.

2.8 Statistical analysis

Proportions carry Wilson 95% intervals. Company-level headline rates use a cluster bootstrap over companies. Country comparisons are made only after standardising across matched category cells with at least five companies per country, and are tested with Kruskal-Wallis (three groups) or Mann-Whitney (two), with effect sizes and Benjamini-Hochberg correction across each family. The English-versus-local comparison is a paired Wilcoxon signed-rank test over companies. Cross-engine agreement is the Jaccard

index over the sets of companies each engine named. Week-to-week stability is reported as the Pearson correlation of company-level mention rates between waves, the mean absolute change, and, at answer level, the conditional probability that a mention in wave 1 recurs in wave 2 for the same prompt and engine. The mean absolute wave-to-wave change (2.5 percentage points) is used as the study's noise floor: a difference smaller than it is not interpreted. No regression modelling was attempted; the per-cell samples do not support it. All statistical code carries a passing test suite.

2.9 Data and code availability

Every number in this report traces to a data file. Company-level scores, category leaders, engine rates, country tests, the language comparison, the invisible list, source categories and top hosts, the reproducibility statistics, the accuracy tables, the parametric grades, the buyer prompt book and the company frame are provided as supplementary files S1 to S17 alongside this document. The external SEO metrics joined in section 3.11 are supplementary file S18, with the traffic-controlled re-analysis and the page-level citation sample as S20 and S21, and the sampling curve in section 3.12 as S19. Raw answers and citations are available on request under the tone rule described in section 5.

3. Results

3.1 How many companies an answer names

Figure 1 shows the number of distinct companies named in each of the 20,200 buying answers, after normalising name variants. The median answer names three companies and the mean is 3.05. 15.3% of answers name none, 61.2% name three or fewer, and 28.9% name one or none. The distribution is unimodal with a long right tail: 13.0% of answers name six or more.

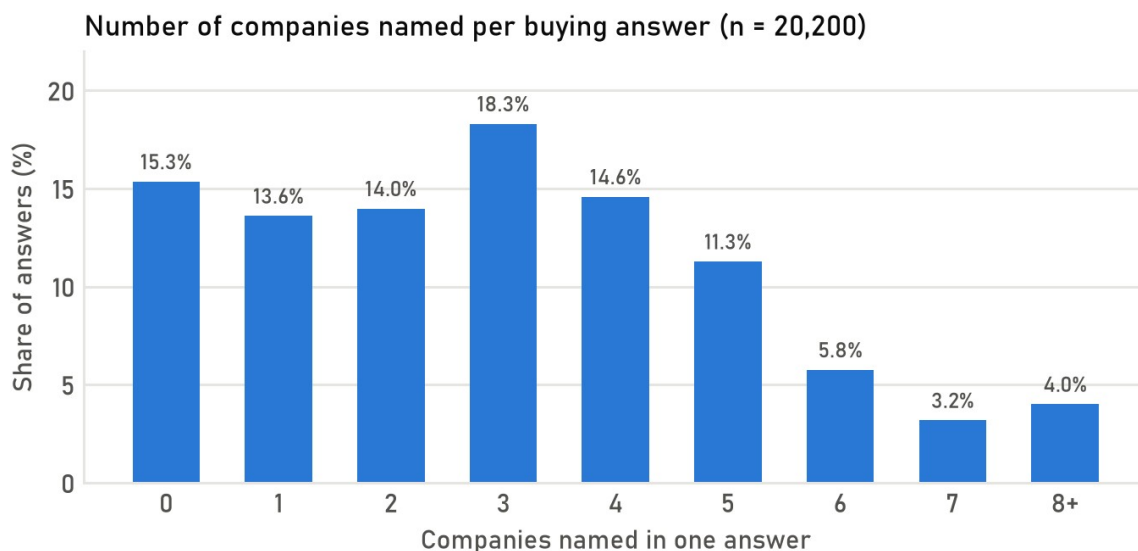


Figure 1. Distribution of the number of distinct companies named per buying answer, all four engines and both waves pooled (n = 20,200). Name variants were normalised before counting.

The engine is the largest determinant of choice-set size (Figure 2). ChatGPT names 4.4 companies per answer on average, Gemini 3.0, Perplexity 2.9 and Google AI answers 1.9; for the Google AI answer the

median is one. The language of the question has a smaller effect: English prompts drew 3.12 names, Estonian 3.11, Lithuanian 2.98 and Latvian 2.83.

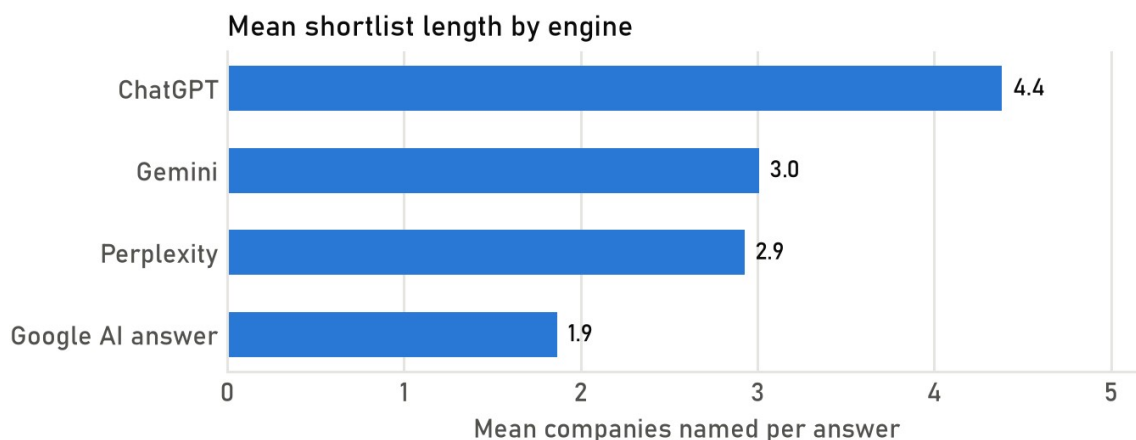


Figure 2. Mean number of distinct companies named per buying answer, by engine (n = 5,050 answers per engine).

3.2 The pool of names and the position of the largest companies

Across the 61,547 name slots in buying answers the engines named 11,136 distinct companies. The distribution is extremely long-tailed: 5,433 companies (48.8%) were named exactly once in the entire study, the ten most-named brands together account for 6.4% of slots, and the top hundred for 26.0%. Within a single country and category, the engines named a median of 106 distinct companies across that cell's 40 answers (minimum 16, maximum 222). The most-named brands overall were Circle K, Rimi, Neste, Bolt, Swedbank, Maxima, Enefit, Lidl, Tele2 and Alexela.

Using the validated matcher to identify which of those names belong to the tracked frame (the largest companies in each country), 27.5% of the name slots went to frame companies and 53.0% of answers named none of the largest companies in the relevant category. This should not be read as the assistants being wrong: the frame covers only the top of the revenue distribution and a smaller specialist is often a correct answer. It should be read as a decoupling of market share from AI share. Within category cells with at least four companies, the Spearman correlation between revenue and mention rate averaged 0.19 (31 cells; pooled $\rho = 0.33$, $p < 0.001$).

Concrete cases illustrate the decoupling. In Estonian banking, Luminor (revenue EUR 684 million, the largest bank in the frame) is named in 25.0% of answers while LHV (EUR 413 million) is named in 93.1%. In Lithuanian online marketplaces, Vinted (EUR 999 million) is named in 30.0% of answers and Pigu (EUR 168 million) in 70.0%. In Latvian energy, Latvenergo (EUR 950 million) is named in 30.0% of answers, behind its own distribution subsidiary Sadales tīkls at 51.9%.

3.3 Visibility tiers and category leaders

Each tracked company has a mention rate: the share of its own category's buying answers, across engines and waves (160 answers per company), in which it is named. Figure 3 groups the 355 companies into tiers. Only 32 (9.0%) are named in more than 80% of answers; 59 (16.6%) in 50 to 80%; 131 (36.9%) in 10 to 50%; 77 (21.7%) in under 10%; and 56 (15.8%, 95% CI 11.8 to 19.4) are never named in any answer for their own category. The never-named group is discussed in section 4.2, because its composition matters for interpretation.

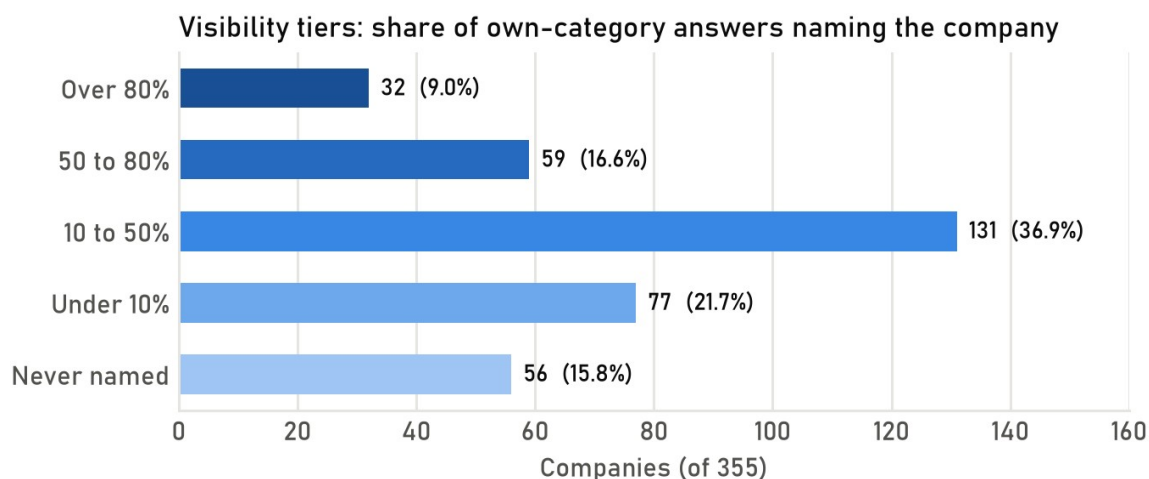


Figure 3. Visibility tiers of the 355 tracked companies, by mention rate in their own category (160 answers per company).

Table 3. The fifteen most visible tracked companies (mention rate in own category, 160 answers each)

Company	Country	Category	Mention rate %	Revenue EUR m
Lhv Pank	Estonia	banks	93.1	413
Seb bankas	Lithuania	banks	92.5	494
Telia Eesti	Estonia	telecom	92.5	363
Bolt Services LV	Latvia	mobility platforms	92.5	25
Bolt Operations	Estonia	mobility platforms	91.2	1465
Elisa Eesti	Estonia	telecom	91.2	223
Bolt Technology	Estonia	mobility platforms	91.2	2270
Telia Lietuva	Lithuania	telecom	90.6	507
Dpd Lietuva	Lithuania	parcel postal	89.4	74
Latvijas Mobilais Telefons	Latvia	telecom	88.1	196
Lmt Retail & Logistics	Latvia	telecom	88.1	137
Enefit	Lithuania	energy retail sales	88.1	249
Citadele banka	Latvia	banks	87.5	142
Air Baltic Corporation	Latvia	airlines aviation	87.5	775
Ignitis	Lithuania	energy retail sales	87.5	900

Most category races are close: the median gap between the first- and second-most-named tracked company in a category is 6.9 percentage points. A minority are one-sided. In Estonian scrap metal, Cronimet Nordic is named in 75.0% of answers and the next tracked company in 1.3%; on Lithuanian railways, LTG Link is named in 73.1% and its sister company LTG Infra in 1.9%; in Lithuanian marketplaces the gap between Pigu and Vinted is 40 points. The full list of category leaders for cells with reliable coverage is Appendix A; the complete per-company table is Supplement S1.

3.4 No country holds a visibility advantage

The crude share of never-named companies is 11.8% in Latvia, 16.9% in Estonia and 18.1% in Lithuania (Figure 4), and mean mention rates are 29.9%, 30.4% and 28.5% respectively. After standardising across

matched category cells, all three pairwise differences fall inside the study's noise floor and none is significant (Benjamini-Hochberg adjusted $p = 0.88$ for every pair; every 95% interval spans zero; Supplement S6). Standardisation also reverses the crude ordering, which is the expected behaviour of a difference driven by category mix rather than by country. We therefore report no national ranking, and we regard any claim that one Baltic country "leads" in AI visibility as unsupported by these data.

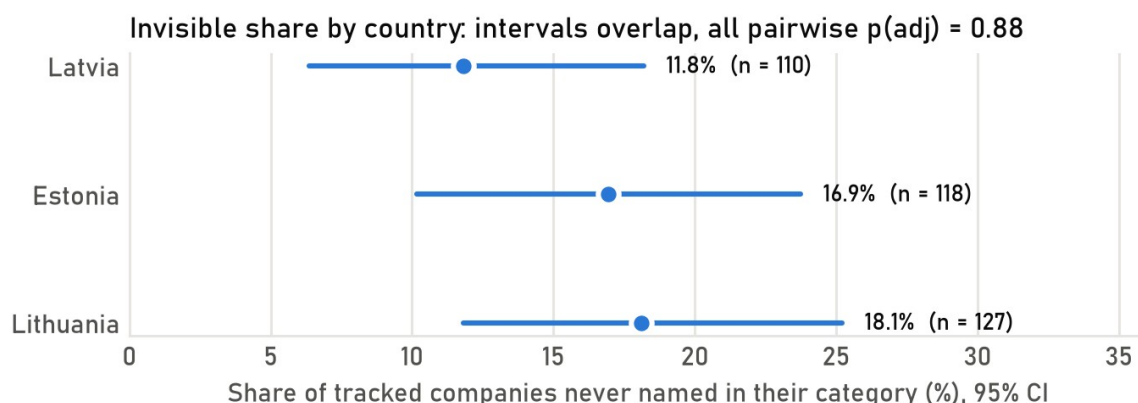


Figure 4. Share of tracked companies never named in their own category, by country, with 95% Wilson intervals. The intervals overlap and no pairwise difference survives correction.

3.5 The engines differ, and disagree with each other

Figure 5 gives each engine's mention rate: the share of the 14,470 company-prompt checks per engine in which the tracked company was named. Google AI answers (34.3%) and Gemini (33.6%) name tracked companies most often, ChatGPT slightly less (30.5%), and Perplexity far less (17.9%). Perplexity's low rate is not explained by shorter answers, since it names 2.9 companies per answer; it names different companies, more of them outside the frame.

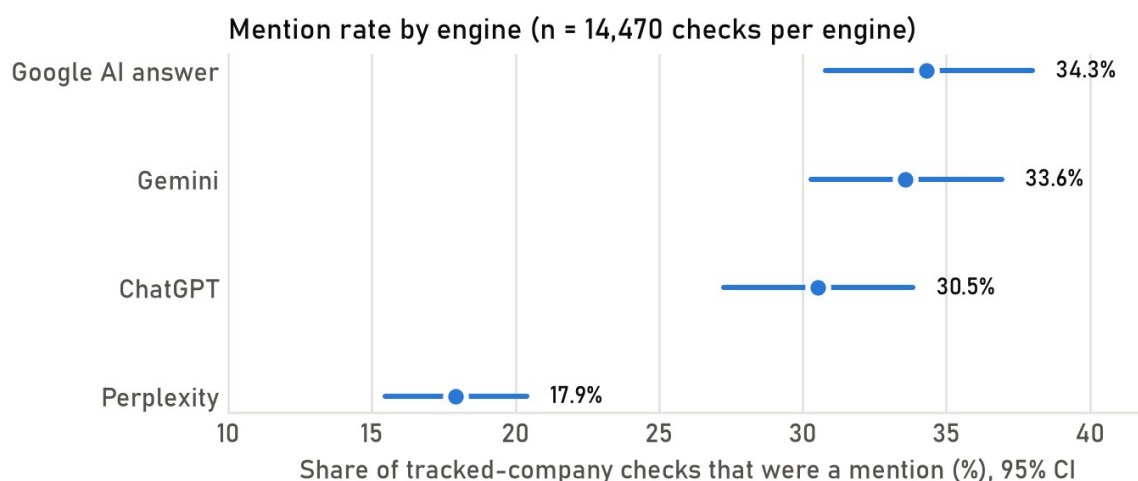


Figure 5. Mention rate by engine with 95% intervals. Each point summarises 14,470 company-prompt checks over both waves.

The engines also disagree on whom to name. At company level (Figure 6), the Jaccard overlap between the sets of tracked companies two engines name ranges from 0.38 (Gemini versus Perplexity) to 0.62 (ChatGPT versus Google AI answers). At the level of a single question on a single day, the overlap between two engines' full name lists is lower still, between 0.14 and 0.23 on normalised brand strings, and for Google AI answers versus Perplexity 60% of questions produced lists with no name in common.

Disagreement of this size is a finding in itself: where a company stands depends on which application its customer opens.

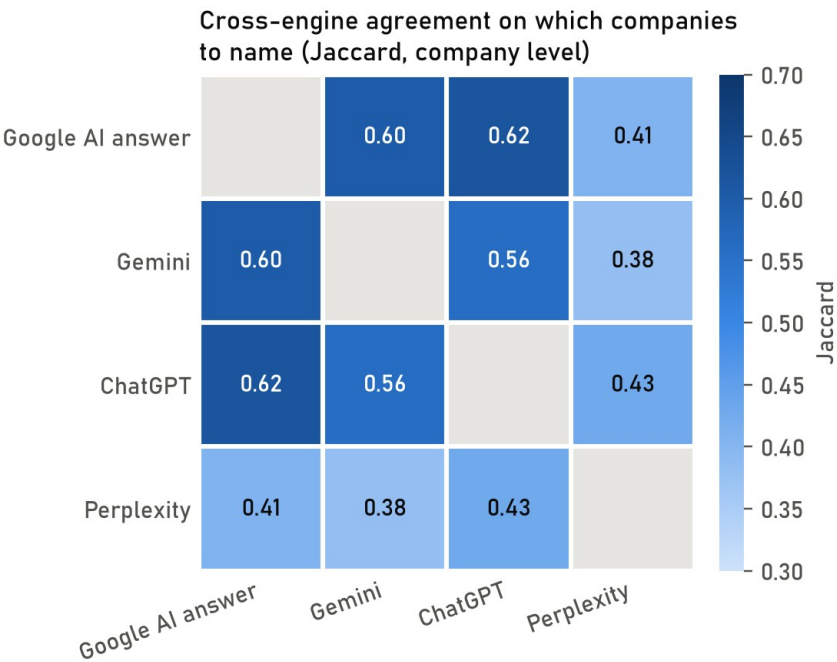


Figure 6. Pairwise agreement between engines on which tracked companies to name (Jaccard index over company sets, n = 14,470 checks per pair).

3.6 Language of the question

Buying questions were asked both in English and in the local language of the country. Local companies were named more often when the question was in English, in all three countries: Estonia 32.8% versus 28.0%, Lithuania 30.5% versus 26.5%, Latvia 32.3% versus 27.4% (Figure 7). The per-country intervals overlap, but the comparison is paired within company, and the paired test is unambiguous: mean difference +4.4 percentage points, 58% of companies individually higher in English, Wilcoxon $p < 0.001$ (n = 361).

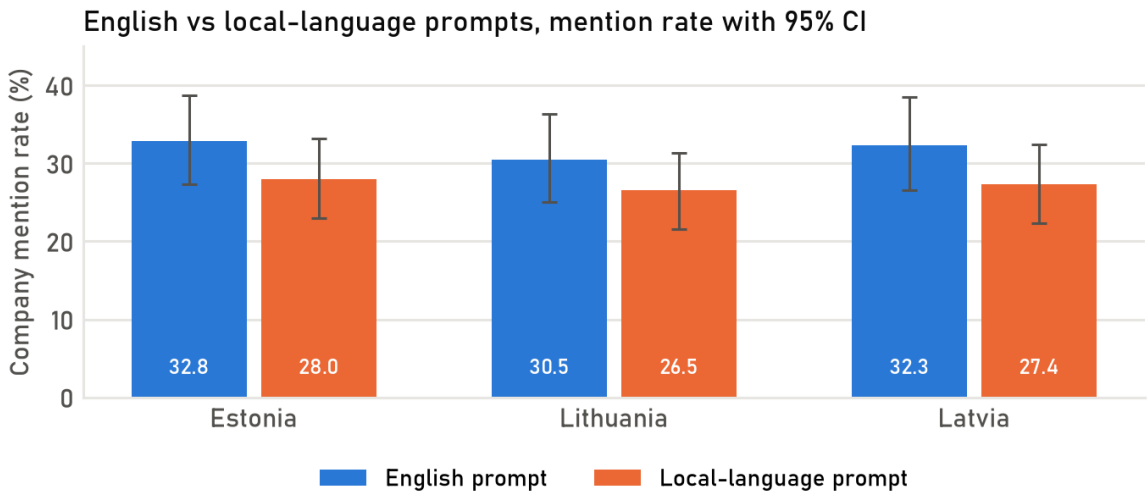


Figure 7. Mention rate for local companies under English and local-language prompts, by country, with 95% intervals. The paired within-company test gives $p < 0.001$.

Part of the mechanism is visible in the background searches the engines run. When the question was asked in a local language, 43% of ChatGPT's background queries, 34% of Gemini's and 7% of Perplexity's were nevertheless in English; Estonian questions drew the most English searching (49% of queries). This is a visibility effect, not an accuracy effect: descriptions given in the local language were slightly more accurate than those in English (90.2% versus 87.6% correct on per-row grades).

3.7 Stability over one week

At company level the measurement is highly reproducible. Mention rates in wave 1 and wave 2 correlate at $r = 0.992$ across 361 companies, with a mean absolute change of 2.5 percentage points and a 95th-percentile change of 7.5 points (Supplement S11). Answer-level agreement on whether a given company is named in a given answer is 89.9% over 28,880 answer pairs.

Individual answers nevertheless churn. For the same prompt and the same engine, a tracked company named in wave 1 was named again in wave 2 in 82.6% of 8,397 cases, and a company not named in wave 1 appeared in wave 2 in 7.1% of 20,483 cases. Retention depends strongly on how visible the company is overall (Figure 8): 90.2% in the top quartile of mention rate, 73.6% and 58.9% in the middle quartiles, 38.7% in the bottom quartile. Comparing the full set of tracked companies named for a prompt-engine pair between waves, 59.8% of lists were identical and 22.7% had no company in common. On raw brand strings, which include the many small companies outside the frame, week-to-week overlap of the name lists averaged only 0.42. The picture is consistent: the leading names in a category are stable and the remainder of the shortlist is re-drawn. A single query is therefore one draw from a distribution, and repeated sampling is the only defensible way to place a company that is not already at the top of its category.

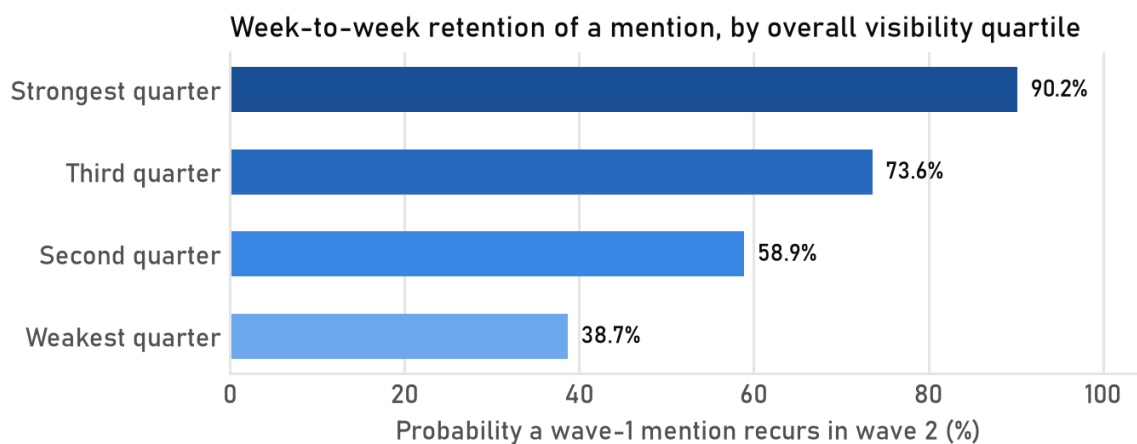


Figure 8. Probability that a company named in wave 1 is named again in wave 2 for the same prompt and engine, by the company's overall visibility quartile ($n = 8,397$ wave-1 mentions).

3.8 What the engines cite

The 176,673 sources cited inside buying answers span 19,615 distinct hosts and are long-tailed (Figure 9): the single most-cited host accounts for 5.1% of citations, the top ten for 13.6%, the top hundred for 27.8% and the top thousand for 58.9%. Company-owned websites make up 8.8% of citations in buying answers. Among third-party hosts (Figure 10), google.com result pages come first (5.1%, an artefact of the retrieval path rather than a source), followed by Reddit (3.3%), the Estonian and Lithuanian business registries ssb.ee and rekvizitai.vz.lt (1.0% each), and the B2B directories ensun.io and clutch.co. 57.2% of cited

hosts sit on a domestic top-level domain; that share is highest for ChatGPT (65 to 67%) and lowest for Google AI answers (42 to 46%).

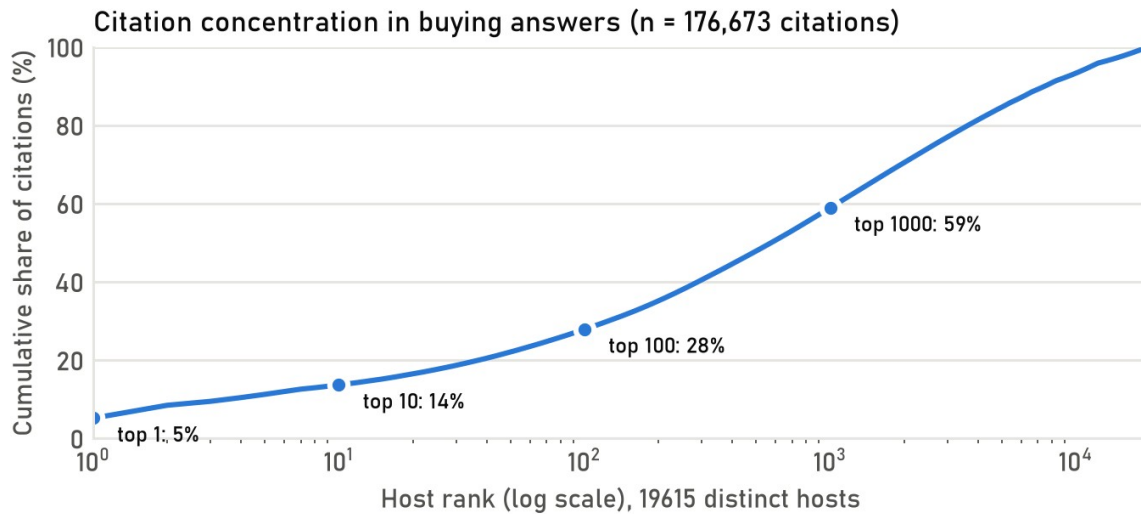


Figure 9. Cumulative share of citations in buying answers against host rank, on a logarithmic axis (n = 176,673 citations over 19,615 hosts).

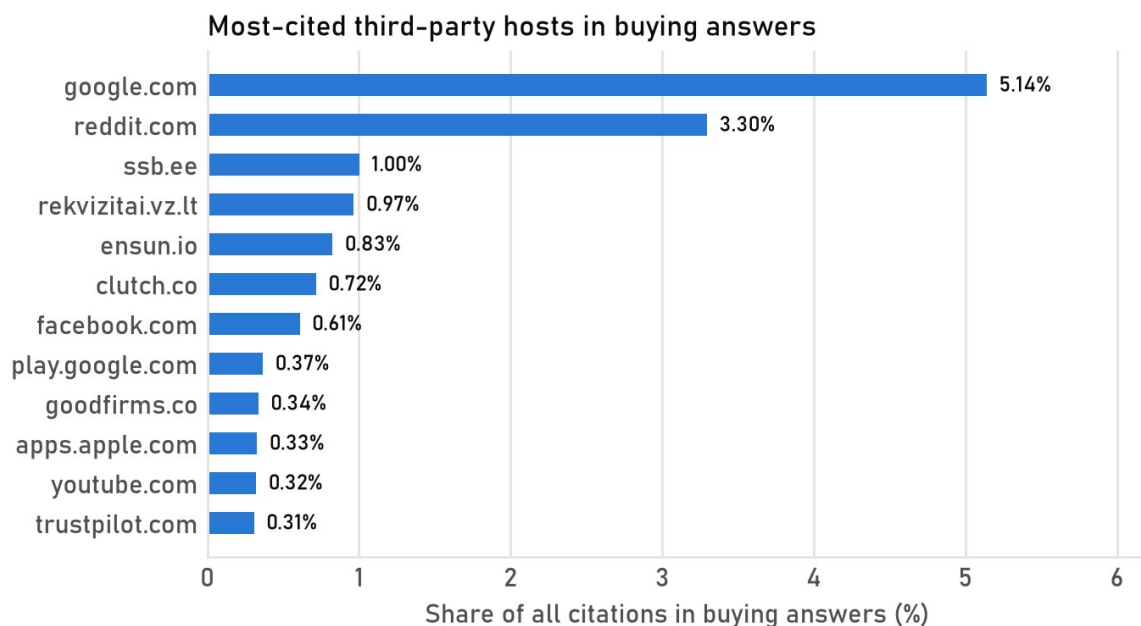


Figure 10. Most-cited third-party hosts inside buying answers, as a share of all such citations. Company-owned domains are excluded from this ranking.

Engines differ in how much they cite and in what kind of source. ChatGPT cites 12.6 sources per buying answer, Google AI answers 10.2, Perplexity 9.8 and Gemini 2.3; Gemini emits any citation at all in only 43.9% of answers, so its source profile rests on a smaller base. By category (Figure 11, wave-1 citations across all prompt types), company-owned sites are the largest identifiable category for ChatGPT, Google AI answers and Gemini (13 to 16%), registries and directories lead for Perplexity (13%), and news and Wikipedia are small for every engine. The "other" category, 61 to 74% of citations, is the long tail of hosts that no simple classifier captures, and it is the honest description of where most of the evidence comes from.

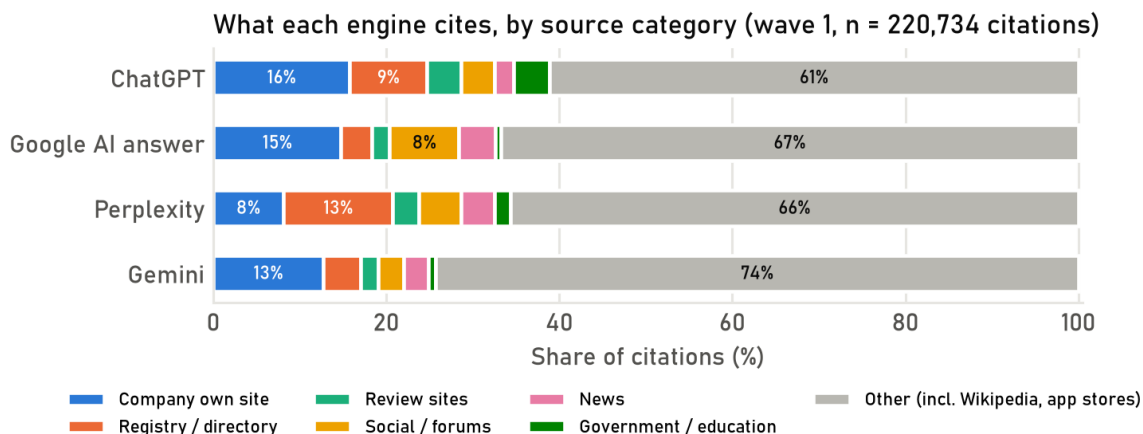


Figure 11. Share of citations by source category and engine, wave 1, all prompt types (n = 220,734). Wikipedia and app stores are folded into "other".

3.9 Accuracy of what the engines say

On the human-anchored grades, 72.2% of AI descriptions of a company were fully correct, 17.8% contained minor errors, 5.3% contained a major error that would mislead a buyer (95% CI 1.7 to 10.4) and in 4.7% the engine did not know the company (Figure 12). The per-row judge grades, which understate major errors, give Google AI answers as the most accurate engine (94.0% correct on 716 answers) against ChatGPT 87.0%, Gemini 87.7% and Perplexity 86.7%, and Estonia (91.4%) and Lithuania (90.1%) ahead of Latvia (84.8%). The characteristic major error is entity blending: facts about a group pasted onto the local legal entity, such as Neste Eesti credited with the Finnish parent's station network or a single county cooperative described as the whole of Coop Eesti.

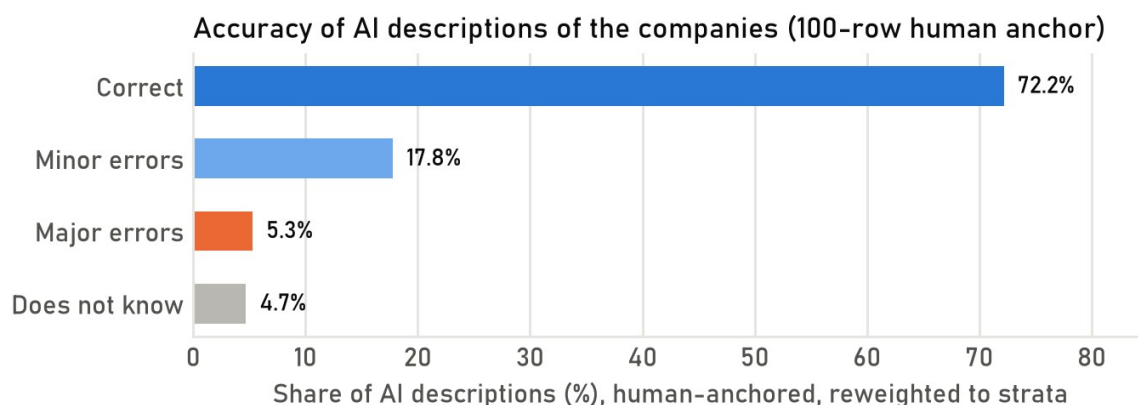


Figure 12. Accuracy of AI descriptions of the tracked companies. Shares from the 100-row human handcheck, graded with live web search and reweighted to population strata.

One methodological result deserves the same prominence as the grades: the error rate is definition-dependent. A judge that treats every claim absent from a two-sentence ground truth as invented reports 15.3% major errors; a judge that counts only contradictions reports 4.4%. The difference is the thinness of the ground truth, not the behaviour of the models (one beverage producer was graded wrong eight times out of eight for correctly mentioning beer and water when its homepage summary said only juice). Any published accuracy figure for AI answers should state which definition it uses.

3.10 What the models know without searching

With web search disabled, ChatGPT held 73.7% of the 358 companies in memory (knows or partial; 95% CI 68.9 to 78.0) and described a wrong entity for 4.2%. The lightweight Gemini model held 60.6% and produced a wrong entity for 19.3%, in several cases confidently substituting a mythological or fictional namesake for a Baltic company. Where the small Gemini model hallucinates, ChatGPT hedges.

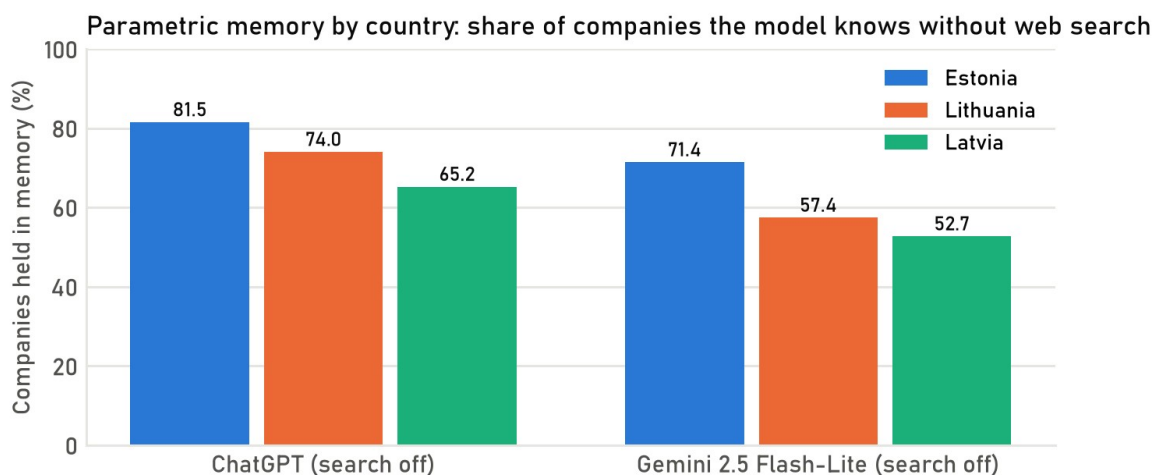


Figure 13. Share of companies held in parametric memory (knows or partial), by country and engine, with web search disabled (n = 358 companies per engine).

This layer contains the only statistically alive country difference in the study. Estonian companies are better represented in memory than Latvian ones in both engines independently: ChatGPT 81.5% versus 65.2% (adjusted $p = 0.015$), Gemini 71.4% versus 52.7% (adjusted $p = 0.006$). Latvia is last in both. Invisibility is, however, mostly not a knowledge problem: 43 of the 56 companies never named in their category are known to at least one engine. Only 13 companies are both unknown and never recommended.

3.11 What predicts being named: SEO authority

The results above describe behaviour without explaining it. To test whether conventional web-authority metrics predict AI visibility, we joined four external measures to every tracked company in October 2026, two months after collection: domain rank (a 0 to 1000 referring-domain-weighted authority score), the count of referring domains, estimated monthly organic search traffic in the home country, and monthly brand-name search volume in the home country, all from DataForSEO. Because company size confounds everything here, each correlation is reported both raw and as a partial Spearman controlling for log revenue.

Pooled across all 355 companies, every metric is positively associated with mention rate, but the ordering is not the one practitioners assume. Organic search traffic is the strongest predictor ($\rho = 0.555$, partial 0.516), followed by brand search volume (0.353, partial 0.286) and domain rank (0.325, partial 0.281). The count of referring domains is weakest (0.229, partial 0.196) and on its own fails to separate the never-named companies from the visible ones at all (median 689 against 950 referring domains, $p = 0.42$). Spam score runs the other way, as expected (partial $\rho = -0.200$, $p = 0.0002$). Within category cells of at least four companies the same ordering holds, with median within-cell correlations of 0.60 for organic traffic, 0.41 for brand volume, 0.40 for domain rank and 0.30 for referring domains, against 0.19 for revenue.

Whose domain is it?

That pooled picture is distorted by a structural feature of the frame. 212 of the 355 companies answer on their own national domain (rimi.lv, telia.ee, ignitis.lt). 102 answer on an international domain, which includes both local arms of foreign groups (Scania Eesti on scania.com, EPAM Sistemas on epam.com) and Baltic-born companies on a .com of their own (Veriff, Oxyllabs, Hostinger). The remaining 41 share a domain with another tracked entity (four Tallink entities on tallink.com, three Bolt entities on bolt.eu, four Enefit, three Coop Eesti). For the international group the authority being measured largely belongs to a global brand rather than to the Baltic company the assistant was asked about.

Splitting on that distinction changes the result substantially. Among the 212 companies on their own national domain every link metric strengthens: domain rank rises to $\rho = 0.481$ (partial 0.403), and referring domains, the weakest metric in the pooled sample, rises to 0.555 (partial 0.471). The gradient is clean and monotonic, from a 16.5% mean mention rate in the lowest authority quartile to 55.1% in the highest, with invisibility falling from 17.0% to zero: not one company in the top authority quartile of its own national domain was never named. Among the 102 companies on an international domain the relationship disappears: domain rank is $\rho = 0.109$ raw and 0.125 partial ($p = 0.21$), and the quartile gradient is flat, running from a 13.1% mean mention rate in the lowest quartile to 20.2% in the highest, with 23.1% of that highest quartile never named at all. The 41 companies on a shared domain behave like the national group rather than the international one (domain rank partial 0.437, referring domains 0.545), so the effect is not about sharing a domain. It is about whether the authority was earned for the entity and the market in the question.

The comparison between the groups is the finding. The national and international groups have effectively the same median domain rank, 312 against 315, and the international group carries more than four times the referring domains, yet it is named less than half as often and is three times as likely never to be named at all. Authority attached to a global brand does not transfer to the local entity. Hostinger's Lithuanian operating company sits on a domain with rank 610 and 118,129 referring domains and was never named in any Lithuanian IT services answer in the study.

Table 4. The three domain groups, with authority and visibility in each

Where the company answers	Companies	Median domain rank	Median referring domains	Mean mention rate	Never named
Its own national domain	212	312	326	34.7%	9.4%
An international domain	102	315	1,422	15.4%	28.4%
A domain shared with a sibling	41	411	1,384	38.1%	17.1%

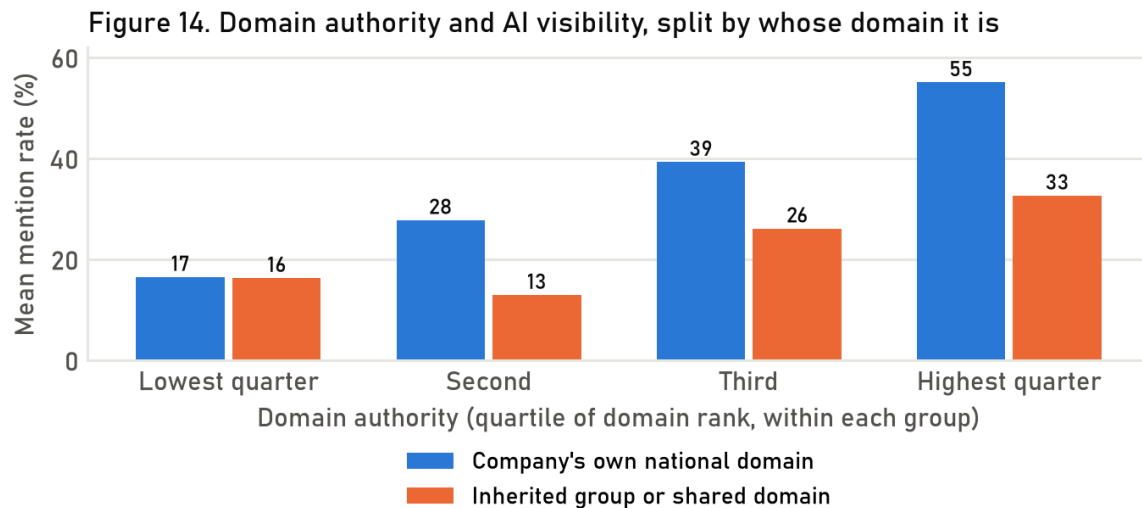


Figure 14. Mean mention rate by domain-authority quartile, split by whether the domain belongs to the company itself or to a parent group or sibling entity. Quartiles are computed within each group. The figure pools the international and shared groups that Table 4 separates.

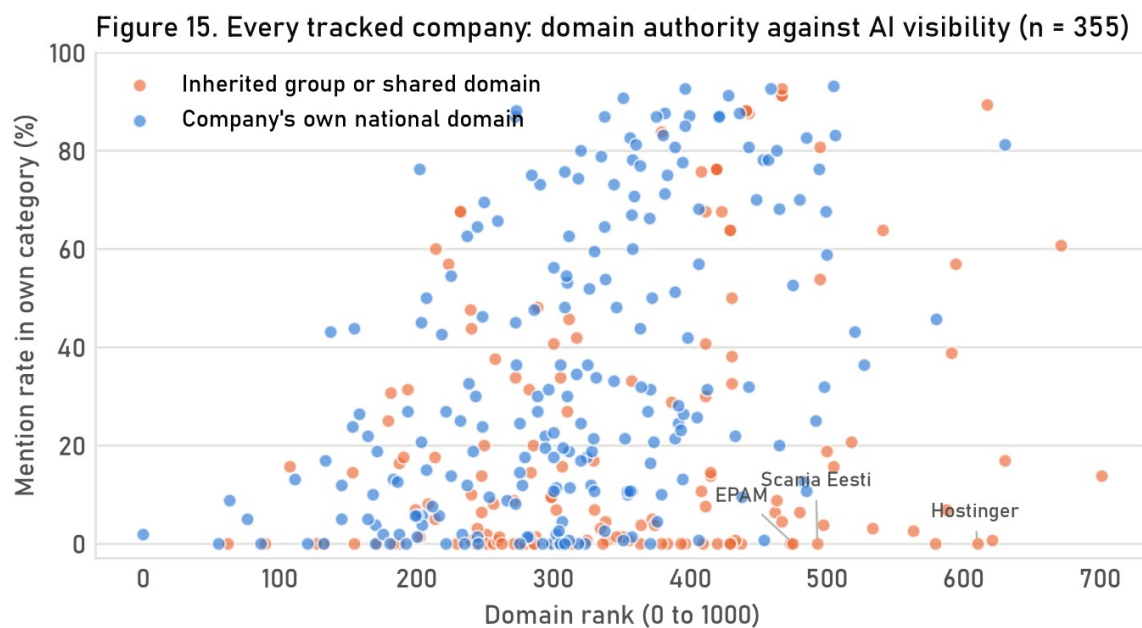


Figure 15. Domain rank against mention rate for all 355 tracked companies. Labelled points are high-authority domains whose Baltic operating company was never named in its own category.

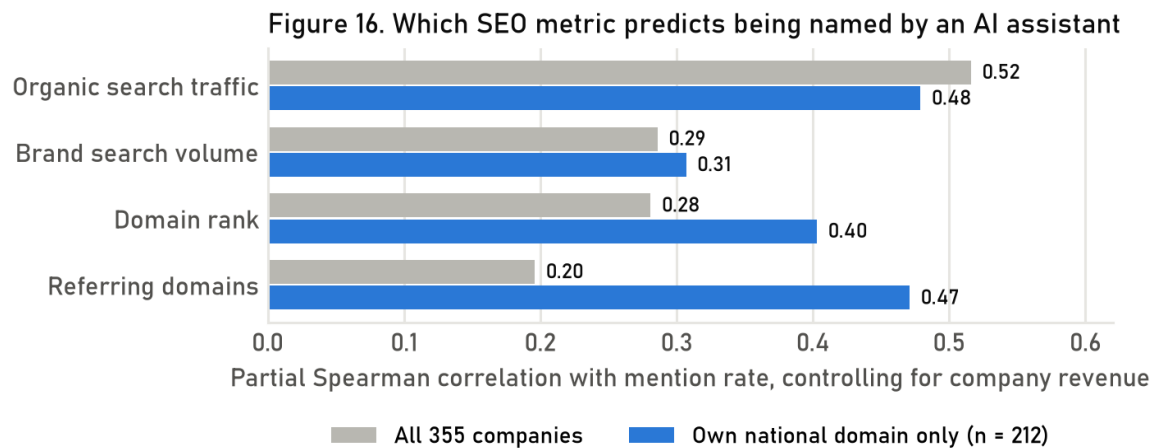


Figure 16. Partial Spearman correlation between each external metric and mention rate, controlling for company revenue, for the full sample and for companies on their own national domain.

A model, with the caveats it needs

With real covariates available, the regression deliberately withheld from the first analysis becomes possible. A logistic model of the probability that a company is never named, on standardised predictors and restricted to companies with their own domain ($n = 212$), gives odds ratios of 0.40 per standard deviation of domain rank ($p = 0.022$), 0.37 for brand search volume ($p = 0.005$) and 0.52 for revenue ($p = 0.048$); organic traffic loses significance once domain rank is in the model, which is collinearity rather than irrelevance. Pseudo R-squared is 0.220, against 0.148 for the same model on the full sample. Roughly four fifths of the variation in who gets named therefore remains unexplained by size, authority and brand demand combined, which is consistent with the shortlist churn documented in section 3.7.

Redundant is not the same as irrelevant

Every correlation above holds company revenue constant. None of them holds organic traffic constant, and that distinction decides what the authority metrics are worth to a practitioner. Links cause rankings and rankings cause traffic, so traffic is not a confounder of the link metrics, it is the channel through which they act. Entering it as a second control therefore answers a narrower question than the one above: does this metric say anything that organic traffic has not already said? A metric that collapses under that control is redundant rather than irrelevant.

A second pass in October 2026 ran that test over eight metrics, adding raw backlink count and the number of keywords a site ranks for to the four already reported. Correlations are partial Spearman against mention rate on the same 355 companies, computed as the Pearson correlation of the rank residuals; the pooled figures earlier in this section correlate the same residuals with Spearman, and the two agree to within 0.01 on every metric. Confidence intervals resample whole niches rather than companies, so firms inside a correlated sector move together, and p-values are Benjamini-Hochberg adjusted across the eight metrics.

Table 5. What each SEO metric predicts before and after controlling organic traffic

Metric	Controlling revenue	Also controlling organic traffic	q
Organic keywords ranked	+0.534	+0.173	0.008
Organic traffic	+0.518	it is the control	-
Backlinks, raw count	+0.322	+0.104	0.085

Metric	Controlling revenue	Also controlling organic traffic	q
Domain rank	+0.289	+0.044	0.523
Brand search volume	+0.282	+0.147	0.022
Referring domains	+0.194	+0.040	0.523
Spam score	-0.204	-0.103	0.085
Registry share of coverage	-0.260	-0.128	0.042

Every link metric collapses. Domain rank falls from 0.289 to 0.044, referring domains from 0.194 to 0.040 and raw backlink count from 0.322 to 0.104, and none of the three is distinguishable from zero after correction. Two metrics survive. Keyword coverage holds at 0.173 ($q = 0.008$) and is also the strongest single metric in the dataset before any control at 0.534. Brand search volume holds at 0.147 ($q = 0.022$), and it is the only measure on the list that is not downstream of the company's own site. Referring domains is weak even on its total effect, since the niche-level bootstrap puts its interval at -0.00 to +0.37. Organic traffic itself holds separately in each country (Latvia 0.321, Estonia 0.508, Lithuania 0.633), so no single market is driving it.

This does not overturn the model in the previous subsection, and the reconciliation is worth setting out because the two results were briefly read as contradicting each other. That logistic model already contains organic traffic, which is why traffic reads as non-significant in it, and its domain rank odds ratio of 0.399 ($p = 0.022$) reproduces exactly on re-running. The collapse in Table 5 is a difference of sample, not a correction. Inside the national-domain group the authority effects start much higher and only partly fade once traffic is controlled: domain rank 0.412 falling to 0.124 ($p = 0.071$), referring domains 0.489 to 0.167 ($p = 0.015$), keyword coverage 0.501 to 0.173 and brand search volume 0.304 to 0.123. Pooled across all 355, which includes the international group where authority predicts nothing, both link metrics land near 0.04. Link authority therefore keeps a small independent association with being named where the domain belongs to the entity in the question, and none where it does not.

One further specification is reported because it fails informatively. Adding keyword coverage to the own-domain logistic model leaves domain rank at 0.399 ($p = 0.030$) but sends organic traffic to an odds ratio of 4.750 ($p = 0.012$) and keyword coverage to 0.214 ($p = 0.017$): the two swap signs. Traffic and keyword coverage correlate at 0.88 in this subset and only 20 of the 212 companies were never named, roughly four events per predictor, so the pair cannot be separated at this sample size. The continuous outcome in the preceding paragraph is the better evidence here, and the two ranking metrics should be read as one signal rather than two.

The pages the engines cite carry almost no link authority

The same question can be put to the pages rather than to the companies. From the 220,735 cleaned citations, those belonging to the identity answers were kept, every citation pointing at the subject company's own site was dropped, and 24,444 distinct third-party URLs remained. A stratified sample of 300 was drawn, weighted so that the hosts dominating citations and the long tail are both represented, and page-level link data was pulled from DataForSEO.

29 of the 300 pages carry any detected backlinks at all: 9.7%, Wilson interval 6.8% to 13.5%, with only 9 of 300 above a page link rank of 100 and a citation-weighted share of 10.0%. What the engines quote is company register records, directory profiles and credit-report subpages, at URLs of the form ssb.ee/en/11411278-TERE-AS/liabilities-debts-court or

rekvizitai.vz.lt/en/company/baltisches_haus/report/. Two checks were run against the obvious objections. Fifteen localised duplicates of the uk.trustpilot.com and glassdoor.ie kind were re-tested at canonical form, of which two gained links, moving the figure to 10.3%. And among the pages that were cited, page link rank does not predict citation count (0.106) or the number of distinct engines citing the page (0.082). Two limits travel with this. A commercial crawler finding no backlinks is not proof that none exist, and this is a sample of cited pages, so it cannot establish that authority has no influence on retrieval; that would need a matched set of comparable pages that were not cited, drawn from each engine's own retrieval rather than from a search index.

A finding withdrawn

The same pass produced, and then withdrew, a result that is recorded here rather than omitted. The share of a company's third-party coverage made up of registry and directory records correlated with mention rate at -0.260 controlling revenue, with a monotonic quartile gradient from a 45.0% mean mention rate in the least registry-dependent quarter of companies to 17.8% in the most. It does not survive controls that measure online presence directly: -0.168 adding brand search volume, -0.128 adding organic traffic, and -0.084 at $p = 0.11$ adding both. The reason is construct overlap. "Most of what exists about this company is its company registration" and "assistants rarely name this company" are both descriptions of being obscure online, and a revenue control does not separate them, because a large company can have very little online presence. The gradient is real and the interpretation was not.

Three limits travel with this layer. The metrics were collected in October 2026 and the answers in August, so a company that changed materially in between is mismeasured. Domain rank and referring domains are third-party crawler estimates, not ground truth. And association is not direction: a well-known company attracts links, traffic and AI mentions at the same time, so these results identify which measurable signals accompany visibility, not which actions produce it.

3.12 How many answers a visibility check needs

Every rate in this report is estimated from 160 answers per company. A practitioner who checks one answer, which is how most AI visibility checks are performed, is estimating the same quantity from a sample of one. The cost of that can be measured directly here, because the full-sample rate supplies a reference value. The 56,800 company-answer observations, 355 companies with exactly 160 answers each, were resampled with replacement, drawing k answers 600 times per company and comparing each draw with that company's reference rate and band.

Table 6. Accuracy of a short visibility check against the 160-answer reference rate

Answers checked	Mean absolute error	Visible company reads as 0%	Wrong visibility band
1	24.2 pts	50.0%	56.6%
2	17.5 pts	31.0%	46.1%
3	14.1 pts	21.3%	40.8%
5	10.9 pts	11.3%	35.2%
10	7.7 pts	3.2%	27.5%
20	5.4 pts	0.3%	19.7%
40	3.8 pts	0.0%	13.1%
80	2.7 pts	0.0%	8.2%

A single answer puts a company in the wrong band 56.6% of the time and reads a genuinely visible company as a flat zero half the time. The error falls with the square root of the sample, so ten answers still misband 27.5% and forty 13.1%. The engine dimension compounds it. The mean gap between a company's best and worst engine is 22.4 points, the median best-to-worst ratio is 2.6 times where both are non-zero, and 27.7% of the companies named at least once are absent on at least one engine, while per-company rates still correlate 0.83 to 0.95 across engines. The engines agree on who is strong and disagree on any single answer, so a one-engine, one-answer check is not a small version of this study. It is a different measurement.

4. Discussion

4.1 Three slots, drawn from thousands

The central descriptive result is the combination of a shallow choice set and a very deep pool. Any single answer offers about three names, and on Google's AI answer usually one; across answers the engines drew on more than eleven thousand companies, half of them only once. The two facts together explain most of what follows. Because the shelf is short, being large does not guarantee a place on it, and fewer than one tracked company in ten is reliably present. Because the pool is deep, the lower positions on the shelf are re-drawn from week to week, while the leading names, which the engines converge on from many sources, stay put. The stability results in section 3.7 are the quantitative form of that argument: retention rises monotonically with overall visibility.

4.2 What did not survive scrutiny

Two candidate findings were removed during analysis and are recorded here so that they are not rediscovered. First, the number of citations to a company's own website correlates with its mention rate at $\rho = 0.83$, which looked like the study's most actionable result. It is circular: the engine cites the company's site in the same answer in which it names the company. Tested cleanly, using citations from the separate identity answers to predict mention rate in buying answers, the correlation falls to $\rho = 0.15$ ($p = 0.004$). The only defensible statement is the extreme case: the 13 companies whose own site was never cited even when the engine was asked about them directly average a 5.5% mention rate. Second, an earlier version of the country comparison, based on crude rates, produced a national ranking; standardisation removed it, and we report the null.

The never-named list also needs a caveat before it is quoted. Of the 56 companies, many are registry entities that no customer would ask an assistant about (a retailer's international sourcing arm, a utility's generation subsidiary, a rail freight operator, several franchise operating companies, shipyards and bunkering firms), and some are category mismatches in which zero is the correct answer (an identity verification vendor and a hosting provider tested against "best IT services company" prompts, a beauty retailer tested against pharmacy prompts). The 15.8% figure is computed correctly under the pre-registered definition, but the interpretable core is the 13 companies that are both unknown and unrecommended, plus the hand-checkable cases in Appendix B.

4.3 Language and country

The language effect is small in absolute terms but consistent, paired and highly significant, and the captured background queries suggest a mechanism: the engines partly search in English regardless of the question's language, so English-language evidence about a company is more likely to be retrieved. That the local-language descriptions are nonetheless slightly more accurate suggests the gap is one of retrieval breadth, not of understanding. The absence of a country effect on recommendation, alongside a clear country effect on parametric memory, is worth stating carefully: the assistants know Estonian companies better than Latvian ones, but once they search, they recommend companies from all three countries at indistinguishable rates.

4.4 Implications for measurement

Three practical conclusions follow for anyone measuring AI visibility. Report per engine, because the engines disagree with each other more than they agree. Sample repeatedly, because a single answer is one draw from a distribution whose week-to-week retention below the top quartile is a coin flip. And state the error definition, because the accuracy figure can move by a factor of three depending on it. The study's own instrument meets these conditions and, with a wave-to-wave correlation of 0.992 at company level, is stable enough to repeat annually as an index.

5. Limitations

- AI answers drift. This is a dated snapshot from August 2026, not a permanent ranking, and it is designed to be repeated.
- Visibility measures how an assistant represents a company, not how good the company is. A low mention rate is not a judgement of quality.
- The frame is the top of the revenue distribution. Smaller companies named by the engines are often correct answers, so the share of slots held by frame companies is a statement about decoupling from size, not about error.
- Some companies were tested under their registry entity rather than the brand customers know, which understates them; the alias table mitigates but does not eliminate this.
- English prompts stand in for domestic consumer categories where a real buyer would ask in the local language, so consumer-category recommendation rates are directional.
- Four engines, not five. Claude was not measured.
- Gemini cites sources in fewer than half its answers, so its source profile rests on a smaller base and source rankings are reported per engine.
- Ground truth is homepage-thin for multi-line companies and registry-thin for 48 companies. Automated accuracy judges under-catch entity-boundary errors, so per-engine and per-country accuracy shares are ceilings and only the human-anchored shares are headline figures.
- Parametric memory grades are single-pass model labels without human calibration, from one wave, and the two engines differ in model tier and access path.
- Background search queries could not be captured for Google AI answers.

- The choice-set counts in section 3.1 rely on normalised brand strings; residual name variants could inflate distinct counts slightly, so 11,136 is best read as an order of magnitude with a floor of several thousand.
- The SEO metrics in section 3.11 were collected in October 2026, two months after the answers, and are third-party estimates. They identify signals that accompany visibility, not causes of it.
- Organic traffic is treated in section 3.11 as a mediator of the link metrics rather than a confounder of them. That is a causal assumption, conventional for links and rankings but still an assumption, and it is what licenses reading the collapse of the link metrics as redundancy rather than as absence of effect.
- Page-level backlink data in section 3.11 records what a commercial crawler detected. No backlinks detected is not proof of none, and the sample covers cited pages only, with no matched comparison group of pages that were not cited.

6. Conclusion

When a Baltic buyer asks an AI assistant for a recommendation, the answer is a shortlist of about three companies, drawn from a pool of more than eleven thousand, in which the largest firms hold roughly a quarter of the slots and the leading names are stable while the rest are re-drawn weekly. The engines differ in how many companies they name, how often they name the large ones, whom they name, and what they cite, so no single figure describes "AI visibility" for a company; four do. Country makes no difference to recommendation once category mix is controlled, though the models know Estonian companies better than Latvian ones. Questions in English surface local companies more often than questions in the local language. Roughly a quarter of AI descriptions contain some error and one in twenty contains a serious one, with the exact figure depending on how an error is defined. Conventional SEO authority does predict visibility, but only for companies that own the domain being measured, and only as a proxy: authority sitting on an international group domain does not transfer to the local entity, and pooled across the sample no link metric survives a control for the organic traffic it produces. Keyword coverage and brand search demand are the two signals that do survive it, and even the best combination of size, authority and brand demand leaves most of the variation unexplained. The instrument that produced these results is reproducible at $r = 0.992$ and is intended to be run annually.

7. Ethical note and reporting rule

Aggregate results, category leaders and per-company mention rates are reported by name, because they are counts of public behaviour by public systems about public companies. Individual misdescriptions of a company by an engine are reported only in anonymised or permissioned form, because reproducing a false claim about a named company, even to criticise it, propagates it. No personal data was collected. The study was funded by WebGlazer, which sells AI visibility services; the analysis code, data files and validation samples are provided so that the results can be checked independently of that interest.

References

1. Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 209–212.
2. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300.
3. Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall.
4. Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2), 37–50.
5. Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621.
6. Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18(1), 50–60.
7. Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83.
8. WebGlazer (2026). *Baltic AI Visibility Index 2026: analysis code and derived data* (scripts/baltic-study, raw/baltic-study/derived). Internal repository; supplementary files S1–S21 accompany this report.

Appendix A. Category leaders

Most-named tracked company per country and category, for cells with reliable coverage (small cells excluded; the complete list including small cells is Supplement S2). Mention rate is the share of the cell's 160 answers per company in which the company is named; the interval is a Wilson 95% interval; n is the number of tracked companies in the cell.

Country	Category	Leader	Rate %	95% CI	n
Estonia	banks	Lhv Pank	93.1	88.1 to 96.1	6
Estonia	car sales import	Auto 100	10.0	6.2 to 15.6	6
Estonia	construction infrastructure	Merko Ehitus Eesti	43.8	36.3 to 51.5	7
Estonia	dairy	Valio Eesti	53.8	46.0 to 61.3	3
Estonia	diy building materials retail	Kesko Senukai Estonia	77.5	70.4 to 83.3	3
Estonia	electrical equipment manufacturing	Harju Elekter	45.6	38.1 to 53.4	5
Estonia	electronics manufacturing	Gpv Estonia	37.5	30.4 to 45.2	6
Estonia	energy retail sales	Enefit	67.5	59.9 to 74.3	3
Estonia	energy utilities grid heating	Elektrilevi	48.1	40.5 to 55.8	5
Estonia	fmcg food distribution	Sanitex	32.5	25.7 to 40.1	3
Estonia	food manufacturing	Eesti Pagar	30.0	23.4 to 37.5	5
Estonia	fuel gas wholesale trading	Jetoil	36.2	29.2 to 43.9	4
Estonia	fuel stations retail	Alexela	83.1	76.6 to 88.1	4
Estonia	grocery retail	Selver	80.0	73.1 to 85.5	10
Estonia	insurance	Swedbank P&c Insurance	53.8	46.0 to 61.3	3
Estonia	metals scrap	Cronimet Nordic	75.0	67.8 to 81.1	3
Estonia	restaurants	Hesburger	0.6	0.1 to 3.5	3
Lithuania	agri commodities grain	Linas Agro	70.6	63.2 to 77.1	3
Lithuania	agri inputs machinery	Baltic Agro	36.2	29.2 to 43.9	4
Lithuania	airlines aviation	GetJet Airlines	5.0	2.6 to 9.6	3
Lithuania	auto parts tyres	Inter Cars Lietuva	31.2	24.6 to 38.8	3
Lithuania	banks	Seb bankas	92.5	87.3 to 95.7	3
Lithuania	dairy	Žemaitijos Pienas	53.1	45.4 to 60.7	3
Lithuania	ecommerce marketplaces	Pigu	70.0	62.5 to 76.6	5
Lithuania	energy retail sales	Enefit	88.1	82.2 to 92.3	3
Lithuania	energy utilities grid heating	Energijos skirstymo operatorius	51.2	43.6 to 58.9	3
Lithuania	fertilizers chemicals plastics	Achema	25.0	18.9 to 32.2	5
Lithuania	fmcg food distribution	Sanitex	50.0	42.3 to 57.7	5
Lithuania	food manufacturing	Biovela-utenos Mėsa	26.9	20.6 to 34.2	7
Lithuania	fuel stations retail	Viada Lt	80.0	73.1 to 85.5	4
Lithuania	furniture	Narbutas International	3.8	1.7 to 7.9	5
Lithuania	grocery retail	Maxima Lt	80.6	73.8 to 86.0	5

Country	Category	Leader	Rate %	95% CI	n
Lithuania	industrial services	Rndv Group	0.6	0.1 to 3.5	3
Lithuania	it electronics distribution	Acc Distribution	60.0	52.3 to 67.3	3
Lithuania	parcel postal	Dpd Lietuva	89.4	83.6 to 93.3	4
Lithuania	pharma medical manufacturing	Thermo Fisher Scientific Baltics	16.9	11.9 to 23.4	3
Lithuania	pharma medical wholesale	Limedika	50.0	42.3 to 57.7	4
Lithuania	pharmacies drugstores	Eurovaistinė	80.6	73.8 to 86.0	3
Lithuania	rail	Ltg Link	73.1	65.8 to 79.4	3
Lithuania	road freight logistics	Girteka Logistics	40.6	33.3 to 48.4	9
Lithuania	telecom	Telia Lietuva	90.6	85.1 to 94.2	3
Latvia	banks	Citadele banka	87.5	82.2 to 91.4	3
Latvia	beverages alcohol	Cēsu alus	17.5	12.4 to 24.1	3
Latvia	car sales import	Sabiedrība ar ierobežotu atbildību Moller Auto Latvia	43.1	35.7 to 50.9	4
Latvia	construction infrastructure	Upb	36.2	29.2 to 43.9	8
Latvia	energy utilities grid heating	Sadales tīkls	51.9	44.2 to 59.5	4
Latvia	food manufacturing	Sabiedrība ar ierobežotu atbildību Orkla Latvija	46.2	38.7 to 54.0	4
Latvia	forestry	Latvijas valsts meži	21.9	16.2 to 28.9	4
Latvia	fuel gas wholesale trading	Orlen Latvija	23.8	17.8 to 30.9	4
Latvia	fuel stations retail	Sabiedrība ar ierobežotu atbildību Neste Latvija	86.9	80.8 to 91.3	3
Latvia	gambling gaming	Alfor	10.0	6.2 to 15.6	3
Latvia	grocery retail	Rimi Latvia	83.8	77.3 to 88.7	3
Latvia	hospitals	Valsts sabiedrība ar ierobežotu atbildību Paula Stradiņa klīniskā universitātes slimnīca	21.9	16.2 to 28.9	3
Latvia	it electronics distribution	Elko Grupa	60.6	52.9 to 67.9	3
Latvia	peat mining	Sabiedrība ar ierobežotu atbildību Laflora	64.4	56.7 to 71.4	3
Latvia	road freight logistics	Kreiss	15.0	10.3 to 21.3	4
Latvia	wood timber	Sabiedrība ar ierobežotu atbildību Gaujas Koks	16.9	11.9 to 23.4	10

Appendix B. Companies never named in their own category

The 56 tracked companies with zero mentions across all 160 own-category answers, with FY2025 revenue in EUR millions. See section 4.2 before quoting this list: many entries are registry entities or category mismatches rather than consumer-facing brands that the assistants overlook.

Country	Company	Category	Revenue EUR m
Estonia	Magnetic Mro	airlines aviation	222

Country	Company	Category	Revenue EUR m
Estonia	Scania Eesti	car sales import	120
Estonia	Auto-Bon Baltic	car sales import	141
Estonia	Connecto Infra	construction infrastructure	119
Estonia	Orien Trade	fmcg food distribution	116
Estonia	Panamir	food manufacturing	175
Estonia	Energia Nord	fuel gas wholesale trading	145
Estonia	Arte Bunkering	fuel gas wholesale trading	262
Estonia	Tulundusühistu Tartu Tarbijate Kooperatiiv	grocery retail	118
Estonia	A1m	grocery retail	99
Estonia	Tallinn Shipyard	industrial services	47
Estonia	LTH-Baas	industrial services	149
Estonia	Mbr Metals	metals scrap	290
Estonia	Riigi Kinnisvara	real estate	177
Estonia	Baltic Restaurants Estonia	restaurants	38
Estonia	Premier Restaurants Eesti	restaurants	42
Estonia	Veriff	software it services	78
Estonia	Hansaliin	staffing business services	45
Estonia	HT Laevateenindus	staffing business services	32
Estonia	Nortal Capacity Services	staffing business services	51
Lithuania	FL Technics	airlines aviation	204
Lithuania	Sct Lubricants	auto parts tyres	205
Lithuania	Orbio world	ecommerce marketplaces	264
Lithuania	Ignitis gamyba	energy utilities grid heating	452
Lithuania	Retal Lithuania	fertilizers chemicals plastics	114
Lithuania	Maxima International Sourcing	fmcg food distribution	295
Lithuania	Uždaroji Vičiūnai ir Ko	fmcg food distribution	444
Lithuania	Roquette Amilina	food manufacturing	222
Lithuania	Vakarų Baltijos laivų statykla	industrial services	68
Lithuania	Vakarų Laivų Remontas	industrial services	71
Lithuania	NT Service	it electronics distribution	227
Lithuania	Vinted Go	parcel postal	283
Lithuania	Intersurgical	pharma medical manufacturing	189
Lithuania	Hollister Lietuva	pharma medical manufacturing	193
Lithuania	Stada Baltics	pharma medical wholesale	163
Lithuania	Ltg Cargo	rail	298
Lithuania	Baltisches Haus	real estate	43
Lithuania	Stiemo	renewable energy construction	118
Lithuania	Premier Restaurants	restaurants	69

Country	Company	Category	Revenue EUR m
Lithuania	Nostrada	road freight logistics	179
Lithuania	oxylabs	software it services	139
Lithuania	Hostinger operations	software it services	271
Lithuania	Epam Sistemas	software it services	85
Latvia	Sabiedrība ar ierobežotu atbildību Latakko	auto parts tyres	91
Latvia	MV Group Distribution LV	beverages alcohol	97
Latvia	Sabiedrība ar ierobežotu atbildību Acbr	construction infrastructure	96
Latvia	Baltreids	forestry	48
Latvia	Valsts Latvijas Loto	gambling gaming	106
Latvia	Drogas	pharmacies drugstores	117
Latvia	Rīgas namu pārvaldnieks	real estate	82
Latvia	Sabiedrība ar ierobežotu atbildību Plesko Real Estate	real estate	46
Latvia	Valsts Valsts nekustamie īpašumi	real estate	47
Latvia	Sabiedrība ar ierobežotu atbildību Premier Restaurants Latvia	restaurants	88
Latvia	toplogistic	road freight logistics	106
Latvia	Imk 365	staffing business services	178
Latvia	Latgran	wood timber	221

Appendix C. Supplementary files

The following files accompany this document in the supplement folder. All are UTF-8 CSV.

File	Contents
S1_company_scores.csv	Per-company mention rate, band, per-engine, per-language and per-wave rates (361 rows)
S2_category_leaders.csv	Most-named company per country-category cell with intervals and cell size (117 rows)
S3_engine_mention_rates.csv	Mention rate per engine with 95% CI
S4_engine_agreement.csv	Pairwise engine agreement, both-named and either-named shares, Jaccard
S5_country_rates.csv	Per-country mean mention rate, invisible share, crude and standardised rates
S6_country_tests.csv	Pairwise country differences, CIs, raw and BH-adjusted p-values
S7_language_gap.csv	English vs local-language mention rate per country with CIs
S8_invisible_companies.csv	All 355 companies with band assignment and shrunk rates
S9_source_categories.csv	Citation share by source category and engine (wave 1)
S10_top_hosts.csv	Most-cited hosts per engine (wave 1)
S11_reproducibility.csv	Wave-to-wave correlation, mean and p95 absolute change, answer-level agreement
S12 to S14	Accuracy by engine, country and language (per-row judge grades)

File	Contents
S15_parametric_grades.csv	No-search probe verdicts per company and engine (716 rows)
S16_buyer_prompt_book.csv	All buyer prompts in four languages
S17_company_frame.csv	Company frame with category, revenue, register source and domain
S18_seo_metrics.csv	Domain rank, referring domains, spam score, organic traffic, keyword coverage and brand search volume per company, with the domain-group flag (collected October 2026)
S19_sampling_curve.csv	Accuracy of a short visibility check against the 160-answer reference rate, k from 1 to 80 (section 3.12)
S20_transfer_map.csv	Partial Spearman of eight SEO metrics against mention rate, with and without the organic-traffic control, with niche-bootstrap intervals (section 3.11)
S21_cited_page_ranks.csv	Page-level link data for the 300 sampled third-party cited pages (section 3.11)