AI VISIBILITY · INTERNATIONAL SEO · B2B STRATEGY · AI SEARCH

Does your AI visibility transfer across geographies?

AI platforms cite different sources in different countries, may run different retrieval configurations or model deployments by market, and carry training-data bias that systematically favours English-language content. AI visibility measured from a single location is a partial picture.

AIVIARA RESEARCH · MAY 2026

A B2B software brand spends six months building AI citation presence through targeted comparison pages and review platform coverage. The citation rate improves measurably. Then someone runs the same queries from a German IP. Different sources. Different brand recommendations. Some of the cited brands don't even operate in Germany.

That gap is not a monitoring bug. AI platforms are designed to behave differently by geography, they were trained on data that is not distributed evenly around the world, and the product that exists in the US market is not always the same product that exists in Germany or Japan. The evidence suggests the gap is structural rather than incidental. AI citation presence in your home market is a starting point. Whether it transfers depends on factors that most visibility tracking programs don't yet touch.

The platforms are built to behave differently by location

The geographic differentiation in AI engines is documented at the API and product level, not just inferred from output patterns.

Perplexity's official documentation confirms an explicit `user_location` API filter. It accepts country, city, region, and coordinates, and adjusts search results accordingly. Perplexity describes the filter as designed for queries involving local businesses, regional news, and location-specific information. OpenAI's web search API accepts geographic context as well, including country, city, region, and timezone, to refine what the retrieval component returns.

© 2026 Aiviara Research

Neither company publishes a detailed specification of how heavily these signals are weighted downstream. But the existence of these parameters is itself informative. Geographic differentiation is a first-class input, not an afterthought in the serving layer.

Google AI Overviews is a further illustration of how far the divergence extends. As of May 2025, Google confirmed AI Overviews live in 200-plus countries across 40-plus languages. But Google stated that the US rollout uses a customized Gemini 2.5 implementation not yet available in all markets. France had not received the AI Overviews rollout as of May 2025, where strict neighboring rights legislation has complicated Google's expansion into that market. According to Stackmatix's March 2026 analysis, the trigger rate varies by country too: the US sees AI Overviews in around 15% of searches, the UK at roughly 12%, Germany at 10%, Japan at 8%, and Brazil at 6%.

A brand that built its citation presence around Google AI Overviews is working with a product that fires less frequently in every market it might want to reach outside the US, may run a different model configuration in many of them, and had not launched at all in France as of May 2025.

The training data starts from an uneven baseline

The geographic differentiation in AI engines is visible in their outputs. Its deeper source is the data those models were trained on.

Manvi, Khanna, Burke, Lobell, and Ermon at Stanford tested GPT-4, Gemini Pro, Mixtral, and Llama 2 on geospatial prediction tasks. They found systematic errors with a correlation up to 0.70 indicating bias against lower-income regions, alongside disproportionate US and Western references throughout model outputs. This was published in February 2024 and revised in October 2024. The bias is not an accident of any particular model choice. It reflects training corpora that skew heavily toward English-language, Western, high-income-country sources. The Common Crawl corpus, which underpins most large model training, is strongly weighted toward English content. A 2025 Nature article characterised AI models as "geared towards the needs of English-speaking people in high-income countries" as a direct consequence.

Janowicz et al. ran 200 queries asking GPT-5.1 "Name a country, please." The model replied with Japan 168 out of 200 times. The study was not identifying a malfunction. It was testing how geographic priors become embedded from training data. When prompted with apparently neutral language, the model's learned distribution of what "country" means surfaces a default. Minor syntactic changes to the prompt produced dramatic shifts in which country was named. Geographic defaults sit in the model's probability structure, not in any

© 2026 Aiviara Research

deliberate design choice that could be cleanly reversed.

For a B2B brand trying to appear in AI-generated answers for buyers in Mexico, South Korea, or Poland, this is the starting condition. The training-data baseline was set before the retrieval system ever ran a query.

What the citation data actually shows by country

The most comprehensive published study on cross-geography AI citation behaviour is Johannes Beus's work at SISTRIX, published in May 2026. The methodology covers 82,619 prompts across 1,548,213 snapshots over 17 weeks, across Germany, the US, the UK, Italy, Spain, and France, testing Google AI Overviews, Google AI Mode, and ChatGPT Search.

Two findings are most relevant for brands tracking AI presence across markets.

First, ChatGPT citation churn varies significantly by country for identical queries. Weekly churn in Germany ran at 74%. In the UK, 60%. In France, 42%. Google AI Mode showed similar churn rates across countries (54–59%), suggesting that platform's volatility is mostly structural rather than geographic. For ChatGPT specifically, the same underlying product is behaving substantially differently by market week over week.

> Weekly ChatGPT citation churn for identical queries: 74% in Germany, 60% in the UK, 42% in France.
>
> SISTRIX / JOHANNES BEUS — 82,619 PROMPTS, 17 WEEKS, 6 COUNTRIES

Second, and more significant for non-English-language markets, is the language-domain finding. For identical German-language prompts, ChatGPT cited sources that were 68% English-language domains. Google AI Mode used 80% German-language domains for the same queries.

A brand that exists primarily in German trade press, German review platforms, and German directories is substantially better positioned in Google AI Mode than in ChatGPT for a German buyer running the same search.

xFunnel's 2025 analysis covering 56,223 citations from six answer engines in four international markets found a wide spread in how much platforms draw from in-market sources. (The xFunnel source URL was not directly accessible for independent verification; these figures are from search-aggregated reporting of the study.) Perplexity and Microsoft Copilot both drew more than half their citations from local sources. ChatGPT drew roughly

© 2026 Aiviara Research

three in ten from in-market sources. Gemini appeared to localize very little, consistently returning global .com sources regardless of where the query originated. The reported figure for Gemini was 5.3%. A European SaaS brand competing in Gemini is competing primarily against the domain authority of US-headquartered players, no matter what country the query comes from.

> Perplexity and Copilot drew more than 50% of citations from local in-market sources. Gemini's reported localization rate was 5.3%.
>
> XFUNNEL 2025 — 56,223 CITATIONS, 6 ANSWER ENGINES, 4 MARKETS

Consider a B2B software company headquartered in Munich, with strong German-language press coverage and regional backlinks, running the same competitive query in Gemini and in Perplexity. In Perplexity, that local coverage is a relevant signal. In Gemini, it is largely irrelevant. The Munich company's regional content investment pays off on one platform and not the other. Which platform matters depends on where the buyer actually searches, which varies by market and demographic.

Three ways AI gets geography wrong for brands

Motoko Hunt documented three specific failure modes in practitioner analysis published by Search Engine Journal in November 2025.

The first is language used as a location proxy. AI models often treat the language of a query as if it indicates the buyer's market, but a Spanish-language query could be coming from Mexico, Colombia, or Spain. A brand serving Mexico may be invisible in responses to Spanish queries if its content is thin relative to Spain-based competitors with more training-data representation. The buyer's language and the buyer's geography are not the same signal.

The second is market aggregation bias. English-dominant training causes AI systems to surface a brand's global or English-language presence ahead of its regional one. Hunt describes it as a winner-takes-the-synthesis dynamic. AI systems tend to gravitate toward the strongest global representation of a brand, and in most cases that is the main English-language website. The model tends to treat that global entity as canonical, and regional pages get absorbed into its profile. A buyer in Singapore searching for local enterprise software vendors may get a list of US-headquartered companies as the default answer.

© 2026 Aiviara Research

The third failure mode is canonical amplification, which operates differently from the first two. When AI synthesises an answer about a company, it gravitates toward the most authoritative, most cited, most referenced instance of that entity in its training data. For a multinational brand, that is the global or US parent entity. Regional subsidiaries with fewer training-data references, fewer press mentions, and fewer inbound links from that market get absorbed into the parent entity's profile.

This is distinct from the language-proxy problem. A German subsidiary publishing in German can still be canonically absorbed into the parent English entity if the parent has overwhelming training-data presence. Hunt offers a case study from practitioner work. A search for "proveedores de químicos industriales" (industrial chemical suppliers in Spanish) returned US-based suppliers that did not serve Mexico at all. The AI retrieved brands with training-data presence, regardless of whether those brands were operationally relevant to the searcher's geography.

What the old localization playbook does and doesn't carry over

Most enterprise brands that operate across regions have an international SEO infrastructure: hreflang tags, country-code domains or subdirectories, translated content libraries, regional sitemaps. That infrastructure was built for traditional search. Its relevance to AI retrieval is partial at best.

Hreflang operates at the serving layer in traditional search, routing users to the appropriate regional URL after a query is processed. In AI-mediated retrieval, as Motoko Hunt's January 2026 Search Engine Land analysis describes, content is evaluated before serving, and geographic routing signals may be evaluated late in the process or bypassed entirely. The mechanism that made hreflang valuable in traditional search does not transfer directly to how AI retrieval works.

Translation-only localization has the same problem. Hunt's January 2026 analysis states that AI models "collapse multilingual content into shared semantic representations." A translated page that contains no new regional entity, no market-specific case study, and no locally relevant context is not semantically distinct to the retrieval system. Two versions of the same content in two languages compress into a single semantic representation. The regional version contributes nothing that the original page did not already provide.

This does not mean regional content investment is worthless. It means the nature of the investment has to change. Translated pages do not add to a brand's regional AI presence.

© 2026 Aiviara Research

Pages with genuinely regional content (local references, local compliance context, regional case studies, locally known customers) do something different. Many brands have invested heavily in translation while remaining semantically identical to their parent entity. From an AI retrieval perspective, that investment has limited return.

What actually moves regional AI visibility

The evidence supports a few specific interventions, though controlled outcome studies in B2B SaaS at the regional level do not yet exist.

Practitioner guidance from Lokalise argues that in-market editorial coverage and local backlinks may act as geographic authority signals in retrieval systems, particularly for Perplexity and Microsoft Copilot where in-market citation rates are highest. A brand that appears in German trade press, Dutch enterprise software directories, or French analyst reports is building the local source presence those platforms draw from. For Gemini, which appeared to localize very little in the xFunnel study, that investment has less clear return.

Translation alone does not seem to create a separate regional footprint in retrieval systems. Per Hunt's January 2026 analysis, regional content that adds genuinely new semantic value (a local case study, a compliance-specific use case, a reference to local regulations or market conditions) is not collapsed into the parent entity's representation. It gives the model something distinct to retrieve. This is why publishing content with genuine regulatory context tends to perform differently from translated marketing copy. ARGEO, an AI visibility agency, documented a case in March 2026 where a SaaS brand publishing EU AI Act compliance documentation rose from 0% to over 60% appearance rate in ChatGPT for "AI compliance software Europe" within eight weeks. ARGEO reports this as a self-described case study, and the methodology is not independently verifiable. But the outcome is consistent with how AI retrieval handles genuinely regional content.

> A SaaS brand publishing EU AI Act compliance documentation rose from 0% to over 60% appearance rate in ChatGPT for "AI compliance software Europe" — within eight weeks.
>
> ARGEO, MARCH 2026

A brand that publishes something genuinely relevant to a regional context creates a semantic presence that a translated brochure page does not.

Whether platform strategy should differ by target market is an open empirical question. The localization rate data from xFunnel's 2025 study suggests that regional content investments

have clearer paths to return in Perplexity and Copilot than in Gemini. Whether to prioritise regional content or global .com domain authority depends on which platform the target buyer actually uses in that market. That data is often not collected.

The measurement problem no one has solved yet

Rand Fishkin at SparkToro published a study in January 2026 covering 2,961 research sessions run by 600 volunteers, testing brand consistency across AI platforms for the same queries. There was less than a 1-in-100 chance of the same brand list appearing twice for the same query run by different users at the same time, even within the same location. That is the baseline volatility before any geographic variable is introduced.

The within-location variance the SparkToro study measured was not designed to test geographic differences. It was documenting inconsistency among users in the same market. That underlying volatility is what cross-geography measurement has to work on top of.

No published, open-access study runs the same B2B vendor queries from UK and US IP addresses and compares the resulting brand recommendation lists directly. DerivateX's 2026 B2B SaaS AI Visibility Benchmark, as reported by Demand Gen Report, covered 1,400 buyer-intent prompts across 50 B2B SaaS companies and found Stripe scoring 65 out of 100 on AI Presence versus Razorpay at 39. That gap is circumstantially consistent with training-data bias favouring US-headquartered brands. But the study did not control for geographic variables, and the gap could equally reflect brand authority, content volume, or press coverage differences. The inference is suggestive, not settled.

Industry data comparing US/UK AI brand visibility exists. Pi Datametrics published a comparison for Q4 2025 covering AI brand visibility by country in the outdoor clothing category. That study is gated behind a lead form and the specific figures are not publicly accessible. Its existence confirms that tracking organisations see the demand for this measurement, but it is not yet a comparable evidence base.

The infrastructure for geographic differentiation is documented. Training-data bias toward English-language, Western sources is established. Platform behaviour by country is measurably different, as the SISTRIX study shows. What remains absent is a controlled study confirming that a B2B brand's citation rate differs when queried from London versus San Francisco. Brands that measure AI presence from a single location are producing a real number. Whether it describes what a buyer in another market sees is an inference.

Brands that have invested in AI visibility measurement are, for the most part, measuring one geography. Usually their home market, usually from a single IP or set of IPs, usually in a

© 2026 Aiviara Research

single language. The SISTRIX study's finding that ChatGPT citations in Germany show 74% weekly churn while citations in France show 42% churn, for identical queries, is a signal that what a brand measures at home and what a buyer sees elsewhere are not reliably the same. Brands that measure AI visibility from a single location are measuring one slice of a more complex picture. Where that slice sits relative to what buyers see in other markets remains, for most brands, an open question.

Aiviara is building infrastructure for monitoring AI brand citations and factual accuracy across LLM platforms. Early access information is available at aiviara.com.

© 2026 Aiviara Research