Gracker.ai

RESEARCH REPORT

AI Search Visibility in Enterprise Compliance Automation

A multi-engine, multi-market benchmark study of how six AI search engines — ChatGPT, Perplexity, Gemini, Microsoft Copilot, Google AI Overview and Google AI Mode — answer buyer-intent questions about compliance automation, governance, risk and compliance software and integrated risk management across the United States, Canada, India and Germany.

Table 1. Study scale by AI engine — responses recorded and sources cited

AI engine	Responses	Cited sources	Markets
ChatGPT	40,000	1,690,000	4
Microsoft Copilot	40,000	152,000	4
Gemini	40,000	124,000	4
Google AI Mode	40,000	735,000	4
Google AI Overview	40,000	438,000	4
Perplexity	40,000	325,000	4
Total	**240,000**	**3,464,000**	**4**

Published: August 2026

Research period: August 2026

Category: Enterprise compliance automation (compliance automation, GRC, IRM)

Scope: 10,000 buyer-intent queries · 240,000 AI engine responses · 6 AI engines · 4 markets

Evidence base: 3,464,000 cited sources · 99 distinct product entries

Prepared by: GrackerAI · AI Search Visibility Benchmark Series

Abstract

Enterprise compliance buyers now begin an evaluation inside a generated response, so the vendor set an AI engine returns forms the consideration set before any vendor is contacted. This study measures how six commercial AI search engines answer buyer-intent questions about enterprise compliance automation, governance, risk and compliance software and integrated risk management. Each engine received 10,000 enterprise buyer-intent queries in each of four markets during a single collection window in August 2026, producing 240,000 responses that contained 3,464,000 cited sources and covered 99 distinct product entries. For every response the study recorded length, cited pages and their source type, page reachability, hedging, recency anchoring, truncation, refusal, and each brand mention with its ordinal position. The engines converge on a core: 13 of 99 brand strings were named by all six engines, led by Vanta at 139,000 mentions and Drata at 123,000. Below that core visibility is thin, with 60 brand strings named by a single engine. The brand named most often opens the response in only 3 of the six engines, and MetricStream is read earlier on average in 4 of them. The engines cite largely separate domains, with a mean pairwise overlap of 0.11, and 41.7% of everything they cite sits on a vendor's own site, a lower share than one engine alone would suggest: ChatGPT contributes 48.8% of all citations and draws 59.2% of them from vendor domains. Dead links are rare, at 1.6% of cited pages. The same two brands lead every market in the same order.

Contents

1. Introduction

1.1 Background

An enterprise buyer evaluating compliance software now begins inside a generated response. The buyer asks an AI search engine which platforms automate evidence collection for SOC 2 and ISO 27001, which map one control set across several frameworks, or which integrated risk management suite fits an organisation already running a workflow platform. The engine

returns a short ordered list of named vendors with a paragraph of justification for each. That list is not a ranking of the market. It is the output of one system reading one collection of web sources at one moment, and it arrives before the buyer has opened a single vendor site.

The structural consequence is narrow. A vendor absent from the returned list is not ranked low; it is not present. A generated response has no second page, so the buyer has no mechanism equivalent to scrolling further. Where a search results page offered ten positions and a set of related queries, a generated response typically names between three and six vendors and closes. Visibility in this setting is closer to inclusion than to ranking, and the difference between being inside and outside the returned set is discontinuous.

The measurement problem that follows is that no single engine describes the market. Engines differ in the sources they read, the length at which they answer, and the order in which they name vendors. A vendor that measures its position in one engine learns about that engine. This study measures six at once, in four markets, over one collection window, so that agreement and divergence between them can be separated from the behaviour of any one system.

1.2 Why this category

Enterprise compliance automation is a suitable instrument for this measurement. The evaluation criteria are stable and externally defined, because the frameworks that structure the category — SOC 2, ISO 27001, GDPR, PCI DSS and HIPAA — are published standards rather than vendor constructs, and 13 of the 99 brand strings recorded were named by every engine. A category with a settled core isolates engine behaviour from genuine market churn.

The category also spans two buying motions that one query set can separate. Compliance automation platforms address framework readiness and continuous evidence collection; integrated risk management suites address enterprise risk workflow at a different price and implementation scale. Across the corpus the engines returned both in answer to overlapping questions, grouping Vanta, Drata, Secureframe and Sprinto for multi-framework coverage and ServiceNow, MetricStream, Archer, IBM and LogicGate for large configurable programmes, which lets the study observe whether the engines distinguish between the two.

Buying behaviour in the category is research-intensive. Purchases are preceded by framework scoping, auditor selection and readiness assessment, and much of that research is now conducted through generated answers. The cost of omission from an early shortlist is correspondingly high: a platform absent from the generated shortlist is absent from the readiness conversation that precedes the purchase.

1.3 Contribution

This study provides four things a single-engine measurement does not. It establishes how far six engines agree on the vendor set in one category, separating the shared core from the long tail each engine holds alone. It measures the overlap between the engines' evidence bases directly, as a pairwise coefficient over cited domains, and classifies every cited page by source type, so that agreement on vendors can be set against what the engines actually read. It reports mention volume, first-mention position and mean ordinal position as three separate quantities, which allows them to be shown moving independently. And it tests the vendor set across four markets in the same window, which distinguishes market-specific results from a set the engines return everywhere. All four rest on a single corpus of 240,000 responses collected under one design.

2. Research Questions

Six questions are addressed, each answerable from the recorded data and each answered explicitly in §5.

1. **RQ1.** Do the six engines converge on the same set of named vendors?
2. **RQ2.** How concentrated is brand visibility across the recorded vendor set?
3. **RQ3.** Do the engines draw on a shared evidence base?
4. **RQ4.** Does ordinal position track mention volume?

5. **RQ5.** Does the vendor set vary between the four markets?
6. **RQ6.** Do the engines cite vendors' own sites or third-party sources?

3. Methodology

3.1 Study design

The study is observational and cross-sectional. Six AI search engines were queried with a fixed query set in four markets during a single collection window in August 2026. No engine was manipulated, no result was solicited, and no intervention was applied between conditions. The design records what the engines returned; it does not test a hypothesis about why they returned it. This edition reprocesses the same collected responses under the instrumentation's current citation, source-type and entity rules; the responses themselves are unchanged.

3.2 Engines under test

Six engines were tested: ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview and Perplexity. All six are commercial services. None is version-pinned, none exposes a stable model identifier to the caller, and all are subject to continuous change by their operators. A result reported here describes the service as it behaved during the August 2026 window, not a fixed artefact that can be re-queried later in the same state. Three of the six — Gemini, Google AI Mode and Google AI Overview — are operated by the same parent, which is relevant to the overlap results in §4.6 and the first-mention results in §4.7.

3.3 Markets

Four markets were tested: United States, Canada, India and Germany. Market variance was tested because engines localise results, and a vendor set returned in one market cannot be assumed to hold in another. The four combine two large English-language markets, one large non-English market, and one market with a substantial domestic buyer base and a separate regulatory environment; in a category defined by regulation, that last property matters. Each query was issued in every market, so market is a within-design condition rather than a separate sample.

3.4 Query construction

The query set contains 10,000 distinct buyer-intent queries. It combines core category questions of the form a buyer would ask when assembling a shortlist — enterprise GRC software, integrated risk management, HIPAA tooling, multi-framework compliance platforms, and SOC 2 options for teams without dedicated staff — with paraphrase variants of those questions, so that the stability of a returned vendor set under rewording can be observed rather than assumed. Queries were written to elicit vendor recommendations rather than definitions.

Every query was issued to every engine in every market. The corpus identity follows directly: 10,000 queries in 4 markets gives 40,000 responses per engine, and 10,000 queries across 6 engines and 4 markets gives 240,000 responses in total.

3.5 Corpus and unit of analysis

The unit of analysis is one engine response to one query in one market. Metrics are computed over responses and aggregated to the engine, the brand, the product entry or the market as the metric requires. The corpus contains 240,000 responses, distributed evenly across the six engines at 40,000 responses each and across the four markets at 60,000 responses each. Every per-engine comparison in §4 therefore rests on the same denominator.

3.6 Metric definitions

Every metric used anywhere in this report, including the appendices, is defined here.

Table 2. Metric definitions — all metrics used in this report

Metric	Definition	Unit
Response	One engine output returned for one query in one market.	count
Mean response length	Arithmetic mean word count of an engine's responses.	words
Hedging rate	Share of an engine's responses containing explicit qualification of a recommendation.	% of responses
Recency-anchored rate	Share of an engine's responses that anchor a claim to an explicit calendar date.	% of responses
Truncation rate	Share of an engine's responses ending mid-sentence, inferred from the terminal characters of the response text.	% of responses
Refusal rate	Share of an engine's responses declining to answer the query.	% of responses
First-mention vendor	The brand most often named before any other brand in an engine's responses.	brand
Brand mention	One occurrence of a vendor brand name in a response. Repeat occurrences within one response are counted once. Spelling variants of one brand are merged.	count
Mean ordinal position	Arithmetic mean of the position at which a brand is named within a response, counting named vendors from one.	position
Mean sentiment	Arithmetic mean of the score the analysis model assigned to how favourably the response describes the brand, on a scale from minus one to one, under a fixed rubric.	score
Cited source	One distinct page a response cites, in its source list or as an in-text link, after removal of page assets, trackers and advertising links. Variants of one page differing only in tracking parameters or text fragments count once.	count
Distinct domains	The number of different registrable domains cited by that engine across its responses. Reported per engine and never summed across engines.	count
Source type	The class of the cited page's domain: vendor site (a domain of a vendor named in the response), review aggregator, news, forum, social, government, or other.	class
Dead-link rate	Share of an engine's cited pages that returned not-found, gone or a server error, or whose host could not be reached, when tested. A page that refused the request is not a dead link.	% of cited sources
Self-citation rate	Share of an engine's cited sources hosted on a domain operated by the engine's own parent.	% of cited sources
Jaccard coefficient	Size of the intersection of two engines' cited-domain sets divided by the size of their union. Zero is disjoint, one is identical.	ratio
Product entry	One product recorded against a brand, after plan, edition and deployment spellings of one product are merged.	count
Mean rank	Arithmetic mean of the ordered position of a product entry within the response that named it.	position

3.7 Data collection and processing

Responses were collected through each engine's public interface and normalised to plain text. Word counts, terminal-character inspection for truncation, and detection of qualifying and date-anchoring language were computed over that normalised text.

Cited sources were taken as the union of the source list each engine attached to a response and the pages the response linked to in its text. Page assets, tracking and advertising links were removed, and variants of one page that differ only in tracking parameters or highlighted-text fragments were counted once. Where an engine wrapped a citation in a redirect, the redirect was

followed to the cited page. Each page was then reduced to its registrable domain for the distinct-domain, source-type and overlap measures. Source type was assigned by rule from the domain. A domain belonging to a vendor the response names is a vendor site. Forums, social and video sites, news publishers, government and academic domains and review aggregators were matched against lists and host-name patterns, and the remainder is other.

Reachability was tested once per page during reprocessing. A page was recorded as a dead link when it returned not-found, gone or a server error, or when its host could not be reached after a retry. A page that refused the request, as sites behind bot protection do, was recorded as reachable, since the page exists.

Brand mentions were detected against the vendor names each response used and recorded with the ordinal position at which the brand appeared among the named vendors, counting from one. Spelling variants of one brand, such as a vendor's name with and without a corporate suffix, were merged, and plan, edition and category-noun variants of one product were merged into one product entry. The first-mention vendor for an engine is the brand that most often held position one. Sentiment was assigned by the analysis model to each brand and product mention, on a scale from minus one to one. The rubric scores how favourably the response describes the entity and is documented with the instrumentation.

One implementation detail is not documented by the instrumentation: the lexicon used to detect qualifying and date-anchoring language. §6.1 states the consequence for how those two rates should be read. Source age was recorded where a page declared a publication date, but fewer than a fifth of cited pages did, so it is not reported.

3.8 Scope exclusions

The study did not paraphrase, summarise or edit engine output before analysis. It did not check any claim an engine made against an external source, and it did not score responses for accuracy. It did not test, benchmark or evaluate any vendor product. No placement was paid for, solicited or arranged. The object of study is how the engines described the category, not whether their descriptions were correct.

4. Results

4.1 Response characteristics by engine

The six engines construct responses differently. Table 3 gives the response-construction measures for each engine over an identical number of responses.

Table 3. Response construction by AI engine — n = 40,000 responses per engine, 240,000 in total

AI engine	Responses	Mean words	Hedging	Recency-anchored	Truncation	Refusal	First-mention vendor
ChatGPT	40,000	542.9	20%	23%	30%	0%	ServiceNow
Microsoft Copilot	40,000	456.9	5%	63%	0%	0%	LogicGate
Gemini	40,000	478.3	7%	3%	28%	0%	Vanta
Google AI Mode	40,000	338.0	5%	0%	3%	0%	Vanta
Google AI Overview	40,000	217.6	13%	17%	5%	0%	Vanta
Perplexity	40,000	132.4	17%	25%	42%	0%	ServiceNow

Mean response length spans a factor of 4.1, from 132.4 words for Perplexity to 542.9 for ChatGPT. Figure 1 shows the distribution. 3 engines exceed 400 words — ChatGPT, Microsoft Copilot and Gemini — and Google AI Overview and Perplexity fall below 250.

Truncation was recorded in 5 of the six engines and is highest for Perplexity at 42%, followed by ChatGPT at 30% and Gemini at 28%; Microsoft Copilot recorded none. Recency anchoring is highest for Microsoft Copilot at 63%, and Google AI Mode recorded none. Hedging runs from 5% for Microsoft Copilot to 20% for ChatGPT; no engine recorded none. No engine refused

a query: the refusal rate is 0% for all six. 3 engines record Vanta as the first-mention vendor, 2 record ServiceNow and 1 records LogicGate; §4.7 reports the position data behind that column.

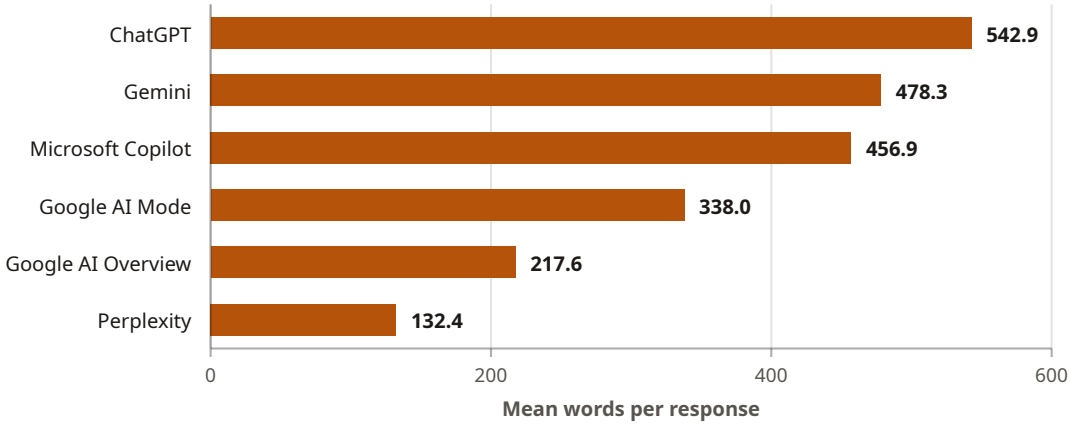

Figure 1. Mean response length by AI engine. Arithmetic mean word count per response; n = 40,000 responses per engine. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

4.2 Brand visibility

Brand visibility was recorded for 99 brand strings, listed in full in Appendix A. Table 4 gives the twelve most-mentioned, with the number of engines that named each and its mean sentiment. Figure 2 shows the same twelve.

Table 4. Brand visibility, twelve most-mentioned brands — of 99 brand strings recorded; mean sentiment where recorded, a dash otherwise; n = 240,000 responses

Brand	Total mentions	Engines recording	Mean sentiment
Vanta	139,000	6	0.69
Drata	123,000	6	—
MetricStream	89,000	6	0.80
ServiceNow	86,000	6	0.81
Secureframe	79,000	6	0.62
Sprinto	65,000	6	0.69
IBM	53,000	6	0.75
LogicGate	49,000	6	0.70
OneTrust	42,000	6	0.68
AuditBoard	39,000	6	0.73
Diligent	39,000	6	0.68
Compliancy Group	37,000	6	0.76

Visibility is concentrated at the top and long-tailed below it. The two most-mentioned brands, Vanta and Drata, account for 22.2% of all mentions recorded across the 99 brand strings, at 139,000 and 123,000 mentions, 16,000 apart. MetricStream follows at 89,000 and ServiceNow at 86,000, 3,000 apart; the four together take 37.0%. Below Secureframe at 79,000 no brand exceeds 65,000.

13 of the 99 brand strings were named by all six engines: Vanta, Drata, MetricStream, ServiceNow, Secureframe, Sprinto, IBM, LogicGate, OneTrust, AuditBoard, Diligent, Compliancy Group and Hyperproof. 60 were named by exactly one engine, none of them above 7,000 mentions. The strings recorded include adjacent categories — healthcare compliance services, audit firms, secure messaging and document platforms — and a standards body, each named in a small number of responses as the engines wrote them. Mean sentiment for the twelve brands in Table 4 occupies a band from 0.62 for Secureframe to 0.81 for ServiceNow, a spread of 0.19.

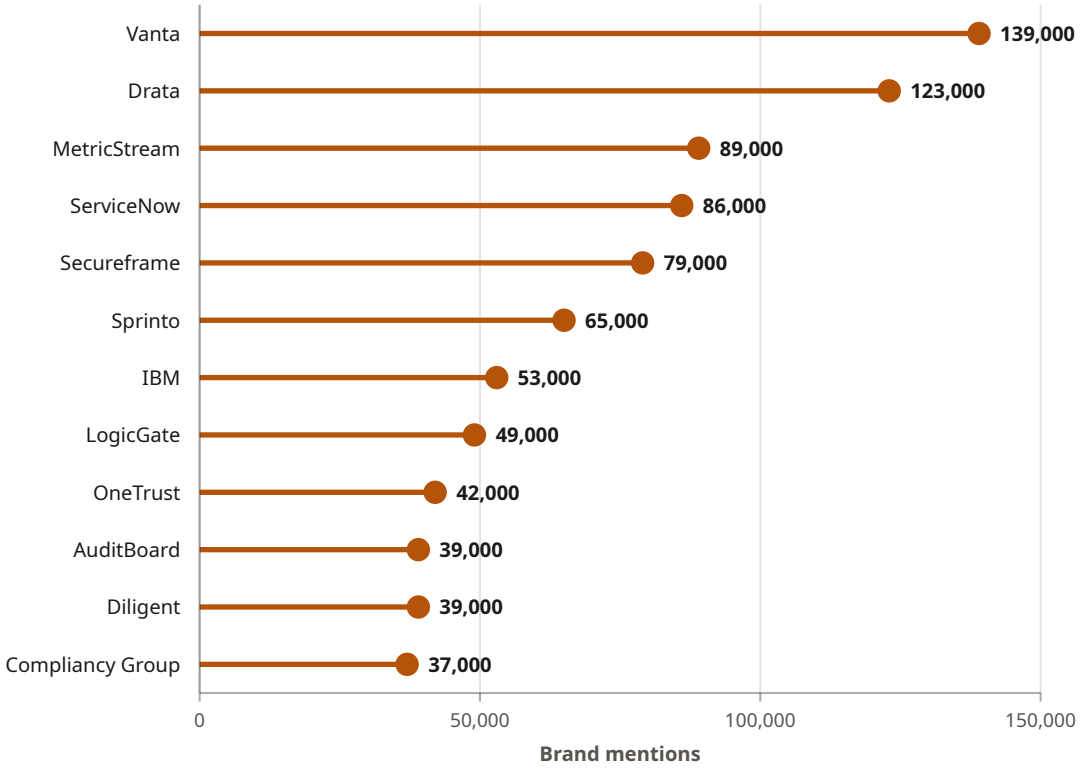

Figure 2. Brand mentions, twelve most-mentioned brands. The 12 most-mentioned of 99 brand strings recorded; n = 240,000 responses. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

4.3 Product-level results

99 distinct product entries were recorded after plan, edition and category-noun spellings of one product were merged. An entry is one product as an engine named it, so a brand with several products holds several entries: MetricStream holds 5, Diligent 5 and ServiceNow 4. Table 5 gives the ten most-mentioned entries.

Table 5. Product entries, ten most-mentioned — of 99 entries; n = 240,000 responses

Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Vanta	Vanta	111,000	2.8	0.71	6	4
Drata	Drata	97,000	3.0	0.70	6	4
MetricStream	MetricStream	68,000	2.3	0.80	6	4
Secureframe	Secureframe	59,000	3.4	0.65	6	4
Sprinto	Sprinto	57,000	3.2	0.71	6	4
Archer	Archer	50,000	3.2	0.75	6	4
ServiceNow IRM	ServiceNow	47,000	1.7	0.84	5	4

Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
IBM OpenPages	IBM	44,000	4.3	0.75	6	4
LogicGate Risk Cloud	LogicGate	43,000	4.4	0.71	6	4
Optro	Optro	38,000	4.2	0.73	5	4

Product entries carry their own brand attributions, and Appendix C lists entries attributed to brands outside the twelve in Table 4. The distribution is long-tailed. 72 of the 99 entries record fewer than 10,000 mentions, 32 record exactly 1,000, and 53 were named by exactly one engine. 11 entries were named by all six engines, 37 were recorded in all four markets, and 39 were recorded in a single market. The full table is in Appendix C.

Mean rank among the ten most-mentioned entries runs from 1.7 to 4.4. The leading entry, Vanta, records 111,000 mentions at a mean rank of 2.8, and the second, Drata, records 97,000 at 3.0. The earliest mean rank among the ten belongs to ServiceNow IRM, at 1.7 on 47,000 mentions.

4.4 Citation volume and source type

The corpus contains 3,464,000 cited sources, distributed unevenly across the six engines, as Figure 3 shows. Table 6 gives volume and domain breadth; Table 7 gives the source-type mix.

Table 6. Cited sources and domain breadth by AI engine — n = 3,464,000 cited sources

AI engine	Cited sources	Share of corpus	Distinct domains	Most-cited domain
ChatGPT	1,690,000	48.8%	218	vanta.com (99,000)
Microsoft Copilot	152,000	4.4%	43	worldmetrics.org (38,000)
Gemini	124,000	3.6%	54	scytale.ai (9,000)
Google AI Mode	735,000	21.2%	257	sprinto.com (39,000)
Google AI Overview	438,000	12.6%	154	vanta.com (24,000)
Perplexity	325,000	9.4%	39	vanta.com (20,000)
Total	**3,464,000**	**—**	**—**	**—**

ChatGPT alone accounts for 48.8% of all cited sources in the corpus, at 1,690,000, and with Google AI Mode at 735,000 the two together account for 70.0%. Google AI Mode cites from the widest domain population at 257 distinct domains, followed by ChatGPT at 218. Perplexity cites from the narrowest at 39. Distinct-domain counts are reported per engine and are not summed across engines.

vanta.com is the most-cited domain for 3 of the six engines, ChatGPT, Google AI Overview and Perplexity. Google AI Mode's is sprinto.com, Gemini's is scytale.ai, and Microsoft Copilot's is worldmetrics.org, a statistics aggregator that no other engine cites in quantity. Appendix B gives the three most-cited domains for each engine.

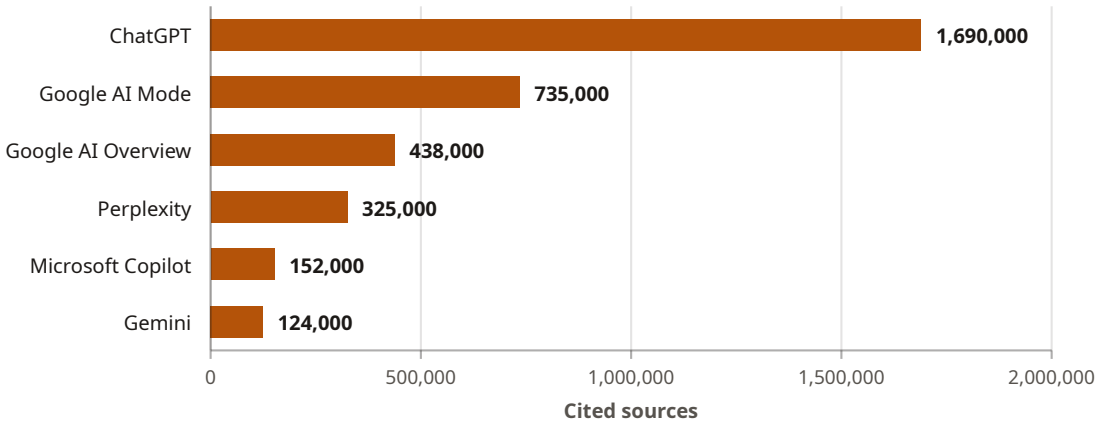

Figure 3. Cited sources by AI engine. Sources cited per engine; n = 3,464,000 cited sources. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

Table 7. Source type by AI engine — share of each engine's cited sources by the class of the cited domain; n = 3,464,000 cited sources

AI engine	Other	Vendor site	Review aggregator	Government	Social	Forum	Wikipedia	Marketplace	Academic	Parent ecosystem
ChatGPT	27.7%	59.2%	8.2%	3.3%	0.4%	0.8%	0.4%	0.1%	0.0%	0.0%
Microsoft Copilot	72.4%	19.7%	7.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gemini	54.8%	37.9%	4.0%	0.0%	0.0%	2.4%	0.0%	0.0%	0.0%	0.8%
Google AI Mode	68.4%	25.2%	2.0%	0.0%	3.5%	0.7%	0.0%	0.0%	0.1%	0.0%
Google AI Overview	68.5%	23.1%	3.4%	0.0%	4.6%	0.5%	0.0%	0.0%	0.0%	0.0%
Perplexity	67.1%	24.9%	8.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
All engines	**48.2%**	**41.7%**	**6.1%**	**1.6%**	**1.5%**	**0.7%**	**0.2%**	**0.1%**	**0.0%**	**0.0%**

Across the corpus, 41.7% of cited sources sit on a vendor site, 1,444,000 of 3,464,000, and 48.2% fall outside every named class. Review aggregators, which here include analyst firms, account for 6.1%. Government domains account for 1.6%, 56,000 citations, of which 56,000 are ChatGPT's; the category's regulatory texts are cited, in one engine. The mix differs by engine. Vendor sites are the largest class for 1 of the six engines, ChatGPT, at 59.2% of its citations. For the other five the residual other class is the largest, from 54.8% to 72.4%, and vendor sites run from 19.7% for Microsoft Copilot to 37.9% for Gemini. Review aggregators reach 8.2% of ChatGPT's citations, the highest share, and 2.0% of Google AI Mode's. Figure 4 shows the vendor-site share by engine, and Appendix B gives the counts behind Table 7.

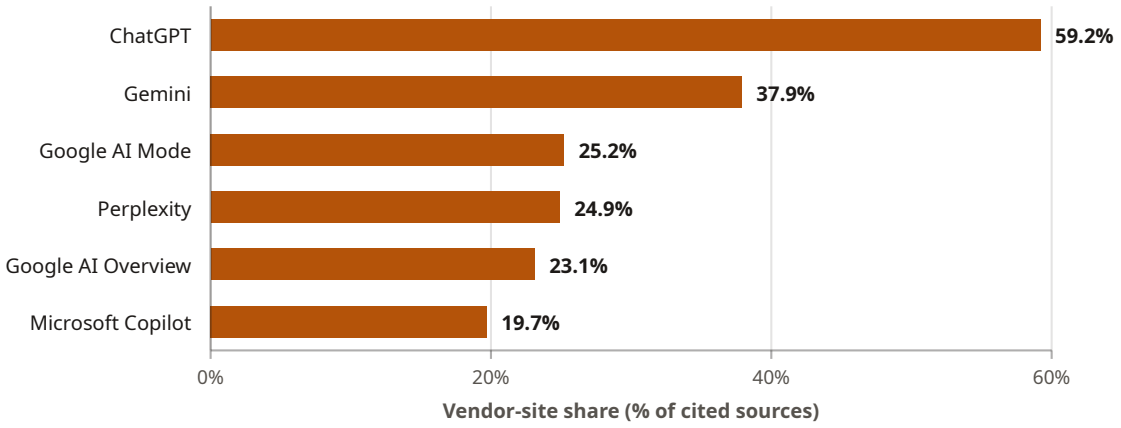

Figure 4. Share of cited sources on vendor sites by AI engine. Cited sources whose domain belongs to a vendor named in the response, as a share of that engine's cited sources; n = 3,464,000 cited sources. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

4.5 Source quality and link decay

A small share of the pages the engines cited did not resolve when tested. Table 8 gives the rate for each engine.

Table 8. Dead links and self-citation by AI engine — n = 3,464,000 cited sources

AI engine	Cited sources	Dead-link rate	Dead links	Self-citation rate
ChatGPT	1,690,000	2%	33,800	0%
Microsoft Copilot	152,000	7.9%	12,000	0%
Gemini	124,000	0%	0	1%
Google AI Mode	735,000	0.5%	3,680	0%
Google AI Overview	438,000	1.4%	6,130	0%
Perplexity	325,000	0%	0	0%
Total	**3,464,000**	**1.6%**	**55,610**	**—**

The overall dead-link rate is 1.6%, 55,610 of 3,464,000 cited pages. 2 engines — Gemini and Perplexity — record no dead links at all. The highest rate is 7.9%, for Microsoft Copilot, on the second-smallest citation volume in the corpus. ChatGPT, at 2%, accounts for 33,800 of the 55,610 dead links, 60.8% of the total, because it cites the most.

The self-citation rate is 1% for Gemini and 0% for the other five engines.

4.6 Cross-engine source overlap

The Jaccard coefficient over each pair of engines' cited-domain sets measures how far the engines read the same web. Table 9 gives the full matrix.

Table 9. Cross-engine cited-domain overlap — Jaccard coefficient over cited-domain sets; 0 is disjoint, 1 is identical; n = 3,464,000 cited sources

	ChatGPT	Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
ChatGPT	**1.00**	0.05	0.10	0.13	0.13	0.07
Copilot	0.05	**1.00**	0.04	0.03	0.05	0.08
Gemini	0.10	0.04	**1.00**	0.12	0.16	0.22

	ChatGPT	Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
Google AI Mode	0.13	0.03	0.12	**1.00**	0.27	0.10
Google AI Overview	0.13	0.05	0.16	0.27	**1.00**	0.14
Perplexity	0.07	0.08	0.22	0.10	0.14	**1.00**

Shading increases with the coefficient: < 0.10 0.10–0.17 0.18–0.25 0.26–0.33 ≥ 0.34

The mean coefficient across the 15 distinct engine pairs is 0.11. The highest value is 0.27, between Google AI Mode and Google AI Overview, two engines operated by the same parent. The second highest is 0.22, between Gemini and Perplexity, two engines operated by different parents; the third same-parent pair, Gemini with Google AI Mode, records 0.12. The lowest is 0.03, for Microsoft Copilot with Google AI Mode.

Microsoft Copilot records the lowest coefficient with every other engine: its five values run from 0.03 to 0.08, the latter with Perplexity. No pair of engines exceeds 0.27, and only 2 of the 15 pairs reach 0.20.

4.7 First-mention position

Mention volume records how often a brand appears. Mean ordinal position records how early it appears among the vendors a response names. The first-mention vendor is the brand that most often opens a response. Table 10 gives all three for the three most-mentioned brands, and Figure 5 shows Vanta and MetricStream.

Table 10. Mean ordinal position of the three leading vendors — lower is earlier in the response; n = 240,000 responses

AI engine	Vanta	Drata	MetricStream	First-mention vendor
ChatGPT	2.4	2.4	3.1	ServiceNow
Microsoft Copilot	2.8	3.1	1.8	LogicGate
Gemini	3.3	3.1	1.8	Vanta
Google AI Mode	3.3	3.4	2.4	Vanta
Google AI Overview	2.9	2.8	2.5	Vanta
Perplexity	2.0	3.0	2.3	ServiceNow

Vanta holds the first-mention position in 3 of the 6 engines, the three operated by the same parent, on 139,000 mentions, the highest total in the corpus. Neither Drata nor MetricStream holds it anywhere. ServiceNow holds it in 2, ChatGPT and Perplexity, on 86,000 mentions, and LogicGate holds it in Microsoft Copilot, on 49,000, the 8th total in the corpus.

Mean ordinal position does not follow the first-mention split either. MetricStream records the earlier mean position than Vanta in 4 engines, Microsoft Copilot, Gemini, Google AI Mode and Google AI Overview, including all three where Vanta is the first-mention vendor: in Gemini MetricStream averages 1.8 against 3.3 for Vanta. Vanta records its earliest mean position in Perplexity, at 2.0. Drata is read earlier than Vanta in Gemini and Google AI Overview and ties with it in ChatGPT; its mean position runs from 2.4 to 3.4.

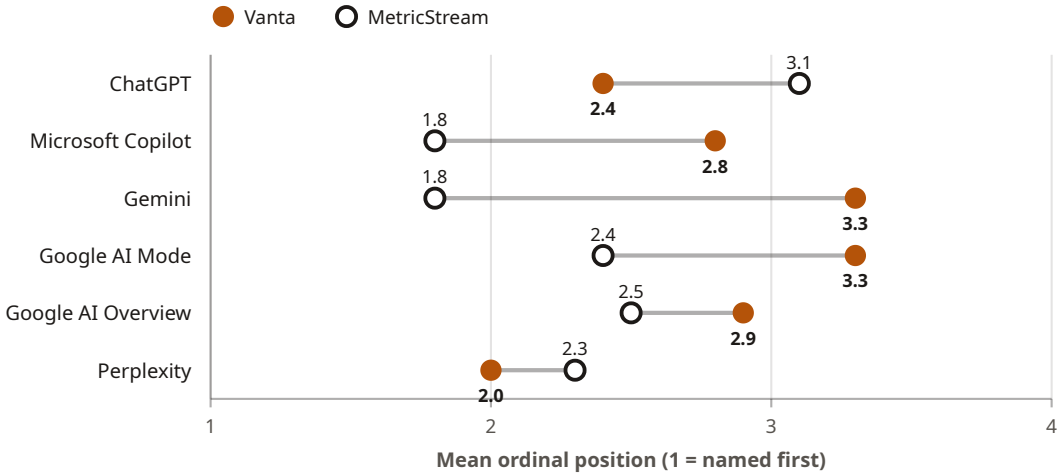

Figure 5. Mean ordinal position of Vanta and MetricStream by AI engine. Lower is earlier in the response; n = 240,000 responses. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

4.8 Geographic variance

Table 11 gives the three leading brands in each market, ranked by mentions within that market.

Table 11. Leading brands by market — ranked by brand mentions within the market; n = 60,000 responses per market

Market	First	Second	Third
United States	Vanta	Drata	MetricStream
Canada	Vanta	Drata	ServiceNow
India	Vanta	Drata	ServiceNow
Germany	Vanta	Drata	MetricStream

The same two brands lead all four markets in the same order: Vanta first and Drata second in the United States, India, Canada and Germany. The third position differs by market: MetricStream in the United States and Germany and ServiceNow in India and Canada. No market returns a vendor absent from the other three in its leading positions, and in a category defined by jurisdiction-specific regulation no market-specific vendor enters those positions anywhere.

At product level, 37 of the 99 product entries were recorded in all four markets, and every entry in Table 5 was recorded in all four.

4.9 Inter-engine disagreement

The engines diverge on three specific points in the category.

The first concerns single-control-set coverage, the claim that one platform's control set satisfies several frameworks without separate mapping work. Microsoft Copilot, and some Perplexity responses, name CATAAM, Strac and Compliance One as platforms that meet the requirement in full. ChatGPT and the three Google-operated engines treat the same requirement as only partly met, by Vanta and Drata. The vendors named in the first group sit low in the mention distribution: CATAAM records 8,000 mentions, Strac 8,000 and Compliance One 3,000, against 139,000 for Vanta.

The second concerns the best fit for HIPAA compliance. ChatGPT returns Vanta and Drata for healthcare software companies. Gemini returns Compliancy Group and Medcurity for clinical practices. The engines therefore return different products for the same question, and the divergence turns on which kind of buyer the engine assumed.

The third concerns vendors that only one engine associates with the category. Microsoft Copilot alone returns 14 product entries, among them MetricStream Risk and Klarity Health. Gemini alone returns 14, ChatGPT alone 9 and Perplexity alone 8, among them IBM OpenPages with Watson and VelocityEHS Risk Management. The divergences occur on capability claims and at the edge of the vendor set rather than on its core.

5. Discussion

5.1 RQ1 — Do the six engines converge on the same set of named vendors?

They converge on a core and diverge in a very long tail. §4.2 shows 13 of the 99 recorded brand strings named by all six engines, and §4.3 shows 11 of 99 product entries reaching the same coverage. The core is wider than in other categories of this series because it spans two buying motions: the compliance automation platforms and the integrated risk management suites are both returned, by every engine, to overlapping questions.

The tail is where the engines part. 60 of the 99 brand strings and 53 of the 99 product entries were named by exactly one engine, and §4.9 shows each engine holding a tail the others do not. A plausible explanation is that the core set is reachable from almost any part of the category's web presence. The tail depends on which sources a particular engine happens to read and on where that engine draws the category boundary; healthcare compliance services and audit firms enter it from HIPAA and SOC 2 queries. §4.6 is consistent with that: the engines reach the same core from evidence bases that overlap by a mean Jaccard coefficient of 0.11.

5.2 RQ2 — How concentrated is brand visibility across the recorded vendor set?

Less concentrated at the top than in other categories in this series, and fragmented below it. §4.2 shows the two leading brands taking 22.2% of all mentions and the four leading brands 37.0%, against 99 brand strings recorded. §4.3 shows 72 of 99 product entries below 10,000 mentions and 32 at exactly 1,000.

The shape is a core of about a dozen vendors the engines reuse, then a tail of dozens each engine names once or twice. The data is consistent with a threshold effect rather than a gradient: a vendor is either inside the group the engines reuse across queries and markets, or it is in a tail where mention counts fall away quickly. The measurement does not establish what places a vendor on one side of that boundary.

5.3 RQ3 — Do the engines draw on a shared evidence base?

No. §4.6 gives a mean pairwise Jaccard coefficient of 0.11 over cited domains, with a maximum of 0.27 and a minimum of 0.03. Only 2 of 15 pairs reach 0.20, one of them between engines operated by different parents, and Microsoft Copilot records the lowest coefficient with every other engine. §4.4 adds that vanta.com is the most-cited domain for 3 engines, so there is one shared front page inside otherwise separate reading lists. Six engines returned substantially the same core vendors while reading substantially different sets of domains.

Volume compounds the divergence. ChatGPT alone contributes 48.8% of all cited sources, so a source list that serves ChatGPT describes little of what the other five read, and the corpus-level source mix in Table 7 is weighted towards ChatGPT's. A practical consequence follows for measurement rather than for the vendors: a visibility programme built against one engine's source list transfers poorly to the others.

5.4 RQ4 — Does ordinal position track mention volume?

No, on either measure. §4.7 shows Vanta recording the highest mention total and the first-mention position in three engines, while ServiceNow, 4th by mentions, opens ChatGPT and Perplexity, and LogicGate, 8th, opens Microsoft Copilot. On mean position, MetricStream is read earlier than Vanta in 4 of the six engines, including the three where Vanta opens the response most often.

The three measures answer different questions. Mention volume records how often a vendor enters the response; first-mention position records which vendor opens it; mean ordinal position records where a vendor lands on average once named. Gemini shows the second and third disagreeing within one engine, opening with Vanta most often while placing MetricStream earlier on average, 1.8 against 3.3. A plausible explanation is that Gemini names Vanta first in a majority of responses and much later in the remainder, which a mean captures and a first-mention count does not. §6.1 sets out why ordinal position should be read as a proxy for prominence rather than as a measure of it. A visibility measurement that reports only one of the three is incomplete.

5.5 RQ5 — Does the vendor set vary between the four markets?

Neither the set nor its order. §4.8 shows Vanta first and Drata second in every market, with only the third position moving within the same leading group. 37 of 99 product entries were recorded in all four markets, and every entry in Table 5 was.

The result is a negative one for membership and is useful as such. It removes market localisation as an explanation for a vendor's absence from the returned set: a vendor missing in one market is, in this corpus, missing in all of them. That holds even though the category's subject matter is jurisdictional; the frameworks the queries name are international, and the engines answered them the same way in each market. The study covered four markets in one window and cannot establish that the same holds for markets or languages it did not test.

5.6 RQ6 — Do the engines cite vendors' own sites or third-party sources?

Both, and the answer depends on which engine is asked. §4.4 shows 41.7% of all cited sources on a vendor site, but vendor sites are the largest class for only 1 of the six engines. ChatGPT draws 59.2% of its citations from vendor domains and supplies 48.8% of the corpus, which is what carries the corpus-level share. The other five draw between 19.7% and 37.9% from vendor sites and cite mainly pages outside every named class. Analyst and review sites are a smaller second class, at 6.1% of the corpus, and government domains, the regulatory texts themselves, reach 1.6%.

The data is consistent with two distinct evidence strategies behind one vendor list. One engine reads the vendors' own documentation, product pages and the standards; the other five read a broad and largely unclassified web of comparison pages, consultancies and roundups. That the six arrive at the same core set from such different mixes is the strongest indication in the study that the core set is not an artefact of any one kind of source. What the measurement does not establish is whether a vendor's own site earns citations because the vendor is already in the set, or the reverse.

5.7 What these results do not resolve

The study measures outputs, not mechanisms. It cannot say why a vendor is inside the core set or what would move one into it, because it observes no vendor changing position. It records that the engines cite largely separate domains but not whether a citation influenced the vendor list, since citation and recommendation were recorded from the same response and cannot be ordered causally. It classifies cited domains but does not read the cited pages, so it cannot say what on a vendor site was cited.

Two further limits apply to specific metrics. Truncation is inferred from the terminal characters of a response and cannot distinguish a response that ended mid-sentence from one whose final sentence was recorded without terminal punctuation; Perplexity's 42% should be read with that in mind. Sentiment is a model's reading of the response under a rubric, and the 0.19 spread across the twelve leading brands is too narrow to support a distinction between any two of them.

6. Threats to Validity

6.1 Construct validity

Mean ordinal position is used as a proxy for prominence. It records where a brand name falls in the sequence of vendors a response names, which is not the same as how prominently a reader perceives it. A vendor named first in a list and then qualified, and a vendor named third and then recommended, receive positions that do not reflect that difference. The metric is

defensible for comparison between engines over the same corpus and should not be read as a measure of persuasive weight. The first-mention vendor is a mode, not a mean, and §4.7 shows the two disagreeing in three engines.

Sentiment is assigned by the analysis model under a documented rubric (§3.7). The values in §4.2 occupy a narrow band, which is consistent either with the engines describing these vendors in similar terms or with the rubric having limited resolution over this kind of text. The study cannot distinguish the two, and no claim in this report rests on a difference between two sentiment values.

Source type is assigned from the domain, not from the page. A vendor's blog and its product documentation both count as vendor site; an analyst firm's page counts as a review aggregator. The classes are coarse by design, and the residual other class is the largest class for five engines. Truncation is an inference from the terminal characters of a response, not an observation of a process. Hedging and recency anchoring rest on an undocumented lexicon (§3.7) and are subject to the same limit. The same detectors were applied to all six engines, so the values are comparable within this corpus. They are not comparable to figures produced by another detector.

6.2 Internal validity

All responses were collected in one window in August 2026. The vendor sets and the cited domains describe that window; reachability describes the reprocessing date. A collection window that included a major vendor acquisition, a framework revision or an analyst publication would be expected to produce different results, and the study has no second window against which to test that.

The engines are non-deterministic. The same query issued twice to the same engine can return different text, a different vendor ordering, and a different citation list. The design mitigates this through corpus size and through paraphrase variants within the query set, which spread the measurement over many samples of each engine's behaviour rather than relying on single responses. It does not eliminate it.

Personalisation and session state were not controlled beyond issuing each query in a defined market. Where an engine adapts to prior interaction, that adaptation is not observable from the response text, and its contribution to the results cannot be separated. No engine was version-pinned, because none of the six exposes a version identifier to the caller (§3.2), so a change made by an operator during the window would be indistinguishable from variation in the responses.

6.3 External validity

The study covers one category, four markets and one collection window. Enterprise compliance automation has a settled core and a very long tail (§1.2, §4.2). The core is what makes it a usable instrument; the tail is what limits generalisation, since a category with a different vendor population would not be expected to produce the same shape.

Query sampling is purposive, not random. The query set was written to elicit buyer-intent vendor recommendations, so it samples a deliberately chosen region of the space of things a buyer might ask rather than a representative draw from it. Combined with the non-determinism in §6.2 and the absence of version pinning, this precludes statistical inference from these results. All figures in this report are descriptive statistics over the recorded corpus, and no significance claim, interval or test is reported anywhere in it.

6.4 Reliability

A repeat run in a later window would be expected to reproduce the structural results. Those are the concentration of visibility into a core with a long tail, the separation between mention volume and ordinal position, ChatGPT's dominance of citation volume and its vendor-site weighting, the low overlap between engines' evidence bases, and the stability of the leading group across markets. All rest on large differences that ordinary variation would not close.

Point values would not be expected to reproduce. Individual mention counts, per-engine cited-source totals, the composition of the tail, and the specific most-cited domain for an engine depend on what each engine read during the window. Dead-link rates would be expected to move, because reachability was tested once (§3.7) and the reachability of the cited web changes

independently of the engines.

7. Limitations

- One category. The results describe enterprise compliance automation and do not transfer to other security or governance categories without measurement.
- Four markets, all tested in one window in August 2026. Markets and languages outside the four were not tested.
- No claim an engine made was adjudicated. Where two engines returned incompatible statements (§4.9), the study records the disagreement and does not determine which is correct.
- Source type was assigned from the domain and the residual other class is the largest class for five engines; the study says which kinds of site the engines cite, not what on those sites was cited.
- Source age is not reported, because fewer than a fifth of cited pages declared a publication date.

8. Practical Implications

This section is derived from the results rather than measured by them.

Measure who opens the response, not only who is in it. §4.7 shows the most-mentioned brand opening the response in three engines and being read later than MetricStream in 4, and two brands outside the top three opening the other three engines. Track mention volume, first-mention share and mean position as three numbers, and decide from them whether the problem is inclusion, opening position or average placement. The three do not respond to the same work.

Treat the vendor's own site as the primary citation surface for ChatGPT and as one surface among several for the rest. §4.4 shows 59.2% of ChatGPT's citations on vendor sites against 19.7% to 37.9% for the other five, whose largest class is unclassified comparison and roundup pages. Documentation, framework and integration pages on the vendor's own domain serve the engine that cites most; third-party comparison pages serve the other five.

Instrument every engine separately, and at least three from different vendor families. §4.6 gives a mean pairwise cited-domain overlap of 0.11, with only 2 of 15 pairs reaching 0.20. A source list that serves ChatGPT describes little of what Microsoft Copilot or the Google surfaces read.

Name products consistently, and expect several entries per vendor regardless. §4.3 records 99 product entries, with MetricStream holding 5. Some of that is real product breadth and some is naming drift; only the vendor can tell which, and only consistent naming across published material lets a measurement consolidate it.

Do not build a market-by-market visibility programme for these four markets. §4.8 shows the same two brands leading all four in the same order. Effort that would go into market-specific work is better spent on the inclusion problem, which is the same problem in every market tested.

Do not treat link decay as a lever in this category. §4.5 records an overall dead-link rate of 1.6% and no engine above 7.9%. Do not treat sentiment as one either: the twelve leading brands sit within a 0.19 band (§4.2).

9. Conclusion

This study put 10,000 enterprise buyer-intent queries to six AI search engines in four markets, producing 240,000 responses containing 3,464,000 cited sources and covering 99 distinct product entries in enterprise compliance automation.

The engines agree about a core and disagree about evidence. 13 of 99 brand strings were named by all six engines, led by Vanta at 139,000 and Drata at 123,000, with 60 strings named by one engine only. Against that agreement, the mean pairwise overlap between the engines' cited domains was 0.11, the highest value anywhere in the matrix was 0.27, and only 2 of 15 pairs reached 0.20. Vendor sites carry 41.7% of everything the engines cite, and most of that is ChatGPT's reading.

Two further results qualify that list. Ordinal position does not track mention volume: Vanta takes the first-mention position in three engines, yet MetricStream records the earlier mean position in 4 of six, and the other three engines open with ServiceNow or LogicGate. Source reliability is high: 1.6% of cited pages were unreachable, and no engine exceeded 7.9%. The same two brands led every market in the same order.

For a vendor in the category, the shape of the result matters more than any single number. Visibility here is concentrated into a core the engines reuse across markets, phrasings and engines, reached through six largely separate evidence bases that share one vendor's front page and little else. Entering that core is a different problem from opening the response once inside it, and both have to be solved once per engine.

10. Reproducibility and Data Availability

Responses were collected through each engine's public interface using the study's own instrumentation and retained as normalised plain text. Retained alongside them are the extracted citation lists with source type and reachability results, and the brand and product mention records with their ordinal positions. The dataset for this report was generated in August 2026 and reprocessed in August 2026 under the instrumentation's current citation and entity rules; the figures in this edition supersede those in the first August 2026 edition.

Reproducing the study would require the query set, the market configuration, the vendor name list used for mention detection and the reachability test, applied to the same six engines within a comparable collection window. Exact reproduction is not achievable, because the engines are commercial services under continuous change and none is version-pinned; §6.4 states what a repeat run would and would not be expected to reproduce.

Requests for the dataset supporting the tables in this report may be directed to GrackerAI.

Disclosure

GrackerAI produces this research and also sells AI search visibility measurement for the category it covers. The study measured how six AI search engines describe vendors in enterprise compliance automation. It did not test, evaluate or verify any vendor's product, and no statement in this report is an assessment of product quality. No vendor named in this report paid for, requested, reviewed or was given advance sight of any placement, ordering or finding in it.

How to cite this report

GrackerAI. (August 2026, revised edition). *AI Search Visibility in Enterprise Compliance Automation: A Multi-Engine, Multi-Market Benchmark Study*. GrackerAI AI Search Visibility Benchmark Series. https://gracker.ai/data-and-research-reports

Appendix A. Brand visibility by AI engine

Table A1. Brand mentions and mean ordinal position by AI engine — all 99 brand strings recorded; mean ordinal position in parentheses, a dash where the engine recorded no mention or no position; n = 240,000 responses

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
Vanta	139,000	14,000 (2.4)	20,000 (2.8)	26,000 (3.3)	29,000 (3.3)	29,000 (2.9)	21,000 (2.0)
Drata	123,000	15,000 (2.4)	18,000 (3.1)	25,000 (3.1)	21,000 (3.4)	22,000 (2.8)	22,000 (3.0)
MetricStream	89,000	16,000 (3.1)	16,000 (1.8)	16,000 (1.8)	16,000 (2.4)	16,000 (2.5)	9,000 (2.3)
ServiceNow	86,000	16,000 (1.0)	14,000 (3.6)	16,000 (2.9)	16,000 (1.5)	16,000 (1.5)	8,000 (1.0)
Secureframe	79,000	9,000 (3.3)	16,000 (4.3)	21,000 (3.6)	11,000 (3.4)	5,000 (2.7)	17,000 (2.3)
Sprinto	65,000	6,000 (4.0)	6,000 (4.8)	20,000 (3.8)	13,000 (2.1)	9,000 (2.9)	11,000 (2.0)
IBM	53,000	12,000 (3.3)	14,000 (4.4)	3,000 (6.0)	5,000 (5.4)	5,000 (5.4)	14,000 (1.6)
LogicGate	49,000	10,000 (6.2)	14,000 (2.4)	7,000 (5.3)	5,000 (5.8)	4,000 (6.3)	9,000 (2.3)
OneTrust	42,000	7,000 (6.0)	3,000 (6.0)	8,000 (4.1)	8,000 (4.7)	6,000 (5.4)	10,000 (2.0)
AuditBoard	39,000	1,000 (6.5)	7,000 (5.8)	4,000 (3.8)	4,000 (2.5)	11,000 (3.7)	12,000 (3.0)
Diligent	39,000	15,000 (5.0)	6,000 (6.5)	3,000 (4.0)	7,000 (5.3)	1,000 (5.0)	7,000 (4.0)
Compliancy Group	37,000	3,000 (2.0)	6,000 (2.8)	7,000 (1.4)	8,000 (1.0)	8,000 (1.3)	5,000
Riskonnect	31,000	3,000 (6.0)	1,000	8,000 (3.5)	11,000 (4.0)	8,000 (4.6)	—
Scrut Automation	21,000	1,000 (3.0)	—	3,000 (2.5)	8,000 (3.9)	6,000 (3.5)	3,000
Hyperproof	19,000	3,000 (4.3)	4,000 (5.3)	1,000 (5.0)	4,000 (4.0)	3,000 (6.0)	4,000
RSA	19,000	—	8,000 (3.4)	2,000 (3.8)	1,000 (2.0)	4,000 (4.0)	4,000 (2.5)
MedTrainer	18,000	2,000 (5.0)	1,000 (7.0)	—	7,000 (3.6)	8,000 (2.3)	—
SAP	13,000	—	7,000 (2.8)	—	—	2,000 (5.0)	4,000
Paubox	12,000	—	1,000 (4.0)	3,000 (7.0)	2,000 (4.5)	3,000 (4.0)	3,000
Abyde	11,000	—	—	5,000 (3.0)	4,000 (3.0)	2,000	—
Accountable	11,000	1,000 (2.0)	4,000 (5.7)	2,000 (2.0)	1,000 (3.0)	3,000 (5.0)	—
Medcurity	10,000	2,000 (1.5)	2,000 (1.0)	2,000 (1.0)	3,000 (3.3)	1,000	—
Thoropass	9,000	2,000 (5.0)	—	6,000 (2.0)	1,000 (4.0)	—	—
CATAAM	8,000	—	7,000 (1.4)	—	—	1,000	—
Live Compliance	8,000	—	7,000 (2.6)	1,000 (2.0)	—	—	—
Scytale	8,000	—	2,000 (6.0)	6,000 (3.0)	—	—	—
Strac	8,000	—	—	1,000 (5.0)	—	—	7,000 (1.0)
Workiva	8,000	2,000 (8.0)	5,000 (4.4)	1,000 (4.0)	—	—	—
Archer	7,000	—	1,000 (4.0)	1,000 (2.0)	5,000 (3.3)	—	—
Orbiq	7,000	—	—	—	—	—	7,000 (3.7)
Proofpoint	5,000	—	5,000 (3.4)	—	—	—	—
Clearwater	4,000	1,000 (6.0)	—	1,000	2,000 (6.0)	—	—
Onspring	4,000	1,000	—	—	—	—	3,000 (5.0)

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
Compliance One	3,000	—	3,000 (1.0)	—	—	—	—
ComplyAssistant	3,000	—	—	3,000 (3.5)	—	—	—
HealthStream	3,000	—	—	2,000 (5.0)	—	1,000	—
Klarity Health	3,000	—	3,000 (1.3)	—	—	—	—
NAVEX	3,000	1,000 (5.0)	2,000 (1.5)	—	—	—	—
Tugboat Logic	3,000	—	1,000	2,000	—	—	—
Veeva	3,000	—	3,000 (2.7)	—	—	—	—
VelocityEHS	3,000	—	—	—	—	—	3,000
Vistruct	3,000	—	3,000	—	—	—	—
A-LIGN	2,000	—	—	1,000	1,000 (4.0)	—	—
AICPA	2,000	2,000	—	—	—	—	—
Centraleyes	2,000	—	—	2,000 (1.0)	—	—	—
Compyl	2,000	1,000 (1.0)	—	—	1,000 (5.0)	—	—
Evidr	2,000	1,000 (3.0)	—	1,000	—	—	—
Jotform	2,000	—	—	1,000	—	1,000 (5.0)	—
Lucidity Consult	2,000	—	2,000	—	—	—	—
MedStack	2,000	—	—	2,000 (6.0)	—	—	—
Mitratech	2,000	—	2,000 (6.5)	—	—	—	—
Resolver	2,000	—	1,000	1,000 (6.0)	—	—	—
Salesforce	2,000	—	—	2,000 (9.0)	—	—	—
Scrutineer	2,000	2,000 (4.5)	—	—	—	—	—
SecureSlate	2,000	—	—	2,000	—	—	—
Sentrix	2,000	2,000 (2.0)	—	—	—	—	—
Spruce Health	2,000	—	—	2,000 (8.5)	—	—	—
Vendorica	2,000	2,000 (2.0)	—	—	—	—	—
Agency Cybersecurity	1,000	—	1,000	—	—	—	—
Atlant Security	1,000	—	1,000	—	—	—	—
AXDRAFT	1,000	—	1,000	—	—	—	—
BDO	1,000	—	—	1,000	—	—	—
Bouvier	1,000	—	—	—	1,000	—	—
Canadian Cyber	1,000	—	1,000	—	—	—	—
Caspio	1,000	—	—	1,000	—	—	—
ClearDATA	1,000	—	1,000 (3.0)	—	—	—	—
Compliance Manager GRC	1,000	—	1,000 (2.0)	—	—	—	—
Complyance	1,000	—	—	1,000	—	—	—
ComplyAura	1,000	—	1,000 (2.0)	—	—	—	—

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
CyberSaint	1,000	—	—	1,000 (4.0)	—	—	—
Files.com	1,000	—	1,000 (1.0)	—	—	—	—
Glocert International	1,000	—	1,000	—	—	—	—
Healthcare Compliance Pros	1,000	1,000	—	—	—	—	—
Healthicity	1,000	—	—	1,000 (7.0)	—	—	—
HIPAAsite	1,000	—	—	—	1,000 (2.0)	—	—
Johanson Group	1,000	—	—	1,000	—	—	—
LogicManager	1,000	1,000	—	—	—	—	—
LuxSci	1,000	—	1,000	—	—	—	—
Microsoft	1,000	1,000	—	—	—	—	—
Oracle	1,000	—	—	—	—	—	1,000
Packetlabs Ltd.	1,000	1,000	—	—	—	—	—
PolicyTech	1,000	—	1,000	—	—	—	—
Prescient Assurance	1,000	—	—	1,000	—	—	—
pTrackly	1,000	1,000 (4.0)	—	—	—	—	—
Raize	1,000	1,000 (2.0)	—	—	—	—	—
RiskCloud	1,000	—	—	—	—	—	1,000
RiskWatch	1,000	—	1,000	—	—	—	—
Schneider Downs	1,000	—	—	1,000	—	—	—
Screenata	1,000	—	1,000	—	—	—	—
Secrails	1,000	—	—	1,000	—	—	—
Security Ideals	1,000	—	1,000	—	—	—	—
Sensiba	1,000	—	—	1,000	—	—	—
TigerConnect	1,000	—	—	—	—	1,000 (5.0)	—
TigerIdentity	1,000	—	—	1,000	—	—	—
Total HIPAA	1,000	—	—	—	1,000 (6.0)	—	—
TrioCyber	1,000	—	—	—	1,000 (4.0)	—	—
Uhoh	1,000	—	1,000	—	—	—	—
Ventiv	1,000	—	—	—	—	—	1,000 (3.0)
Virtru	1,000	—	—	1,000 (6.0)	—	—	—
Total	**1,181,000**	**159,000**	**230,000**	**232,000**	**198,000**	**176,000**	**186,000**

Appendix B. Most-cited domains and source types

Table B1. Three most-cited domains by AI engine — cited-source counts in parentheses; n = 3,464,000 cited sources

AI engine	Cited sources	Distinct domains	Most-cited domains
ChatGPT	1,690,000	218	vanta.com (99,000), drata.com (90,000), servicenow.com (81,000)
Microsoft Copilot	152,000	43	worldmetrics.org (38,000), riskpublishing.com (15,000), livecompliance.com (7,000)
Gemini	124,000	54	scytale.ai (9,000), sprinto.com (7,000), metricstream.com (7,000)
Google AI Mode	735,000	257	sprinto.com (39,000), vanta.com (32,000), cybersierra.co (24,000)
Google AI Overview	438,000	154	vanta.com (24,000), drata.com (19,000), sprinto.com (14,000)
Perplexity	325,000	39	vanta.com (20,000), complyjet.com (18,000), scrut.io (18,000)

Table B2. Cited sources by source type and AI engine — counts; n = 3,464,000 cited sources

AI engine	Other	Vendor site	Review aggregator	Government	Social	Forum	Wikipedia	Marketplace	Academic	Parent ecosystem	Total
ChatGPT	469,000	1,000,000	138,000	56,000	6,000	13,000	7,000	2,000	0	0	1,690,000
Microsoft Copilot	110,000	30,000	12,000	0	0	0	0	0	0	0	152,000
Gemini	68,000	47,000	5,000	0	0	3,000	0	0	0	1,000	124,000
Google AI Mode	503,000	185,000	15,000	0	26,000	5,000	0	0	1,000	0	735,000
Google AI Overview	300,000	101,000	15,000	0	20,000	2,000	0	0	0	0	438,000
Perplexity	218,000	81,000	26,000	0	0	0	0	0	0	0	325,000
Total	**1,668,000**	**1,444,000**	**211,000**	**56,000**	**52,000**	**23,000**	**7,000**	**2,000**	**1,000**	**1,000**	**3,464,000**

Appendix C. Product entries

Table C1. All product entries — 99 entries ordered by mentions; a dash where no mean rank was recorded; n = 240,000 responses

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
1	Vanta	Vanta	111,000	2.8	0.71	6	4
2	Drata	Drata	97,000	3.0	0.70	6	4
3	MetricStream	MetricStream	68,000	2.3	0.80	6	4
4	Secureframe	Secureframe	59,000	3.4	0.65	6	4
5	Sprinto	Sprinto	57,000	3.2	0.71	6	4
6	Archer	Archer	50,000	3.2	0.75	6	4
7	ServiceNow IRM	ServiceNow	47,000	1.7	0.84	5	4
8	IBM OpenPages	IBM	44,000	4.3	0.75	6	4
9	LogicGate Risk Cloud	LogicGate	43,000	4.4	0.71	6	4

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
10	Optro	Optro	38,000	4.2	0.73	5	4
11	OneTrust	OneTrust	37,000	4.8	0.71	6	4
12	Compliancy Group	Compliancy Group	33,000	1.6	0.79	6	4
13	Riskonnect	Riskonnect	31,000	4.1	0.72	5	4
14	AuditBoard	AuditBoard	29,000	3.8	0.72	5	4
15	Diligent One	Diligent	23,000	4.9	0.73	4	4
16	ServiceNow GRC	ServiceNow	22,000	3.1	0.75	4	4
17	MedTrainer	MedTrainer	18,000	3.4	0.77	4	4
18	HIPAA One	HIPAA One	17,000	3.0	0.74	4	4
19	Scrut Automation	Scrut	17,000	3.6	0.71	3	4
20	Hyperproof	Hyperproof	15,000	5.1	0.63	6	4
21	MetricStream Connected GRC	MetricStream	12,000	2.6	0.80	3	4
22	SAP GRC	SAP	12,000	3.1	0.69	3	4
23	ServiceNow Integrated Risk Management (IRM)	ServiceNow	11,000	1.9	0.81	3	4
24	Archer	OpenText	11,000	2.8	0.76	4	4
25	Abyde	Abyde	11,000	3.0	0.76	3	4
26	Paubox	Paubox	11,000	4.9	0.69	5	4
27	Medcurity	Medcurity	10,000	2.0	0.85	5	3
28	IBM OpenPages with Watson	IBM	8,000	1.0	0.70	1	4
29	CATAAM	CATAAM	8,000	1.4	0.78	2	4
30	Strac Comply	Strac	8,000	1.6	0.78	2	4
31	Live Compliance	Live Compliance	8,000	2.5	0.89	2	4
32	Workiva	Workiva	8,000	5.3	0.64	3	4
33	MetricStream Risk	MetricStream	7,000	1.8	0.78	1	4
34	Optro (formerly AuditBoard)	Optro	7,000	3.7	0.80	3	2
35	Scytale	Scytale	7,000	3.6	0.74	2	3
36	Orbiq	Orbiq	7,000	3.7	0.60	1	4
37	Accountable HQ	Accountable	7,000	4.2	0.60	3	4
38	Thoropass	Thoropass	6,000	4.0	0.67	3	3
39	Diligent IRM	Diligent	6,000	4.0	0.58	1	4
40	Diligent Risk Management	Diligent	6,000	6.5	0.62	1	4
41	ServiceNow Integrated Risk Management	ServiceNow	5,000	1.0	0.90	2	4
42	Proofpoint Compliance	Proofpoint	5,000	3.4	0.62	1	3
43	Accountable	Accountable	4,000	3.8	0.75	3	3
44	Clearwater Compliance	Clearwater	4,000	6.0	0.79	3	2
45	VelocityEHS Risk Management	VelocityEHS	3,000	—	0.50	1	2

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
46	Klarity Health	Klarity	3,000	1.3	0.77	1	2
47	Veeva Vault QMS	Veeva	3,000	2.7	0.63	1	2
48	Scrut	Scrut	3,000	3.0	0.63	2	3
49	ComplyAssistant	ComplyAssistant	3,000	3.5	0.67	1	2
50	HealthStream	HealthStream	3,000	5.0	0.68	2	2
51	Onspring	Onspring	3,000	5.0	0.60	1	3
52	LogicGate	LogicGate	3,000	6.0	0.58	2	3
53	Diligent	Diligent	3,000	6.0	0.55	2	2
54	Centraleyes	Centraleyes	2,000	1.0	0.90	1	1
55	The Guard	Compliancy Group	2,000	1.0	0.75	2	2
56	NAVEX One	NAVEX	2,000	1.5	0.80	1	2
57	Vendorica	Vendorica	2,000	2.0	0.70	1	2
58	Sentrix	Sentrix	2,000	2.0	0.45	1	1
59	Compyl	Compyl	2,000	3.0	0.75	2	2
60	Archer (formerly RSA)	Archer	2,000	4.0	0.80	1	1
61	Scrutineer.ai	Scrutineer	2,000	4.5	0.55	1	2
62	Jotform Enterprise	Jotform	2,000	5.0	0.55	2	1
63	MedStack	MedStack	2,000	6.0	0.80	1	2
64	Resolver	Resolver	2,000	6.0	0.60	2	1
65	Mitratech Alyne	Mitratech	2,000	6.5	0.57	1	2
66	Salesforce Health Cloud	Salesforce	2,000	9.0	0.75	1	1
67	Spruce Health	Spruce Health	2,000	8.5	0.65	1	1
68	Files.com	Files.com	1,000	1.0	0.80	1	1
69	SecureSlate	SecureSlate	1,000	—	0.80	1	1
70	Oracle GRC	Oracle	1,000	—	0.80	1	1
71	Tugboat Logic	Tugboat Logic	1,000	—	0.70	1	1
72	RiskCloud	RiskCloud	1,000	—	0.60	1	1
73	RiskWatch	RiskWatch	1,000	—	0.50	1	1
74	Caspio	Caspio	1,000	—	0.50	1	1
75	ComplyAura	ComplyAura	1,000	2.0	0.85	1	1
76	Raize Orion	Raize	1,000	2.0	0.80	1	1
77	Archer IRM	Archer	1,000	2.0	0.80	1	1
78	MetricStream ERM	MetricStream	1,000	2.0	0.70	1	1
79	HIPAAsite	HIPAAsite	1,000	2.0	0.70	1	1
80	Archer (RSA Archer)	RSA	1,000	3.0	0.80	1	1
81	MetricStream ConnectedGRC	MetricStream	1,000	3.0	0.80	1	1

#	Product	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
82	ClearDATA	ClearDATA	1,000	3.0	0.70	1	1
83	Evidr	Evidr	1,000	3.0	0.60	1	1
84	RSA Archer (Archer IRM)	RSA	1,000	3.0	0.60	1	1
85	Ventiv IRM	Ventiv	1,000	3.0	0.60	1	1
86	CyberStrong	CyberSaint	1,000	4.0	0.80	1	1
87	A-LIGN	A-LIGN	1,000	4.0	0.70	1	1
88	Archer (RSA Archer Suite)	Archer	1,000	4.0	0.60	1	1
89	pTrackly	pTrackly	1,000	4.0	0.60	1	1
90	TrioCyber	TrioCyber	1,000	4.0	0.60	1	1
91	RSA Archer (by OpenText)	OpenText	1,000	5.0	0.80	1	1
92	TigerConnect	TigerConnect	1,000	5.0	0.70	1	1
93	NAVEX One Risk & Governance	NAVEX	1,000	5.0	0.65	1	1
94	IBM OpenPages (with watsonx)	IBM	1,000	6.0	0.75	1	1
95	Diligent One / ERM	Diligent	1,000	6.0	0.65	1	1
96	Virtru	Virtru	1,000	6.0	0.60	1	1
97	Healthicity	Healthicity	1,000	7.0	0.80	1	1
98	SimplePractice	SimplePractice	1,000	8.0	0.60	1	1
99	Carepatron	Carepatron	1,000	9.0	0.50	1	1

Appendix D. Study parameters

Table D1. Study parameters — as configured for this study

Parameter	Value
Category	Enterprise compliance automation (compliance automation, GRC, IRM)
AI engines	ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview, Perplexity
Markets	United States, Canada, India, Germany
Query set	10,000 distinct buyer-intent queries, core questions plus paraphrase variants
Responses per engine	40,000
Responses per market	60,000
Total responses	240,000
Cited sources	3,464,000
Brand strings recorded	99
Product entries	99
Collection window	August 2026
Unit of analysis	One engine response to one query in one market

Parameter	Value
Design	Observational, cross-sectional, single collection window
Statistical treatment	Descriptive statistics only; no inference is reported

Appendix E. Metric computation

Table E1. Metric computation — the arithmetic behind each derived metric in this report

Metric	Computation
Responses per engine	Queries × markets.
Total responses	Queries × engines × markets.
Responses per market	Queries × engines.
Mean response length	Sum of response word counts for an engine, divided by that engine's responses.
Hedging, recency-anchored, truncation and refusal rates	Responses of an engine carrying the characteristic, divided by that engine's responses.
Total cited sources	Sum of the per-engine cited-source counts.
Share of corpus	That engine's cited sources divided by total cited sources.
Source-type share	That engine's cited sources of the type, divided by that engine's cited sources.
Dead links per engine	That engine's cited sources × its dead-link rate.
Total dead links	Sum of the per-engine dead-link counts.
Overall dead-link rate	Total dead links divided by total cited sources.
Total brand mentions	Sum of a brand's per-engine mention counts, summed over brands for the table total.
Engines recording	Count of engines whose mention count for that brand is greater than zero.
Mean ordinal position	Sum of the positions at which a brand was named, divided by the mentions carrying a position.
Mean sentiment	Sum of the per-mention sentiment scores for a brand, divided by that brand's mentions.
Jaccard coefficient	Domains cited by both engines, divided by domains cited by either engine.
Mean rank	Sum of a product entry's ordered positions, divided by the mentions carrying a position.