Gracker.ai

RESEARCH REPORT

AI Search Visibility in Enterprise Third-Party Risk Management

A multi-engine, multi-market benchmark study of how six AI search engines — ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview and Perplexity — answer buyer-intent questions about enterprise third-party risk management, vendor risk monitoring and questionnaire automation across the United States, Canada, Germany, India and Australia.

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

AI engine	Responses	Cited sources	Markets
ChatGPT	50,000	535,000	5
Microsoft Copilot	50,000	204,000	5
Gemini	50,000	225,000	5
Google AI Mode	50,000	160,000	5
Google AI Overview	50,000	449,000	5
Perplexity	50,000	25,000	5
Total	**300,000**	**1,598,000**	**5**

Published: September 2026

Research period: September 2026

Category: Enterprise third-party risk management, vendor risk monitoring and questionnaire automation

Scope: 10,000 buyer-intent queries · 300,000 AI engine responses · 6 AI engines · 5 markets

Evidence base: 1,598,000 cited sources · 164 distinct product entries

Prepared by: GrackerAI · AI Search Visibility Benchmark Series

Abstract

Enterprise buyers assembling a shortlist of third-party risk management platforms increasingly begin inside an AI-generated answer, so the vendor set an AI engine returns now forms the consideration set before any vendor is contacted. This study measures how six commercial AI search engines answer buyer-intent questions about enterprise third-party risk management, vendor risk monitoring and questionnaire automation. Each engine received 10,000 buyer-intent queries in each of five markets during a single collection window in September 2026, producing 300,000 responses that contained 1,598,000 cited sources and covered 164 distinct product entries. For every response the study recorded length, cited domains, source reachability, hedging, recency anchoring, truncation, refusal, and each brand mention with its ordinal position. The engines converge on a narrow vendor set: 9 of 15 brands were named by all six engines, and the five leading brands took 63.1% of the mentions recorded across the 15. Mention volume and ordinal position separate: OneTrust is the most-mentioned brand at 191,000 mentions and is the first-mention vendor in none of the six engines, while Bitsight, second on mentions, is the first-mention vendor in three. The engines diverge on evidence: the highest cross-engine cited-domain overlap is a Jaccard coefficient of 0.40, between two engines operated by the same parent, and the mean across the 15 engine pairs is 0.19. Overall, 13.9% of cited URLs did not resolve when tested, ranging from 4.0% to 20.0% by engine, and the leading product entry differed in every one of the five markets.

Contents

1. Introduction

1.1 Background

An enterprise buyer evaluating third-party risk management software now begins the process inside a generated answer. The buyer asks an AI search engine which platforms monitor vendors continuously, automate security questionnaires or connect to an existing service-management and procurement stack. 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 visited a single vendor site.

The structural consequence is narrow. A vendor absent from the returned list is not ranked low; it is not present. The buyer has no mechanism equivalent to scrolling to the second page of results, because a generated answer has no second page. Where a search results page offered ten positions and a set of related queries, a generated answer typically names between three and eight vendors and closes. Visibility in this setting is closer to inclusion than to ranking, and the gap 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 five 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 third-party risk management is a suitable instrument for this measurement. The vendor set is mature and bounded: the engines in this study named a stable core of security-ratings vendors, dedicated third-party risk platforms and broader governance suites, and 9 of the 15 brands recorded appeared in every engine. A category with a settled vendor population isolates engine behaviour from genuine market churn.

The evaluation criteria are also stable and specific. Across the corpus the engines returned the same selection factors: continuous outside-in monitoring and security ratings, automation of vendor security questionnaires, AI-assisted reading of vendor evidence such as audit reports, and onboarding of vendors at volume. To those they add integration with IT service management tools and with source-to-pay procurement systems. These criteria are concrete enough that responses can be compared to each other rather than to a general description of risk software.

The category also carries an internal tension the engines describe directly. The responses contrast outside-in ratings engines with workflow systems of record, dedicated third-party risk platforms with broad governance suites, pre-assessed shared vendor networks with bespoke questionnaires, and quote-only pricing with published tiers. A category in which the engines themselves articulate the trade-offs gives the ordering and first-mention measures something to attach to.

Buying behaviour in the category is research-heavy and the cost of invisibility is high. Third-party risk platforms are bought by security, risk and procurement teams who shortlist before contacting vendors, and contracts are enterprise-scale and multi-year. A vendor omitted from the generated shortlist loses the opportunity at the earliest stage of the cycle, before any conversation in which the omission could be corrected.

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 tail each engine holds alone. It measures the overlap between the evidence bases the engines cite, pair by pair, so that agreement on vendors can be distinguished from agreement on sources. It separates mention volume from ordinal position and shows, for this category, that the two identify different leading vendors. It tests whether the vendor set moves between five markets on different continents, and reports the share of the cited evidence base that no longer resolves.

2. Research Questions

1. **RQ1.** Do the six engines converge on the same set of named vendors for enterprise third-party risk management?
2. **RQ2.** How concentrated is brand visibility, and does brand-level visibility carry through to product-level naming?
3. **RQ3.** Do the engines draw on a shared evidence base, measured by overlap between their cited domains?
4. **RQ4.** Does the vendor an engine names first track the vendor it mentions most?
5. **RQ5.** Does the vendor set vary between the five markets?
6. **RQ6.** Where the engines disagree, on which buyer requirement do they disagree?

3. Methodology

3.1 Study design

The study is observational and cross-sectional. All responses were collected in a single window in September 2026, and no engine was queried outside that window. No experimental manipulation was applied: the engines were asked buyer-intent questions as a buyer would ask them, and their responses were recorded as returned. The design supports description of what the engines did in that window and comparison between them. It does not support inference about a population of queries or about engine behaviour at other times, for the reasons given in §6.

3.2 Engines under test

Six AI search engines were tested: ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview and Perplexity. All six are commercial services. None is version-pinned, and none exposes a stable identifier for the model or retrieval configuration serving a given response. The study therefore describes each engine as a service observed in the collection window, not as a fixed system.

Three of the six — Gemini, Google AI Mode and Google AI Overview — are operated by the same parent. The cross-engine overlap results in §4.6 are read with that in mind.

3.3 Markets

Every query was issued in five markets: the United States, Canada, Germany, India and Australia. The markets were chosen to span three continents, two regulatory environments in Europe and Asia-Pacific alongside North America, and both English-language and non-English-language jurisdictions. Market variance was tested because third-party

risk obligations differ by jurisdiction, so a category-specific localisation of the vendor set was plausible and is reported in §4.8.

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 with paraphrase variants of those questions, so that the stability of a returned vendor set under rewording can be observed rather than assumed. The questions address continuous third-party monitoring, questionnaire automation, integration with service-management and procurement systems, and vendor onboarding at volume. 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 5 markets gives 50,000 responses per engine, and 10,000 queries across 6 engines and 5 markets gives 300,000 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 300,000 responses, distributed evenly across the six engines at 50,000 responses each and across the five 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 or mid-structure, 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
Brand mention	One occurrence of a vendor brand name in a response. Repeat occurrences within one response are counted once.	count
Share of mentions	A brand's mentions divided by the mentions of all brands recorded.	% of brand mentions
Engines recording	Number of engines in which a brand's mention count is greater than zero.	count

Metric	Definition	Unit
Mean ordinal position	Arithmetic mean of the position at which a brand is named within a response, counting named vendors from one. Not recorded where the instrumentation returned no position.	position
First-mention vendor	The brand the instrumentation recorded as most often named before any other brand in an engine's responses.	brand
Mean sentiment	Arithmetic mean of the polarity score assigned to the sentence containing a brand mention, on a zero-to-one scale where higher is more favourable.	score
Cited source	One URL attached to a response by the engine as a reference.	count
Share of cited sources	An engine's cited sources divided by the total cited sources in the corpus.	% of cited sources
Distinct domains	The number of different registrable domains cited by that engine across its responses. Reported per engine and never summed across engines.	count
Most-cited domain	The registrable domain with the highest cited-source count for that engine.	domain
Dead-link rate	Share of an engine's cited URLs that did not resolve to a reachable page when tested.	% 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 distinct product string recorded against a brand. A brand may hold several entries where engines named its products differently.	count
Mean rank	Arithmetic mean of the ordered position of a product entry within the responses that named it.	position
Market top pick	The product entry the instrumentation recorded as leading the responses collected in one market.	product entry

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.

Citations were extracted as the set of URLs the engine attached to a response. Each URL was reduced to its registrable domain for the distinct-domain and overlap measures, and each was tested for reachability; a URL that did not resolve to a reachable page at test time was recorded as a dead link. Reachability was tested once, during the collection window, so the dead-link results describe the state of the cited web at that moment.

Brand mentions were detected against a vendor name list and recorded with the ordinal position at which the brand appeared among the named vendors in that response, counting from one. Product entries were recorded as the distinct product strings the engines used, which is why one brand may hold several entries in Appendix C. Sentiment

was scored on the sentence containing each mention and reported on a zero-to-one scale. For three brand-and-engine cells the instrumentation returned a mention count with no ordinal position; those cells report the count and no position.

Four implementation details are not documented by the instrumentation and are therefore not described here. They are the lexicon used to detect qualifying and date-anchoring language, the model used to assign sentiment polarity, the rule by which the first-mention vendor is selected for an engine, and the rule by which a market top pick is selected. §6.1 states the consequence for how those metrics should be read. Source-type classification was recorded by the instrumentation but did not separate the corpus, 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 = 50,000 responses per engine, 300,000 in total

AI engine	Responses	Mean words	Hedging	Recency-anchored	Truncation	Refusal	First-mention vendor
ChatGPT	50,000	520.0	17%	23%	91%	0%	Whistic
Microsoft Copilot	50,000	475.8	8%	67%	0%	0%	Bitsight
Gemini	50,000	348.1	0%	0%	31%	0%	ServiceNow
Google AI Mode	50,000	292.4	0%	2%	2%	0%	UpGuard
Google AI Overview	50,000	211.0	4%	18%	0%	0%	Bitsight
Perplexity	50,000	318.4	0%	20%	20%	0%	Bitsight

Mean response length spans a factor of 2.5, from 211.0 words for Google AI Overview to 520.0 for ChatGPT. Figure 1 shows the distribution. ChatGPT and Microsoft Copilot sit above 470 words; the two Google surfaces sit below 300.

Truncation was recorded in 4 of the six engines, ranging from 2% to 91%, the latter for ChatGPT, whose responses are also the longest and most often structured as tables. Microsoft Copilot and Google AI Overview recorded none. Recency anchoring is highest for Microsoft Copilot at 67%, more than double the next value, and Gemini recorded none. Hedging is highest for ChatGPT at 17%; Gemini, Google AI Mode and Perplexity recorded none. No engine refused a query: the refusal rate is 0% for all six.

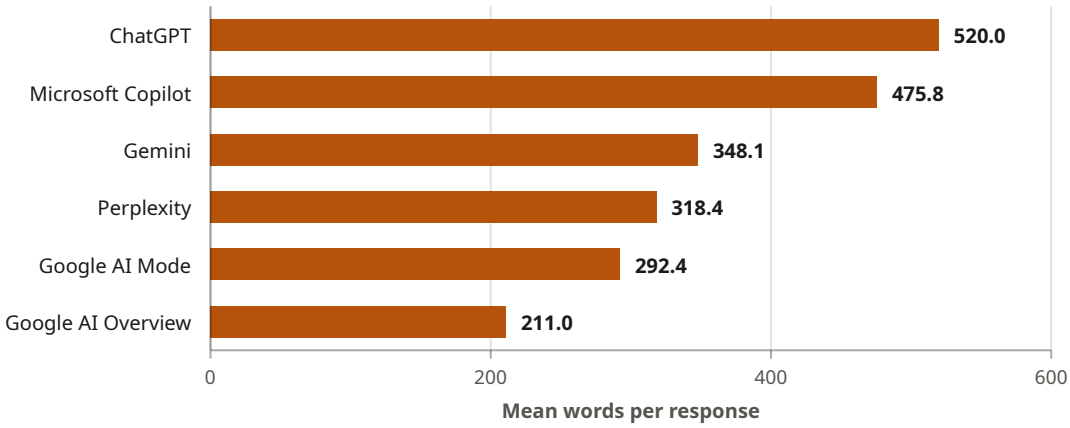

Figure 1. Mean response length by AI engine. Arithmetic mean word count per response; n = 50,000 responses per engine. Source: GrackerAI AI Search Visibility Benchmark Series, September 2026.

4.2 Brand visibility

15 brands were recorded across the corpus, with 1,237,000 brand mentions between them. Table 4 gives each brand's mentions, share, the number of engines recording it, and mean sentiment where the instrumentation scored it.

Table 4. Brand visibility — all 15 brands recorded; n = 300,000 responses, 1,237,000 brand mentions

Brand	Mentions	Share of mentions	Engines recording	Mean sentiment
OneTrust	191,000	15.4%	6	0.77
Bitsight	159,000	12.9%	6	0.80
UpGuard	154,000	12.4%	6	0.80
SecurityScorecard	146,000	11.8%	6	0.75
ProcessUnity	131,000	10.6%	6	0.79
Panorays	104,000	8.4%	6	0.78
ServiceNow	86,000	7.0%	6	0.78
Vanta	59,000	4.8%	6	0.79
Whistic	51,000	4.1%	5	0.80
MetricStream	34,000	2.7%	5	0.74
Drata	29,000	2.3%	5	0.77
Riskonnect	27,000	2.2%	6	—
Venminder	24,000	1.9%	4	—
Aravo	22,000	1.8%	5	—
Black Kite	20,000	1.6%	4	—
Total	**1,237,000**	**100.0%**		

OneTrust records the highest mention count at 191,000, 15.4% of all brand mentions, followed by Bitsight at 159,000, UpGuard at 154,000, SecurityScorecard at 146,000 and ProcessUnity at 131,000. Those five brands account for 63.1% of all brand mentions, and the ten most-mentioned for 90.1%. The fifteenth brand, Black Kite, records 20,000. Figure 2 shows the full distribution.

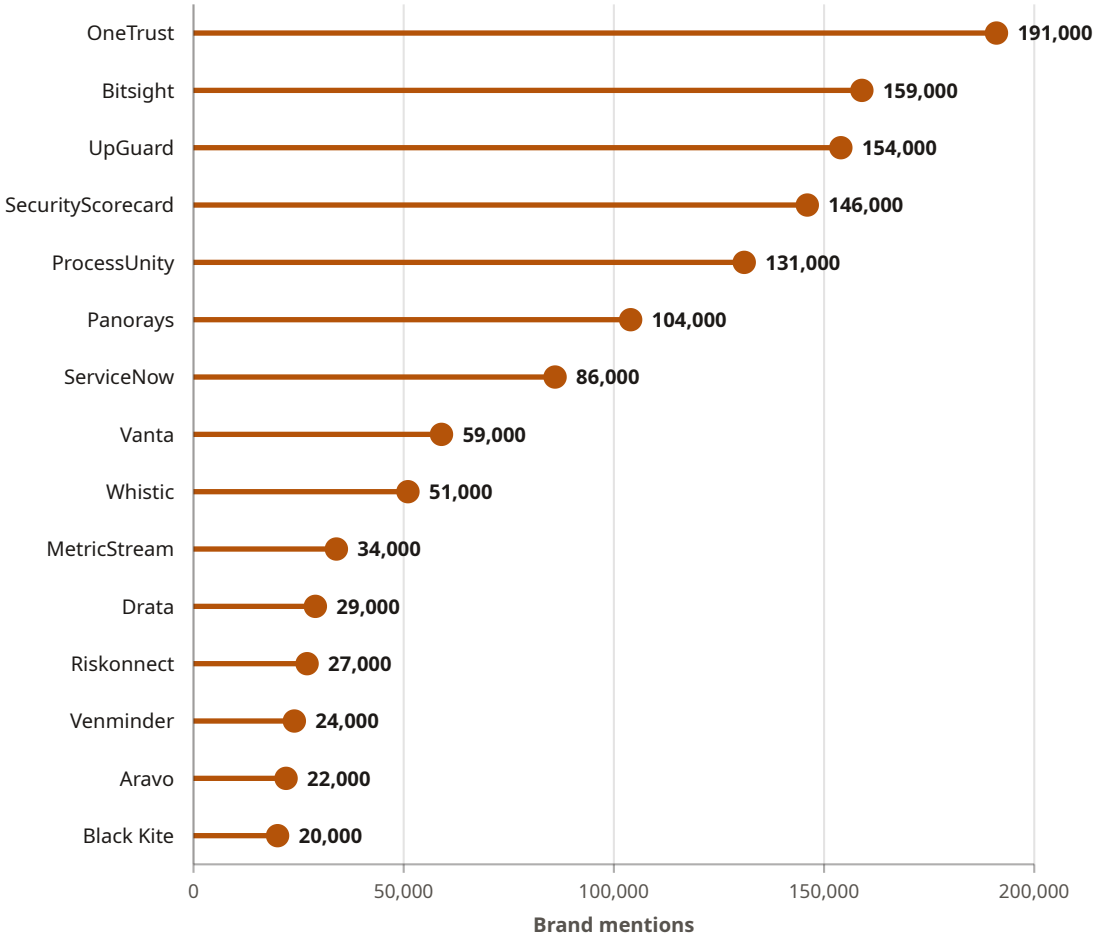

Figure 2. Brand mentions by vendor. All 15 brands recorded; n = 300,000 responses. Source: GrackerAI AI Search Visibility Benchmark Series, September 2026.

9 of the 15 brands were recorded by all six engines: OneTrust, Bitsight, UpGuard, SecurityScorecard, ProcessUnity, Panorays, ServiceNow, Vanta and Riskonnect. The remaining six were recorded by four or five engines each; no brand was recorded by fewer than four. Whistic, MetricStream, Drata and Aravo were absent from Perplexity; Venminder was absent from ChatGPT and Gemini; Black Kite was absent from Google AI Mode and Google AI Overview.

OneTrust is the most-mentioned brand in three engines — ChatGPT, Gemini and Google AI Mode — and Bitsight is the most-mentioned brand in the other three. Mean sentiment for the 11 brands scored ranges from 0.74 to 0.80, a span of 0.06.

4.3 Product-level results

The engines named products using 164 distinct product strings, recorded here as product entries. Table 5 gives the ten most-mentioned entries; Appendix C gives all 164.

Table 5. Ten most-mentioned product entries — of 164 recorded; n = 300,000 responses

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Bitsight	Bitsight	136,000	2.8	0.82	6	5
UpGuard	UpGuard	129,000	2.8	0.82	6	5
SecurityScorecard	SecurityScorecard	120,000	3.8	0.79	6	5
OneTrust	OneTrust	115,000	3.5	0.79	6	5
ProcessUnity	ProcessUnity	101,000	4.4	0.83	6	5
Panorays	Panorays	89,000	3.9	0.81	6	5
Prevalent	Prevalent	49,000	4.6	0.77	6	5
Whistic	Whistic	42,000	3.2	0.80	5	5
Vanta	Vanta	38,000	4.3	0.82	6	5
UpGuard Vendor Risk	UpGuard	22,000	1.8	0.84	5	5

At product level the order differs from the brand level. The most-mentioned product entry is the bare string "Bitsight" at 136,000, then "UpGuard" at 129,000 and "SecurityScorecard" at 120,000; the bare string "OneTrust" is fourth at 115,000. OneTrust's brand mentions are spread over 12 product entries, more than any brand except ServiceNow, which holds 19. All ten entries in Table 5 were recorded in all five markets, and 8 of the ten in all six engines.

The tail is long. 103 of the 164 entries were named by exactly one engine, and 80 entries record exactly 1,000 mentions. 9 entries were recorded by all six engines, and 28 by all five markets. Three entries carry a recorded price point; the remainder record none.

4.4 Citation volume

The engines attached 1,598,000 cited sources to their responses. Table 6 gives each engine's citation volume, the number of distinct domains it drew on, and its most-cited domain.

Table 6. Citation volume by AI engine — n = 1,598,000 cited sources

AI engine	Cited sources	Share of cited sources	Distinct domains	Most-cited domain	Citations to it	Self-citation
ChatGPT	535,000	33.5%	64	whistic.com	65,000	0%
Microsoft Copilot	204,000	12.8%	23	zipdo.co	37,000	0%
Gemini	225,000	14.1%	31	spycloud.com	87,000	0%
Google AI Mode	160,000	10.0%	51	gartner.com	16,000	0%
Google AI Overview	449,000	28.1%	65	bitsight.com	49,000	0%
Perplexity	25,000	1.6%	19	upguard.com	3,000	0%
Total	**1,598,000**	**100.0%**				

ChatGPT accounts for 33.5% of all cited sources and Google AI Overview for 28.1%; together the two hold 61.6%. Perplexity accounts for 1.6%. ChatGPT and Google AI Overview also draw on the widest domain populations, at 64 and 65 distinct domains, while Perplexity and Microsoft Copilot draw on 19 and 23.

Each engine has a different most-cited domain, and no engine's most-cited domain is the most-cited domain of any other. Four of the six are vendor sites in the category; one is a statistics aggregator and one an analyst firm. The self-citation rate is 0% for all six engines. Appendix B lists the three most-cited domains for each engine.

4.5 Source quality and link decay

Each cited URL was tested for reachability once, during the collection window. Table 7 gives the dead-link rate and the resulting dead-link count for each engine.

Table 7. Link decay by AI engine — cited URLs that did not resolve when tested; n = 1,598,000 cited sources

AI engine	Cited sources	Dead-link rate	Dead links	Share of all dead links
ChatGPT	535,000	20.0%	107,000	48.2%
Microsoft Copilot	204,000	5.9%	12,036	5.4%
Gemini	225,000	12.4%	27,900	12.6%
Google AI Mode	160,000	18.1%	28,960	13.1%
Google AI Overview	449,000	10.0%	44,900	20.2%
Perplexity	25,000	4.0%	1,000	0.5%
Total	**1,598,000**	**13.9%**	**221,796**	**100.0%**

221,796 of the 1,598,000 cited URLs did not resolve, an overall dead-link rate of 13.9%. The rate ranges from 4.0% for Perplexity to 20.0% for ChatGPT, a factor of 5. Google AI Mode records 18.1%, Gemini 12.4%, Google AI Overview 10.0% and Microsoft Copilot 5.9%. Figure 4 shows the rates in rank order.

ChatGPT holds 48.2% of all dead links in the corpus, because it combines the highest dead-link rate with the largest citation volume. Google AI Overview holds 20.2%. No other engine holds more than 13.1%.

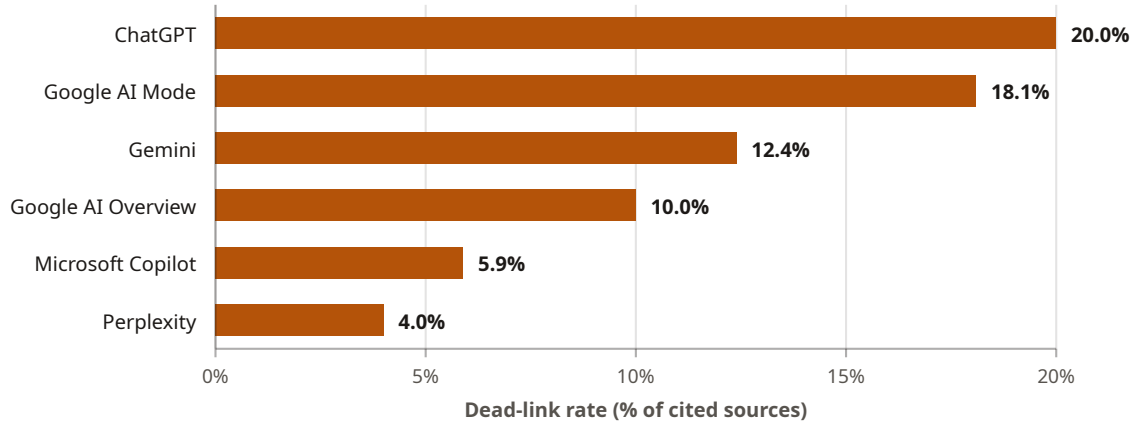

Figure 4. Dead-link rate by AI engine. Share of cited URLs that did not resolve when tested; n = 1,598,000 cited sources. Source: GrackerAI AI Search Visibility Benchmark Series, September 2026.

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 8 gives the full matrix.

Table 8. Cross-engine cited-domain overlap — Jaccard coefficient over cited-domain sets; 0 is disjoint, 1 is identical; n = 1,598,000 cited sources

	ChatGPT	Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
ChatGPT	**1.00**	0.09	0.22	0.21	0.21	0.15
Copilot	0.09	**1.00**	0.10	0.10	0.09	0.17
Gemini	0.22	0.10	**1.00**	0.28	0.30	0.22
Google AI Mode	0.21	0.10	0.28	**1.00**	0.40	0.15
Google AI Overview	0.21	0.09	0.30	0.40	**1.00**	0.14
Perplexity	0.15	0.17	0.22	0.15	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 0.19. The highest value is 0.40, between Google AI Mode and Google AI Overview. The lowest is 0.09, recorded twice, for Microsoft Copilot with ChatGPT and with Google AI Overview. The three engines operated by the same parent record coefficients of 0.28, 0.30 and 0.40 with each other, the three highest values in the matrix.

Microsoft Copilot records a coefficient between 0.09 and 0.17 with every other engine, the lowest row in the matrix; its highest value is with Perplexity. No pair of engines outside the same-parent group exceeds 0.22, recorded for ChatGPT with Gemini.

4.7 First-mention position

Table 9 gives, for each engine, the brand recorded as its first-mention vendor alongside the brand it mentioned most, with each brand's mean ordinal position in that engine.

Table 9. First-mention vendor and most-mentioned brand by AI engine — mean ordinal position counts named vendors from one; n = 50,000 responses per engine

AI engine	First-mention vendor	Its mean position	Most-mentioned brand	Its mentions	Its mean position
ChatGPT	Whistic	1.89	OneTrust	43,000	3.23
Microsoft Copilot	Bitsight	3.33	Bitsight	38,000	3.33
Gemini	ServiceNow	4.83	OneTrust	36,000	3.57
Google AI Mode	UpGuard	1.74	OneTrust	38,000	3.40
Google AI Overview	Bitsight	2.67	Bitsight	36,000	2.67
Perplexity	Bitsight	3.25	Bitsight	5,000	3.25

Bitsight is the first-mention vendor in three engines: Microsoft Copilot, Google AI Overview and Perplexity. Whistic, ServiceNow and UpGuard hold the position in ChatGPT, Gemini and Google AI Mode respectively. OneTrust, the most-mentioned brand in the corpus and in three engines, is the first-mention vendor in none. Its mean ordinal position ranges from 2.75 to 3.57 across the six engines; Bitsight's ranges from 2.47 to 3.33. Figure 3 plots the two side by side.

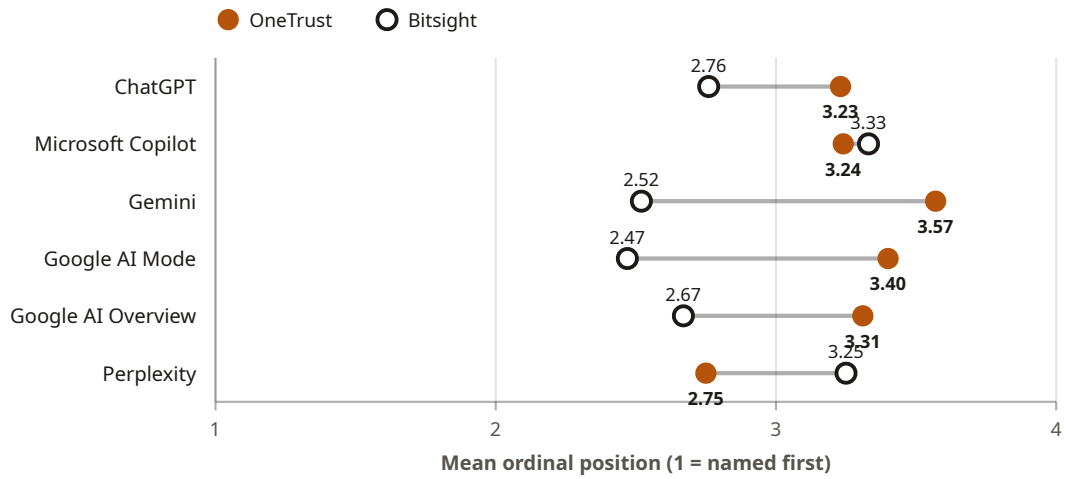

Figure 3. Mean ordinal position of OneTrust and Bitsight by AI engine. Position at which each brand is named, counting from one; lower is earlier. OneTrust is the most-mentioned brand; Bitsight is the first-mention vendor in three engines. Source: GrackerAI AI Search Visibility Benchmark Series, September 2026.

The first-mention vendor and the earliest mean ordinal position do not always coincide. In Gemini the first-mention vendor is ServiceNow, whose mean position there is 4.83, while UpGuard records 2.48. In ChatGPT the two measures agree: Whistic is the first-mention vendor and records the earliest mean position in that engine at 1.89, from 20,000 mentions. The earliest mean position recorded for any brand in any engine is UpGuard in Google AI Mode at 1.74, and UpGuard also records the widest range for a brand present in all six engines, from 1.74 to 4.75.

4.8 Geographic variance

Table 10 gives the product entries the instrumentation recorded as leading each market.

Table 10. Market top picks — leading product entries by market as recorded; n = 60,000 responses per market

Market	Leading product entry	Brand	Further picks recorded
United States	Bitsight Third-Party Risk Management	Bitsight	UpGuard Vendor Risk; OneTrust Third-Party Management
Canada	ProcessUnity Vendor Risk Management	ProcessUnity	UpGuard
Germany	ServiceNow Third-Party Risk Management	ServiceNow	Bitsight
India	Whistic TPRM	Whistic	SecurityScorecard
Australia	Panorays	Panorays	OneTrust

The leading product entry belongs to a different brand in each of the five markets: Bitsight in the United States, ProcessUnity in Canada, ServiceNow in Germany, Whistic in India and Panorays in Australia. Across the further picks, UpGuard is recorded in two markets and Bitsight and OneTrust in two each, counting the leading entry. At the brand level the instrumentation recorded the same three brands — Bitsight, UpGuard and OneTrust — as present in the responses of all five markets.

At the product level, 28 of the 164 product entries were recorded in all five markets, including every one of the ten most-mentioned. No entry in Table 5 is confined to a single market. The responses recorded no market-specific regulatory framing: the instrumentation found no prioritisation of jurisdiction-specific compliance tooling in the German or Australian responses.

4.9 Inter-engine disagreement

The instrumentation recorded one disagreement between engines on a factual claim, concerning which platform fits questionnaire automation. ChatGPT names Whistic and Vanta as the leading choices for reducing questionnaire workload with AI-assisted evidence handling. Microsoft Copilot and Google AI Overview name UpGuard and Bitsight, on the basis of combined security ratings and questionnaire workflows. The disagreement is consistent with the first-mention results in Table 9, where Whistic holds the position in ChatGPT and Bitsight in Microsoft Copilot and Google AI Overview.

5. Discussion

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

Yes, on the core. §4.2 shows that 9 of the 15 brands recorded were named by all six engines and that no brand was named by fewer than four. The five leading brands took 63.1% of all brand mentions, and the ten leading brands 90.1%. The engines agree on who is in the category.

The agreement weakens as the tail is approached. Whistic, which holds the first-mention position in ChatGPT, does not appear in Perplexity at all, and Black Kite appears in neither Google surface. A plausible explanation is that the engines share a core drawn from widely republished vendor lists and diverge where a brand's presence depends on a narrower set of sources. The data is consistent with that reading but does not establish it, because the study did not trace individual mentions to individual sources.

5.2 RQ2 — How concentrated is brand visibility, and does it carry through to product level?

Brand visibility is concentrated but not to a single vendor: OneTrust's 15.4% share is followed closely by Bitsight at 12.9% and UpGuard at 12.4%, and the five leading brands sit within 60,000 mentions of each other (§4.2). This is a category with a broad leading group rather than a single leader.

Brand-level visibility does not carry through cleanly to product level. §4.3 shows that OneTrust, the most-mentioned brand, holds only the fourth most-mentioned product entry, because the engines distributed its mentions over 12 product strings. Bitsight, whose bare name is the most-mentioned entry, is named more consistently. The interpretation offered here is that the engines name Bitsight and UpGuard as products and OneTrust as a portfolio.

The consequence for measurement is that a product-string count and a brand count rank the same vendors differently. A vendor with several product names will appear weaker in any product-level measurement than in a brand-level one.

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

Largely not. §4.6 shows a mean cited-domain overlap of 0.19 across the 15 engine pairs, with the three highest values all recorded between engines operated by the same parent and no cross-parent pair above 0.22. §4.4 adds that no engine's most-cited domain was the most-cited domain of any other. Six engines that agree on the vendor set reached it by reading substantially different parts of the web.

Microsoft Copilot is the outlier, with a coefficient of no more than 0.17 with any other engine and a most-cited domain that is a statistics aggregator rather than a vendor or analyst site. A plausible explanation is a retrieval index with a different population from the others; the data does not distinguish that from a different ranking of a shared index.

5.4 RQ4 — Does the vendor named first track the vendor mentioned most?

No. §4.7 shows OneTrust as the most-mentioned brand overall and in three engines, and as the first-mention vendor in none, with a mean ordinal position between 2.75 and 3.57 everywhere. Bitsight, second on mentions, holds the first-mention position in three engines. Whistic holds it in ChatGPT from 20,000 mentions, fewer than half OneTrust's 43,000 in the same engine.

The data is consistent with the engines treating OneTrust as a reliable member of the list and Bitsight, UpGuard or a specialist as the answer to the specific question asked. The Gemini result, where the first-mention vendor's mean position is later than several other brands', also shows that the two ordering measures capture different things: being named first most often is not the same as being named early on average. Mention volume, first-mention frequency and mean ordinal position are three separate quantities, and a vendor's standing on one does not predict its standing on the others.

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

The membership of the set does not; the leading entry does. §4.8 shows the same three brands present across all five markets and 28 product entries, including all of the ten most-mentioned, recorded in every market. It also shows a different leading product entry in every market, spread over five brands. The two results are compatible: the engines draw the same shortlist everywhere and vary which member of it they put forward.

The absence of jurisdiction-specific framing in the German and Australian responses is the more notable finding for a category defined partly by regulation. A plausible explanation is that the sources the engines read describe the category globally, so market-specific obligations do not reach the generated answer. The study did not examine the query phrasings by market and cannot rule out that different phrasing would have produced localisation.

5.6 RQ6 — On which buyer requirement do the engines disagree?

Questionnaire automation. §4.9 records ChatGPT positioning Whistic and Vanta, and Microsoft Copilot and Google AI Overview positioning UpGuard and Bitsight, for the same requirement. The disagreement aligns with each engine's first-mention vendor and with the two vendor groupings the responses describe: dedicated questionnaire and evidence platforms on one side, ratings-led platforms with questionnaire modules on the other. A vendor in either group is being recommended for the requirement by some engines and passed over by others.

5.7 What these results do not resolve

The study cannot say why an engine names a vendor first, because it did not trace mentions to the sources that produced them. It cannot say whether the leading product entry in a market reflects the market's sources or the engine's ordering of a shared set, because the market top pick is recorded per market and not per engine. It cannot say what truncation removed, only that it occurred. It does not distinguish an engine that reads a small index from one that reads a large index narrowly, which is the difference between the two explanations offered for Microsoft Copilot in §5.3. And it cannot say whether a brand named in every engine converts that presence into buyer action, because the study ends at the generated answer.

6. Threats to Validity

6.1 Construct validity

Ordinal position is used as a proxy for prominence. It captures order within a list and does not capture emphasis, length of treatment or the framing of the sentence in which a vendor appears. A vendor named fourth with a paragraph of justification may be more prominent to a reader than one named first in a comma-separated list. The first-mention vendor and the market top pick are recorded by rules the instrumentation does not document, so they are reported as recorded and read as coarse indicators. §4.7 shows that the first-mention vendor need not hold the earliest mean position.

Sentiment is scored on a single sentence by an undocumented model. The brand-level range in §4.2 spans 0.06, and the study makes no claim that differences within that range separate the brands. Truncation is inferred from terminal characters rather than observed, so a response that ends cleanly at a structural boundary after being cut off would not be counted, and one that ends with an unconventional character might be. The rates are indicative of a construction behaviour, not a precise measure of lost content.

6.2 Internal validity

All responses were collected in one window. Engine behaviour changes between windows as models, retrieval indexes and interface defaults are updated, none of which is announced or version-pinned. The engines are also non-deterministic: the same query can return different vendor sets on repeated issue. The study issued each query once per engine per market and did not measure within-engine variance. Personalisation and session state were controlled to the extent the public interfaces allow, which is incompletely.

6.3 External validity

The study covers one category, five markets and one window. The query set was constructed purposively to represent buyer-intent questions in the category, not sampled randomly from any population of real queries, and the paraphrase variants were written by the study rather than collected from users. No inferential statistic is therefore reported, because there is no sampling frame against which a confidence interval would be meaningful. The results describe the corpus; generalisation to other categories, markets or windows is a hypothesis, not a finding.

6.4 Reliability

A repeat run in a later window would be expected to reproduce the structural results: a shared core of vendors, a long single-engine tail, low cross-parent source overlap, a same-parent cluster at the top of the overlap matrix, and a separation between mention volume and ordinal position. It would not be expected to reproduce the exact mention counts, the exact coefficients, the leading product entry in a given market, or the dead-link rates, which describe the cited web at one moment. The point estimates in this report are the record of one window and are presented as such.

7. Limitations

- One category: the results describe enterprise third-party risk management and do not extend to adjacent categories such as governance suites or security-ratings services bought for other purposes.
- Five markets: the United States, Canada, Germany, India and Australia. No market in South America, Africa, East Asia or the Middle East was tested.
- One collection window in September 2026. The dead-link results in particular describe the cited web at that moment.
- Engine claims were not adjudicated. Where an engine attributed a capability, integration or price to a vendor, the study recorded the attribution and did not test it.
- Four processing steps are undocumented by the instrumentation and are named in §3.7; the metrics they produce are reported as recorded.

8. Practical Implications

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

Separate the inclusion problem from the ordering problem (§4.2, §4.7). 9 of 15 brands are named by every engine, so for a vendor outside that group the first task is inclusion, and nothing about ordering applies yet. For a vendor inside it, the task is ordering, and the OneTrust result shows that the most-mentioned brand can be named first nowhere. Measure mentions, first-mention frequency and mean position as three numbers, not one.

Publish one product name and repeat it (§4.3). OneTrust's mentions are spread across 12 product strings and ServiceNow's across 19, while Bitsight and UpGuard are named by their bare brand. A product-level measurement will understate any vendor whose products are named in several ways. Choose the canonical string, use it in documentation, pricing and analyst material, and expect the engines to reuse it.

Measure each engine separately (§4.4, §4.6). With a mean cited-domain overlap of 0.19 and six different most-cited domains, a citation earned on a domain one engine reads is not evidence of visibility in another. Treat Microsoft Copilot as a distinct source population.

Answer the questionnaire-automation question directly (§4.9). The one recorded disagreement between engines is about which platform automates questionnaires and evidence handling. A vendor with a position on that requirement can state it in the terms the engines reuse: automated questionnaire completion, AI-assisted reading of audit evidence, and shared vendor profiles. A vendor without one is being placed in the argument by the engines regardless.

Document integrations in the engines' vocabulary (§1.2, §3.4). The responses select on integration with service-management and source-to-pay procurement systems. Published, specific connector documentation gives an engine something to extract; a general integration claim does not.

Redirect what you retire (§4.5). 13.9% of cited URLs did not resolve. Some are vendors' own retired pages. A citation to a dead page carries the brand name and not the reader, and a redirect is the cheapest visibility fix available.

Do not localise the vendor story yet (§4.8). The same shortlist was present in all five markets and the responses carried no jurisdiction-specific framing. A vendor absent from the set in one of these markets is absent in all of them; a market-specific page will not be what changes that.

9. Conclusion

This study issued 10,000 buyer-intent queries about enterprise third-party risk management to six AI search engines in five markets, producing 300,000 responses, 1,598,000 cited sources and 164 product entries in a single window in September 2026. It measured what the engines named, in what order, on what evidence, and whether that evidence still resolved.

The engines agree on the category's membership and disagree on almost everything else. 9 of 15 brands were named by all six, and the five leading brands took 63.1% of all brand mentions. Behind that agreement, the engines shared a mean of 0.19 of their cited domains, no two engines had the same most-cited domain, and 13.9% of everything they cited was unreachable when tested. The most-mentioned brand, OneTrust, was the first-mention vendor in no engine, while Bitsight held that position in three; and the leading product entry differed in every market while the shortlist behind it did not.

For a vendor in the category, three things change. Presence in the shortlist is a prerequisite that 9 brands have met and six have not, and it has to be established once per engine because the engines do not read the same web. Standing within the shortlist is a separate quantity from presence, measured in first-mention frequency and mean position, and the vendor with the most mentions holds neither. And the product name the engines reuse is a variable the vendor controls, which the 12 OneTrust entries and 19 ServiceNow entries show is currently controlling the vendor.

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, the brand and product mention records with their ordinal positions, and the URL reachability results. The dataset for this report was generated in September 2026.

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 third-party risk management. 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. (September 2026). *AI Search Visibility in Enterprise Third-Party Risk Management: 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 by AI engine — mentions, with mean ordinal position in parentheses; n/r = position not recorded; n = 50,000 responses per engine

Brand	Total	ChatGPT	Microsoft Copilot	Gemini	Google AI Mode	Google AI Overview	Perplexity
OneTrust	191,000	43,000 (3.23)	36,000 (3.24)	36,000 (3.57)	38,000 (3.40)	34,000 (3.31)	4,000 (2.75)
Bitsight	159,000	18,000 (2.76)	38,000 (3.33)	28,000 (2.52)	34,000 (2.47)	36,000 (2.67)	5,000 (3.25)
UpGuard	154,000	21,000 (3.11)	35,000 (3.24)	31,000 (2.48)	28,000 (1.74)	34,000 (2.63)	5,000 (4.75)
SecurityScorecard	146,000	31,000 (4.52)	34,000 (5.04)	22,000 (2.67)	32,000 (3.13)	24,000 (3.52)	3,000 (6.50)
ProcessUnity	131,000	36,000 (2.59)	17,000 (5.31)	24,000 (5.46)	29,000 (4.79)	21,000 (3.90)	4,000 (6.25)
Panorays	104,000	10,000 (5.25)	24,000 (3.68)	29,000 (3.52)	14,000 (4.14)	23,000 (3.78)	4,000 (5.25)
ServiceNow	86,000	16,000 (3.73)	24,000 (3.07)	18,000 (4.83)	14,000 (3.25)	13,000 (2.40)	1,000 (n/r)
Vanta	59,000	15,000 (2.73)	2,000 (4.00)	13,000 (6.27)	9,000 (4.56)	19,000 (2.42)	1,000 (7.00)
Whistic	51,000	20,000 (1.89)	13,000 (4.50)	5,000 (4.00)	4,000 (3.50)	9,000 (3.22)	—
MetricStream	34,000	6,000 (4.00)	6,000 (3.00)	7,000 (3.17)	7,000 (6.43)	8,000 (4.13)	—
Drata	29,000	12,000 (3.33)	2,000 (6.00)	11,000 (7.50)	1,000 (4.00)	3,000 (3.50)	—
Riskonnect	27,000	4,000 (8.00)	12,000 (3.80)	1,000 (7.00)	5,000 (3.60)	4,000 (3.00)	1,000 (4.00)
Venminder	24,000	—	16,000 (6.42)	—	5,000 (7.60)	2,000 (6.50)	1,000 (3.00)
Aravo	22,000	8,000 (4.00)	10,000 (5.70)	1,000 (4.00)	2,000 (5.00)	1,000 (7.00)	—
Black Kite	20,000	3,000 (n/r)	7,000 (4.00)	9,000 (6.50)	—	—	1,000 (n/r)
Total	**1,237,000**	**243,000**	**276,000**	**235,000**	**222,000**	**231,000**	**30,000**

Appendix B. Most-cited domains

Table B1. Three most-cited domains by AI engine — citations to each domain; distinct domains per engine

AI engine	Distinct domains	Most cited	Citations	Second	Citations	Third	Citations
ChatGPT	64	whistic.com	65,000	upguard.com	51,000	servicenow.com	50,000
Microsoft Copilot	23	zipdo.co	37,000	servicenow.com	30,000	upguard.com	27,000
Gemini	31	spycloud.com	87,000	panorays.com	35,000	gartner.com	23,000
Google AI Mode	51	gartner.com	16,000	bitsight.com	12,000	servicenow.com	11,000
Google AI Overview	65	bitsight.com	49,000	panorays.com	36,000	spycloud.com	24,000
Perplexity	19	upguard.com	3,000	mitratech.com	2,000	optro.ai	2,000

Source-type classification was recorded by the instrumentation but placed almost every cited source in a single residual class, so it did not separate the corpus and is not reported.

Appendix C. Product entries

Table C1. All product entries — 164 entries in descending order of mentions; n = 300,000 responses

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Bitsight	Bitsight	136,000	2.8	0.82	6	5
UpGuard	UpGuard	129,000	2.8	0.82	6	5
SecurityScorecard	SecurityScorecard	120,000	3.8	0.79	6	5
OneTrust	OneTrust	115,000	3.5	0.79	6	5
ProcessUnity	ProcessUnity	101,000	4.4	0.83	6	5
Panorays	Panorays	89,000	3.9	0.81	6	5
Prevalent	Prevalent	49,000	4.6	0.77	6	5
Whistic	Whistic	42,000	3.2	0.80	5	5
Vanta	Vanta	38,000	4.3	0.82	6	5
UpGuard Vendor Risk	UpGuard	22,000	1.8	0.84	5	5
OneTrust Third-Party Risk Management	OneTrust	22,000	2.2	0.82	4	5
ServiceNow Third-Party Risk Management	ServiceNow	21,000	2.7	0.85	5	5
Riskonnect	Riskonnect	21,000	4.0	0.77	6	5
Drata	Drata	21,000	5.7	0.78	5	5
OneTrust Vendor Risk Management	OneTrust	20,000	3.4	0.81	4	5
MetricStream	MetricStream	20,000	4.7	0.81	5	4
Venminder	Venminder	20,000	6.5	0.73	4	5

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Aravo	Aravo	19,000	5.1	0.75	4	5
ServiceNow TPRM	ServiceNow	18,000	3.3	0.80	4	5
OneTrust Third-Party Management	OneTrust	18,000	3.6	0.81	4	5
ServiceNow Vendor Risk Management	ServiceNow	18,000	4.6	0.84	3	5
Archer	Archer	17,000	5.2	0.72	5	5
ServiceNow	ServiceNow	16,000	3.4	0.81	3	5
ProcessUnity Vendor Risk Management	ProcessUnity	13,000	3.1	0.80	4	5
Vanta TPRM	Vanta	12,000	2.7	0.82	4	4
OneTrust TPRM	OneTrust	11,000	3.3	0.77	5	5
Black Kite	Black Kite	11,000	5.8	0.78	2	5
Hyperproof	Hyperproof	11,000	5.9	0.80	5	3
Prevalent TPRM	Prevalent	10,000	4.7	0.76	4	4
Bitsight TPRM	Bitsight	9,000	3.3	0.81	4	4
VISO TRUST	VISO TRUST	8,000	3.5	0.85	5	5
Diligent	Diligent	8,000	6.6	0.81	2	5
RiskRecon	Mastercard	7,000	6.3	0.78	3	4
Bitsight Vendor Risk Management	Bitsight	5,000	2.2	0.86	3	3
UpGuard TPRM	UpGuard	5,000	2.4	0.80	3	4
LogicGate Risk Cloud	LogicGate	5,000	2.6	0.82	2	5
Panorays TPRM	Panorays	5,000	2.8	0.82	4	4
ComplyScore	ComplyScore	5,000	3.0	0.78	1	3
Enlighta	Enlighta	5,000	3.2	0.86	2	3
OneTrust Third-Party Governance	OneTrust	4,000	2.0	0.85	2	3
Exalate	Exalate	4,000	2.0	0.75	1	3
Vanta Third-Party Risk Management	Vanta	4,000	2.5	0.82	2	3
ProcessUnity Third-Party Risk Management	ProcessUnity	4,000	3.5	0.85	3	3
Prevalent (Mitratech)	Mitratech	4,000	4.0	0.78	3	3
Scrut Automation	Scrut Automation	4,000	4.8	0.85	1	4
SecurityScorecard TPRM	SecurityScorecard	4,000	5.0	0.82	3	3
ProcessUnity TPRM	ProcessUnity	4,000	5.0	0.78	3	3
Nudge Security	Nudge Security	4,000	8.0	0.65	1	4
ServiceNow GRC: Third-Party Risk Management (TPRM)	ServiceNow	3,000	1.0	0.90	1	3

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
ZigiOps	ZigiOps	3,000	3.0	0.67	1	2
Drata TPRM	Drata	3,000	3.3	0.80	2	2
OneTrust Vendorpedia	OneTrust	3,000	4.0	0.83	2	3
MetricStream Third-Party Risk Management	MetricStream	3,000	4.3	0.80	1	3
Prewave	Prewave	3,000	5.3	0.83	2	2
Scrut Automation Vendor Risk Management	Scrut Automation	3,000	5.7	0.83	1	3
LogicGate	LogicGate	3,000	7.7	0.73	2	2
SpyCloud Supply Chain Threat Protection	SpyCloud	3,000	8.7	0.87	1	3
SAP Ariba Supplier Risk	SAP	3,000	8.7	0.80	2	3
Vantavanta.com	Vanta	2,000	1.0	0.85	1	2
VISOTRUST	VISOTRUST	2,000	1.0	0.85	1	2
Whistic TPRM	Whistic	2,000	1.5	0.90	1	2
ServiceNow Third-Party Risk Management (TPRM)	ServiceNow	2,000	1.5	0.80	2	1
Dratadrata.com	Drata	2,000	2.0	0.80	1	2
Archer (RSA Archer)	RSA	2,000	2.0	0.60	1	2
OneTrust Third-Party Risk Exchange	OneTrust	2,000	2.5	0.90	2	2
Bitsight Third-Party Risk Management	Bitsight	2,000	2.5	0.80	1	2
MetricStream CyberGRC	MetricStream	2,000	3.0	0.90	2	2
ServiceNow Vendor Risk Management / TPRM	ServiceNow	2,000	3.0	0.90	2	2
Whisticwhistic.com	Whistic	2,000	3.0	0.85	1	2
MetricStream TPRM	MetricStream	2,000	3.0	0.80	2	2
Prevalent Third-Party Risk Management Platform	Prevalent	2,000	3.5	0.90	1	2
3rdRisk	3rdRisk	2,000	3.5	0.70	2	2
OneTrustonetrust.com	OneTrust	2,000	4.0	0.65	1	2
ServiceNow IRM	ServiceNow	2,000	4.5	0.85	2	1
Inference Systems Custom AI Workflow	Inference Systems	2,000	4.5	0.75	1	2
Fairmarkit	Fairmarkit	2,000	4.5	0.70	1	2
Prevalent / Mitratech	Prevalent	2,000	4.5	0.60	1	1
OneTrust Vendor Risk	OneTrust	2,000	5.0	0.75	2	2
ServiceNow VRM	ServiceNow	2,000	5.5	0.80	2	2
SecurityPal	SecurityPal	2,000	5.5	0.75	2	1
Conveyor	Conveyor	2,000	7.5	0.85	2	2

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Sphera	Sphera	2,000	7.5	0.75	1	2
Secureframe	Secureframe	2,000	8.0	0.70	2	2
Cyber Sierra	Cyber Sierra	2,000	9.0	0.80	1	1
VISO TRUST Third-Party Risk Management	VISO TRUST	1,000	1.0	0.90	1	1
Bitsight (with Bitsight AI)	Bitsight	1,000	1.0	0.80	1	1
Panorays TPRM Platform	Panorays	1,000	1.0	0.80	1	1
ServiceNow Third-Party Risk Management (TPRM)	ServiceNow	1,000	1.0	0.80	1	1
ServiceNow Vendor Risk Management (VRM)	ServiceNow	1,000	1.0	0.80	1	1
UpGuard Vendor Risk Management	UpGuard	1,000	1.0	0.80	1	1
MetricStream Third-Party Cyber Risk	MetricStream	1,000	2.0	0.85	1	1
TemplarShield Third-Party Risk	TemplarShield	1,000	2.0	0.80	1	1
Whistic AI-First TPRM	Whistic	1,000	2.0	0.80	1	1
BitSight Cyber Risk Intelligence	BitSight	1,000	2.0	0.80	1	1
Bitsight Integrated Third-Party Risk Management	Bitsight	1,000	2.0	0.80	1	1
Prevalent TPRM	Mitratech	1,000	2.0	0.80	1	1
Riskonnect Third-Party Risk Management	Riskonnect	1,000	2.0	0.80	1	1
ServiceNow TPRM / Vendor Risk Management	ServiceNow	1,000	3.0	1.00	1	1
AuditBoard	AuditBoard	1,000	3.0	1.00	1	1
ServiceNow GRC	ServiceNow	1,000	3.0	0.90	1	1
UpGuard Questionnaire Module	UpGuard	1,000	3.0	0.90	1	1
UpGuard Risk Automations	UpGuard	1,000	3.0	0.80	1	1
Bitsight Framework Intelligence	Bitsight	1,000	3.0	0.80	1	1
Archer TPRM	Archer	1,000	3.0	0.80	1	1
Ciphrix	Ciphrix	1,000	3.0	0.80	1	1
UpGuard Risk Automation	UpGuard	1,000	3.0	0.80	1	1
OneTrust GRC	OneTrust	1,000	3.0	0.70	1	1
ProcessUnity Vendor Risk Assessments	ProcessUnity	1,000	3.0	0.70	1	1
SecurityScorecard Vendor Risk	SecurityScorecard	1,000	3.0	0.70	1	1
Diligent HighBond	Diligent	1,000	4.0	0.90	1	1
SecurityScorecard Atlas	SecurityScorecard	1,000	4.0	0.90	1	1
ServiceNow Supplier Lifecycle & Procurement Operations	ServiceNow	1,000	4.0	0.80	1	1
Archer Third-Party Risk Management	Archer	1,000	4.0	0.80	1	1

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
Scrut Automation	Scrut	1,000	4.0	0.80	1	1
VISO TRUST TPRM	VISO TRUST	1,000	4.0	0.80	1	1
Viso Trust TPRM Automation	Viso Trust	1,000	4.0	0.80	1	1
ServiceNow GRC / IRM	ServiceNow	1,000	4.0	0.80	1	1
VISO TRUST Third-Party Risk Management Automation	VISO TRUST	1,000	4.0	0.80	1	1
Whistic Vendor Risk Management	Whistic	1,000	4.0	0.75	1	1
Aravo Vendor Risk Management	Aravo	1,000	4.0	0.75	1	1
ComplyScore TPRM	ComplyScore	1,000	4.0	0.70	1	1
Coupa Risk Assess	Coupa	1,000	4.0	0.70	1	1
OpsHub	OpsHub	1,000	4.0	0.70	1	1
ServiceNow Risk Assessments Integration for Procurement	ServiceNow	1,000	4.0	0.70	1	1
ConnectALL	ConnectALL	1,000	4.0	0.60	1	1
Coverbase	Coverbase	1,000	4.0	0.60	1	1
Exiger	Exiger	1,000	5.0	0.90	1	1
Drata Third-Party Risk Management	Drata	1,000	5.0	0.90	1	1
Inference Systems	Inference Systems	1,000	5.0	0.80	1	1
ServiceNow IRM / Vendor Risk Management	ServiceNow	1,000	5.0	0.80	1	1
Scrut	Scrut	1,000	5.0	0.80	1	1
VRM GRC	VRM GRC	1,000	5.0	0.80	1	1
Third-Party Risk Management for Jira	Atlassian	1,000	5.0	0.70	1	1
Whistic Questionnaire Exchange	Whistic	1,000	5.0	0.70	1	1
SecurityScorecardsecurityscorecard.com	SecurityScorecard	1,000	5.0	0.60	1	1
Vendict	Vendict	1,000	5.0	0.60	1	1
ServiceNow IRM / VRM	ServiceNow	1,000	6.0	0.80	1	1
Coupa Supplier Risk	Coupa	1,000	6.0	0.80	1	1
Resilinc	Resilinc	1,000	6.0	0.80	1	1
Third Party Risk Management for Jira	Atlassian	1,000	6.0	0.80	1	1
Hyperproof Risk Management	Hyperproof	1,000	6.0	0.80	1	1
RSA Archer	RSA Archer	1,000	6.0	0.70	1	1
Panorays Third-Party Security	Panorays	1,000	6.0	0.70	1	1
Prevalentprevalent.ai	Prevalent	1,000	6.0	0.60	1	1

Product entry	Brand	Mentions	Mean rank	Mean sentiment	Engines	Markets
CoreStream GRC	CoreStream	1,000	6.0	0.60	1	1
Prevalent Third-Party Risk Management	Prevalent	1,000	7.0	0.90	1	1
Diligent Third-Party Risk Management	Diligent	1,000	7.0	0.80	1	1
ServiceNow Vendor Risk	ServiceNow	1,000	7.0	0.70	1	1
SAP Business Network	SAP	1,000	7.0	0.70	1	1
RiskImmune	RiskImmune	1,000	7.0	0.70	1	1
OneTrust Third-Party Due Diligence	OneTrust	1,000	7.0	0.70	1	1
Coupa	Coupa	1,000	8.0	0.90	1	1
SpyCloud	SpyCloud	1,000	8.0	0.90	1	1
CheckFirst	CheckFirst	1,000	8.0	0.80	1	1
Workiva Supplier Risk Management	Workiva	1,000	8.0	0.80	1	1
Interos	Interos	1,000	8.0	0.80	1	1
VRM GRC (TrustArc)	TrustArc	1,000	8.0	0.70	1	1
Hyperproof TPRM	Hyperproof	1,000	8.0	0.60	1	1
Craft	Craft	1,000	9.0	0.70	1	1
IBM OpenPages	IBM	1,000	9.0	0.60	1	1
RiskProfiler (Knyx Vendor AI)	Knyx	1,000	10.0	0.80	1	1
Archer GRC	Archer	1,000	10.0	0.70	1	1
LSEG/Avetta	LSEG/Avetta	1,000	10.0	0.60	1	1
Coupa Risk	Coupa	1,000	11.0	0.70	1	1

Appendix D. Study parameters

Table D1. Study parameters — as configured for this study

Parameter	Value
Category	Enterprise third-party risk management, vendor risk monitoring and questionnaire automation
AI engines	ChatGPT, Microsoft Copilot, Gemini, Google AI Mode, Google AI Overview, Perplexity
Markets	United States, Canada, Germany, India, Australia
Query set	10,000 distinct buyer-intent queries, core questions plus paraphrase variants
Query themes	Continuous third-party monitoring; questionnaire automation; integration with service-management and procurement systems; vendor onboarding at volume
Responses per engine	50,000

Parameter	Value
Responses per market	60,000
Total responses	300,000
Cited sources	1,598,000
Brands recorded	15
Product entries	164
Collection window	September 2026
Unit of analysis	One engine response to one query in one market
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 cited sources	An engine's cited sources divided by total 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.
Share of all dead links	An engine's dead links divided by total dead links.
Total brand mentions	Sum of a brand's per-engine mention counts, summed over brands for the table total.
Share of mentions	A brand's mentions divided by total brand mentions.
Engines recording	Count of engines whose mention count for that brand is greater than zero.

Metric	Computation
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 sentence-level polarity scores for a brand's mentions, 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.