Gracker.ai

RESEARCH REPORT

AI Search Visibility in Enterprise Attack Surface Management

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 enterprise attack surface management platforms across the United States, Canada, India and Germany.

Table 1. Study scale by AI engine — responses recorded and mean response length

AI engine	Responses	Mean words	Markets
ChatGPT	40,000	647.4	4
Microsoft Copilot	40,000	507.4	4
Gemini	40,000	376.2	4
Google AI Mode	40,000	284.2	4
Google AI Overview	40,000	232.3	4
Perplexity	40,000	524.0	4
Total	**240,000**	**—**	**4**

Published: August 2026

Research period: August 2026

Category: Enterprise attack surface management (ASM, EASM, CAASM)

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

Evidence base: 1,254,000 cited sources · 185 distinct product entries

Prepared by: GrackerAI · AI Search Visibility Benchmark Series

Abstract

Enterprise security buyers increasingly begin vendor evaluation inside an AI-generated answer rather than on a search results page, which means the shortlist an AI engine returns now defines the consideration set before a vendor is contacted. This study measures how six commercial AI search engines answer buyer-intent questions about enterprise attack surface management. Between them the study issued 10,000 enterprise buyer-intent queries to each engine across four markets, producing 240,000 AI engine responses containing 1,254,000 cited sources and covering 185 distinct product entries. For every response the study recorded length, citation count, cited domains, source reachability, hedging, truncation, refusal, and each brand mention together with its ordinal position. The engines converge on a narrow vendor shortlist: CyCognito and Palo Alto Networks lead all six engines, with CyCognito recording 160,000 mentions against 115,000 for Palo Alto Networks. They diverge sharply on evidence: the highest cross-engine source overlap observed is a Jaccard coefficient of 0.37, between Google AI Mode and Google AI Overview, while Perplexity shares no cited domains with any other engine. Source reliability varies by a factor of three, from a 9.4% unreachable-URL rate for Google AI Overview to 27.6% for Microsoft Copilot. No market-level variation in vendor recommendation was observed. The results indicate that vendor visibility and the evidence base producing it are governed by separate mechanisms, and that single-engine measurement does not generalise across engines.

Contents

1. Introduction

1.1 Background

An AI search engine does not return a list of documents. It returns a composed answer that names a small number of vendors, places them in an order, and attaches reasons to each. For a high-consideration enterprise security purchase, that single response performs work a buyer previously distributed across many sources: it defines the category boundary, establishes the evaluation criteria, and nominates the candidates.

The consequence for a vendor is structural rather than competitive. A vendor absent from the returned shortlist is not losing on price, capability or references, because it is not present at the moment the consideration set is formed. Conventional search visibility metrics do not detect this condition. A vendor can hold strong organic rankings for its category terms and still be omitted from the answers that increasingly precede a click.

Measuring the condition requires observing the engines directly: issuing the questions a buyer would ask, recording the responses as returned, and analysing which vendors appear, in what order, and on the strength of which cited evidence. This study applies that method to one category.

1.2 Why this category

Enterprise attack surface management is a suitable instrument for three reasons. The vendor set is mature and well documented, so engines have abundant published material to draw on. The category has a stable and publicly articulated capability vocabulary, organised around external attack surface management, cyber asset attack surface management, and continuous threat exposure management, which makes engine responses comparable across markets and phrasings. Buying behaviour is research-intensive and procurement cycles are long, so the cost of omission from an early shortlist is high and difficult to reverse later in the cycle.

1.3 Contribution

This study contributes a directly measured, cross-engine baseline for a single enterprise security category. It reports vendor visibility and ordinal position separately rather than collapsing them into a single share-of-voice figure. It quantifies the degree to which engines draw on shared or disjoint evidence bases, which determines whether a measurement taken on one engine generalises to another. It records the characteristics of engine output itself — response length, citation density, hedging and cited-source reachability — as properties of the category's AI search layer rather than as incidental observations.

2. Research Questions

1. **RQ1.** To what extent do the six AI engines converge on the same vendor shortlist for enterprise attack surface management?
2. **RQ2.** How is brand visibility distributed across the vendor population, and how concentrated is it?
3. **RQ3.** To what extent do the engines draw on a shared evidence base when producing those recommendations?
4. **RQ4.** Does ordinal position within a response track mention volume, or are the two governed independently?
5. **RQ5.** Does vendor recommendation vary across the four markets tested?

Each question is answered explicitly in Section 5, with reference to the results section supporting the answer.

3. Methodology

3.1 Study design

The study is observational and cross-sectional. Six commercial AI search engines were issued an identical set of buyer-intent questions across four markets within a single collection window in August 2026. No engine was manipulated, primed, or given conversational context beyond the question itself. Each response was captured as returned and analysed without modification.

3.2 Engines under test

Six engines were tested: ChatGPT, Perplexity, Gemini, Microsoft Copilot, Google AI Overview and Google AI Mode. All six are commercial services. None exposes a pinned model version to callers, and none guarantees deterministic output. Findings therefore describe engine behaviour during the collection window and are not attributable to a specific model release.

3.3 Markets

Four markets were tested: the United States, Canada, India and Germany. Market was varied to test whether engines localise vendor recommendations in response to differing regulatory environments, data-residency expectations or regional vendor presence. Query text was held constant across markets; only the market context of the request varied.

3.4 Query construction

The study used a set of 10,000 enterprise buyer-intent queries about attack surface management platforms. Every query was issued to all six engines in all four markets, producing 40,000 responses per engine and 240,000 responses in total. The query set comprised core category questions together with paraphrase variants, included to test whether vendor recommendations are stable under rewording or sensitive to surface phrasing.

3.5 Corpus and unit of analysis

The unit of analysis is a single engine response to a single query in a single market. Across the six engines the study recorded 240,000 responses, distributed evenly at 40,000 per engine and 10,000 per engine per market.

3.6 Metric definitions

Every metric used in this report is defined below. Definitions precede their first use in the results.

Table 2. Metric definitions

Metric	Definition	Unit
Mean response length	Mean word count of an engine's responses	words
Citation density	Citations per 100 words of response text	per 100 words
Total cited sources	Count of cited references recorded for an engine across the corpus	count

Metric	Definition	Unit
Distinct domains cited	Number of different domains appearing in an engine's citations. Not a corpus-wide total and not additive across engines	count
Dead-link rate	Share of an engine's cited URLs that did not resolve when tested	%
Median source age	Median age, in days, of the material an engine cited at time of collection	days
Self-citation rate	Share of an engine's citations pointing at domains owned by its own parent ecosystem	%
Source variety	Mean number of distinct cited domains appearing within a single response	mean
Hedging rate	Share of responses containing explicit uncertainty or qualifying language	%
Recency anchoring	Share of responses explicitly referencing a date or recency signal	%
Truncation rate	Share of responses ending mid-sentence or omitting an expected list item. Inferred from response structure, not observed directly	%
Refusal rate	Share of responses declining to answer the question	%
First-mention vendor	The vendor named first in the body of an engine's response	label
Brand mention	One occurrence of a vendor's brand name in a response. Counted separately from product mentions and not additive with them	count
Mean ordinal position	Mean index of a brand's first appearance within a response, where 1.0 is first named	mean
Product mention	One occurrence of a named product in a response. A single vendor may hold many product entries, and a product name may be recorded without its brand name	count
Mean product rank	Mean ordinal position of a product entry within a response	mean
Mean sentiment	Mean polarity of the language surrounding a brand or product mention, on a scale from −1.0 to 1.0	−1.0–1.0
Jaccard source overlap	Size of the intersection of two engines' cited-domain sets divided by the size of their union. 1.00 denotes identical domain sets, 0.00 no shared domains	0.00–1.00

3.7 Data collection and processing

Queries were issued to each engine's public surface for the specified market and the returned response captured in full. Captured responses were normalised to plain text before analysis. Citations were extracted from the response body, resolved to their target domain, and classified by source type. Each cited URL was requested to determine whether it still resolved; URLs that did not resolve were recorded as dead links and counted toward the engine's

dead-link rate. Brand and product mentions were detected against a category vocabulary and recorded together with their ordinal position within the response, allowing mention volume and position to be analysed independently. Sentiment was scored on the language surrounding each mention.

Two aspects of processing are not documented in the source dataset and are therefore not described here: the precise detection rule used to classify a response as truncated, and the specific model and scoring scale used to assign sentiment polarity. Both are treated as construct-validity concerns in Section 6.1 rather than described speculatively.

3.8 Scope exclusions

The study did not paraphrase or edit engine output. It did not fact-check engine claims against external sources, and it did not score responses for technical accuracy. It is not a product evaluation: vendor and product names appear only as observed in engine output, their ordering reflects engine behaviour rather than any assessment by the authors, and no placement was paid for or solicited. Where engines disagree on a factual claim, this report records the disagreement without adjudicating it.

4. Results

This section reports what was observed. Interpretation of these observations is deferred to Section 5, and implications for practice to Section 8.

4.1 Response characteristics by engine

Response length, citation behaviour and answer construction differ substantially between engines.

Table 3. Response characteristics by AI engine — n = 240,000 AI engine responses

AI engine	Responses	Mean words	Citation density	Source variety	Hedging	Recency anchoring	Truncation	Refusals	Self-citation
ChatGPT	40,000	647.4	1.6	1.4	10%	15%	97%	0%	0%
Microsoft Copilot	40,000	507.4	0.8	0.8	17%	60%	0%	0%	0%
Gemini	40,000	376.2	1.4	0.9	0%	3%	23%	0%	0%
Google AI Mode	40,000	284.2	3.6	1.2	0%	13%	0%	0%	0%
Google AI Overview	40,000	232.3	6.5	1.8	0%	7%	0%	0%	0%
Perplexity	40,000	524.0	1.6	0.0	33%	0%	33%	0%	0%

Mean response length spans a factor of 2.8, from 647.4 words for ChatGPT to 232.3 words for Google AI Overview. Citation density runs in the opposite direction: Google AI Overview records 6.5 citations per 100 words against 1.6 for ChatGPT and 0.8 for Microsoft Copilot. Google AI Overview therefore produces the shortest responses and the most citation-dense ones.

Microsoft Copilot anchors 60% of its responses to an explicit date or recency signal, against 15% for ChatGPT and 3% for Gemini. Hedging is concentrated in Perplexity at 33% and Microsoft Copilot at 17%; Gemini, Google AI Mode and Google AI Overview record no hedging language. No engine refused any query, and no engine cited its

own parent ecosystem.

Response completeness also varies. ChatGPT and Perplexity, the two engines with the longest mean responses at 647.4 and 524.0 words, are the two most likely to end an ordered list before it concludes. Microsoft Copilot, Google AI Mode and Google AI Overview complete their lists in every recorded response.

4.2 Brand visibility

Brand mentions were recorded per engine together with the mean ordinal position at which each brand first appeared.

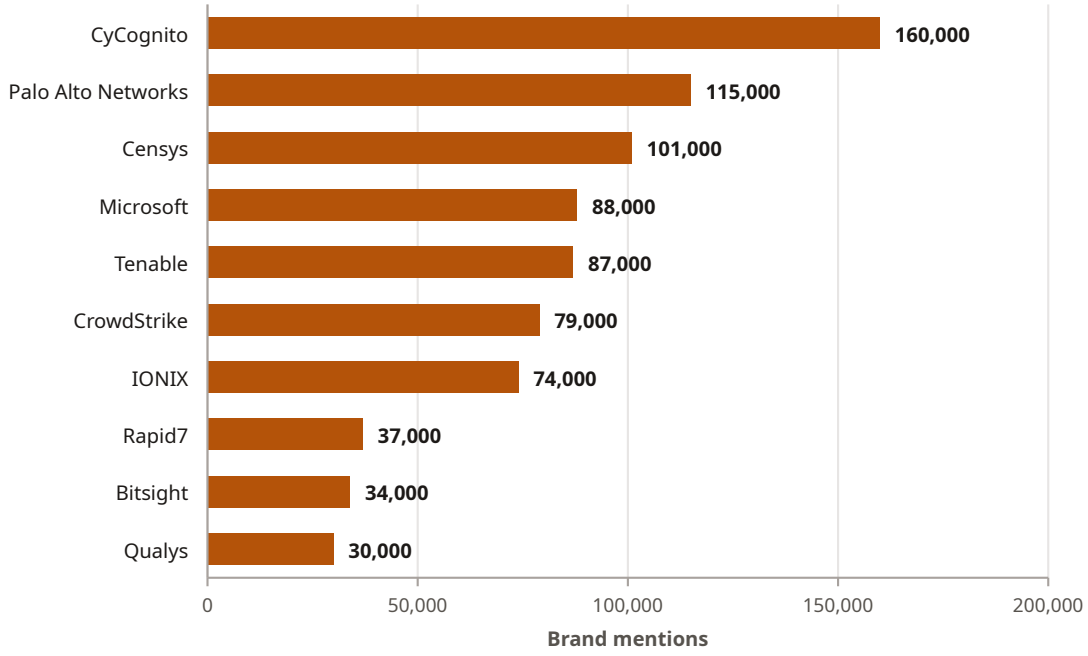

Figure 1. Brand mentions by vendor. Ten most-mentioned vendors; n = 240,000 AI engine responses. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

Table 4. Brand mentions by AI engine — mentions, with mean ordinal position in parentheses. Ten most-mentioned vendors

Brand	Total	ChatGPT	Copilot	Gemini	AI Mode	AI Overview	Perplexity
CyCognito	160,000	26,000 (4.5)	32,000 (2.9)	31,000 (2.2)	32,000 (2.2)	36,000 (1.8)	3,000 (3.3)
Palo Alto Networks	115,000	28,000 (1.4)	13,000 (1.5)	29,000 (3.2)	25,000 (1.7)	19,000 (2.0)	1,000 (2.0)
Censys	101,000	24,000 (3.8)	15,000 (4.2)	27,000 (3.0)	21,000 (3.0)	14,000 (3.7)	—
Microsoft	88,000	23,000 (3.3)	19,000 (3.3)	20,000 (3.9)	13,000 (3.0)	11,000 (3.8)	2,000 (2.0)
Tenable	87,000	22,000 (2.8)	10,000 (3.5)	21,000 (5.0)	18,000 (3.9)	15,000 (4.2)	1,000 (3.0)
CrowdStrike	79,000	13,000 (4.2)	21,000 (2.8)	10,000 (4.3)	20,000 (4.0)	14,000 (4.5)	1,000 (1.0)
IONIX	74,000	12,000 (6.1)	11,000 (3.9)	11,000 (3.1)	14,000 (3.6)	25,000 (3.2)	1,000 (4.0)
Rapid7	37,000	12,000 (5.0)	7,000 (4.6)	3,000 (7.0)	8,000 (4.4)	7,000 (4.5)	—
Bitsight	34,000	6,000 (4.3)	—	18,000 (3.2)	8,000 (5.5)	1,000 (5.0)	1,000 (8.0)
Qualys	30,000	7,000 (7.1)	5,000 (3.4)	4,000 (4.3)	6,000 (5.2)	8,000 (4.4)	—

The ten most-mentioned vendors account for 805,000 brand mentions. CyCognito records 160,000, ahead of Palo Alto Networks at 115,000 and Censys at 101,000. Six vendors are mentioned by all six engines: CyCognito, Palo Alto Networks, Microsoft, Tenable, CrowdStrike and IONIX. Censys, Rapid7 and Qualys are recorded in five engines, as is Bitsight.

Coverage is uneven within the leading group. Bitsight records 18,000 mentions in Gemini and 1,000 in Google AI Overview, an eighteenfold difference across two engines for the same vendor. IONIX records 25,000 mentions in Google AI Overview against 11,000 to 14,000 in each of the other four engines that name it.

Mean ordinal position does not track mention volume. CyCognito holds the highest total at 160,000 mentions but a mean position of 4.5 in ChatGPT, where Palo Alto Networks records fewer mentions at 28,000 and a mean position of 1.4. Within Google AI Overview the ordering reverses: CyCognito records both the highest mention count at 36,000 and the earliest mean position at 1.8.

Table 5. Mean sentiment by vendor — scale −1.0 to 1.0

Brand	Mean sentiment	Engines naming the vendor
Palo Alto Networks	0.84	6 of 6
CyCognito	0.81	6 of 6
CrowdStrike	0.81	6 of 6
IONIX	0.81	6 of 6
Bitsight	0.80	5 of 6
Wiz	0.80	4 of 6
Tenable	0.79	6 of 6
Censys	0.77	5 of 6
Microsoft	0.77	6 of 6
Qualys	0.71	5 of 6
UpGuard	0.71	4 of 6
Rapid7	0.70	5 of 6

Mean sentiment occupies a narrow band from 0.70 to 0.84 across all twelve vendors recorded. No vendor in the category carries negative mean sentiment. The spread between the highest and lowest values is 0.14, against a possible range of 2.0.

4.3 Product-level results

The corpus records 185 distinct product entries. Product mentions are counted separately from brand mentions and are not additive with them: a single vendor holds multiple product entries, and a product name may be recorded in a response that does not name its parent brand.

Table 6. Ten most-mentioned product entries

#	Product	Brand	Mentions	Mean product rank	Mean sentiment	Markets
1	Cortex Xpanse	Palo Alto Networks	149,000	1.8	0.86	4
2	CyCognito	CyCognito	132,000	2.7	0.81	4
3	Microsoft Defender EASM	Microsoft	73,000	3.6	0.79	4
4	IONIX	IONIX	57,000	4.1	0.81	4
5	CrowdStrike Falcon Exposure Management	CrowdStrike	51,000	3.1	0.81	4
6	Censys	Censys	46,000	4.0	0.77	4
7	Tenable One	Tenable	40,000	3.1	0.80	4
8	Censys ASM	Censys	38,000	2.7	0.83	4
9	CrowdStrike Falcon Surface	CrowdStrike	36,000	4.7	0.80	4
10	Microsoft Defender External Attack Surface Management	Microsoft	27,000	3.0	0.79	4

The product distribution is more concentrated than the brand distribution. Cortex Xpanse records 149,000 mentions at a mean rank of 1.8, the earliest of the ten most-mentioned entries. Six low-volume entries elsewhere in the table record a mean rank of 1.0. The 185 product entries include substantial naming fragmentation: Tenable appears as fifteen separate entries, Qualys as eleven, Bitsight and Rapid7 as ten each, Palo Alto Networks as nine, Censys as seven and CyCognito as six. Of the 185 entries, 104 record 1,000 mentions or fewer.

> Product mentions for a vendor may exceed that vendor's brand mentions. Cortex Xpanse records 149,000 product mentions while Palo Alto Networks records 115,000 brand mentions. The two metrics count different objects, as defined in Table 2, and are not additive.

4.4 Citation volume

Across the five engines for which citation totals were recorded, 1,254,000 cited sources were captured. Citation totals were not recorded for Perplexity.

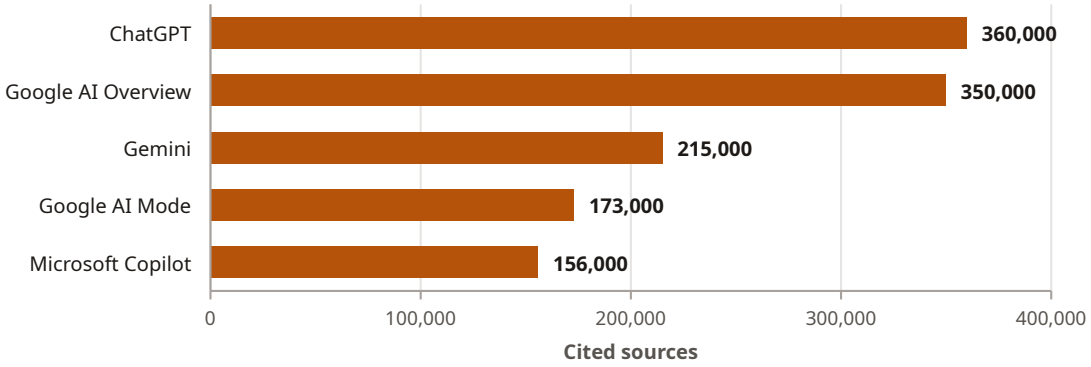

Figure 2. Citation volume by AI engine. Total cited sources across five engines; n = 1,254,000 citations. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

ChatGPT supplies 360,000 citations, 28.7% of the recorded evidence base, and Google AI Overview 350,000, or 27.9%. Microsoft Copilot supplies 156,000, or 12.4%, from a comparable number of responses. The ratio between the highest and lowest recorded citation volume is 2.3.

4.5 Source quality and link decay

Table 7. Citation behaviour and source quality by AI engine

AI engine	Total cited sources	Distinct domains cited	Dead-link rate	Unreachable citations	Median source age
ChatGPT	360,000	56	17.8%	64,100	0 days
Microsoft Copilot	156,000	33	27.6%	43,100	0 days
Gemini	215,000	36	9.8%	21,100	0 days
Google AI Mode	173,000	46	20.2%	34,900	0 days
Google AI Overview	350,000	72	9.4%	32,900	0 days
Total	**1,254,000**	**—**	**15.6%**	**196,100**	**0 days**

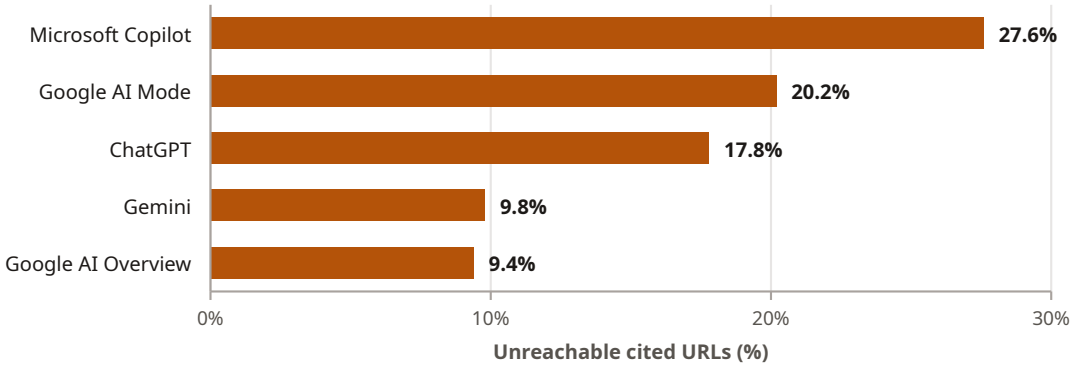

Figure 3. Unreachable cited URLs by AI engine. Share of each engine's cited URLs that no longer resolve; n = 1,254,000 citations. Source: GrackerAI AI Search Visibility Benchmark Series, August 2026.

Of the 1,254,000 cited URLs tested, 196,100 did not resolve, an aggregate dead-link rate of 15.6%. The rate varies by a factor of 2.9 between engines, from 9.4% for Google AI Overview to 27.6% for Microsoft Copilot. Microsoft Copilot records both the lowest citation volume and the highest dead-link rate.

Median source age is recorded as 0 days for every engine. The recorded value is uniform across all five engines, which is consistent with either same-day source material or an undated default in the extraction, and the available metadata does not separate the two. No engine cited any domain belonging to its own parent ecosystem.

Distinct domain counts range from 33 for Microsoft Copilot to 72 for Google AI Overview. These counts are per engine and are not additive: the same domain cited by two engines is counted once within each. Source variety, the mean number of distinct domains within a single response, ranges from 0.8 for Microsoft Copilot to 1.8 for Google AI Overview, indicating that most responses draw on one or two domains regardless of how many citations they carry.

Table 8. Most-cited domains by AI engine

AI engine	Most-cited domains, with citation counts
ChatGPT	paloaltonetworks.com (59,000), gartner.com (42,000), tenable.com (30,000)
Microsoft Copilot	uinat.com (24,000), eastbaycyber.com (17,000), ionix.io (13,000)
Gemini	synack.com (42,000), bitsight.com (27,000), guideflow.com (20,000)
Google AI Mode	gartner.com (33,000), ionix.io (19,000), cycognito.com (13,000)
Google AI Overview	cycognito.com (53,000), ionix.io (30,000), paloaltonetworks.com (27,000)

The most-cited domain sets differ by engine. Vendor-owned domains appear in the top three for four of the five engines. Two engines cite an analyst domain in their top three: ChatGPT at 42,000 citations and Google AI Mode at 33,000. Microsoft Copilot's top three consists entirely of third-party sites that do not appear in any other engine's top three.

4.6 Cross-engine source overlap

Jaccard source overlap measures the proportion of cited domains two engines hold in common. A value of 1.00 denotes identical domain sets and 0.00 denotes no shared domains.

Table 9. Jaccard source overlap between AI engines

	ChatGPT	Copilot	Gemini	AI Mode	AI Overview	Perplexity
ChatGPT	**1.00**	0.07	0.12	0.19	0.17	0.00
Copilot	0.07	**1.00**	0.04	0.10	0.08	0.00
Gemini	0.12	0.04	**1.00**	0.28	0.27	0.00
Google AI Mode	0.19	0.10	0.28	**1.00**	0.37	0.00
Google AI Overview	0.17	0.08	0.27	0.37	**1.00**	0.00
Perplexity	0.00	0.00	0.00	0.00	0.00	**1.00**

Shading denotes overlap magnitude: 0.00–0.05 | 0.06–0.14 | 0.15–0.22 | 0.23–0.30 | 0.31–0.40

The highest overlap recorded between two different engines is 0.37, between Google AI Mode and Google AI Overview. The next highest values are 0.28 between Gemini and Google AI Mode, and 0.27 between Gemini and Google AI Overview. All three pairs sit within the same vendor's family of surfaces.

Overlap across vendor families is markedly lower. ChatGPT records 0.07 with Microsoft Copilot, 0.12 with Gemini, 0.19 with Google AI Mode and 0.17 with Google AI Overview. Microsoft Copilot records the lowest values overall, between 0.04 and 0.10 against every other engine. Perplexity records 0.00 against all five, indicating no cited domain in common with any other engine, on a base of 3,000 responses.

4.7 First-mention position

The first vendor named in a response was recorded for each engine.

Table 10. First-mention vendor by AI engine

AI engine	First-mention vendor	That vendor's mentions in this engine	Its mean ordinal position
ChatGPT	Palo Alto Networks	28,000	1.4
Microsoft Copilot	Palo Alto Networks	13,000	1.5
Gemini	CyCognito	31,000	2.2
Google AI Mode	CyCognito	32,000	2.2
Google AI Overview	CyCognito	36,000	1.8
Perplexity	Microsoft	2,000	2.0

Three engines open with CyCognito, two with Palo Alto Networks, and one with Microsoft. The split does not follow overall mention volume. Palo Alto Networks holds the opening position in Microsoft Copilot on 13,000 mentions, the lowest count of any first-mention vendor in the table, while CyCognito holds 32,000 mentions in the same engine without holding the opening position.

Microsoft is the first-mention vendor in Perplexity on 2,000 mentions, ranking fourth by total mentions across the corpus at 88,000. This observation rests on 3,000 analysed responses.

4.8 Geographic variance

The same vendors surface as leading recommendations in all four markets. The per-market leading recommendations recorded were:

Table 11. Leading product recommendations by market

Market	First	Second	Third
United States	Cortex Xpanse (Palo Alto Networks)	CyCognito Platform (CyCognito)	Tenable One (Tenable)
Canada	Cortex Xpanse (Palo Alto Networks)	CyCognito (CyCognito)	Microsoft Defender EASM (Microsoft)
India	Tenable One (Tenable)	Cortex Xpanse (Palo Alto Networks)	IONIX EASM (IONIX)

Market	First	Second	Third
Germany	CyCognito (CyCognito)	Cortex Xpanse (Palo Alto Networks)	Bitsight EASM (Bitsight)

Cortex Xpanse appears in the leading three in all four markets, and CyCognito in three of four. Ordering varies between markets while membership of the leading group does not. No engine introduced a market-specific vendor absent from the other three markets, and no engine referenced country-specific regulatory or data-residency requirements in its recommendations.

4.9 Inter-engine disagreement

One substantive disagreement between engines was recorded in the corpus. On the question of which platform to use for merger-and-acquisition and multi-subsidiary asset discovery, ChatGPT and Microsoft Copilot named Palo Alto Networks Cortex Xpanse, citing its global IPv4 scanning scale, while Gemini and Google AI Mode named CyCognito, citing its seedless graph attribution approach.

Two single-engine claims were recorded that no other engine reproduced. Microsoft Copilot alone described Vulcan Cyber exposure orchestration as a native integration milestone within Tenable One. ChatGPT alone identified runZero as the primary agentless internal cyber asset attack surface management complement to external attack surface management. This report records both as observed without assessing their accuracy.

5. Discussion

This section answers each research question from the results reported in Section 4. Statements about mechanism are marked as interpretation.

5.1 RQ1 — Do the engines converge on the same vendor shortlist?

They converge at the level of membership and diverge at the level of ordering. CyCognito and Palo Alto Networks are named by all six engines and hold the two highest mention totals in the corpus at 160,000 and 115,000 respectively (§4.2). Censys, Microsoft, Tenable, CrowdStrike and IONIX follow in a second tier named by five or six engines each. No engine nominates a leading vendor that the others omit entirely.

Ordering is not shared. Three engines open with CyCognito and two with Palo Alto Networks (§4.7). A buyer consulting Google AI Overview and a buyer consulting ChatGPT receive the same candidate set in a different order, with a different vendor occupying the position that carries most weight.

5.2 RQ2 — How concentrated is brand visibility?

Visibility is concentrated at the vendor level and fragmented at the product level. The ten most-mentioned vendors account for 805,000 mentions, with CyCognito alone holding 160,000 (§4.2). Against this, the corpus records 185 distinct product entries, 96 of which record 1,000 mentions or fewer (§4.3).

The gap between these two views is partly an artefact of naming. Tenable appears as fifteen product entries, Qualys as eleven, Bitsight and Rapid7 as ten each, and Censys as seven. A plausible interpretation is that a vendor whose product naming is inconsistent across its own published material has its recorded visibility distributed across several entries rather than consolidated into one, though this study measures the fragmentation without establishing its cause.

Sentiment does not discriminate within the category. All twelve vendors recorded fall between 0.70 and 0.84 on a scale spanning −1.0 to 1.0 (§4.2). Sentiment carries little information here, and a measurement programme treating it as a primary signal in this category would be tracking a metric with almost no variance.

5.3 RQ3 — Do the engines share an evidence base?

Largely they do not. The highest overlap between two different engines is 0.37, between Google AI Mode and Google AI Overview, and the three highest values all fall within one vendor's family of surfaces (§4.6). Across vendor families, overlap falls to between 0.04 and 0.19. Perplexity records no shared cited domain with any other engine.

The result is consistent with engines reading substantially different portions of the web while arriving at similar conclusions. This has a direct methodological consequence: a visibility programme instrumented on a single engine observes a source population largely unrelated to the one another engine reads, and cannot be assumed to generalise. The convergence on vendors reported under RQ1 does not imply convergence on evidence.

5.4 RQ4 — Does ordinal position track mention volume?

It does not. Palo Alto Networks holds the opening position in Microsoft Copilot on 13,000 mentions, while CyCognito records 32,000 mentions in the same engine without holding the opening position (§4.7). Within ChatGPT, CyCognito holds a higher mention count than several vendors while recording a mean ordinal position of 4.5, against 1.4 for Palo Alto Networks.

The two measures behave independently, which is consistent with their being produced by different mechanisms: breadth of mention plausibly reflects how widely a vendor is documented across the sources an engine reads, while ordinal position reflects which source the engine drew on first when composing a given response. On this interpretation they are separate objectives requiring separate treatment, and a programme reporting only aggregate share of voice will not detect movement in either. The data establishes the independence; it does not establish the mechanism.

5.5 RQ5 — Does recommendation vary by market?

Membership of the leading group does not vary; ordering within it does. Cortex Xpanse appears in the leading three in all four markets and CyCognito in three of four (§4.8). No market-specific vendor appears, and no engine referenced country-specific regulatory or data-residency requirements. The ordering differences observed — Tenable One leading in India, CyCognito in Germany — occur within a stable candidate set rather than producing a different one.

The absence of localisation is itself a finding with resource consequences, discussed in Section 8.

5.6 What these results do not resolve

The study observes engine behaviour but not its causes. It cannot say why an engine cites one domain rather than another, why the two Google surfaces overlap more with each other than with Gemini, or whether the ordering differences between markets reflect genuine localisation signals. It cannot distinguish a vendor absent because it is poorly documented from one absent because the engine's retrieval did not reach its documentation. Because the engines are not version-pinned, the study also cannot separate stable engine characteristics from conditions specific to August 2026.

6. Threats to Validity

6.1 Construct validity

Several metrics are proxies for the concepts they represent, and two carry documented problems.

Response completeness is inferred from response structure rather than observed directly, as noted in Section 3.7. A response ending before its list concludes is recorded as incomplete, but the study cannot distinguish an output-length limit from an engine legitimately concluding its answer.

Sentiment scoring uses an undocumented model and scale (§3.7). Given that all observed values fall within a 0.14 band, the metric may be measuring the generally positive register of vendor descriptions rather than differential sentiment toward specific vendors.

Mean ordinal position is used as a proxy for prominence. This assumes earlier mention confers more influence on a reader, which is plausible but not measured by this study.

Distinct domain counts are recorded per engine and are not additive, as stated in Table 2. Summing them would count a domain cited by three engines three times.

6.2 Internal validity

All data comes from a single collection window in August 2026. Engine output is non-deterministic, so a repeat of the identical query set would not reproduce identical responses. The engines are commercial services without pinned model versions, and any of them may have changed behaviour during or after the window without notice. Personalisation and session state were not controlled beyond issuing each query without prior conversational context; residual market-level personalisation cannot be excluded.

6.3 External validity

The study covers one category, four markets and one collection window. The query set is purposive rather than randomly sampled from a defined population of buyer questions. For this reason no statistical inference is performed anywhere in this report: no significance tests, no confidence intervals, and no claims of generalisation beyond the observed corpus. All figures are descriptive statistics of what was observed.

Findings for enterprise attack surface management should not be transferred to other security categories without measurement. Categories differ in vendor-set maturity, volume of published material and stability of terminology, all of which plausibly affect engine behaviour.

6.4 Reliability

A repeat study would be expected to reproduce the broad structure of these results: a narrow leading vendor group, low cross-family source overlap, and substantial variation in dead-link rates. It would not be expected to reproduce exact mention counts, exact ordinal positions, or the specific most-cited domains, all of which depend on retrieval conditions at the moment of collection. Categorical findings such as first-mention vendor by engine are more stable than continuous ones, but have not been tested across windows.

7. Limitations

- One category, four markets, one collection window.
- Citation totals were recorded for five of the six engines, so the 1,254,000-source evidence base covers those five.
- Engine responses were not fact-checked; where engines disagree on a factual claim, the disagreement is recorded but not adjudicated.
- Product-entry naming is fragmented, and no entity resolution was applied to consolidate variant product names.
- The sentiment scoring method is not documented in the source dataset.

8. Practical Implications

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

- **Instrument at least three engines drawn from different vendor families.** Cross-family source overlap runs between 0.04 and 0.19 (§4.6), so a programme measuring one engine observes a source population largely unrelated to the others. Measuring ChatGPT, one Google surface and Microsoft Copilot covers three substantially disjoint evidence bases.
- **Track mention volume and ordinal position as separate objectives.** The two do not move together (§4.7, §5.4). A single share-of-voice number conceals movement in both.
- **Treat link integrity as a visibility task.** An aggregate 15.6% of cited URLs did not resolve, rising to 27.6% for Microsoft Copilot (§4.5). The sources carrying a vendor's claims decay faster than most editorial calendars refresh them, and a dead source cannot carry visibility.
- **Consolidate product naming.** The corpus records 185 product entries for a far smaller number of actual products, with Tenable alone spread across fifteen (§4.3). Consistent product naming across published material is the prerequisite for consolidated recorded visibility.
- **Do not fund market-specific variants on the assumption that engines localise.** Membership of the leading group is identical across all four markets and no engine referenced country-specific regulatory requirements (§4.8). Resources directed at locale variants would return more if redirected into documentation depth in the primary language.
- **Do not treat sentiment as a primary signal in this category.** All twelve vendors measured fall within a 0.14 band (§4.2). The metric has too little variance here to support a decision.

9. Conclusion

Across 240,000 AI engine responses and 1,254,000 cited sources, six AI search engines converge on a narrow leading group of vendors for enterprise attack surface management, disagree systematically on which of them to name first, and disagree almost entirely on which evidence to read.

Convergence on vendors sits alongside near-total divergence on sources. The highest overlap between any two different engines is 0.37, and that pair consists of two surfaces from the same vendor; across vendor families overlap falls to between 0.04 and 0.19. Similar answers are being produced from substantially different evidence, which means the shortlist is stable while the mechanism producing it is not shared.

For a vendor inside the leading group the position appears durable across engines, markets and phrasings. For a vendor outside it, absence is consistent across all of them. The measurable factors associated with position are documentation depth, consistent product naming, and the reachability of the sources carrying a vendor's claims — the last of which is degrading at an aggregate 15.6% across the corpus.

10. Reproducibility and Data Availability

Queries were issued to each engine's public surface through a commercial data collection layer, with each response, its cited sources, and the per-response analysis persisted to a columnar analytical database. Retained records include the response text, the extracted citations with their resolved domains and reachability status, the brand and product mentions with ordinal position, and the per-engine aggregate metrics reported here.

Reproducing the study requires the 10,000-query set, the four market contexts, access to the six engine surfaces during a comparable window, and the metric definitions in Table 2. Because the engines are not version-pinned and their output is non-deterministic, an exact reproduction of these figures is not achievable; the appropriate reproduction target is the structure of the findings rather than their values.

The underlying dataset for this report was generated in August 2026. Requests for the aggregate dataset supporting the tables in this report may be directed to GrackerAI.

Disclosure

GrackerAI publishes this research and sells AI search visibility measurement for the category it covers. The study measured how AI engines describe vendors in their responses. It did not test, evaluate, or rank any vendor's product, and no vendor paid for or requested placement. Vendor and product names appear only as observed in engine output.

How to cite this report

GrackerAI. (August 2026). *AI Search Visibility in Enterprise Attack Surface Management: A Multi-Engine, Multi-Market Benchmark Study*. GrackerAI AI Search Visibility Benchmark Series. Available at: https://gracker.ai/data-and-research-reports

Appendix A. Brand visibility by AI engine

Figures are brand mentions, with mean ordinal position in parentheses. An em dash indicates the vendor was not recorded in that engine. Mean ordinal position is stated on a scale where 1.0 denotes the first vendor named in a response.

Table A1. Brand mentions and mean ordinal position, ten most-mentioned vendors

Brand	Total	ChatGPT	Copilot	Gemini	AI Mode	AI Overview	Perplexity
CyCognito	160,000	26,000 (4.5)	32,000 (2.9)	31,000 (2.2)	32,000 (2.2)	36,000 (1.8)	3,000 (3.3)
Palo Alto Networks	115,000	28,000 (1.4)	13,000 (1.5)	29,000 (3.2)	25,000 (1.7)	19,000 (2.0)	1,000 (2.0)

Brand	Total	ChatGPT	Copilot	Gemini	AI Mode	AI Overview	Perplexity
Censys	101,000	24,000 (3.8)	15,000 (4.2)	27,000 (3.0)	21,000 (3.0)	14,000 (3.7)	—
Microsoft	88,000	23,000 (3.3)	19,000 (3.3)	20,000 (3.9)	13,000 (3.0)	11,000 (3.8)	2,000 (2.0)
Tenable	87,000	22,000 (2.8)	10,000 (3.5)	21,000 (5.0)	18,000 (3.9)	15,000 (4.2)	1,000 (3.0)
CrowdStrike	79,000	13,000 (4.2)	21,000 (2.8)	10,000 (4.3)	20,000 (4.0)	14,000 (4.5)	1,000 (1.0)
IONIX	74,000	12,000 (6.1)	11,000 (3.9)	11,000 (3.1)	14,000 (3.6)	25,000 (3.2)	1,000 (4.0)
Rapid7	37,000	12,000 (5.0)	7,000 (4.6)	3,000 (7.0)	8,000 (4.4)	7,000 (4.5)	—
Bitsight	34,000	6,000 (4.3)	—	18,000 (3.2)	8,000 (5.5)	1,000 (5.0)	1,000 (8.0)
Qualys	30,000	7,000 (7.1)	5,000 (3.4)	4,000 (4.3)	6,000 (5.2)	8,000 (4.4)	—

The matrix makes coverage gaps explicit. A vendor may hold a strong position in one engine while being absent from another, a condition that a single aggregate mention total conceals. Bitsight is absent from Microsoft Copilot while holding 18,000 mentions in Gemini; Censys, Rapid7 and Qualys are absent from Perplexity.

Appendix B. Source mix and most-cited domains

Table B1. Source-type distribution by AI engine — citations by classified source type

AI engine	Forum	Marketplace	Other	Total
ChatGPT	—	—	360,000	360,000
Microsoft Copilot	—	—	156,000	156,000
Gemini	—	—	215,000	215,000
Google AI Mode	1,000	—	172,000	173,000
Google AI Overview	3,000	1,000	346,000	350,000
Total	**4,000**	**1,000**	**1,249,000**	**1,254,000**

Source-type classification assigned 99.6% of citations to the residual "other" category. The classification scheme as applied does not discriminate usefully within this corpus, and the source-type dimension is therefore not analysed in Section 4.

Table B2. Three most-cited domains by AI engine

AI engine	First	Second	Third	Distinct domains
ChatGPT	paloaltonetworks.com (59,000)	gartner.com (42,000)	tenable.com (30,000)	56
Microsoft Copilot	uinat.com (24,000)	eastbaycyber.com (17,000)	ionix.io (13,000)	33
Gemini	synack.com (42,000)	bitsight.com (27,000)	guideflow.com (20,000)	36

AI engine	First	Second	Third	Distinct domains
Google AI Mode	gartner.com (33,000)	ionix.io (19,000)	cycognito.com (13,000)	46
Google AI Overview	cycognito.com (53,000)	ionix.io (30,000)	paloaltonetworks.com (27,000)	72

Appendix C. Product entries

The corpus records 185 distinct product entries. The twenty most-mentioned are listed. Mean product rank is the mean ordinal position of the product entry within a response, where 1.0 denotes first named.

Table C1. Twenty most-mentioned product entries

#	Product	Brand	Mentions	Mean product rank	Mean sentiment	Engines	Markets
1	Cortex Xpanse	Palo Alto Networks	149,000	1.8	0.86	6	4
2	CyCognito	CyCognito	132,000	2.7	0.81	6	4
3	Microsoft Defender EASM	Microsoft	73,000	3.6	0.79	6	4
4	IONIX	IONIX	57,000	4.1	0.81	5	4
5	CrowdStrike Falcon Exposure Management	CrowdStrike	51,000	3.1	0.81	6	4
6	Censys	Censys	46,000	4.0	0.77	5	4
7	Tenable One	Tenable	40,000	3.1	0.80	5	4
8	Censys ASM	Censys	38,000	2.7	0.83	5	4
9	CrowdStrike Falcon Surface	CrowdStrike	36,000	4.7	0.80	5	4
10	Microsoft Defender External Attack Surface Management	Microsoft	27,000	3.0	0.79	6	4
11	Tenable Attack Surface Management	Tenable	19,000	4.4	0.81	5	4
12	Censys Attack Surface Management	Censys	17,000	3.6	0.85	3	4
13	UpGuard	UpGuard	16,000	5.0	0.72	4	4
14	runZero	runZero	15,000	7.9	0.74	4	4
15	Mandiant ASM	Google Cloud	13,000	4.1	0.80	2	4
16	Synack	Synack	12,000	4.3	0.84	2	3
17	Rapid7 Surface Command	Rapid7	12,000	4.6	0.78	5	3
18	Detectify	Detectify	12,000	5.1	0.72	2	4
19	Bitsight	Bitsight	11,000	2.6	0.86	2	4
20	Wiz	Wiz	11,000	6.8	0.77	3	4

Of the 185 entries recorded, 104 register 1,000 mentions or fewer. Naming fragmentation is substantial: Tenable holds fifteen separate entries, Qualys eleven, Bitsight and Rapid7 ten each, Palo Alto Networks nine, Censys seven and CyCognito six.

Appendix D. Study parameters

Table D1. Collection parameters

Parameter	Value
AI engines	ChatGPT, Perplexity, Gemini, Microsoft Copilot, Google AI Overview, Google AI Mode
Markets	United States, Canada, India, Germany
Buyer-intent queries	10,000
Responses per engine	40,000
Responses per engine per market	10,000
Total responses recorded	240,000
Total cited sources	1,254,000
Query topic	Selection of enterprise attack surface management platforms for 2026
Collection window	August 2026
Dataset generated	August 2026

Appendix E. Metric computation

Table E1. Formulas for derived metrics

Metric	Computation
Citation density	cited sources per 100 words of response text
Unreachable citations	total cited sources × dead-link rate, per engine
Aggregate dead-link rate	Σ unreachable citations ÷ Σ total cited sources
Total cited sources	Σ total cited sources across engines reporting the metric
Brand mention total	Σ brand mentions across all six engines, per brand
Jaccard source overlap	$\|A \cap B\| \div \|A \cup B\|$, where A and B are two engines' cited-domain sets
Mean ordinal position	mean index of a brand's first appearance within a response