THE AI SALES OUTCOME GAP 2026

A public-claims audit of 100 AI sales and revenue products

Headline finding

On the public pages reviewed, 69 of 100 products made sales execution a strongly evidenced core capability. Only eight did the same for learning from downstream commercial outcomes.

Research snapshot: 7 September 2026

Commissioned and authored by Go7

Go7 was excluded from scoring.

Full row-level coding, sources and stage evidence are supplied in the accompanying master audit workbook.

Research note and disclosure

This report evaluates what 100 products publicly claim, describe and evidence on the pages reviewed. It is not hands-on product testing, a vendor ranking or a declaration that undocumented capabilities do not exist.

Publication-safe interpretation

Where this report says “no public evidence located”, it means the reviewed pages did not provide sufficient support for the coded capability. It does not mean the product cannot perform it.

Go7 is developing a platform in the same broad category and therefore has a commercial interest in the market. To reduce that conflict, Go7 was excluded from scoring, the methodology is published, every coded conclusion has source URLs in the workbook, and a corrections log is included.

Contents

- Executive summary
- Why this matters now
- Methodology
- Ten strongest findings
- Segment analysis
- Implications for buyers and vendors
- Limitations and corrections
- Appendix: sample, chart specifications and references

Executive summary

The audit identifies a clear public-positioning gap in the AI sales market: products are generally much more explicit about the work they can execute than the commercial outcomes they can learn from.

100 products audited	**69** strongly evidenced Act	**8** strongly evidenced Learn
61 point Act–Learn gap	**57** no public approval-control evidence	**63** no public explainability evidence

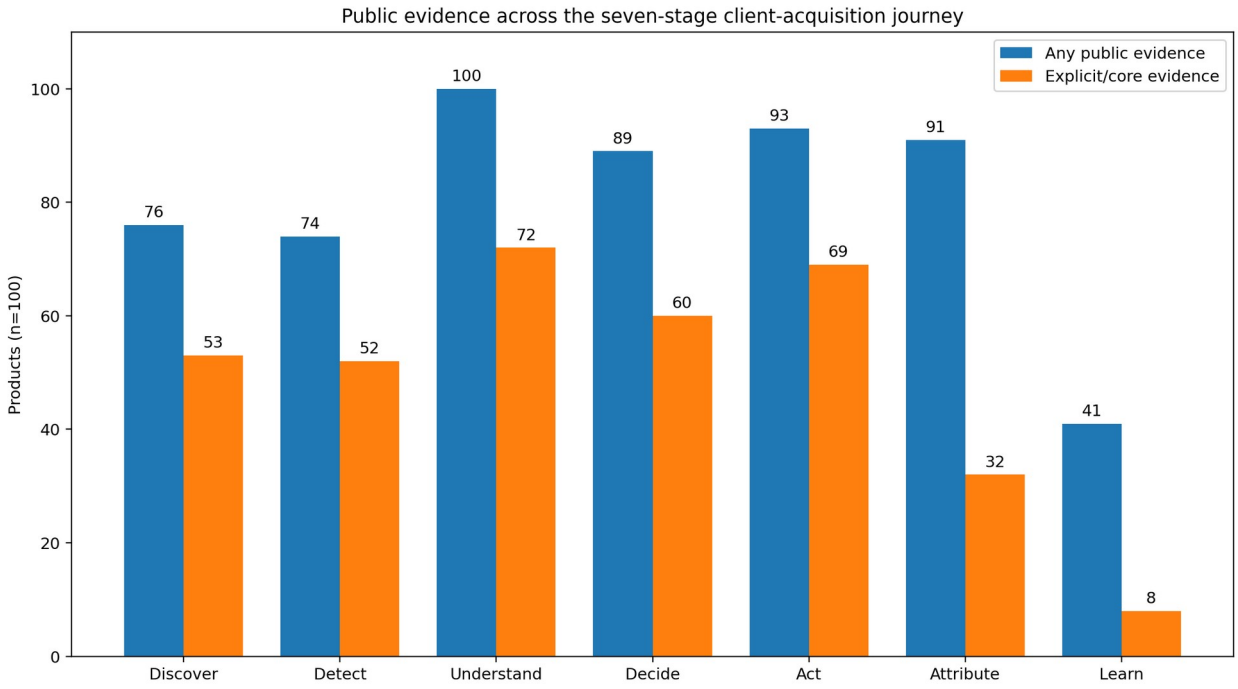


Figure 1. Any evidence and explicit/core evidence across the seven-stage client-acquisition journey. Source: Go7 public-claims audit, 7 September 2026.

Five headline conclusions

1. Action is ahead of learning — 69 products strongly evidenced execution, compared with only eight that strongly evidenced learning from downstream commercial outcomes.

2. The gap is widest in outbound execution — All 25 sales-engagement and outbound products strongly evidenced Act. None strongly evidenced Learn.

3. The market is split upstream and downstream — Agentic SDR products are strongest at action; revenue-intelligence products are stronger at attribution and learning but rarely discover new accounts.

4. Proxy metrics still dominate public messaging — 55 products led primarily with activity, efficiency, replies or meetings. Twenty-two led with revenue, win rate or forecast outcomes.

5. Controls and explainability are not prominent — 57 products had no public evidence of human approval controls in the reviewed pages, and 63 had no public evidence of explainability or audit trails.

Why this matters now

AI agents are moving rapidly from demonstrations into sales workflows. Gartner predicts that agents could outnumber human sellers ten to one by 2028, while fewer than 40% of sellers may say those agents improved productivity. Gartner's warning is that fragmented systems can simply lead agents to scale the fragmentation.[1]

The governance gap is also visible beyond sales. Gartner reported in April 2026 that only 13% of organisations believed they had the right AI-agent governance in place.[2]

McKinsey's 2026 B2B sales research argues that value comes from embedding agentic AI into end-to-end commercial workflows rather than layering isolated tools onto existing processes.[3] Salesforce says nine in ten sales teams use agents or expect to within two years.[4] HubSpot has even shifted parts of its agent pricing towards completed outcomes rather than output volume.[5]

The category question

If AI sales products become increasingly capable of taking action, the next competitive question is whether they can connect those actions to opportunity, revenue, governance and learning.

The audit therefore focuses on seven stages: Discover, Detect, Understand, Decide, Act, Attribute and Learn. The stages are not intended as a universal product taxonomy. They are a practical way to test how much of the commercial loop a product publicly describes.

Methodology

Sample

A purposive sample of 100 active, publicly marketed English-language products was assembled across four segments:

- Agentic SDR and revenue agents: 30
- Sales engagement and outbound execution: 25
- Prospecting, data and signal intelligence: 25
- Revenue and conversation intelligence: 20

Products were identified through official vendor pages, public category searches, pricing pages, help centres, product documentation, case-study libraries and current market roundups. The sample is structured but not random and must not be described as representative of every AI sales product worldwide.

Source hierarchy

Official product, pricing, documentation and case-study pages were preferred. Third-party sources were used only where official pages were inaccessible or where a current product rename or public pricing reference required verification. Source URLs and review dates appear in the master workbook.

Stage scoring

Score	Meaning
0	No public evidence located in the pages reviewed. This is not evidence that the feature is absent.
1	Partial or supporting evidence. The capability assists the stage but is not clearly positioned as a core product capability.
2	Explicit/core evidence. The capability is clearly described, documented or positioned as central to the product.

Stage	What was coded
Discover	Market/account identification, list building or contact discovery.
Detect	Buying signals, timing, intent, behavioural or trigger detection.
Understand	Account research, CRM/history use, conversation context or account intelligence.
Decide	Prioritisation, scoring, routing, next-best action, strategy or explainable recommendation.
Act	Autonomous or workflow execution across email, calls, CRM actions, routing or other sales tasks.
Attribute	Connection to qualified opportunities, pipeline, forecast, win rate or revenue outcomes.
Learn	Use of downstream replies, opportunities, wins, losses or revenue to improve future decisions or agent behaviour.

Evidence and governance coding

The audit also coded historical first-party context, human approval controls, explainability/audit trails, public pricing, public case studies and the primary metric used in headline positioning. Commercial evidence was graded from qualitative/unquantified evidence through named quantified customer cases.

Ten strongest findings

1. Execution is far more strongly evidenced than downstream learning.

69 of 100 products strongly evidenced Act; 8 of 100 strongly evidenced Learn.

This is the clearest public-positioning gap in the audit.

Workbook calculation: COUNTIF(Master Audit[Act score],2) versus COUNTIF(Master Audit[Learn score],2)

2. The gap remains even when partial evidence is included.

93 products had at least some public evidence of Act; 41 had any public evidence of Learn.

Most products can take or trigger action, but fewer describe a feedback loop from commercial outcomes.

Workbook calculation: 93 minus 41 = 52-point gap

3. AI SDR and revenue-agent products are positioned as execution engines.

28 of 30 agentic SDR/revenue-agent products strongly evidenced Act; none strongly evidenced Learn.

The category's public story is autonomy and execution, not closed-loop commercial learning.

Workbook calculation: Segment='Agentic SDR and revenue agents'; Act score=2; Learn score=2

4. Sales-engagement platforms show the widest outcome gap.

All 25 sales-engagement/outbound products strongly evidenced Act; none strongly evidenced Learn.

Execution is universal in the segment, but learning from downstream outcomes is not a core public capability.

Workbook calculation: Segment='Sales engagement and outbound execution'; Act score=2 and Learn score=2

5. Revenue-intelligence products are stronger downstream but weak upstream.

14 of 20 strongly evidenced Attribute and 7 strongly evidenced Learn, while only 2 had any Discover evidence.

The market is structurally split between tools that generate activity and tools that analyse outcomes.

Workbook calculation: Revenue/conversation segment stage counts

6. Attribution is frequently mentioned but less often a core capability.

91 products had some attribution evidence; only 32 made it a strongly evidenced core capability.

Reporting or CRM visibility is common; explicit opportunity/revenue attribution is less central.

Workbook calculation: Attribute score >0 versus Attribute score=2

7. Public messaging still leans towards proxy metrics.

55 products used activity/efficiency, replies or meetings as the primary headline metric; 22 used revenue, win rate or forecast.

Many products lead with immediate, measurable outputs rather than commercial endpoints.

Workbook calculation: Metric bands: 24 activity + 13 replies + 18 meetings versus 22 revenue/win/forecast

8. Human approval and control are not prominent in most public positioning.

57 products had no public evidence of human approval controls in the pages reviewed; 14 made controls an explicit/core capability.

Autonomy is marketed more visibly than operational guardrails.

Workbook calculation: Human approval controls score 0 versus score 2

9. Explainability and audit trails are even less visible.

63 products had no public evidence of explainability or audit trails in the pages reviewed; 12 made them explicit/core.

The public market narrative gives limited attention to why agents act and how decisions can be reviewed.

Workbook calculation: Explainability/audit score 0 versus score 2

10. Historical first-party context is emerging but not universal.

34 products strongly evidenced historical first-party context; 35 had no public evidence of it in the pages reviewed.

Business memory and CRM history may be a key differentiator as models commoditise.

Workbook calculation: Historical context score 2 versus score 0

The market is split between execution and intelligence

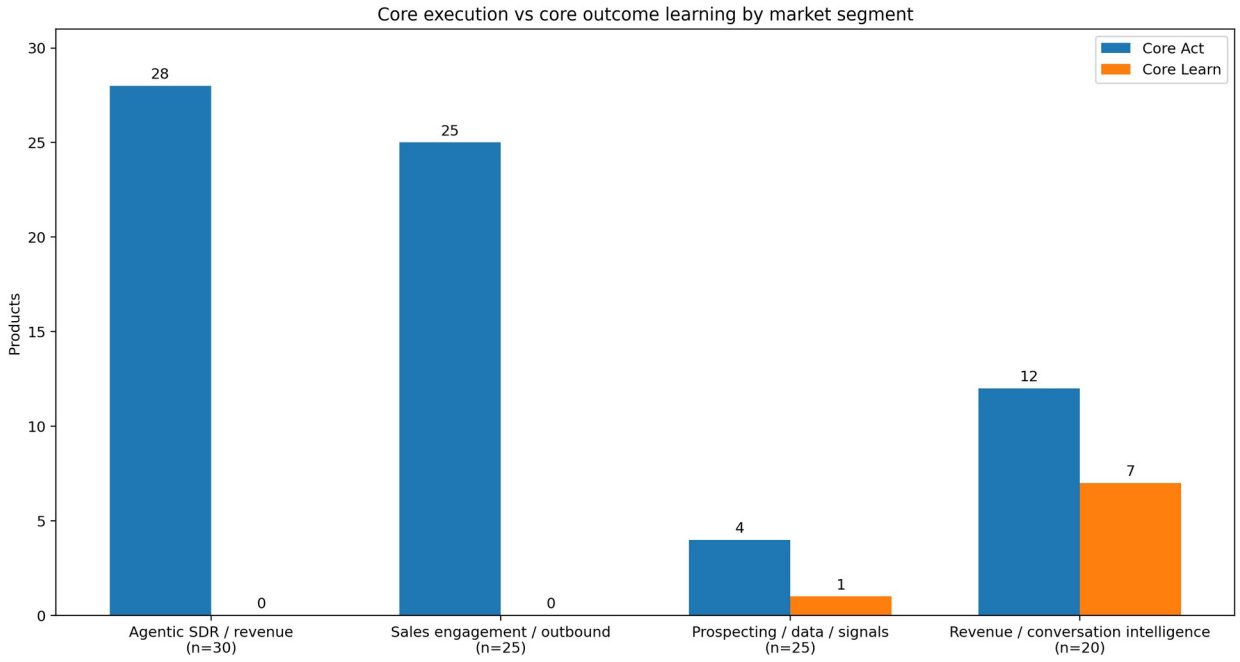


Figure 2. Strongly evidenced execution and learning by segment. Source: Go7 audit.

The strongest outcome gap appears in sales engagement and outbound execution: all 25 products strongly evidenced Act, while none strongly evidenced Learn. Agentic SDR and revenue-agent products show a similar pattern: 28 of 30 strongly evidenced Act, while none made outcome learning an explicit/core public capability.

Revenue and conversation intelligence products are the exception. Fourteen of 20 strongly evidenced attribution and seven strongly evidenced learning. However, only two showed any discovery capability. This supports an interpretation that the market remains divided between upstream systems that create activity and downstream systems that inspect outcomes.

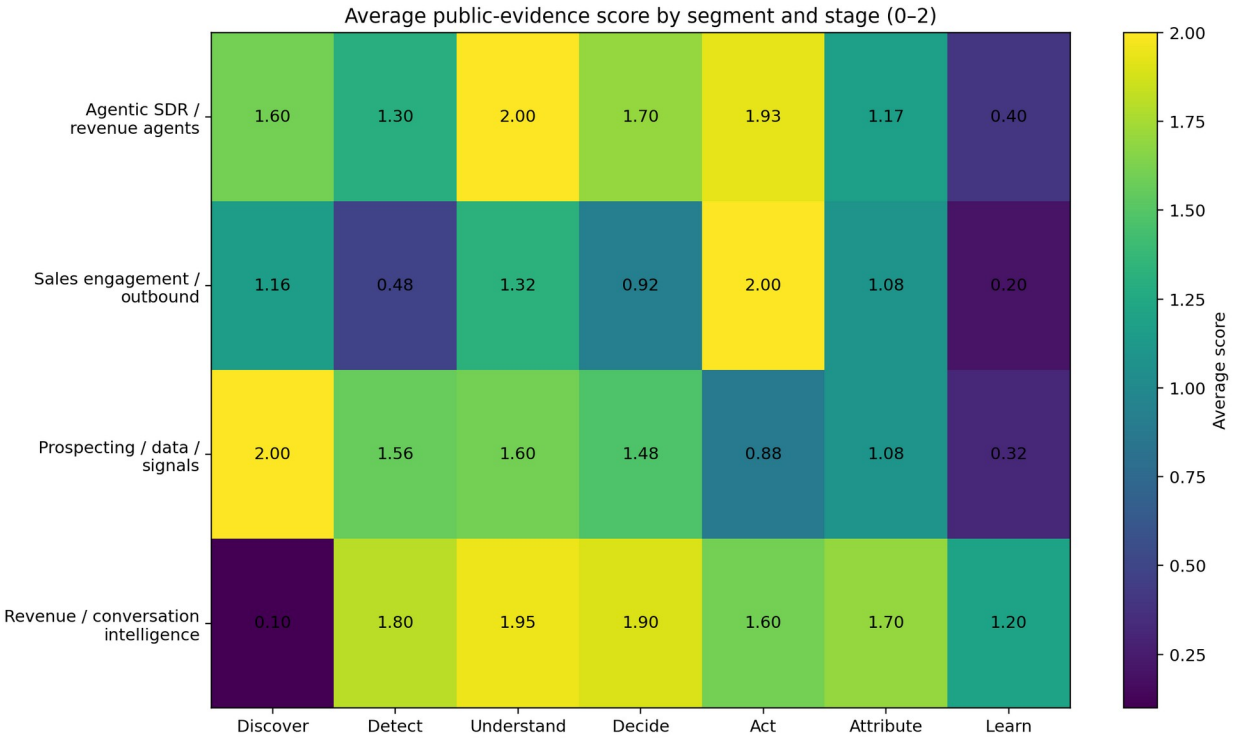


Figure 3. Average public-evidence score by segment and stage. Source: Go7 audit.

Public messaging still leans towards proxy metrics

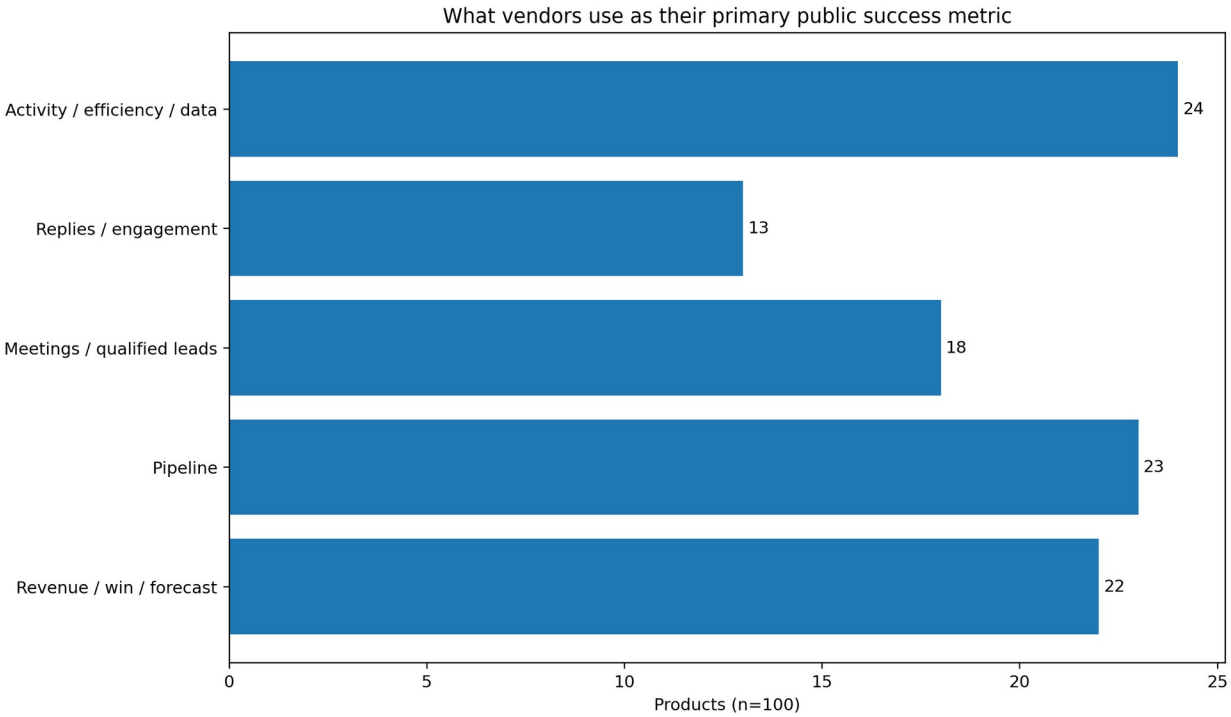


Figure 4. Primary public success metric used in headline positioning. Source: Go7 audit.

Fifty-five products led primarily with activity, efficiency, data outputs, replies or meetings. Twenty-two led with revenue, win rate or forecast outcomes, while 23 led with pipeline.

This does not mean activity or meeting metrics are unimportant. They are immediate and measurable. The risk is that they become the objective rather than a diagnostic step on the route to qualified opportunity and revenue.

Interpretation

As agents gain autonomy, metric selection becomes product design. A system optimising a proxy can become extremely efficient without necessarily improving the outcome the business actually values.

Attribution is common; strong outcome attribution is less central

Ninety-one products had at least some public evidence of attribution, reporting or CRM visibility. Only 32 made attribution to qualified opportunities, pipeline, forecast, win rate or revenue a strongly evidenced core capability.

The 59-product gap between “some evidence” and “core evidence” suggests that many platforms can report activity or sync it to a CRM without making closed-loop commercial attribution central to their public value proposition.

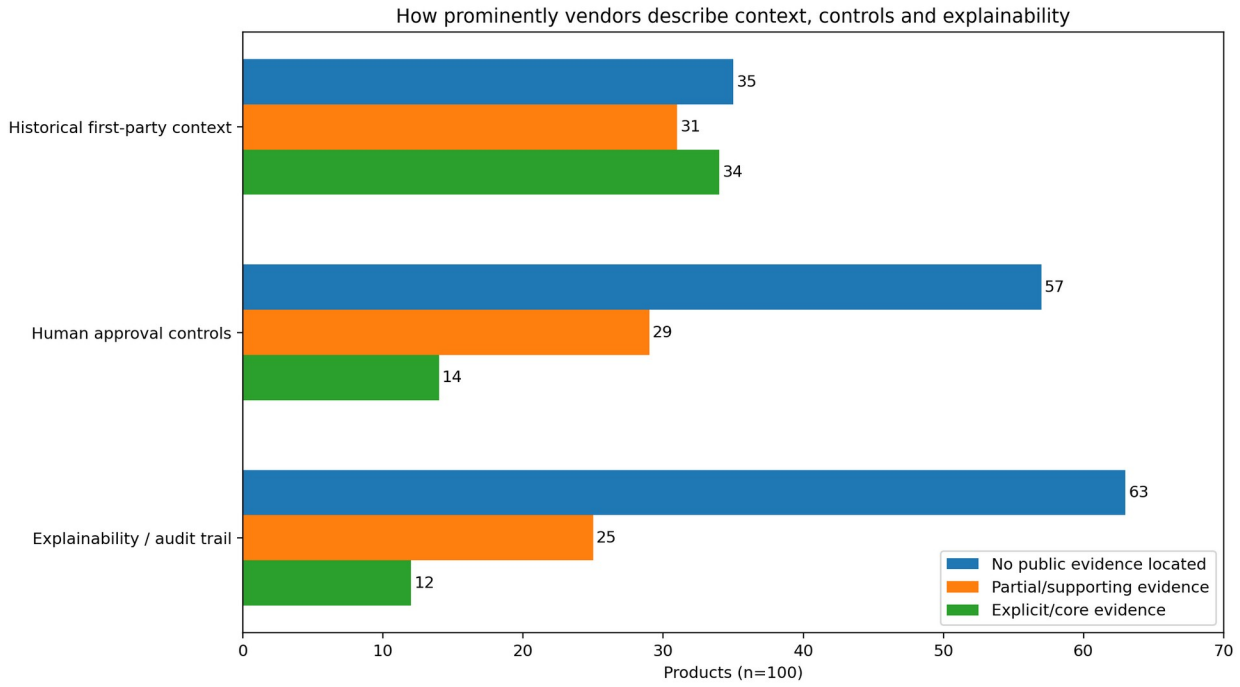


Figure 5. Public visibility of historical context, human controls and explainability. Source: Go7 audit.

Historical first-party context is more visible than controls or explainability: 34 products made it explicit/core, compared with 14 for human approval controls and 12 for explainability/audit trails. In the reviewed pages, 57 products showed no public evidence of approval controls and 63 showed no public evidence of explainability or audit trails.

Important caveat

These figures measure public visibility. Some vendors may provide controls, logs or context inside the product or in private enterprise documentation that was not present on the pages reviewed.

Commercial evidence and pricing

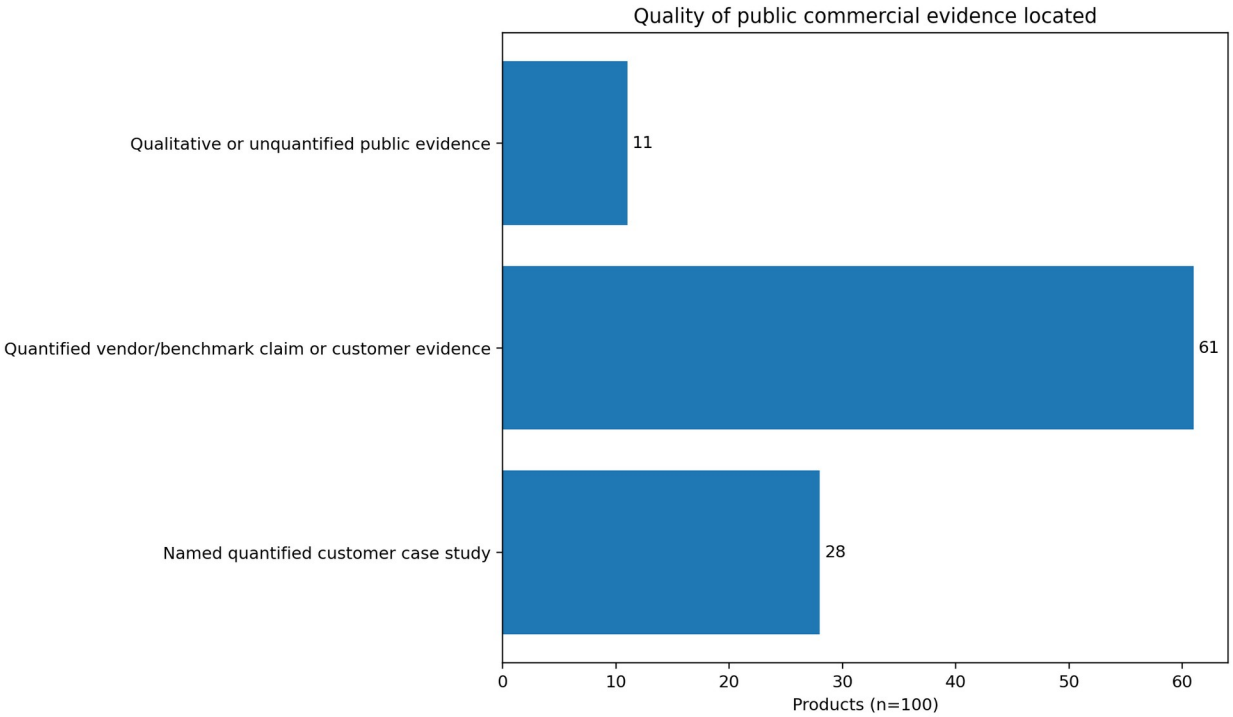


Figure 6. Quality of public commercial evidence located in the reviewed pages. Source: Go7 audit.

Twenty-eight products had a named quantified customer case located in the reviewed pages. Sixty-one had quantified vendor, benchmark or customer evidence, while 11 relied on qualitative or unquantified public evidence in the pages reviewed.

Vendor case studies are useful but should not be treated as independent evaluation. The workbook records evidence type so journalists and analysts can distinguish named customer outcomes from aggregate or vendor-level claims.

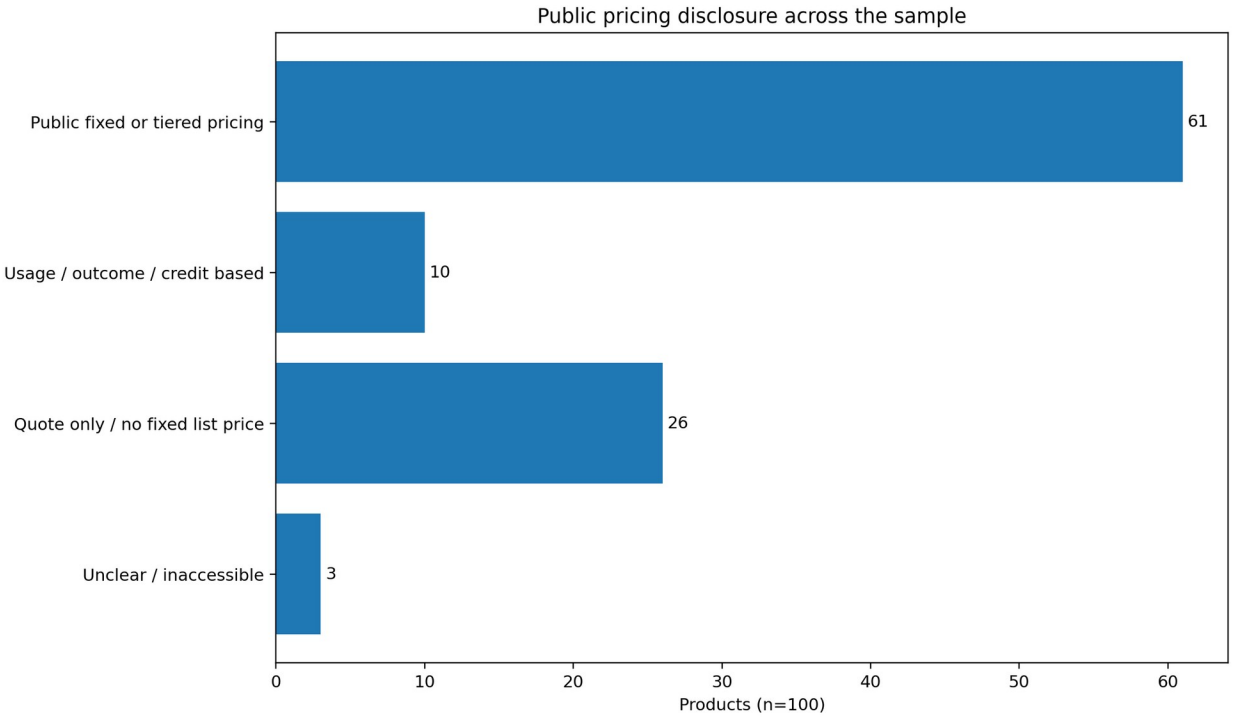


Figure 7. Public pricing disclosure across the sample. Source: Go7 audit.

Sixty-one products showed public fixed or tiered pricing, 26 were quote-only or did not disclose a fixed list price, ten used usage, credit or outcome-based pricing, and three had unclear or inaccessible current pricing.

Outcome-based pricing remains a minority even as vendors increasingly describe agents in outcome-led language. HubSpot is one notable example of pricing a prospecting agent by leads recommended for outreach rather than recurring contact enrolment.[5]

Implications for buyers

Start with the commercial loop, not the demo

Map where the product begins and ends across discovery, signals, context, decisioning, execution, attribution and learning.

Ask how outcomes change future behaviour

A dashboard that reports pipeline is different from a system that uses pipeline, wins and losses to improve the next decision.

Inspect approval boundaries

Buyers should know which actions require human approval, what happens when confidence is low and how external actions can be stopped.

Separate activity from business value

Emails, calls and meetings can be useful leading indicators, but they are not substitutes for qualified opportunity, pipeline quality and revenue.

Demand evidence at the right level

Ask whether a claim is a vendor benchmark, a named customer outcome or an independently evaluated result.

Implications for vendors

Outcome learning is an open category position

Only eight products made downstream commercial learning an explicit/core public capability. Vendors that can prove this safely have a differentiated story.

Controls should be part of positioning

Human approvals, decision logs and explainability are not merely compliance features; they help users trust autonomous systems.

Attribution needs to move from reporting to product behaviour

Connecting activity to revenue matters more when the result changes what the system does next.

Context may outlast model advantage

As foundation models become more accessible, historical business context, workflow state and outcome memory may become more durable sources of value.

Evidence quality will become a competitive asset

Named, quantified and methodologically clear customer outcomes are more defensible than generic productivity claims.

Limitations, risk controls and corrections

Limitation	How it is handled
Public-claims audit	The research evaluates public evidence, not all product capability.
No product testing	Researchers did not log into every platform or independently test performance.
Non-random sample	The 100 products form a structured sample, not a census or statistically representative market sample.
Dynamic market	Product pages, names, prices and functionality can change quickly.
Evidence interpretation	A score reflects how explicit a capability was in reviewed material, not whether one product is better than another.
Commercial conflict	Go7 operates in the category. The commissioner disclosure, exclusion of Go7 and source-backed coding should appear in every release.
Vendor corrections	Vendors should be invited to submit source-backed corrections. Changes must be logged rather than silently overwritten.

Safe language

- **Use:** “No public evidence was located in the pages reviewed.” **Avoid:** “The product cannot do this.”
- **Use:** “Eight products strongly evidenced outcome learning.” **Avoid:** “Only eight products can learn.”
- **Use:** “In this structured sample of 100 products...” **Avoid:** “The AI sales market as a whole...”
- **Use:** “Vendor-published named customer case.” **Avoid:** “Independently proven result.”

Appendix A: products included

Agentic SDR and revenue agents

Product	Product
11x Alice	Artisan Ava
AiSDR	Agent Frank
Jason AI	Regie.ai AI Agents
Duo	Unify
Persana AI	Alta
Piper	Revenue Digital Assistants
SalesAPE	B2B Rocket
Jeeva AI	Telescope AI
Rox	Lindy
Relevance AI	Coldreach
Lyzr AI SDR	Agentforce SDR
Microsoft Sales Agent	Breeze Prospecting Agent
Sendr	Autobound
Ora	Apollo AI
1mind Superhuman	Landbase GTM-2 Omni

Sales engagement and outbound execution

Product	Product
Outreach AI Revenue Agent	Salesloft
Instantly	Smartlead
lemlist	Mailshake
Woodpecker	QuickMail
Saleshandy	Snov.io
Hunter Campaigns	Klenty
Outplay	La Growth Machine
Waalaxy	Dripify
Expandi	Skylead
Overloop	Growbots
Mixmax	Close
PersistIQ	Yesware
SmartReach.io	

Prospecting, data and signal intelligence

Product	Product
Clay	Common Room
6sense Revenue AI	Demandbase
ZoomInfo Copilot	Cognism
Lusha	Seamless.AI
LeadIQ	Kaspr
RocketReach	UpLead
Wiza	Surfe
Ocean.io	Dealfront
Leadfeeder	Factors.ai
Warmly	Trigify
RB2B	FullEnrich
BetterContact	Databar.ai
Albacross	

Revenue and conversation intelligence

Product	Product
Gong	Clari
Backstory (formerly People.ai)	Aviso
Attention	Momentum
Rattle	Sybill
Avoma	Fireflies.ai
Fathom	MeetRecord
Jiminny	Salesroom
Revenue Grid	Orum
Nooks	Dialpad AI Sales
JustCall AI	Aircall AI

Appendix B: chart specifications

Chart	Format	X / rows	Y / columns	Series	Editorial use
1. Seven-stage coverage	Grouped column chart	Stage	Number of products	Any evidence vs explicit/core evidence	Primary launch visual; shows the Act-to-Learn drop.
2. Act vs Learn by segment	Grouped column chart	Four market segments	Number of products	Core Act vs Core Learn	Supports the execution-engine headline.
3. Primary metric orientation	Horizontal bar	Metric band	Number of products	Five bands	Shows proxy metrics versus commercial endpoints.
4. Context, controls and explainability	Grouped horizontal bar	Three dimensions	Number of products	No/partial/core evidence	Governance and trust angle.
5. Evidence quality	Horizontal bar	Evidence grade	Number of products	Three evidence levels	Shows named quantified cases are not the norm.
6. Pricing disclosure	Horizontal bar	Pricing model	Number of products	Four pricing categories	Supports outcome-pricing follow-up story.
7. Segment-stage heatmap	Heatmap	Seven stages	Four segments	Average 0–2 score	Explains market bifurcation visually.

Appendix C: references

[1] Gartner, 28 July 2026 — AI agents will outnumber sellers 10 to 1; agent-sprawl warning.
https://www.gartner.com/en/newsroom/press-releases/2026-07-28-gartner-predicts-ai-agents-will-outnumber-sellers-10-to-1-by-2028-yet-fewer-than-40-percent-of-sellers-will-say-agents-improved-productivity

[2] Gartner, 28 April 2026 — Six steps to manage AI-agent sprawl; 13% governance readiness.
https://www.gartner.com/en/newsroom/press-releases/2026-04-28-gartner-identifies-six-steps-to-manage-artificial-intelligence-agent-sprawl

[3] McKinsey, July 2026 — How agentic AI transforms B2B sales growth.
https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-future-of-b2b-sales-how-growth-champions-rewire-their-playbooks-with-ai

[4] Salesforce, 2026 — State of Sales. https://www.salesforce.com/ap/sales/state-of-sales/

[5] HubSpot, 2026 — Outcome-based pricing for Customer Agent and Prospecting Agent.
https://www.hubspot.com/company-news/hubspots-customer-agent-and-prospecting-agent-now-you-pay-when-the-task-is-complete

About Go7

Go7 is a UK B2B client-acquisition company developing a platform around seven specialist agents and a shared Business Brain. Go7 commissioned this research to examine the gap between AI sales activity and commercial outcome learning. Go7 was excluded from scoring.

Research enquiries and corrections: contact@go7.ai

Full audit workbook: supplied separately with this report.