Rev-Lens
GTM INSIGHTS

WHITEPAPER

Index Graph vs Evidence Graph

Why fields are not evidence, and what that means for every AI answer about your pipeline

Plus four tests a fact should pass before a revenue leader acts on it. Rev-Lens, November 2026.

IN BRIEF

- A language model can read a few hundred calls at once. Something has to choose what it reads, and that choice decides the answer.
- Most graphs built for AI are index graphs: maps that help a model find text. They do not prove anything.
- An evidence graph ties every fact to the customer's exact words, who said them, the call and the minute, and the deal and its outcome, so a question can be answered by counting across every deal.
- CRM fields cannot do this job. In one 24 month review, the loss reason and competitor fields were empty on every row.
- Four tests separate a fact you can act on from a plausible sentence: sourced, verified, attributed and labelled.

Point a large language model at two years of your sales calls and ask why you keep losing to your biggest competitor. You will get an answer in seconds, and it will be fluent, specific and plausible. Whether it is right depends almost entirely on something the model never sees: how the data underneath it was organized before the question was asked.

What a graph is

A graph is the simplest way to write down how things are connected. The nodes are the things: accounts, people, opportunities, calls, and the individual statements made on those calls. The edges are the connections: this person works at this account, joined this call, said this sentence about this competitor, on this opportunity, which closed lost in March. Your CRM holds some of these links as fields. Your call recordings hold the rest, unwritten, inside hours of transcript.

The graph matters because of how a model reads. Even the largest model can hold only a few hundred calls' worth of transcript at once. The company in the example later in this paper recorded 39,258 calls in 24 months. So every AI answer over a dataset that size starts with a step

nobody sees: something decides which few pieces of text the model reads. Without a graph, that step is similarity search, which finds the passages that sound most like the question. That works for questions that live in one place. It fails on the questions revenue leaders ask, such as which lost deals raised a competitor before the proposal went out, because the pieces of those answers sit across a call, a contact, an opportunity and an outcome, never in one passage.

What goes wrong without one

The model answers anyway. It reads the passages that sounded closest, fills the gaps with what is typical, and presents the result with the confidence of a verified fact. None of the following is visible in the answer.

It counts what it retrieved, not what happened. Twenty passages about price become "price is the main reason we lose," whether or not price came up on the lost deals at all.

It cannot tell whose words it is reading. A rep's summarizing question ("so security is really the main driver here?") reads exactly like a customer's statement once it is cut out of the call. In the example review, the first automated pass made exactly this mistake on 30 records. A check caught it. Without a record of who said what, nothing would have.

It cannot connect a sentence to an outcome. Knowing that a competitor was mentioned is trivia. Knowing it was mentioned on deals that later closed lost, before it ever appeared in a CRM field, is a finding.

Paraphrase looks exactly like a quote. The risk is rarely outright invention. It is the fluent paraphrase, close enough to pass a skim and not what the customer said. In the same review, every quote was checked against its source: 97.9 percent matched exactly, 1.3 percent were faithful but joined two statements, 0.4 percent were near misses and 0.4 percent matched nothing. Without the check, those misses would sit among the rest, indistinguishable from evidence. And because summaries are written fresh each time, two people asking the same question on different days can get two versions of the same deal. That is fine for call prep. It is not fine for the number a CEO takes to the board.

> **A model can only reason over the connections someone made before the question was asked.**

Index graph or evidence graph

The obvious fix is to give the model a graph, and much of the industry has. Most graphs built for AI today are index graphs. In the best known method, GraphRAG, a language model reads every chunk of text, pulls out the things it mentions (a company, a person, a competitor, a topic) and the relationships between them, and writes a short description of each. Revenue tools are moving the same way, under names like revenue graph and context graph. It is a real improvement on

similarity search: it helps the model find the right passages and answer broad questions across a whole collection.

An index graph is built to find text. Its nodes are things a model decided were mentioned, its edges are relationships a model inferred, and its descriptions are new writing. Ask it which lost deals raised a competitor and it can point you to calls where the competitor came up. On its own, it cannot tell you who said it, whether the quote is exact, or how many lost deals raised it compared with won ones.

An **evidence graph** is built to prove things. Its basic unit is a single statement: the exact words, who said them and on which side of the deal, the call, the date and the minute, and the opportunity and outcome it belongs to. Nothing in it is paraphrased, and its links come from the records rather than from a model's guess. Because every statement sits on a deal with a known result, a question can be answered by counting across every deal, and every number opens to the sentences behind it.

	Index graph	Evidence graph
Built to	Find relevant text for a model to read	Prove and count the facts a leader acts on
Basic unit	An entity or topic a model extracted	One verbatim statement from a call
Text it holds	Model-written descriptions and summaries	The customer's exact words, checked against the transcript
Who said it	Not part of the design	Recorded on every statement: customer or seller
Link to the result	Not required	Required: statement, speaker, call, opportunity, outcome
What you get back	A fluent answer about what was found	A count across all deals, with each quote one click away
Ask again next week	The answer is written fresh and can shift	The same facts and the same count

The two are not rivals, and neither are the tools you already run: conversation platforms, CRM copilots and dashboards each do their own job well inside their own boundary. An evidence graph can use an index to find candidate passages. The difference is what is allowed to count. In an index graph, what the search found is already the answer. In an evidence graph it is only a lead until it passes the four tests below. When a vendor describes a graph, the fair questions are what its nodes are made of, how it matches people across the CRM and the calls, and whether any number it gives you can be traced to a sentence. A graph that skips provenance cannot be turned into an evidence graph later, because the link to the source was never kept.

Fields are not evidence

A graph only connects what it is given. Build it from CRM fields and it inherits their gaps. Consider the questions a forecast call turns on. Is there a confirmed budget? Who signs? When is the decision, and who else is in it? The answers come from fields the rep entered, usually after the conversation and often to satisfy a stage requirement. Meanwhile the customer stated the budget, named the competitor and described the approval path on a recorded call. Every deal has two records, and the one leaders inspect is the one that cannot be audited.

This is not a hygiene problem that stricter required fields will fix. In one 24 month review of a company's sales calls and CRM data, the loss reason, competitor, pain point and volume fields were empty on every exported row. Three more fields held one constant value on every row, so a fill rate check passed them while they carried no information. That was not a careless team. It is what happens when fields exist to move a deal between stages and nothing downstream reads them. Every competitive and loss finding in that review had to come from the calls.

> **The CRM tells you what someone recorded. The call tells you what the customer said. Only one of them can be checked.**

The evidence standard

Four tests separate a fact a leader can act on from a plausible sentence.

> **Sourced**
>
> The fact links to a verbatim customer quote with the call, the date and the minute. A fact with no sentence behind it does not enter the graph as a fact.

Verified

The quote is checked against the transcript before it is counted. Exact matches pass. Joined statements are marked as joined. Near misses are reported separately and never counted, because close enough is how paraphrase gets laundered into evidence.

Attributed

Who said it decides what it can prove. A customer's words can establish a fact; a rep's words can support one but never establish it.

Labelled

Stated facts, inferences and gaps are kept apart and shown as such. Where the evidence is thin, the answer says so instead of filling the space with something fluent.

None of these tests is new. Auditors, researchers and good journalists apply versions of all four. What is new is applying them to every recorded call a company has, in 60 days rather than a year of analysts reading transcripts.

What one fact looks like

Here is a single deal from that 24 month review, with names changed. The CRM recorded the use case as operations monitoring. The calls say otherwise, and they also show the evidence arguing with itself, which is exactly what a summary would smooth over.

Arvane Motors · $640K · Closed Won

Electric vehicle manufacturer · 125 day cycle · 5 calls on the deal · names and figures changed

CRM use case	FIELD Operations monitoring
Customer, Oct	STATED *"I'm going to forward the plant logs into Kelvara, that's the security copy, correct?"* `customer · Oct 2025 · 05:01`
Customer, Sep	STATED *"We, we never talked about security."* `customer · Sep 2025 · 38:59`
Reading	INFERRED Security entered the deal between September and October. The October statement is what makes this a security deal; the September one stays on the record rather than being smoothed away.
What nearly killed it	STATED *"I wonder how we went through the professional services deploy and both those things got missed."* `call · Sep 2025 · 71:07`

One deal from an illustrative engagement output. Every quote carries its source; the inference is labelled as one.

Multiply that card across every opportunity and the questions a revenue leader can ask change character. In the same review, 24 opportunities the first pass could not resolve were settled only by going back to calls a keyword screen had dismissed; four were already closed and won. A separate screen for risky claims by reps returned 992 hits, and reading each in its sentence left 488. A screen that loses half of itself on reading is a finding about the screen, and about every keyword tracker built the same way.

What it changes for a new revenue leader

The value is sharpest in a new revenue leader's first ninety days, when they inherit a pipeline, a forecast and a set of beliefs about why deals are won, all built from the same records.

The inherited pipeline can be inspected on evidence. Before the first forecast call, a leader can see which commit deals have a budget the customer stated, a named decision maker and a dated next step, and which have those only in a field.

Win and loss trace to the conversations that decided them. The trailing 12 months can be read for what customers actually raised: the competitor that appeared in calls before any field, the objection that loses more than it wins, the request the product could not meet.

CRM gaps and coaching start from the call. A missing budget or competitor surfaces with the quote attached, and a manager can point to the minute and the sentence instead of an opinion.

For the CEO, the same graph answers a different question: whether the plan the new leader brings in ninety days rests on what customers said, or on the same reporting that produced the results the leader was hired to change.

> **The plan you get in ninety days will be built from the records you have today. It is worth knowing what they can prove.**

Where Rev-Lens fits

Rev-Lens is a services firm, and a new one. In a 60 day, founder-delivered engagement we build an evidence graph from your CRM and your call recordings inside your own infrastructure, apply the four tests above, and run the analysis your planning needs: where you win, where you lose and to whom, what customers ask for, and which fields can be trusted. The graph and the analysis are yours. Your team runs it after the engagement, nothing is licensed from us, and no fee continues.

The first conversation is a review of your own trailing 12 months of won and lost deals. It can conclude that the evidence does not support a change to the plan. That is a useful answer, and you keep the analysis either way.

The figures cited come from an illustrative Rev-Lens engagement output in which company, product and personal names have been changed and counts shifted. Rev-Lens · rev-lens.ai · sales@rev-lens.ai