CUSTOMER LIFECYCLE MANAGEMENT · BOOTH CRO PROGRAM

Confidently Wrong

An executive playbook for knowing when to trust AI across the revenue lifecycle

Why GenAI gives persuasive but unreliable answers about deals and renewals + a practical verification architecture.

DAVID LAHEY · DAVIDLAHEY.COM · OCTOBER 15, 2026

The most dangerous AI answer isn't obviously wrong.

It's 95% right, beautifully written, and wrong about the **5%** that determines your outcome.

GenAI produces answers. It does not inherently establish truth.

THE COST OF THE 5%

The 5% isn't random. It lands on your biggest decisions.

LOW STAKES

Call summaries,
follow-up emails

COST WHEN WRONG

Minutes

The error surfaces as you use it.

›

OPERATIONAL DECISIONS

Health scores, deal priority,
coverage plans

COST WHEN WRONG

Quarters of effort

Resources aim at the wrong fix.

›

STRATEGIC DECISIONS

Forecast commit, renewal pricing,
NRR to the board

COST WHEN WRONG

ARR + credibility

Committed decisions that compound.

A 5% error rate is not 5% risk. When plausible-but-wrong answers reach judgment-heavy questions, the cost can compound.

THE REVENUE LIFECYCLE

Where the 5% hides across the customer lifecycle

CASE A · FORECAST

CASE B · CHURN

ACQUIRE

“Which deals will close this quarter?”

DECISION IT FEEDS

Forecast commit

A missed number on the board call.

ONBOARD

“Is this rollout on track?”

DECISION IT FEEDS

Time-to-value

Risk surfaces after the window to fix it.

ADOPT

“How healthy is this account?”

DECISION IT FEEDS

Health score + CSM coverage

Attention goes to the wrong accounts.

EXPAND

“Who is ready to buy more?”

DECISION IT FEEDS

Expansion pipeline

Reps chase upsell that isn’t there.

RENEW

“Why are we at risk of losing them?”

DECISION IT FEEDS

GRR / NRR forecast

The wrong save play, too late.

Each stage hands a decision to the next. An unverified answer early in the funnel becomes the input to everything downstream.

THE FAILURE MODE · BOTH ENDS OF THE FUNNEL

Coherent, but unverified

CASE A · ACQUIRE

“Why will this deal close this quarter?”

THE SYSTEM CONFIDENTLY CONCLUDES

- Champion is engaged
- Budget is approved
- Competitor has been eliminated
- Legal is reviewing the contract

SOURCE CHECK FINDS

- Champion’s last reply was six weeks ago
- Budget is for next fiscal year
- Competitor is still in the evaluation
- Legal hasn’t received the contract

CASE B · RENEW

“Why did this strategic customer churn?”

THE SYSTEM CONFIDENTLY CONCLUDES

- Product adoption declined
- The executive sponsor departed
- Competitor activity increased
- Renewal was in 90 days

SOURCE CHECK FINDS

- Adoption remained stable
- Sponsor change was 18 months earlier
- Competitor signal came from one comment
- Renewal was nine months away

The output is coherent. The evidence is not.

WHY IT HAPPENS

Plausibility meets confirmation bias

Plausibility

A coherent answer can be generated without independently establishing the claims inside it.

Sycophancy

Responses can shift toward the assumptions or preferences expressed in the interaction.

WHEN WE PROMPT

“This deal is going to close this quarter. Look at these notes and tell me why.”

INSTEAD OF

“What does the evidence say about this deal’s timing?”

The first prompt supplies a question, a hypothesis, and a preferred conclusion. The model now has an anchor.

AI Sycophancy × Confirmation Bias → Persuasive Error

WHY SMART TEAMS STILL GET TRAPPED

The Confidence Trap

THE MODEL

Plausibility

Looks right, like a clean deal summary

Sycophancy

Agrees with the rep's framing

Lost nuance

Drops the objection from call three

THE CONTEXT

Ambiguity

"At risk" was never defined

Missing evidence

Blank CRM fields filled by inference

Misplaced emphasis

Loudest notes ≠ most important

US

Benchmark bias

Averages hide your top-10 accounts

Confirmation bias

Commit bias: we test what we hope

Fluency bias

Readable feels like a forecast

Eloquence ≠ Evidence. Failure emerges from the model, the context, and us — and the last three we can catch ourselves doing.

TWO FIXES THAT CAN FAIL

Two fixes that feel right — and still fail

BAD FIX #1

Ask AI to check itself

Like asking the rep to inspect their own forecast. The model carries the same framing and may polish its reasoning rather than overturn it.

Self-review can produce consistency without correctness.

BAD FIX #2

Just add more AI models

Like a deal review where everyone read the same rep notes. Independent models help, but overlapping training data can share blind spots.

Agreement can be insight — or shared error. From the outside they look identical.

Both can add signal. Neither is the same thing as independent evidence.

What verification actually requires

One model → **answer** Multiple independent models → **challenge** Independent evidence → **verification**

1 PRODUCE

AI Role: Analyst

Turn the answer into explicit claims; separate what is known from what is inferred.

Facts · Assumptions · Inferences

IN YOUR DEAL REVIEW
The rep's call

2 ATTACK

AI Role: Adversary

Challenge the claims without inheriting the first model's rationale.

Evidence gaps · Hidden assumptions · Alternatives

IN YOUR DEAL REVIEW
Manager inspection

3 VERIFY

Role: Verifier

Test material claims against evidence that exists independently of the model.

Sources · Tools · Deterministic checks

IN YOUR DEAL REVIEW
CRM, telemetry, billing

4 SYNTHESIZE

Role: Arbiter

Reconcile the competing outputs with the evidence; classify each conclusion.

Reconcile · Classify · Escalate

IN YOUR DEAL REVIEW
The CRO's forecast call

Model agreement is not evidence. Different vendors ≠ independent reasoning.

LIVE DEMONSTRATION

Multi-Model Adversarial Verification in practice.

1 PRODUCE

AI Role: Analyst

PROMPT

```
You’re an expert in [ROLE]. Help me build a ____ for ____. The primary goal is to [GOAL1]. The secondary goal is to [GOAL2]. Format the output into ____ for a v1. Ask me any clarifying questions.
```

2 ATTACK

AI Role: Adversary

PROMPT

Attached is v1 of ____ for ____. The primary goal is to ____. The secondary goal is to ____. **Rate it on a scale of 0-10, 10 being perfect to the two goals. Show your reasoning and suggest any needed improvements.** Ask me any clarifying questions.

3 VERIFY

Role: Verifier

PROMPT

Is the attached v2 better or worse than or the v1 we created on the last turn? Show your reasoning and quantify the differences. Also rate v2 on a scale of 0-10, 10 being perfect to the two goals. Show your reasoning and suggest any needed improvements.

4 SYNTHESIZE

Role: Arbiter

PROMPT

I agree/disagree with ____. Update v2 of the document and create a v3 based on ____. Highlight any changes in v3 so they standout and are obvious.

Which model was most confident? Irrelevant. **What determined the conclusion?** Evidence.

MONDAY MORNING

Five rules for high-stakes AI

1

Separate fact from inference. Make the model label both — every claim carries its status.
In CLM: Health scores and deal summaries show which drivers are measured and which are inferred.

2

Don't unnecessarily reveal your preferred answer. Reduce anchoring and sycophancy before the model responds.
In CLM: Ask “what does the evidence say about this renewal?” — not “why are we losing it?”

3

Engineer disagreement. Independent models in different roles — and different vendors don't guarantee independence.
In CLM: One model builds the rep’s case; another is tasked with breaking it before the forecast call.

4

Verify against source evidence. Models are not evidence for one another.
In CLM: Telemetry, CRM history, billing and contracts — not another model’s summary.

5

Scale rigor with consequence. Bigger decisions get stronger verification.
In CLM: A call summary gets a glance. A commit, a renewal price or the NRR number gets full MMAV.

CONCLUSION

Confidence is not evidence.
Consensus is not evidence.
Independent evidence earns trust.

MULTI-MODEL ADVERSARIAL VERIFICATION

Multiple models create challenge. Independent evidence creates verification.

APPENDIX

Course materials

Reference material for participants, not presented in the session

A1 Evidence map · A2 Worksheet · A3 Prompt templates · A4 Churn example · A5 Case A walkthrough

APPENDIX A1 · LIFECYCLE EVIDENCE MAP

What to verify against, stage by stage

STAGE	WHAT THE AI CLAIMS	VERIFY AGAINST · INDEPENDENT OF THE MODEL
ACQUIRE	*“This deal will close this quarter.”*	Last buyer reply · mutual action plan dates · contract and legal status
ONBOARD	*“Implementation is on track.”*	Go-live milestones · provisioning and integration logs
ADOPT	*“This account is healthy.”*	Product telemetry by seat and feature · usage vs. entitlement
EXPAND	*“This account is ready to buy more.”*	Usage at or over entitlement · budget cycle · new-team activation
RENEW	*“They’re leaving because adoption fell.”*	Telemetry trend · sponsor timeline in CRM · renewal quote vs. budget

APPENDIX A2 · EXERCISE

Lifecycle verification worksheet

Pick one high-stakes decision from your lifecycle. List what the AI claims, label each claim, name the independent source and who checks it, then record what the evidence says.

Decision: _____ **Stage:** Acquire / Onboard / Adopt / Expand / Renew **Tier:** Low / Operational / Strategic

#	What the AI claims	Fact, assumption or inference?	Independent source	Who checks	Verified / Unsupported / Uncertain
1					
2					
3					
4					
5					

MMAV prompt templates

1 PRODUCE

AI Role: Analyst

PROMPT

```
You're an expert in [ROLE]. Help me build a ____ for ____. The primary goal is to [GOAL1]. The secondary goal is to [GOAL2]. Format the output into ____ for a v1. Ask me any clarifying questions.
```

2 ATTACK

AI Role: Adversary

PROMPT

Attached is v1 of ____ for ____. The primary goal is to ____. The secondary goal is to ____. **Rate it on a scale of 0-10, 10 being perfect to the two goals. Show your reasoning and suggest any needed improvements.** Ask me any clarifying questions.

3 VERIFY

Role: Verifier

PROMPT

Is the attached v2 better or worse than or the v1 we created on the last turn? Show your reasoning and quantify the differences. Also rate v2 on a scale of 0-10, 10 being perfect to the two goals. Show your reasoning and suggest any needed improvements.

4 SYNTHESIZE

Role: Arbiter

PROMPT

I agree/disagree with ____. Update v2 of the document and create a v3 based on ____. Highlight any changes in v3 so they standout and are obvious.

Fill in the blanks for your own decision. Run each role in a fresh session or a different model.

APPENDIX A4 · CHURN EXAMPLE

Multi-Model Adversarial Verification in practice.

1 PRODUCE

AI Role: Analyst

PROMPT

```
You're a revenue analyst. Using the attached account notes, CRM history and support tickets, explain why [Account] did not renew. State each hypothesis as an explicit claim and label it fact, assumption or inference.
```

2 ATTACK

AI Role: Adversary

PROMPT

```
You did not write the attached analysis. For each claim, state what evidence would prove it, what would disprove it, and the strongest alternative explanation. Do not propose a better story.
```

3 VERIFY

Role: Verifier

PROMPT

```
Test each claim only against the attached sources: product telemetry, CRM sponsor history, and the renewal quote and emails. For each claim, cite the source and mark it supported, contradicted or not found.
```

4 SYNTHESIZE

Role: Arbiter

PROMPT

```
Reconcile the analyst's claims, the adversary's objections and the source checks. Classify each hypothesis VERIFIED, UNSUPPORTED or UNCERTAIN, and flag anything that needs a human decision.
```

Which model was most confident? Irrelevant. **What determined the conclusion?** Evidence.

The forecast commit, run through MMAV

1 PRODUCE

AI Role: Analyst

CLAIMS

- Champion is engaged
- Budget is approved
- Competitor has been eliminated
- Legal is reviewing the contract
- The buyer has a real need

2 ATTACK

AI Role: Adversary

OBJECTIONS

- When did the champion last reply?
- Approved for which fiscal year?
- Who said the competitor is out?
- Has legal received the paper?
- What shows the need is real?

3 VERIFY

Role: Verifier

SOURCE CHECKS

- Email log → last reply six weeks ago
- Budget memo → next fiscal year
- Call transcript → competitor still evaluating
- Legal queue → no contract received
- Discovery notes + pilot usage → need confirmed

4 SYNTHESIZE

Role: Arbiter

VERDICT

- Champion engaged — **UNSUPPORTED**
- Budget approved — **UNSUPPORTED**
- Competitor out — **UNSUPPORTED**
- Legal review — **UNSUPPORTED**
- Real need — **VERIFIED**

→ Move from Commit to Pipeline; re-engage the economic buyer.

Which model was most confident? Irrelevant. **What moved the deal out of commit?** Evidence.

CONTINUE THE CONVERSATION

David Lahey

Advisor to growth-stage B2B CEOs.

SIX EXITS AS FOUNDER & OPERATOR · AI DISRUPTION & GTM TRANSFORMATION

davidlahey.com · linkedin.com/in/davidlahey