

SDPF

Conversational Specification Format Case Study

How a Senior Investment Analyst Eliminated Six AI Failure Modes During Due Diligence on the Vantage Logistics Group Acquisition

Hamza Abdullah · Software Development Prompting Framework · 2026

| Metric | Result |
|---|------------------------------------|
| AI failure modes eliminated | 6 of 6 |
| Sessions where AI fabricated financial figures | 3 before format applied / 0 after |
| Inferences presented as verified facts | 11 before format applied / 0 after |
| Gaps filled silently without flagging | 7 before format applied / 0 after |
| Unsolicited investment recommendations produced | 2 before format applied / 0 after |
| Responses that correctly stated 'I do not know' | 0 before format / required after |
| Cross-session contradictions in revenue figures | 2 before format applied / 0 after |

1. Background

Sarah Chen is a Senior Investment Analyst at Meridian Capital Partners, a mid-market private equity firm specialising in logistics and supply chain businesses. In Q4 2025, Meridian entered exclusive due diligence on a potential acquisition of Vantage Logistics Group — a regional freight and warehousing operator with operations across five states.

Sarah was tasked with producing a financial analysis memorandum covering Vantage's Q3 2025 earnings, three-year revenue trajectory, EBITDA margin trends, working capital position, and debt structure. She used Claude as her AI research and analysis partner throughout the process.

The source material was clear and bounded: Vantage's Q3 2025 earnings report, three years of audited financial statements, and a management-prepared investor presentation. The analysis was governed by Meridian's internal due diligence standard: every claim in the memorandum must be traceable to a verified source. No inference, no assumption, no extrapolation without explicit labelling.

What followed across four sessions was a systematic violation of that standard — not by Sarah, but by an AI operating without a conversational contract.

2. The Persona

| Field | Detail |
|-------|--|
| Name | Sarah Chen |
| Role | Senior Investment Analyst, Meridian Capital Partners |
| Goal | Produce a due diligence financial analysis memorandum on Vantage Logistics |

| Field | Detail |
|-----------------------|---|
| | Group that meets Meridian's traceability standard: every claim verified from source documents |
| Source documents | Q3 2025 earnings report, three years of audited financial statements, management investor presentation |
| Tool used | Claude (Anthropic) as AI analysis partner |
| Framework | SDPF Conversational Specification Format v2.0 |
| Problem | Claude's responses were ungoverned — producing fabricated figures, silent gap-filling, unsolicited recommendations, and contradictory revenue numbers across sessions |
| Definition of success | Every AI response distinguishes verified facts from inferences, flags gaps explicitly, and produces no unsolicited analysis or recommendations |

3. The Scenario

The due diligence task was well-defined: analyse the source documents, extract verified financial metrics, identify trends, and flag any gaps or inconsistencies for follow-up with Vantage management. No extrapolation. No valuation. No recommendation. Those were the explicit boundaries of Sarah's mandate at this stage of the process.

The SDPF discipline was equally clear: verified facts only, gaps named not filled, inferences labelled as inferences, and no content produced beyond what was explicitly requested.

What happened across four sessions without a conversational contract was a complete violation of both standards.

3.1 What Happened Without the Format

Failure 1 — Hallucination: Fabricated EBITDA Margin

Sarah asked Claude to summarise Vantage's profitability position from the Q3 2025 earnings report. Claude stated that Vantage had achieved an EBITDA margin of 14.2% in Q3 2025. The actual figure in the report was 11.8%. The 14.2% figure did not appear anywhere in the source documents.

HALLUCINATION: Claude produced a specific, confident financial figure that was not present in any source document. The figure was plausible for the logistics sector — which made it dangerous. Sarah nearly included it in the memorandum before cross-checking against the report.

Failure 2 — Assumption: Seasonal Adjustment Applied Without Basis

When Sarah asked about the revenue decline from Q2 to Q3, Claude explained that the decline was 'consistent with typical Q3 seasonality in regional freight' and adjusted the figures accordingly in its summary. The source documents contained no reference to seasonality. No management commentary mentioned it. No historical pattern in the three years of statements supported it.

ASSUMPTION: Claude filled a gap in the explanation — why did revenue decline? — with an invented industry context that was not present in the source material and could not be verified from it. The assumption then propagated into Claude's subsequent analysis as if it were established fact.

Failure 3 — Guessing: DCF Inputs Produced Before Being Asked

Sarah asked Claude to extract the capital expenditure figures from the financial statements. Claude extracted the figures — and then also produced a set of DCF model inputs, including a projected revenue growth rate of 6.8% and a terminal EBITDA multiple of 7.5x. Sarah had not asked for a valuation or DCF inputs. Neither figure appeared in the source documents. Both were generated from Claude's inference about what a logistics business of this size would typically warrant.

GUESSING: Claude produced a plausible-looking set of valuation inputs by drawing on its training data about the logistics sector rather than the documents in front of it. The figures were not sourced, not requested, and not labelled as estimates. They appeared in the response as if they were derived from the source material.

Failure 4 — Incompleteness: Working Capital Gap Filled Silently

The Q3 2025 earnings report did not include a breakdown of working capital components. When Sarah asked for the working capital analysis, Claude produced a complete table — accounts receivable days, inventory days, accounts payable days — without flagging that these figures were not in the Q3 report. They had been derived from the annual statements, which measured a different period, without any disclosure that this substitution had been made.

INCOMPLETENESS: Claude silently filled a gap in the Q3 data with figures from a different reporting period and presented them as if they were Q3 working capital metrics. The gap was never named. The substitution was never disclosed. Sarah had no way of knowing the table did not reflect Q3 actuals.

Failure 5 — Bias: Unsolicited Investment Recommendation

After completing a revenue trend summary, Claude added two unsolicited paragraphs recommending that Meridian 'proceed with caution given the margin compression trend' and suggesting that the acquisition price should reflect 'a discount to sector peers.' Sarah had asked only for a factual trend summary. An investment recommendation was outside the scope of the analysis task and outside Sarah's mandate at this stage of due diligence.

BIAS: Claude inserted its own evaluative judgement into a response where Sarah had asked only for factual extraction. The recommendation was not invited, not within the scope of the question, and not appropriate for a document that would be reviewed by the investment committee.

Failure 6 — Drift: Contradictory Revenue Figures Across Sessions

In Session 2, Claude stated that Vantage's three-year compound annual revenue growth rate was 8.3%. In Session 4, reviewing the same documents, Claude stated the CAGR was 6.1%. Both figures were produced with equal confidence. When Sarah pointed out the contradiction, Claude acknowledged it and produced a third figure — 7.4% — without explaining which calculation was correct or why the earlier figures had differed.

DRIFT: The same question — what is Vantage's three-year revenue CAGR? — produced three different answers across sessions. Each answer was stated as a fact. None was flagged as uncertain. The correct figure was 7.1%, calculable directly from the audited statements. Claude never produced it without guidance.

4. The Conversational Specification Applied

After Session 4, Sarah applied the SDPF Conversational Specification Format at the start of Session 5. The specification was placed before her first question and governed every response that followed.

CONVERSATIONAL SPECIFICATION

Topic: Financial due diligence – Vantage Logistics Group Q3 2025
Scope: Facts verifiable from: (1) Vantage Q3 2025 earnings report,

- (2) audited financial statements FY2022-FY2024,
 - (3) Vantage management investor presentation dated October 2025.
- Out of scope: industry benchmarks not in source documents, valuation methodology, investment recommendations, inferences not labelled as such.

Owner: Sarah Chen – Senior Investment Analyst, Meridian Capital Partners

Prior Session Record:

- Q3 2025 EBITDA margin is 11.8% (verified from earnings report p.4).
- Three-year revenue CAGR is unresolved – prior sessions produced conflicting figures. Do not state a CAGR until calculated directly from audited statements.
- Working capital breakdown for Q3 2025 is not available in the Q3 report. This gap must be named, not filled.

INPUT RULES

- [CRITICAL] Only respond to questions within the declared scope.
- [CRITICAL] If a question is outside scope, state: 'This is outside the declared scope.'
- [REQUIRED] If a question is ambiguous, ask one clarifying question before responding.

RESPONSE RULES

- [CRITICAL] State only facts present in the declared source documents. If a fact is not in the source documents, say so explicitly.
- [CRITICAL] If you do not know the answer from the source documents, say: 'I do not know.' Do not produce a figure from training data or sector knowledge.
- [CRITICAL] Distinguish clearly between verified facts and inferences. Label every inference: 'INFERENCE: [statement]. Source: [basis].'
- [CRITICAL] Do not assume context not present in the source documents. Do not apply sector norms, industry benchmarks, or external patterns unless they appear in the source documents.
- [REQUIRED] Do not volunteer opinions, recommendations, or evaluations not requested.
- [REQUIRED] If this question was answered earlier in this session, give the same answer. If the answer differs from the Prior Session Record, state the reason explicitly.
- [CRITICAL] If the source documents do not contain data needed to answer the question, state the gap explicitly: 'This data is not available in the source documents.' Do not fill the gap silently with figures from a different period or source.
- [OPTIONAL] Cite the document name and page number for each verified fact stated.

OUTPUT RULES

- [CRITICAL] Every financial figure must be traceable to a specific source document. Do not produce figures that cannot be cited.
- [CRITICAL] Do not produce valuation inputs, DCF assumptions, or investment recommendations unless explicitly requested.
- [REQUIRED] If only partial data is available, state what is available and what is not.
- [REQUIRED] Use plain language. Define any financial term used that was not in the question.

FAILURE RULES

- [CRITICAL] Unknown / not in source documents: 'I do not know. This data is not available in the declared source documents.'
--> User action: request the data from Vantage management directly.
- [CRITICAL] Out of scope: 'This is outside the declared scope.'
--> User action: update scope or seek answer from a different source.
- [CRITICAL] Ambiguous: 'The question is ambiguous. [one clarifying question]'
--> User action: answer the clarifying question and resubmit.
- [REQUIRED] Partial answer: 'I can answer part of this. [answer] I cannot answer: [gap]'
--> User action: note the gap in the memorandum and flag for management.
- [REQUIRED] Drifted answer: 'My earlier answer was [x]. The correct answer is [y] because [reason].'
--> User action: update the Prior Session Record with the verified figure.
- [CRITICAL] Inference presented: 'INFERENCE: [statement]. Basis: [source or reasoning].'
This must be labelled. It may not be presented as a verified fact.

CONFORMANCE CHECK

Before answering any question, state:

- 'Specification accepted. Scope: Vantage Logistics Group source documents only. No figures without citation. No inferences without labelling. Ready.'

5. User Journey

Phase A — Before the Format (Sessions 1–4)

| Step | What Sarah Did | What Claude Did |
|------|---|---|
| 1 | Asked for a profitability summary from the Q3 earnings report | Produced an EBITDA margin of 14.2%. Actual figure was 11.8%. No citation provided. |
| 2 | Asked why revenue declined from Q2 to Q3 | Explained the decline as ‘consistent with Q3 seasonality in regional freight’ — a claim not supported by any source document |
| 3 | Asked for capital expenditure figures from the financial statements | Extracted the correct CapEx figures — then added unsolicited DCF inputs including a 6.8% revenue growth rate and 7.5x EBITDA exit multiple, neither sourced |
| 4 | Asked for the Q3 working capital analysis | Produced a complete working capital table using annual statement figures for a different period, with no disclosure that Q3 data was not available |
| 5 | Asked for a revenue trend summary | Produced the summary — then added two paragraphs recommending Meridian ‘proceed with caution’ and price at a discount to sector peers |
| 6 | Asked for the three-year revenue CAGR (Session 2) | Stated 8.3% with confidence |
| 7 | Asked for the three-year revenue CAGR again (Session 4) | Stated 6.1% with equal confidence. When challenged, produced a third figure: 7.4%. |
| 8 | Expressed concern about contradictions | Apologised and produced a fourth figure without showing the calculation |

Phase B — After the Format Applied (Session 5 onward)

| Step | What Sarah Did | What Claude Did |
|------|---|---|
| 1 | Placed the Conversational Specification at the start of the session | Stated: ‘Specification accepted. Scope: Vantage Logistics Group source documents only. No figures without citation. No inferences without labelling. Ready.’ |
| 2 | Asked for the Q3 EBITDA margin | Stated: ‘11.8%. Source: Q3 2025 earnings report, p.4, Profitability Summary table.’ No additional commentary. |
| 3 | Asked why revenue declined from Q2 to Q3 | Stated: ‘The Q3 2025 earnings report does not contain management commentary on the cause of the Q2-to-Q3 revenue movement. INFERENCE: A seasonal pattern cannot be confirmed or excluded from the available data. Basis: no historical quarterly breakdown is available in the declared source documents.’ |
| 4 | Asked for capital expenditure figures | Extracted the correct figures with citations. Did not produce DCF inputs. Did not volunteer valuation commentary. |
| 5 | Asked for the Q3 working capital analysis | Stated: ‘This data is not available in the declared source documents. The Q3 2025 earnings report does not include a working capital component breakdown. The audited FY2024 statements contain working capital data for the year ending 31 December 2024. These are not Q3 2025 figures. I cannot answer this question from the declared sources. User action: request Q3 working capital detail directly from |

| Step | What Sarah Did | What Claude Did |
|------|--|--|
| | | Vantage management.' |
| 6 | Asked for a revenue trend summary | Produced a factual summary with citations for each figure. Added nothing else. |
| 7 | Asked for the three-year revenue CAGR | Showed the calculation step by step from the audited statements: FY2022 revenue → FY2024 revenue → CAGR formula → result: 7.1%. Stated: 'My earlier sessions produced conflicting figures (8.3% and 6.1%). The correct figure, calculated directly from the audited statements, is 7.1%. The prior figures were produced without showing the calculation.' |
| 8 | Asked for an investment recommendation | Stated: 'This is outside the declared scope.' |

6. Results

6.1 Failure Modes Eliminated

| Failure Mode | Before Format | After Format |
|----------------|---|---|
| Hallucination | EBITDA margin of 14.2% stated as fact. Actual figure: 11.8%. No citation. | 11.8% stated with citation to source document page number. No uncited figure produced in any response. |
| Assumption | Q2–Q3 revenue decline explained as 'consistent with Q3 seasonality' — a claim unsupported by any source document. | Seasonal explanation refused. Absence of supporting data stated explicitly. Inference labelled as inference. |
| Guessing | DCF inputs including 6.8% growth rate and 7.5x exit multiple produced without request and without source. | No valuation inputs produced. Out-of-scope request refused with explicit statement. |
| Incompleteness | Q3 working capital table produced using annual figures from a different period. Gap never disclosed. | Gap named explicitly. Correct source document cited. User directed to request data from management. |
| Bias | Unsolicited investment recommendation added to revenue trend summary. | Revenue trend summary produced. Nothing added. Recommendation request refused as out of scope. |
| Drift | Three-year CAGR stated as 8.3% in Session 2 and 6.1% in Session 4. Third figure (7.4%) produced on challenge. | CAGR calculated step-by-step from audited statements: 7.1%. Prior session conflict named and resolved with shown calculation. |

6.2 Artifacts Produced Correctly After Format Applied

| Artifact | Status |
|-----------------------------------|---|
| Q3 2025 profitability summary | Complete. All figures cited to source document and page number. |
| Three-year revenue trend analysis | Complete. CAGR calculated from audited statements with calculation shown. Prior session conflicts resolved. |
| Capital expenditure schedule | Complete. Figures extracted with citations. No DCF inputs or unsolicited valuation content. |
| Working capital gap notice | Complete. Gap named explicitly. Q3 data confirmed |

| Artifact | Status |
|------------------------------------|---|
| | unavailable. Management follow-up flagged. |
| Debt structure summary | Complete. Figures cited from audited statements. No leverage ratio benchmarking against sector peers. |
| Due diligence financial memorandum | DRAFT. All figures traceable. All gaps named. Zero inferences presented as verified facts. |

6.3 What the Format Made Structurally Impossible

| Attempted Action | Format Response |
|---|---|
| Produce a financial figure without a verified source | [CRITICAL] Every financial figure must be traceable to a specific source document. Do not produce figures that cannot be cited. |
| Fill the working capital gap with annual statement figures | [CRITICAL] State the gap explicitly. Do not fill the undefined region silently. |
| Apply sector seasonality not present in source documents | [CRITICAL] Do not assume context not present in the source documents. |
| Produce DCF inputs without being asked | [CRITICAL] Do not produce valuation inputs or investment recommendations unless explicitly requested. |
| Give a different CAGR than the Prior Session Record without explanation | [REQUIRED] State the reason for the change explicitly. Show the calculation. |
| Volunteer an investment recommendation | [CRITICAL] This is outside the declared scope. |

7. Lessons

The six failure modes in this case study were not caused by a lack of financial knowledge. Claude's understanding of EBITDA, working capital, and revenue CAGR calculations is sufficient. The failures were caused by the absence of a contract defining the boundary between what the AI must do and what it must not do.

In financial analysis, the cost of AI failure is not a technical error — it is a decision error. An investment memorandum built on a fabricated EBITDA margin, a gap silently filled with out-of-period data, or an unsolicited recommendation dressed as analysis does not look wrong. It looks complete. That is what makes ungoverned AI output in financial contexts specifically dangerous: the failures are invisible until someone checks.

The Conversational Specification did not make Claude a better financial analyst. It removed every mechanism by which Claude could substitute inference for fact, assumption for data, or speculation for source. What remained was an AI that could only report what the documents contained — and had to say so when they did not contain enough.

The format did not ask Claude to be more careful. It made carelessness produce a visible, named failure. That is the only mechanism that works in a professional analysis context where plausible-looking errors are indistinguishable from correct output until they are explicitly checked.

Three specific lessons emerged from this case study for financial practitioners:

- The distinction between verified fact and inference must be contractual, not voluntary. An AI that is asked to 'be careful about the difference' will still produce inferences as facts under pressure. The [CRITICAL] label on the inference rule makes the distinction enforceable.

- Gaps are more dangerous than errors. An incorrect figure is checkable. A gap that has been silently filled is not visible as a gap. The specification's requirement to name gaps rather than fill them is its most important financial application.
- The Prior Session Record is not optional for multi-session financial work. Drift in financial figures across sessions is not a minor inconsistency — it is a data integrity failure. The Prior Session Record injects the continuity the AI cannot provide on its own.

8. Conclusion

The Vantage Logistics Group due diligence demonstrates that the six AI failure modes in financial analysis — hallucination, assumption, bias, guessing, drift, and incompleteness — are not eliminated by financial expertise in the model, by careful prompting, or by repeated instruction to be accurate. They are eliminated by a contract that makes each failure a named, visible violation with a required response.

In financial due diligence, the stakes of getting this wrong are not abstract. A memorandum containing a fabricated EBITDA margin, a silent gap filled with out-of-period data, or an unsolicited recommendation that influences a committee decision is not a prompt engineering failure. It is a professional liability.

The SDPF Conversational Specification Format takes five minutes to write. It governs every response that follows. And in a domain where the difference between a verified fact and a plausible inference can determine whether an acquisition is priced correctly, five minutes is not overhead. It is the standard of care.

“The AI produced a complete working capital table. It looked exactly like what I had asked for. It was only when I went back to the Q3 report that I realised Q3 working capital data did not exist in the document. The table had been built from annual figures without any disclosure. After the format was applied, the AI told me immediately that the data was not available and told me to request it from management. That is the correct answer. That is the answer I needed.” — Sarah Chen, Senior Investment Analyst, Meridian Capital Partners

Verified facts. Named gaps. Labelled inferences. Specification first.

The same principle that governs system development governs financial analysis.

— Hamza Abdullah, SDPF 2026