The Science of Better Prompts:
Research-Backed Fixes for 5 Common Errors

Authors:

Maya Boeye, Head of AI Research, ToltIQ

Steiner Williams, Private Equity AI Researcher

Poor prompting is the silent productivity killer in PE diligence, and most don’t even realize it’s happening. A badly constructed prompt may produce results that appear adequate on the surface, but underneath, the analysis is shallow and the insights are limited.

Figuring out why a prompt isn't working can be surprisingly difficult. You might get an answer, just not the one you’re looking for, or it misses the nuance that matters, surfacing the obvious while burying what you actually need to know.

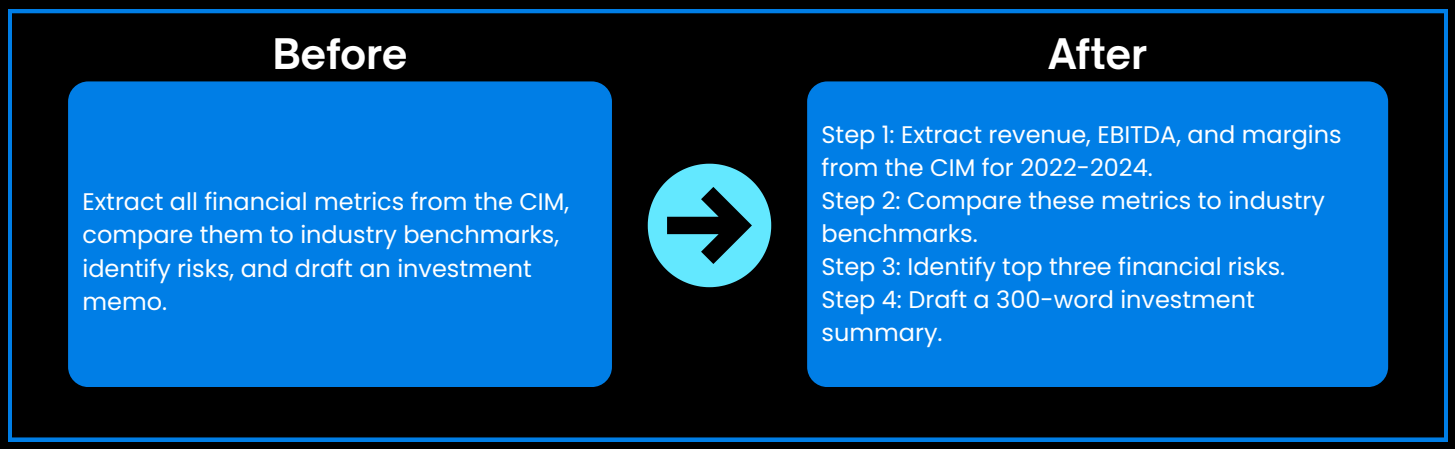
Through our work with thousands of users, we've identified the five most common patterns that limit prompt quality. For each pattern, we'll show why it fails and how to fix it, with before-and-after examples that demonstrate the difference. These patterns look reasonable on the surface. The problem is they introduce constraints that block the kind of analysis you're actually after.

1 - Overburdened, Multi-Task Prompts

When a single prompt attempts to handle extraction, analysis, comparison, and decision-making simultaneously, something has to give. Packing five or ten requests into one prompt fragments the model's attention and produces lower-quality results across the board.

In order to fix this, break the workflow into clear, sequential steps using a prompt playlist. Think: extract, then prioritize, then structure, then expand, then draft. Each prompt should optimize a single retrieval or analysis strategy.

Example



The bundled prompt produces surface-level analysis that misses critical details. When extraction, comparison, and analysis happen simultaneously, the AI prioritizes breadth over depth and each task gets lower-quality results.

Sequential prompts ensure each analytical layer is precise, producing more accurate extraction and higher-quality outputs at each stage of your diligence process. Breaking down tasks into focused steps helps, but even well-structured prompts fail when they lack the next critical element: context.

If you're on the ToltIQ platform, use prompt playlists to chain these steps together and run them with a single click. If you're working elsewhere, simply run each prompt manually in sequence.

2 - Insufficient Context and Purpose

Prompts that jump straight to "do X" without explaining a purpose, document scope, or decision can lead to generic or misaligned outputs.

In order to fix this, use a framework. A very simple one to remember is Context + Task + Output. Provide background on why this analysis matters, define the exact objective, and specify the deliverable format and length.

Example

Without context, the AI defaults to generic contract analysis. With context specifying SaaS investment criteria, it focuses on signals such as; contract duration, termination rights, expansion mechanisms, and non-standard terms that affect churn assumptions.

This is the difference between surface-level contract summaries and investment-relevant risk assessments. Structure and context sets the stage, but it's not enough on its own.

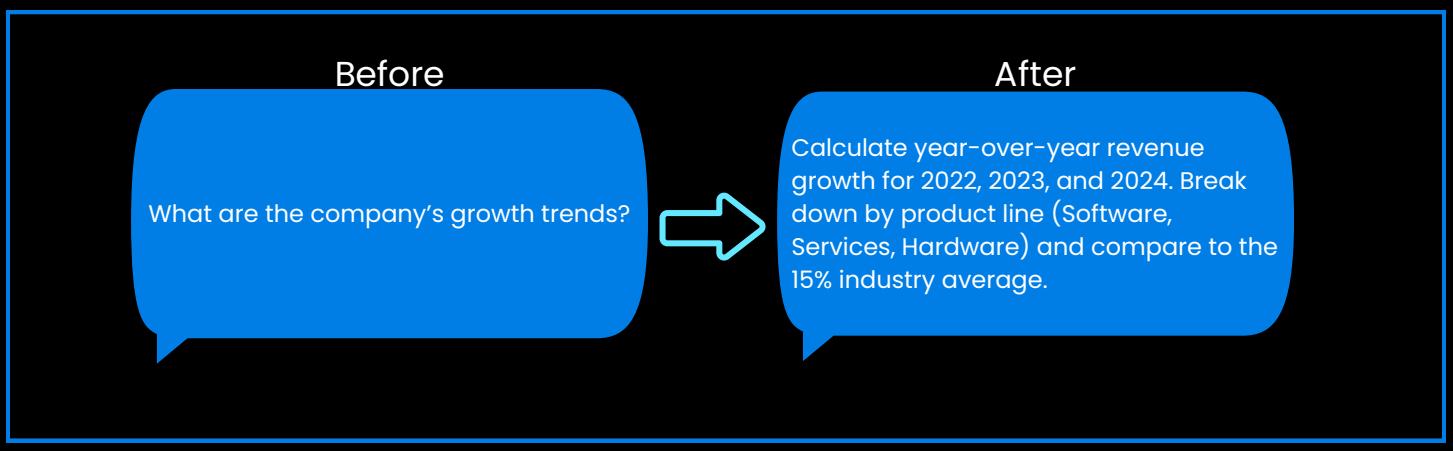
Your prompts also need precision in what you're actually asking for.

3 - Vague Asks Lacking Specificity

Requests that omit entities, metrics, definitions, or timeframes hamper precise retrieval and analysis. "Analyze the financials" is too vague to produce useful results.

In order to fix this, name specific entities, metrics, calculations, and time periods. Include comparative anchors like year-over-year or quarter-over-quarter to tighten the scope and improve precision.

Example



The vague prompt forces the AI to guess what you mean, which years, which metrics, compared to what. This produces outputs that require multiple follow-up prompts to become usable while wasting time iterating.

A specific prompt with defined metrics, timeframes, and comparative benchmarks produces decision-ready analysis on the first run, eliminating the back-and-forth that turns a 5-minute task into a 45-minute process.

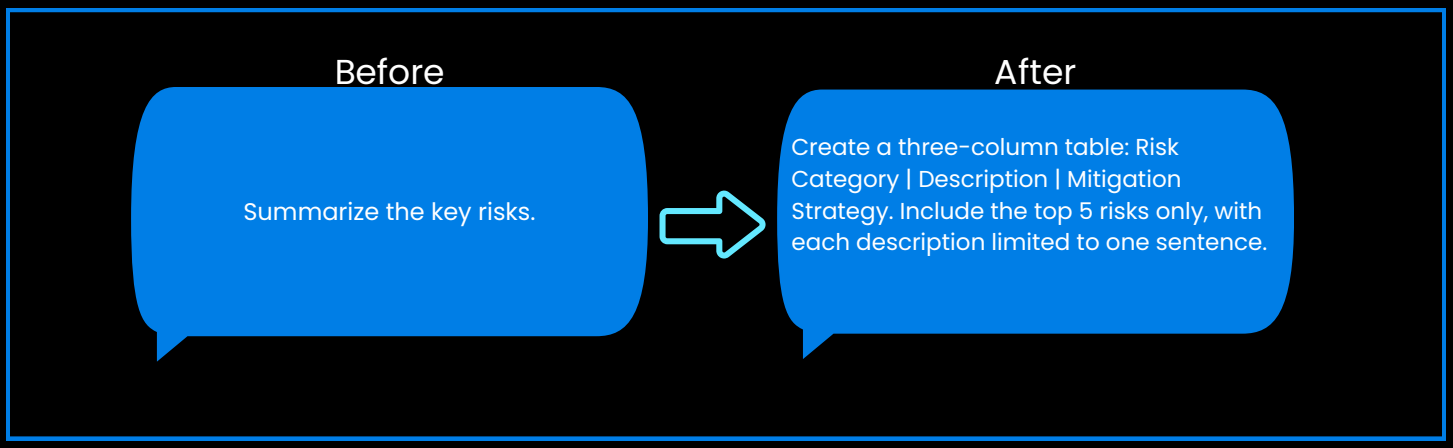
Specific inputs drive better analysis, but the output also matters.

4 - Missing Output-Format Guidance

When you don't specify whether you want bullets, tables, or memos, you increase iterations and rework.

In order to fix this, state the desired structure and length upfront. Specify table columns, bullet points, and sentence limits to produce immediately usable outputs.

Example



When you do not have format specification, you can get unpredictable outputs. Sometimes the LLM will output paragraphs, sometimes bullets, and sometimes unstructured lists that require manual reformatting.

Specifying the exact table structure, column headers, and items, produces immediately usable outputs that can be dropped directly into your memo, eliminating rework and ensuring consistency across your deal team's analyses.

The last layer that shapes both tone and depth of the response is all about defining who the AI is speaking as.

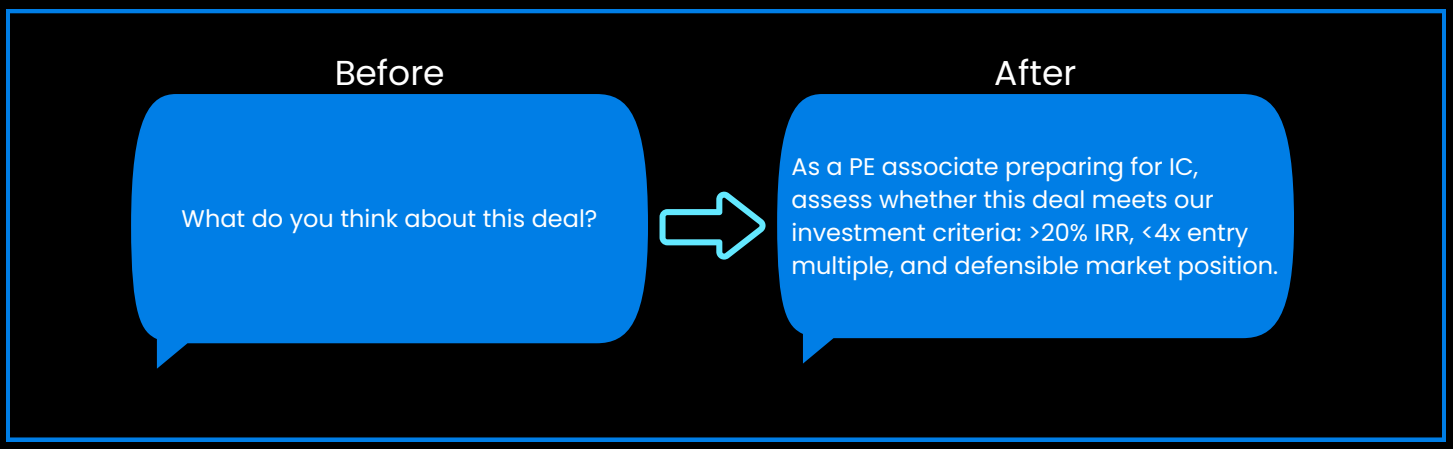
5 - Undefined Role and Audience

Without a persona or audience, tone and depth often miss professional standards.

In order to fix this, tell the LLM who it is acting as. Assign it a clear role and audience to calibrate language, depth, and structure.

If you're using ToltIQ, the system already positions the LLM as a Private Equity Analyst. If you're working elsewhere, explicitly state the role: "You are a credit analyst preparing materials for an investment committee" or "You are a PE associate summarizing findings for senior partners."

Example



When you don’t define the LLM’s role and audience, you get generic analysis that doesn't align with your specific diligence objectives.

The role-defined prompt calibrates the analysis to your exact use case, ensuring the output directly supports your current diligence task rather than requiring you to extract relevant insights from broad datasets.

Putting It All Together

The difference between a weak prompt and a refined one isn't subtle. It's the difference between spending 45 minutes iterating on a task that should take 5, between surface-level summaries, and analysis that catches the detail that kills or saves a deal.

Every badly constructed prompt compounds. One weak prompt leads to generic output, which leads to follow-up questions, which leads to more iterations, which leads to hours spent extracting insights that should have been delivered in the first run. Across a deal team, across multiple deals, the productivity loss is significant.

The fix is straightforward:

- Start with your next prompt. Use the context + task + output framework. Explain why the analysis matters, define the exact objective, and specify the format you need. If the task is complex, break it into sequential steps. If you're on ToltIQ, use prompt playlists to chain them together.
- Standardize what works. Take the prompts that worked and turn them into templates your team can reuse. Add role definitions so outputs calibrate to your specific use case.
- Track what prompts require iteration. If a prompt consistently requires follow-ups, refine it.

The goal is recognizing what's limiting your prompt quality and knowing how to fix it.

You now have a diagnostic framework: if output is shallow, check for multi-tasking. If it's generic, add context. If it requires multiple iterations, add specificity or format guidance. If the tone is off, define the role.

Start with your next diligence task. Look at the prompt you would have written yesterday, spot which pattern is limiting it, and fix it. The output will be noticeably better, and you'll spend less time refining it.