

Innovation Mindset

Prompt Engineering

Overview

In today's rapidly changing business environment, understanding technology is vitally important. Every profession, from healthcare to entertainment, has been reshaped by technological advancements, and accounting is no exception. Accounting is undergoing a paradigm shift as technology, including artificial intelligence (AI), is changing the day-to-day activities of accountants.

Central to this transformation is the rise of sophisticated large language models (LLMs), like OpenAI's ChatGPT, Google's Bard, and Meta's Llama models. These models can understand, generate, and even reason with human-like text. While their application has been widespread—from chatbots to content generation—their potential in the realm of accounting remains relatively untapped.

Knowing how to work with AI LLMs, called prompt engineering, has become an important emerging technological skill. At its core, prompt engineering is about harnessing the power of language models to perform specific tasks by providing them with clear, concise, and effective instructions, or “prompts.” Prompt engineering is knowing how to ask the right question so the AI gives the desired answer. In the context of accounting, this could mean such things as automating data entry, deriving insights from financial data, or enhancing communication with stakeholders.

As developing accountants, understanding and leveraging prompt engineering skills can equip you with a competitive edge. It can help you design and streamline processes, and open doors to developing innovative solutions. This case will briefly review the history of automation in accounting, help you learn about prompt engineering, and give you practice working with a large language model in accounting settings.

Background on Accounting and Automation

The discipline of accounting has undergone several transformations since its inception. From the abacus to the modern-day enterprise resource planning (ERP) systems, every innovation has aimed to make the accounting process more relevant, efficient, accurate, real-time and insightful. Automation in accounting isn't a novel concept. Over the past few decades, software solutions have greatly reduced the manual labor associated with tasks like data entry, reconciliation, and report generation. Tools such as QuickBooks, SAP, and Oracle Financials have become staples in many accounting departments, streamlining operations and improving accuracy.

However, while these tools are powerful, they primarily operate based on predefined rules and lack the flexibility to handle ambiguous or unstructured data. They're excellent for routine tasks but fall short when it comes to understanding context, generating insights, or handling non-standardized data sources. This is where AI can help. Here are examples of challenges faced by accountants that can be solved with AI.

- **Data Overload:** With the significant increase in the amount of data that is generated and stored, accountants have to understand what data is relevant and what data is not useful. Manually extracting meaningful insights from this sea of data is time-consuming and prone to mistakes.
- **Unstructured Data:** Invoices, receipts, and memos come in varied formats. Traditional systems struggle with such unstructured data, leading to manual interventions and potential errors.

- **Complex Decision-making:** Accounting isn't just about number crunching. Accountants are involved in complex decisions based on financial data, regulations, market conditions and human judgments. AI can assist by providing data-driven insights and recommendations. AI helps remove many of the routine tasks that humans previously performed so that humans can perform more interesting and value-adding tasks.
- **Client Communication:** Addressing client queries, explaining financial regulations, or advising on financial decisions requires clear and effective communication. AI can assist in drafting communications, increasing clarity, and even answering routine queries autonomously.
- **Fraud Detection:** With financial fraud becoming increasingly sophisticated, traditional rule-based systems often fall short. AI, with its pattern recognition capabilities, can help in detecting anomalies and potential fraudulent activities.

By integrating AI into accounting processes, we're not just looking at incremental improvements but a fundamental shift in how accounting functions are approached and executed.

Basics of Prompt Engineering

Prompt engineering is the practice of crafting effective instructions to guide language models in generating desired outputs. While it might sound simple, deriving the most accurate and relevant answers from an AI model requires a nuanced approach to question formulation.

Definition and importance

Prompt engineering can be defined as the methodical process of designing, testing, and refining prompts to elicit specific and accurate responses from language models. It's not just about asking questions but asking the right questions. At its core, prompt engineering isn't a new concept; it's about understanding the needs, context and outcome precision of phrasing questions. Given the vast knowledge and capabilities of AI models, the way a question or instruction is phrased can dramatically alter the output.

In the context of accounting, the stakes are high. Misinterpretation or inaccurate data can have significant financial implications. Thus, to be useful in accounting settings, prompts must be carefully crafted to create useful output.

The art and science behind effective prompting

Creating an effective prompt is both an art and a science. Here's a glimpse into the process.

- **Clarity:** The prompt should be clear and free of ambiguity. For accounting tasks, this often means providing context or specifying the format of the desired answer and sometimes naming the end user.
- **Brevity:** While providing context is crucial (sometimes even providing examples can help significantly), overly verbose prompts can confuse the model. Striking a balance is key.
- **Iterative Refinement:** Rarely will the first prompt yield the perfect result. Prompt engineering involves tweaking the instruction, analyzing the output, and refining the approach. The finished deliverable may be the result of multiple prompts and not just a single statement.
- **Feedback Loops:** Continually training the model with feedback on its responses enhances its accuracy over time, especially for domain-specific tasks like accounting.
- **Pattern:** While prompts will vary in complexity based on the end user and the desired outcome, a basic template often includes:

[Objective] [Context] [Criteria] [Output Indicators].

For example:

“Give me a list of beautiful, but less popular, shorter hikes located in or near Glacier National Park. I’m a moderate skilled hiker, who enjoys the outdoors. Make sure the list contains at least 10 hikes, avoiding the most popular, and matching my preferences.”

or

“Give me a list of at least five top-rated kosher restaurants in the Dallas metro area that have an average price of \$15 a meal.”

Rarely is copying and pasting the result from a prompt sufficient for a professional deliverable. Prompt engineering should be seen as part of a process. After the AI develops the solution based on prompting, a human should carefully review and edit the solution. A good rule of thumb is that prompt engineering will get you 80% to 90% to completion, but a human needs to review and complete the final 10% to 20% of the work to finalize the deliverable.

Different Types of Prompt Techniques

Below are examples of three different types of prompt techniques that can enhance your prompting effectiveness.

1. Zero-shot prompting is when a model is asked to perform a task without being given any specific examples or “shots” in the prompt. The model relies entirely on its pre-existing knowledge and training to comprehend and answer the query. For example, you might type all of the following as a prompt: “Explain the elements of the audit risk model.”
2. Few-shot prompting is providing a model with a small number of examples (or “shots”) to instruct the model on the desired task and outcome. This helps the model understand the context and provide answers accordingly. For example, you might type all of the following as a prompt:
Scenario: John earned \$50,000 in salary this year.
Tax Treatment: John needs to report \$50,000 as earned income on his tax return.
Scenario: Sarah received \$10,000 as an inheritance from her grandmother.
Tax Treatment: Generally, inheritances are not considered taxable income, so Sarah does not need to report the \$10,000 on her tax return.
Scenario: Michael sold stocks and made a capital gain of \$5,000.
Tax Treatment: Michael needs to report the \$5,000 as capital gains on his tax return and pay appropriate capital gains tax.

Given the above examples, please provide the tax treatment for the following scenario:
Scenario: Lisa received \$3,000 as a gift from her parents.
Tax Treatment: ?.
3. Chain-of-thought prompting is a prompting technique that allows complex reasoning capabilities through intermediate reasoning steps. Chain-of-thought prompting can be written as a no-shot or few-shot prompt. Here is an example of no-shot, chain-of-thought prompt: “
ABC Ltd manufactures toys. Each toy requires raw materials costing \$2. The labor cost associated with producing one toy is \$1. Monthly overhead costs amount to \$2,000. In a particular month, they produced 1,000 toys. Calculate the total cost of producing 1 toy and the contribution margin per toy. Show your work.

Here is an example of a few-shot, chain-of-thought prompt:

Scenario: ABC Ltd manufactures toys. Each toy requires raw materials costing \$2. The labor cost associated with producing one toy is \$1. Monthly overhead costs amount to \$2,000. In a particular month, they produced 1,000 toys.

Calculation & Conclusion:

Variable cost per toy: \$3 ($2 + 1$)

Fixed cost for that month: \$2,000

Total cost of producing one toy: \$5 ($\$3 + \$2,000/1,000$)

If each toy is sold for \$8, contribution margin per toy: \$3 ($\$8 - \5)

Scenario: DEF Corp makes pens. Each pen uses raw materials that cost \$0.50. The labor cost for one pen is \$0.20. Monthly overheads are \$1,000, and they manufactured 5,000 pens that month.

Calculation & Conclusion:

Variable cost per pen: \$0.70 ($\$0.50 + \0.20)

Fixed cost for that month: \$1,000

Total cost of producing one pen: \$0.90 ($\$0.70 + \$1,000/5,000$)

If each pen is sold for \$1.50, contribution margin per pen: \$0.60 ($\$1.50 - \0.90)

Given the above examples, please provide calculations and conclusions for the following scenario:

Scenario: XYZ Inc. manufactures widgets. Each widget requires raw materials costing \$5. The labor cost associated with producing one widget is \$3. The monthly overhead costs for the factory amount to \$10,000. In the month of August, the company produced 5,000 widgets.

Calculation & Conclusion:

a) Variable cost per widget: ?

b) Fixed cost for August: ?

c) Total cost of producing one widget in August: ?

d) If XYZ Inc. sells each widget for \$15, contribution margin per widget: ?

In the examples, the chain-of-thought prompting for no-shot included the phrase “Show your work.” This was intended to have the model reason through the steps that would produce a correct answer. In the few-shot example, the examples were very similar to the final question. While this will improve model output, even providing questions with the similar structure can help. For example, if you prompt with multiple choice questions—even if the topic is different than the question of interest—it will enhance accuracy.

Examples of basic prompts for accounting-related tasks

1. Data Entry:

- Ineffective Prompt: “Enter this data.”
- Effective Prompt: “Extract and categorize the financial data from the attached invoice into debit and credit columns.”

2. Financial Analysis:

- Ineffective Prompt: “Analyze the report.”

- Effective Prompt: “Examine the quarterly financial report and highlight any anomalies in the revenue and expense accounts compared to the previous quarter.”
3. Client Communication:
- Ineffective Prompt: “Send a message to the client.”
 - Effective Prompt: “Draft a professional email to my client, Sam Wilson, at ABC Widgets, summarizing the key findings from their annual financial statement audit and suggesting a meeting to discuss further.”

These examples demonstrate the importance of precision in prompting. Even with precision, the output of these prompts will need further refinement. Viewing prompt engineering as an iterative process will help you produce higher quality output.

There are many other resources online that give examples of different types of prompts. For example, <https://www.promptingguide.ai/introduction/examples> provides additional examples of different types of prompts. Consider reviewing this resource and searching online (or using a large language model) to learn other examples of prompting.

Applications in Accounting

The blend of accounting with AI LLMs, facilitated through prompt engineering, opens a treasure trove of opportunities. In general, AI LLMs facilitate the move towards analytical tasks, including important big picture tasks like financial analysis. The potential application of using AI to complete lower level, repetitive accounting tasks is vast and transformative. Let's explore a few potential applications.

Assessing Internal Controls:

Well-designed and operating internal controls are essential to maintaining an effective control environment, producing materially accurate financial statements, preventing financial fraud and establishing accounting processes that are efficient and reliable. Companies can utilize AI to regularly review transactions, journal entries, management approvals, and a host of other activities.

- Prompt Example: “Based on the provided transaction logs, identify any suspicious or repeated transactions that might indicate a potential control weakness or fraudulent activity. Compare the frequency and volume of such transactions to the historical average.”

This prompt could help identify risky transactions that should be investigated more in-depth. The initial screening of transactions could save accountants significant time by focusing their attention on the higher risk transactions.

Enhancing Data Analysis

The real value in accounting data often lies in the insights it can provide. Here are examples of prompts to generate insights from financial statements.

- Prompt Example: “Review the annual balance sheet and income statement. Identify any significant year-over-year variations in revenue streams or expense categories.”

Such prompts can assist accountants in quickly pinpointing areas that need further investigation or that may indicate emerging business trends.

Improving Communication

Effective communication is critical in accounting, whether it's with clients, regulators, or internal stakeholders. Here is an example of a prompt to generate a response to a client.

- Prompt Example: "Draft a response to the client's query about the discrepancies in their tax report, ensuring clarity and professionalism. Use the provided data summary for reference."

AI can help draft initial responses or provide templates, ensuring consistent quality and freeing up accountants for more complex client interactions.

Challenges and Considerations

While the potential of integrating prompt engineering and AI into accounting is vast, it's essential to approach this intersection with a critical eye. As with any technology, there are challenges and considerations to be mindful of, ensuring that the adoption is responsible, ethical and effective.

Accuracy is non-negotiable in accounting. A slight misinterpretation or oversight can lead to significant financial discrepancies, with serious implications to the financial markets. This fact places immense importance on the clarity and precision of prompts and on having a qualified human review the AI's output. Even a small ambiguity in a prompt can lead to incorrect data interpretation or categorization by the AI. Regular training sessions for accounting professionals on effective prompt engineering techniques can help mitigate this risk. Organizations will need to implement a review system to double-check AI's output, especially for high-stake tasks.

Having a qualified human review any AI output before relying on it is also especially important because AI can "hallucinate." A hallucination refers to an instance where the model produces an output that is factually incorrect, nonsensical, or not aligned with the input it received. This can happen for various reasons, including biases in training data, the model misinterpreting the prompt, or the inherent challenges in training large models to understand context perfectly. Consider the following example:

- Prompt: "Calculate the net profit given: Total Revenue of \$1,000,000; Cost of Goods Sold (COGS) of \$500,000; Operating Expenses of \$300,000."
- Hallucinated Response: "The net profit is \$250,000."

The response is incorrect and considered to be a hallucination because the correct answer should be

Net Profit = Total Revenue - (COGS + Operating Expenses)

= \$1,000,000 - (\$500,000 + \$300,000)

= \$200,000.

Any AI answer could suffer from hallucinations. Hallucinations typically happen because of:

1. **Ambiguous prompts:** If a prompt is vague or lacks clear context, the model may generate outputs that don't align with the user's expectations.
2. **Lack of domain-specific training:** While many models have vast general knowledge, they might not be as refined in very niche or specialized areas. If the model hasn't been trained sufficiently on specific domain data, it might produce incorrect or nonsensical outputs.
3. **Longer responses:** The probability of hallucinating can increase with the length of the response. The longer the generated output, the more chances there are for the model to deviate from accurate or relevant information.
4. **Complex queries:** If a question involves multiple steps or requires deep reasoning, the chances of the model making an error in one of those steps increases.
5. **Uncommon scenarios:** If a prompt describes a scenario that is rare or very uncommon in the model's training data, the AI might struggle to provide a relevant or accurate response.

6. **Biased or inaccurate training data:** The model's outputs are only as good as its training data. If the AI has been trained on biased, outdated, or inaccurate data, it's more likely to produce hallucinated responses.
7. **Over-reliance on model confidence:** Sometimes, even if the model is highly confident in its response, it can still be wrong. Blindly trusting the model's confidence can lead to overlooking hallucinations.
8. **Iterative prompting:** In situations where users iteratively refine their prompts based on the model's feedback, there's a chance of "leading" the model down a path that results in hallucinated or nonsensical outputs.

Understanding these potential pitfalls and always verifying the model's output is crucial for AI users. Having a solid understanding of these limitations can help to mitigate the risks associated with model hallucinations.

Practical Application

The best way to learn how to be a successful prompt engineer is to practice and to experiment. Use the following scenarios to practice your prompt engineering. Use a resource like ChatGPT, Bard, or other model to perform this task. Unlike many accounting assignments, there is not a single, right answer to some of these prompts. In assessing the quality of your prompts, review the output to make sure it is accurate and communicates using the appropriate tone.

Also, note that different language models have different abilities. Some models allow for uploading documents (like PDFs or Excel workbooks). Some models allow connecting to the internet. Some models have plugins that allow special abilities (like the model being able to access a calculator or write computer code). You will need to learn about the specific model strengths and weaknesses to improve your prompting abilities.

Scenario 1

The following text comes from the 2022 10-K filing for Airbnb (see <https://www.sec.gov/ix?doc=/Archives/edgar/data/1559720/000155972023000003/abnb-20221231.htm>). This text is included as part of "Item 7. Management's Discussion and Analysis of Financial Condition and Results of Operations".

Overview

We are a community based on connection and belonging—a community that was born in 2007 when two Hosts welcomed three guests to their San Francisco home, and has since grown to over 4 million Hosts who have welcomed over 1.4 billion guest arrivals to over 100,000 cities and towns in almost every country and region across the globe. Hosts on Airbnb are everyday people who share their worlds to provide guests with the feeling of connection and being at home. We have five stakeholders and we have designed our company with all of them in mind. Along with employees and shareholders, we serve Hosts, guests, and the communities in which they live. We intend to make long-term decisions considering all of our stakeholders because their collective success is key for our business to thrive.

We operate a global marketplace, where Hosts offer guests stays and experiences on our platform. Our business model relies on the success of Hosts and guests (collectively referred to as "customers") who join our community and generate consistent bookings over time. As Hosts become more successful on our platform and as guests return over time, we benefit from the recurring activity of our community.

Initial Public Offering

Our initial public offering ("IPO") was completed on December 14, 2020. Our consolidated financial statements as of December 31, 2020 and for the year then-ended reflect the sale by us of an aggregate of 55,000,000 shares in our IPO, including the exercise of the underwriters' option to purchase additional shares, at the public offering price of \$68.00 per share, for net proceeds to us of approximately \$3.7 billion, after underwriting discounts and commissions and offering expenses, and the conversion of all outstanding shares of our redeemable convertible preferred stock into an aggregate of 240,910,588 shares of Class B common stock, including 1,286,694 shares of Class B common stock issuable pursuant to the anti-dilution adjustment provisions relating to our Series C redeemable convertible preferred stock.

Our consolidated financial statements as of December 31, 2020 and for the year then-ended include stock-based compensation expense of \$2.8 billion associated with the vesting of RSUs in connection with our IPO for which the requisite service-based vesting condition was met as of December 31, 2020. The liquidity-based vesting condition for RSUs was satisfied upon the effectiveness of our Registration Statement on Form S-1 on December 9, 2020.

2022 Financial Highlights

In 2022, revenue grew by 40% to \$8.4 billion compared to 2021, primarily due to a 31% increase in Nights and Experiences Booked of 93.0 million combined with higher average daily rates driving a 35% increase in Gross Booking Value of \$16.3 billion. The growth in revenue demonstrated the continued strong travel demand. On a constant-currency basis, revenue increased 46% in 2022 compared to 2021.

We ended 2022 with net income of \$1.9 billion, an improvement from a net loss of \$352.0 million in 2021, and our first profitable year to date. Our net profit margin increased to 23% from a negative 6% in 2021, primarily due to our revenue growth outpacing the growth in our operating expenses and cost management.

Adjusted EBITDA increased 82% to \$2.9 billion in 2022 demonstrating the continued strength of our business and disciplined management of our cost structure.

Our net cash provided by operating activities was \$3.4 billion in 2022, up from \$2.3 billion in 2021, and we generated Free Cash Flow of \$3.4 billion. The increase was driven by our revenue growth, net margin expansion, and significant growth in unearned fees.

In August 2022, our board of directors approved a share repurchase program with authorization to purchase up to \$2.0 billion of our Class A common stock at management's discretion. During 2022, we repurchased and retired 13.8 million shares of common stock for \$1.5 billion.

With this data, prepare a prompt (or a few prompts) to generate the following:

- A summary, in bullet form, of the most important financial highlights in 2022.
- A post for LinkedIn highlighting the most important items to attract and recruit new employees to work at Airbnb. Remember, LinkedIn posts should be concise and professional and are more effective if they use hashtags.
- A post for Twitter/X that will appeal to potential renters to make them want to stay at Airbnb. It should reference something in the data above. This post should be highly creative and have a significant chance of going viral.

Scenario 2

Copy the income statement for Airbnb (see link above) and then write the prompts to generate the following:

- Create two bulleted lists, one of the company's greatest financial improvements and the second of the greatest financial concerns based on the data from 2020 to 2022.
- Based on the financial concerns, generate a list of possible risks, including fraud risks, that could have led to the financial concern. For one of these concerns, generate a list of audit procedures that would be most likely to mitigate these risks and/or potentially discover a fraud.

Scenario 3

Airbnb reports the following risk factor in Item 1A of their 2022 financial statements, "If we fail to retain existing Hosts or add new Hosts, or if Hosts fail to provide high-quality stays and experiences, our business, results of operations, and financial condition would be materially adversely affected." Based on this, develop the following:

- Thinking as a managerial accountant, develop measures using the balanced score card approach to mitigate this risk.
- For the financial area of the balanced score card, specifically define any metrics that you would track. Although you may not be familiar with SAP (a large database that is used by many companies), use the AI to develop a list of the SAP tables where the data for those metrics is mostly likely to be found. Generate the SQL code that would likely extract the information from the SAP tables (you can assume a general SAP install, and we recognize you will not likely be able to audit or execute if this code is correct).

Final Deliverable

Based on what you learned from these scenarios, prepare a memo (without help from AI) listing what you learned about prompt engineering. Make sure to discuss at least the following:

- What made your prompts more or less successful? Give specific examples.
- How accurate were your prompts? Identify any hallucinations and how you discovered they were hallucinations.

Prepare a second memo based solely on prompting the AI that does the same thing as your non-AI memo. Be prepared to discuss what you learned based on generating each of the memos in class.