

Class 4: Prompt Engineering

DPI-681M

The Science and Implications of Generative AI

Profs. Goel, Levy, and Svoronos

Harvard Kennedy School

1. Introduction to Prompt Engineering
2. Prompting Extensions
3. Personas

Welcome to Unit 2!

Three big questions we aim to answer together:

- Unit 1: How does generative AI work?
- Unit 2: How can I use it effectively?**
- Unit 3: How will it affect the people and issues I care about?

1. Introduction to Prompt Engineering

From science to practice

You now have an understanding of the mechanics of how LLMs work, from initial **training** to LLM **alignment**.

However, the vast majority of us will primarily interact with existing LLMs, not the creation of new ones (at least for now!)

From science to practice

You now have an understanding of the mechanics of how LLMs work, from initial **training** to LLM **alignment**.

However, the vast majority of us will primarily interact with existing LLMs, not the creation of new ones (at least for now!)

The primary way that we interact with LLMs is through a text entry box. What we type in that box is called **prompting**.

Today, we will focus on best practices of creating effective prompts, and practice using LLMs to accomplish a variety of tasks.

Gen AI: What is it good for?

Gen AI: What is it good for?

- Explaining and teaching
- Writing and rewriting
- Creative brainstorming
- Analyzing and summarizing
- Organizing and classifying

ARTIFICIAL INTELLIGENCE

Google DeepMind used a large language model to solve an unsolved math problem

They had to throw away most of what it produced but there was gold among the garbage.

By Will Douglas Heaven

December 14, 2023

ARTIFICIAL

Google model

They had to thr
gold among the garbage.

By Will Dougl

December 14, 2

The researchers started by sketching out the problem they wanted to solve in Python, a popular programming language. But they left out the lines in the program that would specify how to solve it. That is where FunSearch comes in. It gets Codey to fill in the blanks—in effect, to suggest code that will solve the problem.

A second algorithm then checks and scores what Codey comes up with. The best suggestions—even if not yet correct—are saved and given back to Codey, which tries to complete the program again. “Many will be nonsensical, some will be sensible, and a few will be truly inspired,” says Kohli. “You take those truly inspired ones and you say, ‘Okay, take these ones and repeat.’”

age
blem

ARTIFICI

Google

mod

They had
gold amor

By Will De

December 11, 2014

The researchers started by sketching out the problem they wanted to solve in Python, a popular programming language. But they left out the lines in the program that

After a couple of million suggestions and a few dozen repetitions of the overall process—which took a few days—FunSearch was able to come up with code that produced a correct and previously unknown solution to the cap set problem, which involves finding the largest size of a certain type of set. Imagine plotting dots on

to complete the program again. “Many will be nonsensical, some will be sensible, and a few will be truly inspired,” says Kohli. “You take those truly inspired ones and you say, ‘Okay, take these ones and repeat.’”

"Prompt Engineers" - the job of the future?

"Prompt engineering is the process where you **guide** generative artificial intelligence (generative AI) solutions to generate desired outputs.

Even though generative AI attempts to mimic humans, it **requires detailed instructions to create high-quality and relevant output**. In prompt engineering, you choose the most appropriate formats, phrases, words, and symbols that guide the AI to interact with your users more meaningfully. Prompt engineers use **creativity plus trial and error** to create a collection of input texts, so an application's generative AI works as expected."

F Forbes

The Hot, New High-Paying Career Is An AI Prompt Engineer

As AI job listings have surged 42%, prompt engineering is quickly becoming a hot job with the ascendancy of artificial intelligence.

3 weeks ago



F Forbes

The Hot, New High-Paying Career Is An AI Prompt Engineer

As AI job listings have surged 42%, prompt engineering is quickly becoming a hot job with the ascendancy of artificial intelligence.

3 weeks ago

#d Diginomica

Prompt Engineer, the hottest new job in tech, is already almost obsolete

Prompt Engineer, the hottest new job in tech, is already almost obsolete ... It was foretold last year that tech jobs would shift to include...

1 month ago



Anatomy of a Prompt (TIC)

Though prompts can get very complex, they tend to boil down to three components:

- **Task:** What you want the AI to do
- **Instructions:** How you want the AI to do it
- **Context:** What you want the AI to know [examples?]

Writing an admissions essay

I am applying to a Harvard Kennedy School masters program in public policy. I would like you to produce a first draft of one of my essays (250 word limit). The essay prompt is the following: "Describe a time when interactions with others and/or an experience caused you to change your mind or expanded your point of view." I would like you to write about my experience changing my mind about the potential of online learning to transform education. I used to believe that it did not have much potential and now I believe that it has. An interaction with two faculty members and 3 students made me realize that I was wrong and they were right. Please produce an essay that is personal, crisp and persuasive. Please don't ask me any questions; just write the essay.

[]: Task

(): Instructions

{ }: Context

Sometimes helpful: rearranging

Task: I will give you an essay prompt and you will write a first draft of an admissions essay.

Instructions: Produce an essay responding to the essay prompt, "Describe a time when interactions with others and/or an experience caused you to change your mind or expanded your point of view." It should be personal, crisp and persuasive. Please don't ask me any questions; just write the essay. Keep it under 250 words.

Context: I am applying to the Harvard Kennedy School's master's program in public policy. I used to believe that online learning did not have the potential to transform online education. A conversation with two faculty members and three students made me realize I was wrong and they were right. [Add additional context]

[]: Task

(): Instructions

{ }: Context



How else can we improve this prompt?

Effective prompting: General guidelines

1. Start **simple** and **build** (via conversing)
2. Be **specific** about what you want the LLM to do/produce
3. Say **what to do** rather than what not to do
4. Ask for **several** options/examples
5. **Use symbols** (like `##` or `"""`) to separate parts of your prompt (e.g., pasting in a lot of background information)



Job Search

- Imagine(?) you are looking for a job. As a group, use PingPong to brainstorm the types of positions you might want to apply for given your interests. After choosing some types of positions, ask for suggestions of organizations that you should look into.
- If instead your group already has a job in mind, write a cover letter for it (like in PCE 0!), this time using the principles of effective prompting when doing so.

2. Two Prompting Extensions

2. Two Prompting Extensions

Intermediate prompting

Usually, LLMs perform quite well at the tasks we ask them to do, provided we supply a well-structured and concise prompt.

Intermediate prompting

Usually, LLMs perform quite well at the tasks we ask them to do, provided we supply a well-structured and concise prompt.

However, for some tasks, LLMs might need a little help:

- **Complex** analytical tasks
- Tasks where the output should follow a certain **format/structure**
- Tasks that are hard to describe, but **easy to show** via example

Here are some strategies that have been shown to help!

Extension 1: Chain of Thought (CoT)

Just like for humans, it is often effective to ask an LLM to break down a problem into component parts. This was originally demonstrated for math problems:

Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Chain-of-Thought Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

Source: Wei (2022) Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, <https://arxiv.org/abs/2201.11903>



From your PCE...

GPT-4 Classic ▾

Classify the emotional intensity of each of the 40 comments below, which are surrounded by triple quotes """". They should be classified as intense, moderate, or indifferent. For each, explain why you classified them the way you did.

""""Trevor C.

1/24/24 - 11:23AM



I'm glad Cambridge is building a protected bicycle lane network and doing it in a timely manner. I wish Boston would get on board; it's needlessly difficult to cycle in Boston.





From your PCE...

4. **lem144's comment "You plan for them. That's what happens.":** Moderate. The tone is somewhat assertive but not overly intense.
5. **Comment on the city acting on a dangerous stretch due to crashes:** Moderate. Expresses concern and approval for the project's progress without intense language.
6. **Jane Ruby's complaint about bike and bus lanes:** Moderate. The comment expresses frustration and personal experiences but does not resort to extreme language.



From your PCE...

4. **lem144's comment "You plan for them. That's what happens.":** Moderate. The tone is somewhat assertive but not overly intense.
5. **Comment on the city acting on a dangerous stretch due to crashes:** Moderate. Expresses concern and approval for the project's progress without intense language.
6. **Jane Ruby's complaint about bike and bus lanes:** Moderate. The comment expresses frustration and personal experiences but does not resort to extreme language.



Moderate?

Jane Ruby

1/24/24 - 12:28PM

As a 20+ year resident of Cambridge who drives and walks, the bike and bus lanes are a mess! We don't have the infrastructure in place to support a non-car driving city. The T is barely usable (twice I walked from Newbury Street to Harvard Square because I gave up waiting for a bus). We have an empty bus lane and cars that slowly move through the city, adding pollution to our air.

Extension 2: Few-shot prompting

When you are asking an LLM to judge things, it is often useful to give examples to help motivate it. This is particularly true when it's hard to describe the criteria you want it to use.



You

The examples below demonstrate translations of English questions into Blablarian.

English: How are you?

Blablarian: Zim quex voobi?

English: How old are you?

Blablarian: Zim glax voobi?

English: How is he?

Blablarian: Zim quex vopli?

Based on the examples, translate the following into Blablarian:

How old is he?



ChatGPT

Based on the provided examples and the established patterns in the translations from English to Blablarian, the translation of "How old is he?" into Blablarian would be:

Blablarian: Zim glax vopli?

This translation follows the observed rules:

- "Zim" is used to start questions related to "How."
- "glax" specifically refers to inquiries about age.
- "vopli" is the Blablarian equivalent of "he," used to refer to a third person.



From your PCE...



theodore_svoronos@hks.harvard.edu

Classify the emotional intensity of the below comments, surrounded by triple quotes """". They should be classified as intense, moderate, or indifferent. Here are some examples:

Comment: I don't like this proposal Answer: Moderate

Comment: I guess the city just doesn't care about businesses Answer: Intense



From your PCE...



theodore_svoronos@hks.harvard.edu

Classify the emotional intensity of the below comments, surrounded by triple quotes """". They should be classified as intense, moderate, or indifferent. Here are some examples:

Comment: I don't like this proposal Answer: Moderate

Comment: I guess the city just doesn't care about businesses Answer: Intense



GPT-4 Classic

Classifying the emotional intensity of the comments:

3. **"As a 20+ year resident of Cambridge who drives and walks, the bike and bus lanes are a mess! We don't have the infrastructure in place to support a non-car driving city."**

- Intense: Expresses frustration and dissatisfaction with the current infrastructure, presenting strong opinion based on personal experience.



Activity



Let's say you got a job interview at a climate change think tank called the Climate Action Institute. You want to get a list of possible **substantive interview questions** about climate topics that they study and that you are less familiar with, such as carbon capture.

Using PingPong and your new prompting tools, generate a set of interview questions and possible responses that might be useful to you. Use this to **learn about the topics that you might be asked about!**

Personas

Giving AI personality

One of the most powerful uses of LLMs is to have it take on a **persona** with a particular set of goals and perspectives.

These personas can help you **step outside of your own perspective** by discussing with a virtual partner.

As you will see next class, this also the basis of most **chatbots**, which are instructed to take on specific roles (provided the instructions are aligned!).

Some examples of Personas

Persona	Role	Use case
Research Assistant	Gathers/processes information	Summarization, data analysis
Mentor	Provides feedback & suggestions	Improve an essay
Critic	Critiques your work	Find weak points in arguments
Tutor	Provides personalized instruction	Learn about a new topic
Student	Asks questions about content	Deepen your understanding and ability to explain ideas
Interviewer	Probes your thinking and asks for introspection	Help make a difficult decision
Simulator	Takes on a persona in a simulated environment	Practice a skill or interaction

Source: Modified from "[Assigning AI: Seven Approaches for Students, with Prompts](#)" by Mollick & Mollick, 2023



Example from API-201

"I want you to take on the role of a policymaker who is curious, passionate, but not well-versed in statistics. I would like you to ask me, a student of statistics, to explain a particular statistical concept to you in words that are accessible. You can choose from the following list of concepts:

[List of concepts]

Start by providing me with a policy-related scenario that you, the policymaker, have encountered. Be specific about the context and the variables you are thinking about. Then use one of the concepts from the list above in a way that has some subtle misunderstanding of the idea. I will then answer by correcting your understanding. You will then evaluate whether my explanation was (a) correct, (b) explained in accessible terms, and (c) using an example specific to your scenario's context. When I say "accessible," I mean understandable by someone with a college education but no background in statistics."



Sample response from PingPong

As a policymaker concentrating on improving education outcomes, I've been analyzing the recent standardized test scores from schools across our region. We've noticed that some schools, particularly those in lower-income areas, are performing below the national benchmarks. To address this, we aim to allocate more resources and support to the schools in need. However, we're grappling with how to accurately identify those schools.

In our meetings, there's been a lot of talk about using the "average" test score of a school to make these decisions. We thought about comparing these averages to the national benchmark and then providing extra help to the schools falling below this line.

I've heard about using the "median" as well, but I'm not sure how that differs really, or why it would matter in our case. It seems to me that the middle score might not show the whole picture if most students are doing worse than that. Could you clarify this for me?



Activity



Congratulations! You have received an offer letter for a Policy Analyst position from the Climate Action Institute (read it at the link on the left). You'll be having a call with the Head of Policy Research to negotiate your offer in a week.

Write a prompt instructing PingPong to take on the persona of this person so you can practice negotiations with it. Think about incorporating some of the techniques that we've discussed.



What part of your prompting did you find most helpful? What was challenging?

- The importance of starting with "you are a representative of the Climate Action Institute charged with negotiating with me on a job offer..."
- Explain "I will propose changes to the job offer and you will respond"
- Imbue it with characteristics, "you are a tough negotiator and concerned about budgeting, so be resistant to my proposals"
- Ask for feedback after the negotiation ends

Key Takeaways

Key Takeaways

Takeaways

- **Prompting** is the primary way that most people interact with an LLM.
- Remember **TIC**: Task, Instructions, Context
- Do not be satisfied with an LLM's first response to a prompt - **engage with it** to get better, more personalized responses.
- For more complex tasks, ask the LLM to break down tasks into **component parts**, and provide it with **examples** for what effective responses look like.
- **Be creative** with the kinds of personas an LLM can play in your life!

Glossary

- **Prompt:** Text that you provide to an AI tool. Prompts generally consist of a task, instructions, and context.
- **Chain of Thought (CoT) prompting:** Prompts that instruct an LLM to break down their responses into intermediate steps before arriving at an answer.
- **Single/Few-shot learning:** Prompts that provide sample answers to guide an LLM toward more effective answering
- **Personas:** Instructing an LLM to take on a particular role that the user interacts with.
- **Chatbots:** LLMs that have been given a custom set of instructions to act in a particular way and respond to a particular set of queries.