DATA COUCH

Part II -
Prompt Engineering

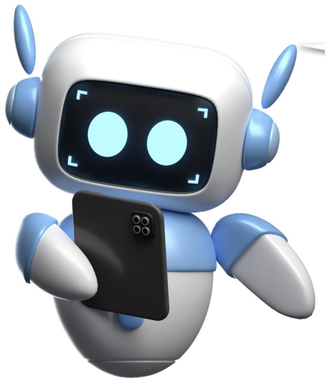
Agenda

- **Recap of LLMs**
- Understanding Prompt Design
- Context, Examples, & Tuning Prompt Parameters
- Prompt Design Best Practices

What is a Language Model?

A language model is like a ***smart machine that helps you finish sentences***. It's trained to predict and complete sentences, making it a super useful tool when you're typing or texting

Listen to your _____



Heart **10.5%**

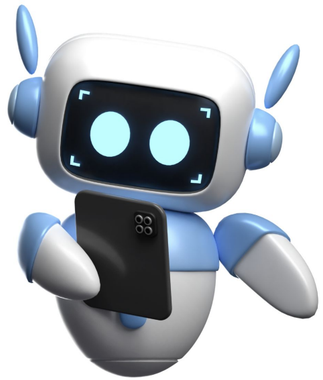
Gut 9.8%

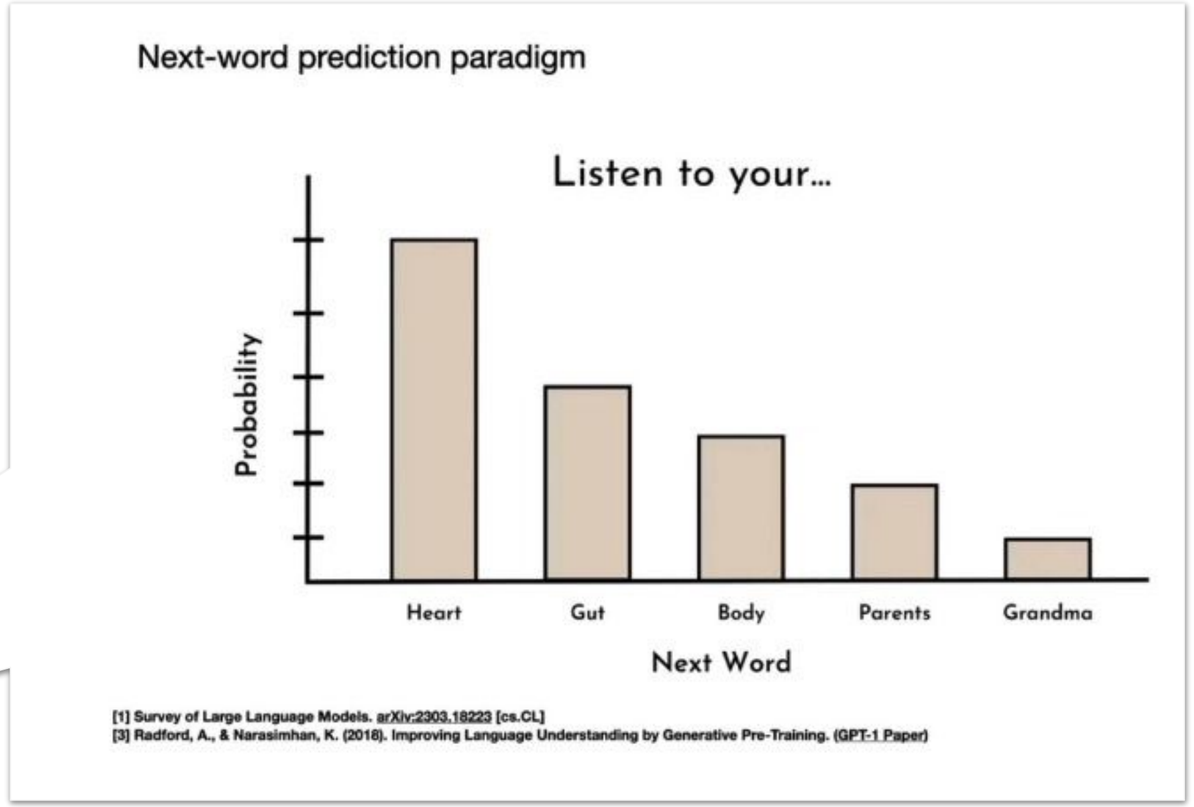
Body 5.6%

Parents 3.5%

…

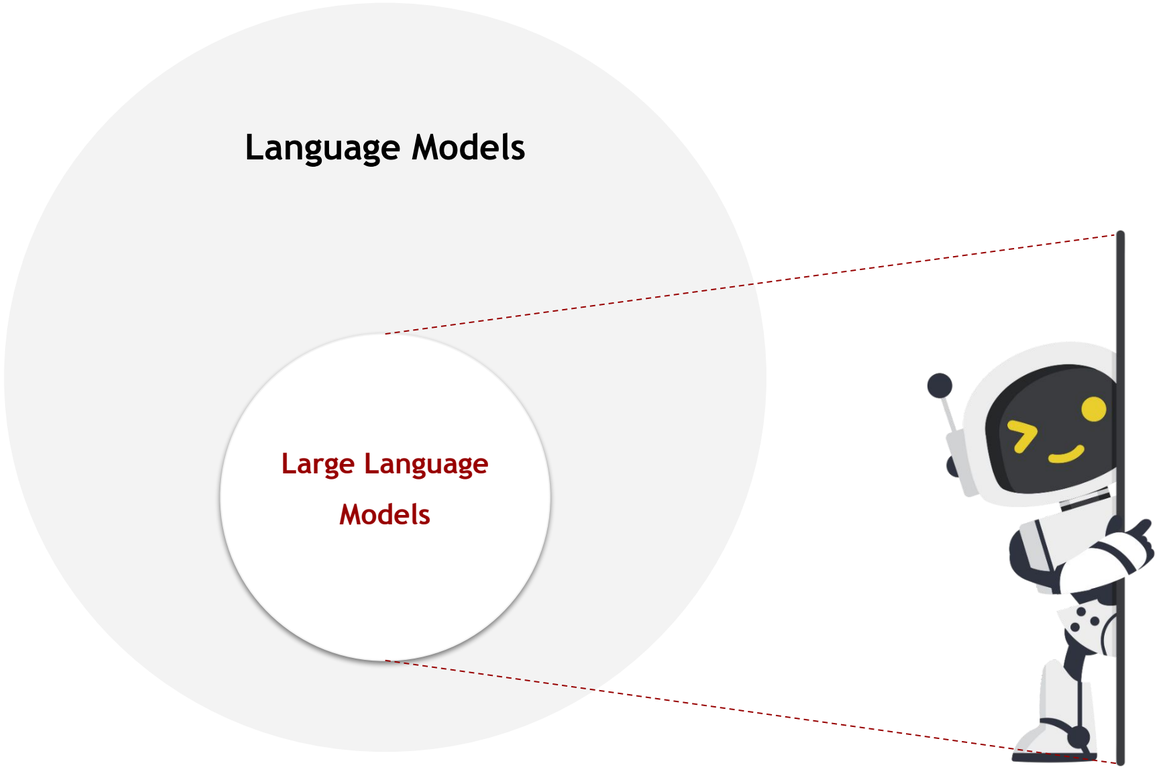
What is a Language Model?





[1] Survey of Large Language Models. arXiv:2303.18223 [cs.CL]
[3] Radford, A., & Narasimhan, K. (2018). Improving Language Understanding by Generative Pre-Training. (GPT-1 Paper)

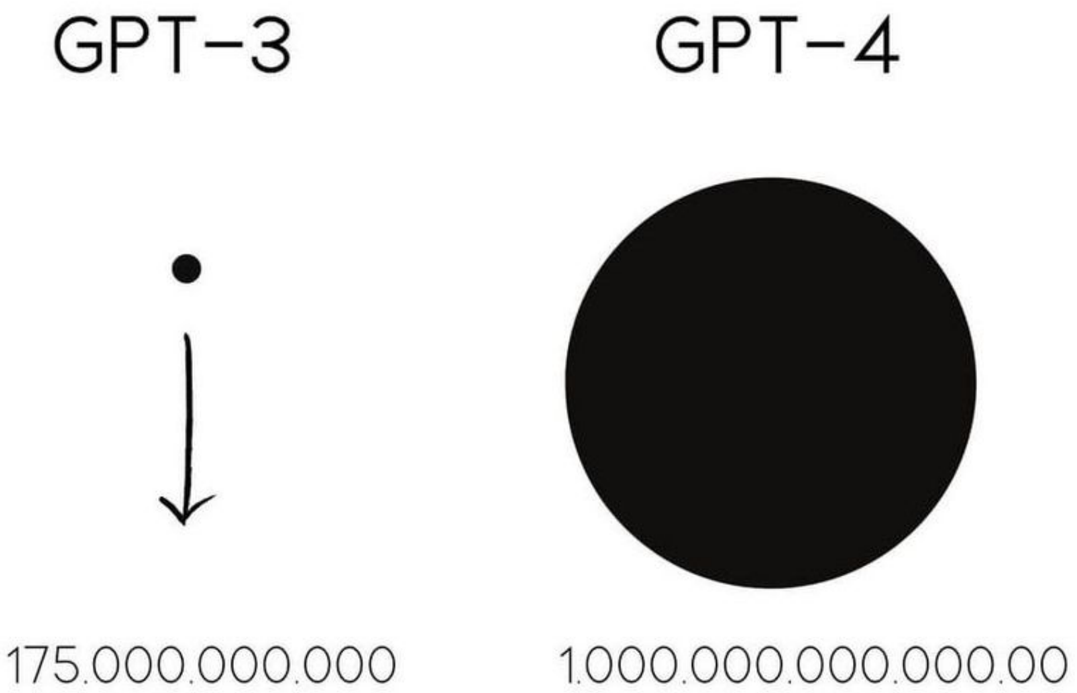
What are Large Language Models?



Models trained on large amounts of all type of publicly-available data that can perform a wide variety of tasks, including text/image/document summarization, generation, and categorization. These models can perform generative tasks like text/image/document generation and so there is some overlap between LLMs and Generative Al.

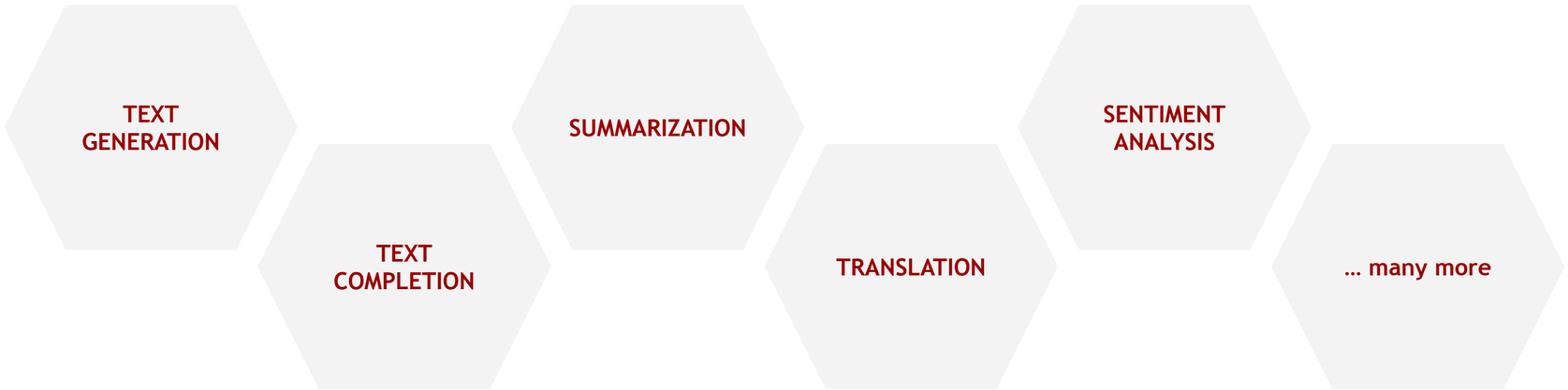
What is “**Large**” in Large Language Models?

- The term "**large**" refers to the size of the model in terms of its parameters and the volume of the training data
- Parameters are a part of the model that learns from historical training data
- They help in predicting the next word in a sequence
- One Trillion Parameters have been used in GPT 4



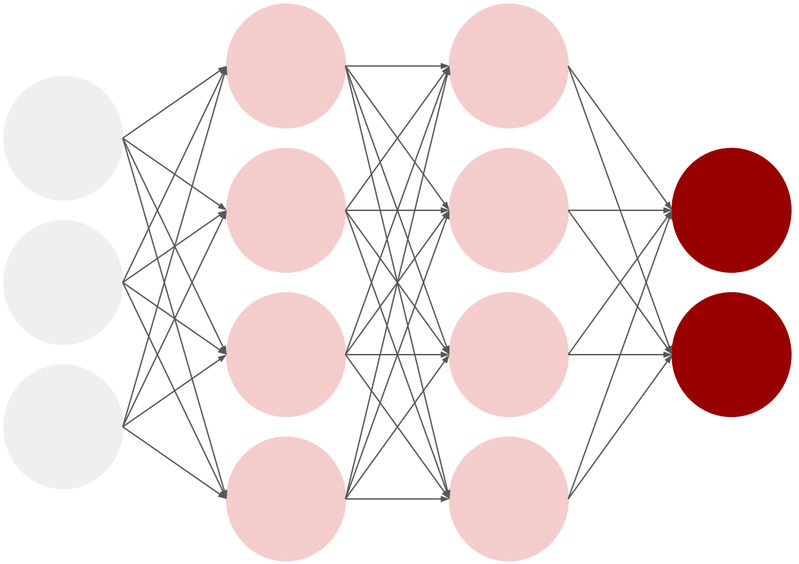
What is “**Language**” in Large Language Models?

- Human-understandable languages like English
- The models learn statistical patterns of the data
  - Example: Probability of "the" preceding the noun "cat" are very high
- These models can generate new text, mimicking Humans
- Applications:-



What is a “**Model**” in Large Language Models?

- A **model** refers to a mathematical representation of a real-world process
- It is essentially a program or an algorithm
  - Trained on a dataset to recognize patterns
  - Uses these patterns to make predictions or decisions without being explicitly programmed to perform the task
- ChatGPT uses GPT3 or GPT4 models



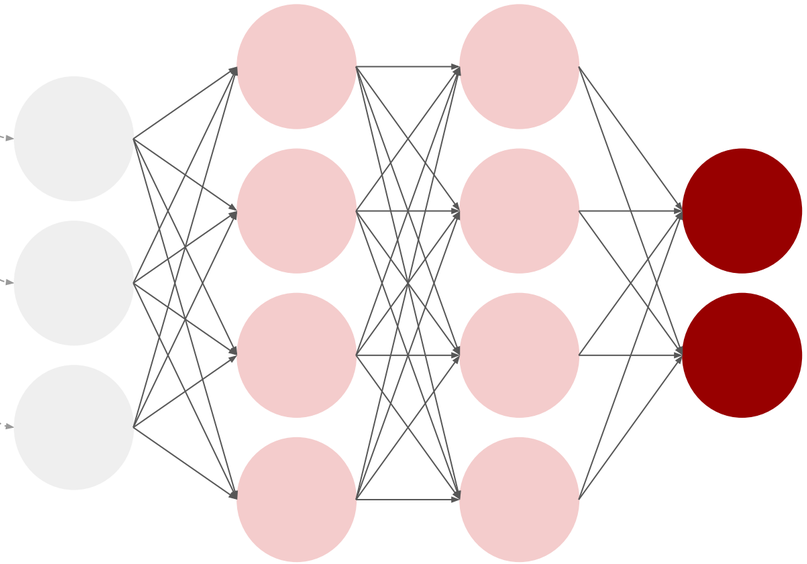
How a Large Language Model Works?

PRE-PROCESSING *(**Example Text**: She Sells Seashells)*

1) Normalize	2) Tokenize	3) Stop Words	4) Stemming & Lemmatization
she sells seashells	*[“she”, “sells”, “sea”,”sh”,”ells”]*	*[“sells”, “seashells”]*	*[“sell”, “seashell”]*

- **Text Normalization** is the process of converting text to a standard format, such as to lowercase, removing special characters, and converting numbers to their written form
- **Tokenization** is the process of breaking down text into individual units, such as words or phrases. This is an important step in preparing text data for NLP tasks
- **Stop Words** are common words that are usually removed during text processing, as they do not carry much meaning and can introduce noise or affect the results of NLP tasks. Examples of stop words include "the," "a," "an," "in," and "is"
- **Stemming and Lemmatization** are techniques used to reduce words to their base form. This helps reduce the dimensionality of the data and improve the performance of models

- LLMs are trained using a process called ***unsupervised learning***
- This involves ***feeding the model massive amounts of text data***, such as books, articles, and websites, and having the model learn the patterns and relationships between words and phrases in the text
- The model is then ***fine-tuned on a specific task***, such as language translation or text summarization



Challenges with Large Language Model

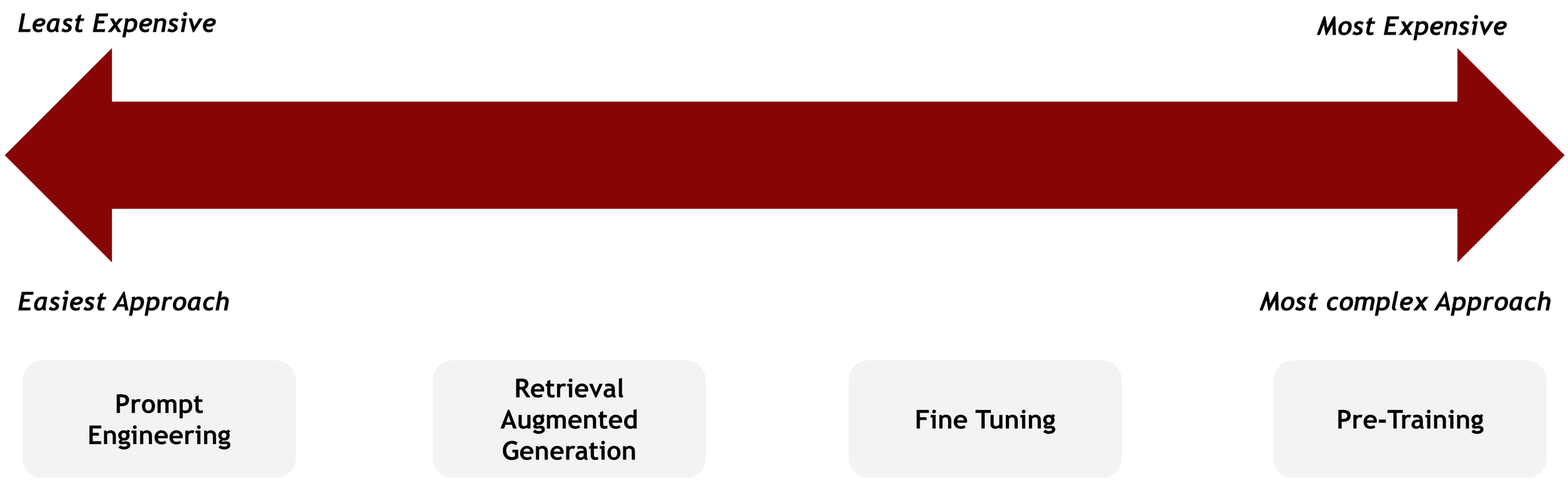
- One of the main challenges with LLMs is the **potential for offensive language**, as the models learn from the patterns found in the training data
- Unethical considerations, such as **gender and racial biases**
- **Amount of computational resources** needed to train and run LLMs, which can be expensive and energy-intensive
- **Making up things** which are not even facts
- While large language models have shown impressive performance on a variety of NLP tasks, they may not perform as well on specific tasks, such as those that require a **deeper understanding of the underlying context**

Agenda

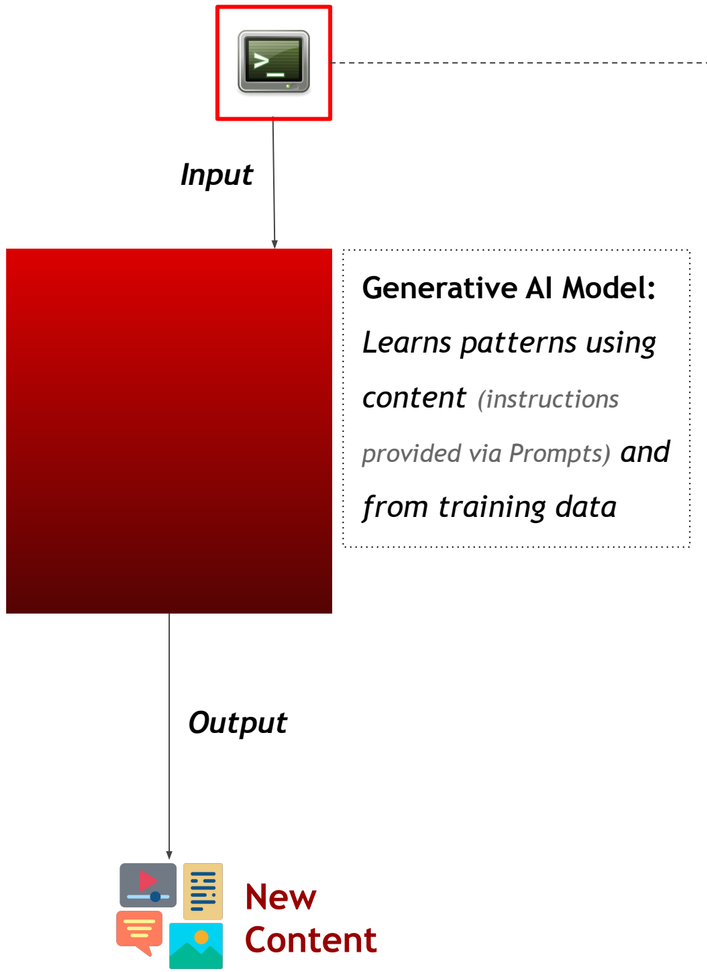
- Recap of LLMs
- **Understanding Prompt Design**
- Context, Examples, & Tuning Prompt Parameters
- Prompt Design Best Practices

Ways to customize Large Language Models

Accuracy Requirements, Skillset, Compute Resources, & Data Availability influence the choice of approach

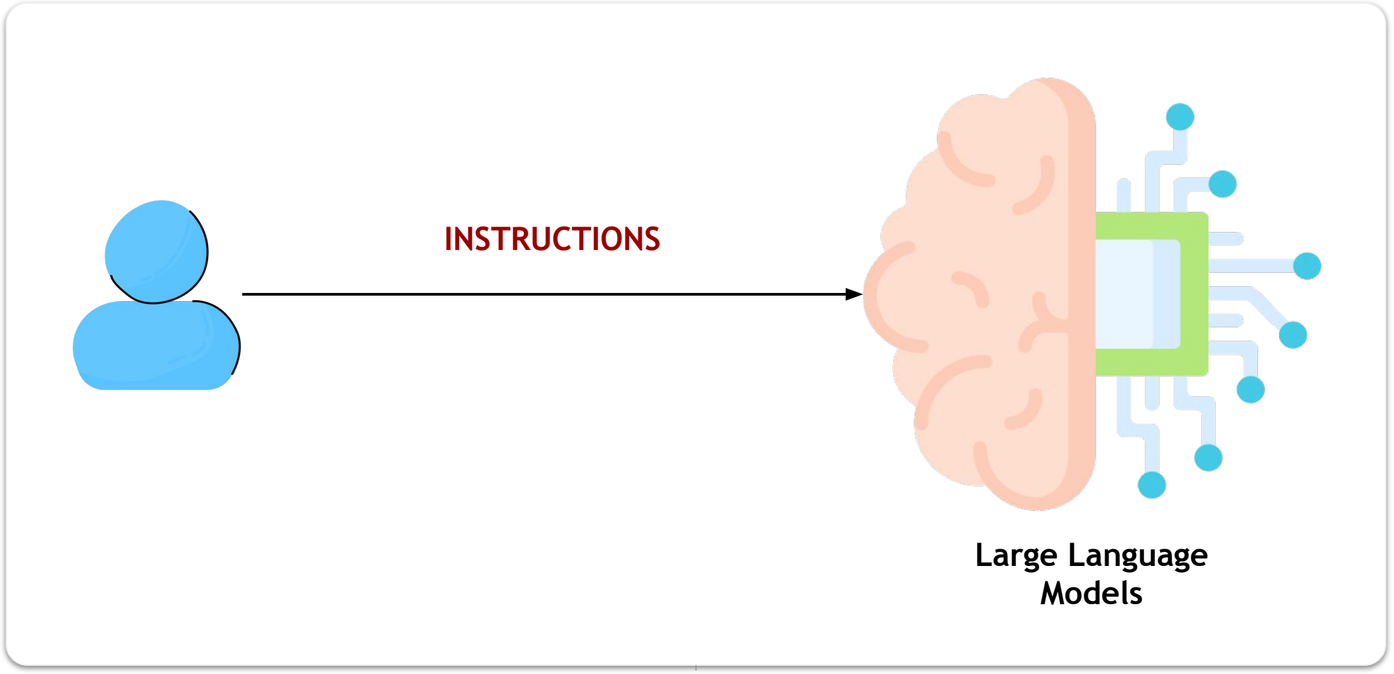


What is a Prompt?



- *A prompt refers to a specific instruction, query, or input given to an AI model to generate a response or perform a task*
- *It serves as cue, guidance, or a starting point for the AI system to produce an output based on its learned knowledge and language comprehension*
- *The form and tone of a prompt varies depending on the AI application and the desired outcome*
- *Well-defined, context-driven, and unambiguous prompts bring AI output closer to the intended results*

What is Prompt Engineering?

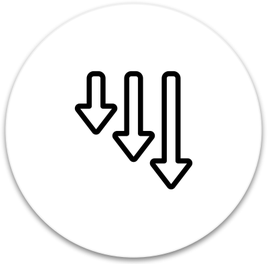


Prompt Engineering is an art of asking the right question to get the best response from an LLM

Why is Prompt Engineering Important?

Improve the Accuracy

Improve the usefulness of AI generated content



Reduce the risk of generating harmful and biased content

Key elements of a Prompt

Let’s look at aspects that make up a good prompt:

Task

Clearly specify the desired action or response from the model

Context

Provide relevant background information to help the model understand the setting or scenario

Exemplar

Offer a sample of the desired output to guide the model

Persona

Define the character or role the model should respond to

Format

Specify the structure or style of the response

Tone

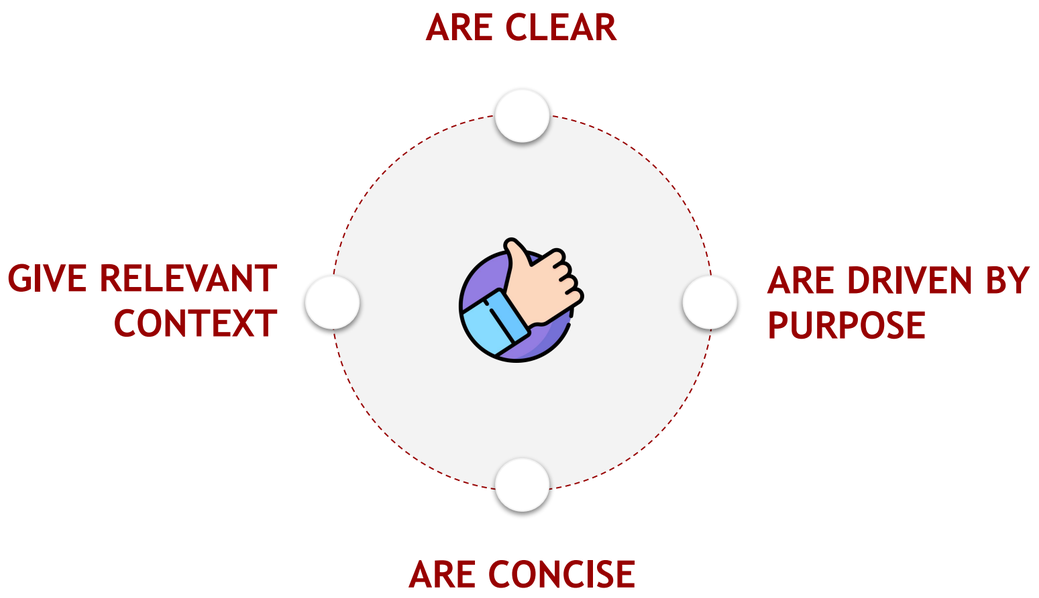
Clearly communicate the desired emotional or stylistic approach

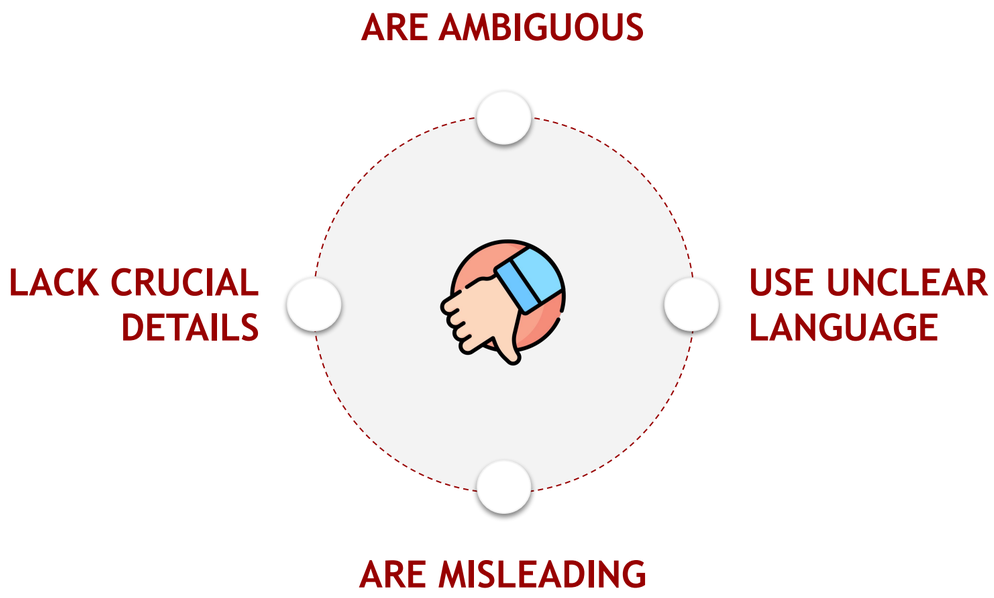
More Important

Optional

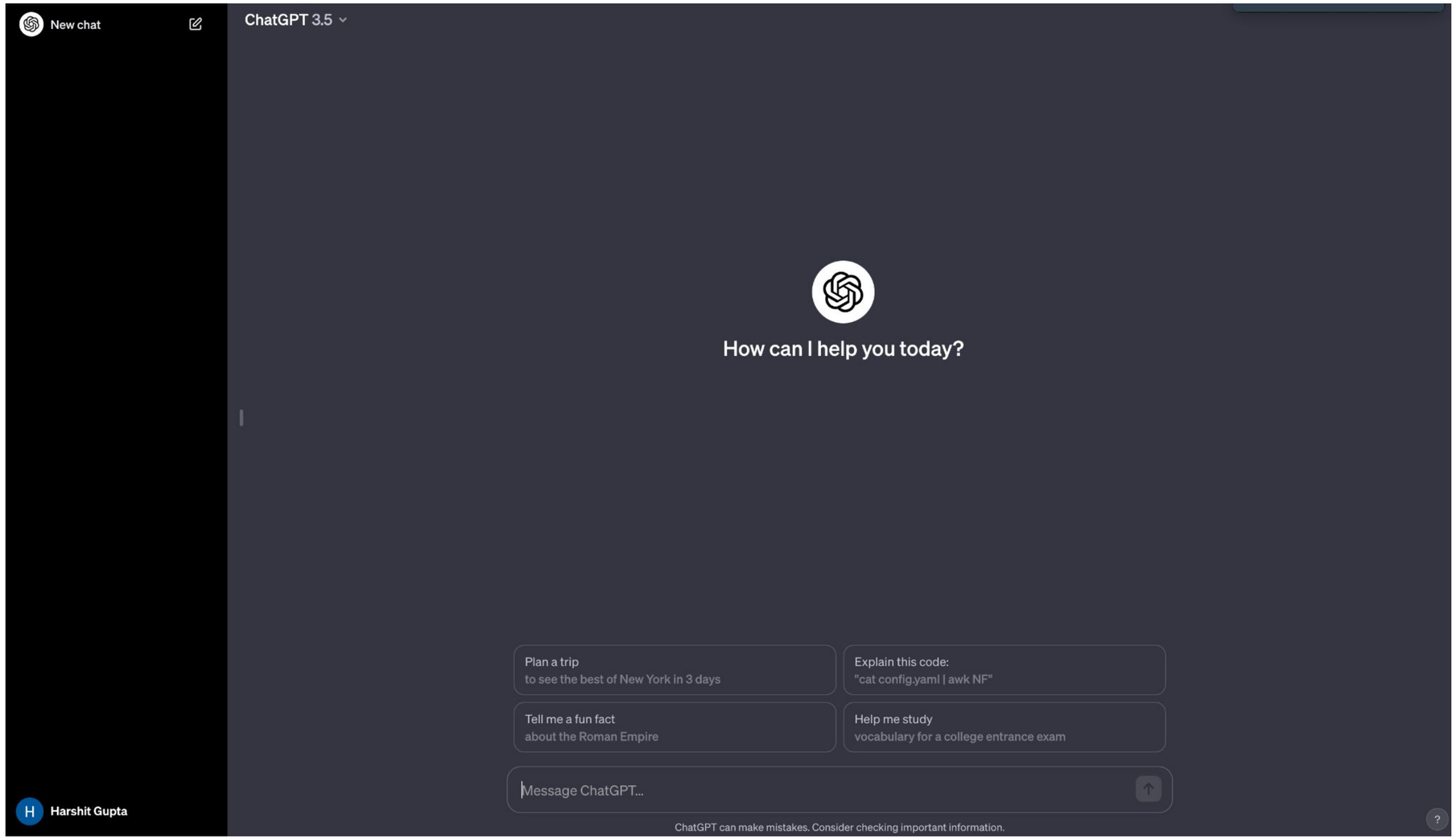
Good Prompts vs. Bad Prompts

The quality of the prompt significantly influences the output generated by the model. Here are the traits of Good & Bad Prompts:





ChatGPT UI


New chat
ChatGPT 3.5
How can I help you today?
Plan a trip
to see the best of New York in 3 days
Explain this code:
"cat config.yaml | awk NF"
Tell me a fun fact
about the Roman Empire
Help me study
vocabulary for a college entrance exam
Message ChatGPT...
ChatGPT can make mistakes. Consider checking important information.
Harshit Gupta

Tips for designing Prompts

Write Clear Instructions

Tell me something about Generative AI

Change your prompt as per following to make it clearer:

- *Include details in your query to get more relevant answers*
- *Ask the model to adopt a persona*
- *Use delimiters to clearly indicate distinct parts of the input*
- *Specify the steps required to complete a task*
- *Provide examples*
- *Specify the desired length of the output*

Act as an AI Engineer working for a multinational Consumer Packaged Goods (CPG) organization. One of the CXOs have asked you to explain about "What is Generative AI?" through a concise article, just to keep up with the latest trends in the market. Now provide me with an explanation of Generative AI using following instructions:

1. *It should be easy to understand and must be tweaked as per CPG Industry*
2. *Overall length of article should not be more than 500 words*
3. *Add relevant examples and some cases with the source links*

Tips for designing Prompts

Ask to provide reference text when required

Share some case studies related to the adoption of Generative AI in consumer packaged goods field.

Share some case studies related to the adoption of Generative AI in consumer packaged goods domain. You can take references from websites like gartner.com, mckinsey.com, forbes.com

Also add proper source links for the case studies shared in the response.

Change your prompt like this to make it clearer:

- *Instruct the model to answer using a reference text/website*
- *Instruct the model to answer with citations from a reference text*

Tips for designing Prompts

Split a complex task into simpler subtasks

In order to propel our Consumer Packaged Goods (CPG) company into the next era of operational excellence, we are seeking a comprehensive strategy to evolve from a Data-Driven Organization to an AI-Driven Organization. Please provide a detailed plan

Use the following tactics to improve the prompt:

- *Use intent classification to identify the most relevant instructions for a user query*
- *For dialog-driven applications that require very long conversations, summarize or filter previous dialog(s)*
- *Summarize long documents piecemeal and construct a complete summary, recursively*

To propel our CPG organization towards an AI-driven future, we seek a detailed strategy for transitioning from a Data-Driven to an AI-Driven Organization. Key focus areas include:

- *Data Integration and Enrichment:*
  - *Propose methods for enhanced data integration, ensuring seamless interoperability and enrichment from external sources*
- *Advanced Analytics Capabilities:*
  - *Specify tools to extract actionable insights, surpassing traditional analytics, and gaining predictive capabilities*
- *Machine Learning Applications:*
  - *Identify use cases for machine learning in decision-making, supply chain management, and demand forecasting, with a scalable implementation plan*
- *Automation and Process Enhancement:*
  - *Propose AI-driven automation solutions to streamline tasks, enhance efficiency, and address integration challenges*
- *Cultural Shift and Skill Development:*
  - *Devise a plan for fostering AI adoption culture, promoting awareness, enthusiasm, and workforce skill development*
- *Performance Monitoring and Iterative Improvement:*
  - *Establish KPIs for measuring AI initiative success, implement a monitoring framework, and detail a feedback loop for continuous improvement*

Provide a comprehensive, phased roadmap aligning with organizational goals, emphasizing smooth transition and sustained evolution. The plan should prioritize data integration, advanced analytics, machine learning, automation, cultural adaptation, and iterative improvement for a successful shift to an AI-Driven Organization

Tips for designing Prompts

Test changes systematically. Playground is a nice way to test for Developers

Evaluate model outputs with reference to gold-standard answers

You will be provided with text delimited by triple quotes that is supposed to be the answer to a question. Check if the following pieces of information are directly contained in the answer:

- Neil Armstrong was the first person to walk on the moon.
- The date Neil Armstrong first walked on the moon was July 21, 1969.

For each of these points perform the following steps:

1 - Restate the point.
2 - Provide a citation from the answer which is closest to this point.
3 - Consider if someone reading the citation who doesn't know the topic could directly infer the point. Explain why or why not before making up your mind.
4 - Write "yes" if the answer to 3 was yes, otherwise write "no".

Finally, provide a count of how many "yes" answers there are. Provide this count as {"count": <insert count here>}.

"""Neil Armstrong is famous for being the first human to set foot on the Moon. This historic event took place on July 21, 1969, during the Apollo 11 mission."""

Techniques in Prompt Engineering

Here, we see more intricate strategies that require a deeper understanding of the model's behavior

ZERO-SHOT PROMPTING

This technique involves providing the model with a task it hasn't seen during its training. It tests the model's ability to generalize and produce relevant outputs without relying on prior examples

Classify the text into neutral, negative or positive.

Text: I think the vacation was okay.

Sentiment:

Neutral

Techniques in Prompt Engineering

Here, we see more intricate strategies that require a deeper understanding of the model's behavior

FEW-SHOT PROMPTING

Here, the model is given a few examples (shots) to guide its response. By providing exemplars, the model can better understand and generate the desired output. For example, showing a model several examples of translated sentences before asking it to translate a new one

Craft a compelling tagline for social media introducing our upcoming product, a sustainable packaging initiative for our soft-drink line. Highlight the eco-friendly features and emphasize the positive impact on the environment.
Example1: Every sip is a step towards a greener future
Example2: Good for you, great for the Earth
Example3: Because flavor shouldn't cost the Earth. Share the taste, share the planet

Sip sustainably, savor responsibly

Techniques in Prompt Engineering

Here, we see more intricate strategies that require a deeper understanding of the model's behavior

CHAIN OF THOUGHTS (COT)

It involves guiding the model through a series of reasoning steps. By breaking down a complex task into intermediate steps or a "chains of reasoning," the model can achieve better language understanding and more accurate outcomes

Process the problem step-by-step and answer the question asked:

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?

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


DATA COUCH
Thank You!!
Scan to Stay Connected!