



# HUMANISING HUMANS

How can AI augment our meetings?

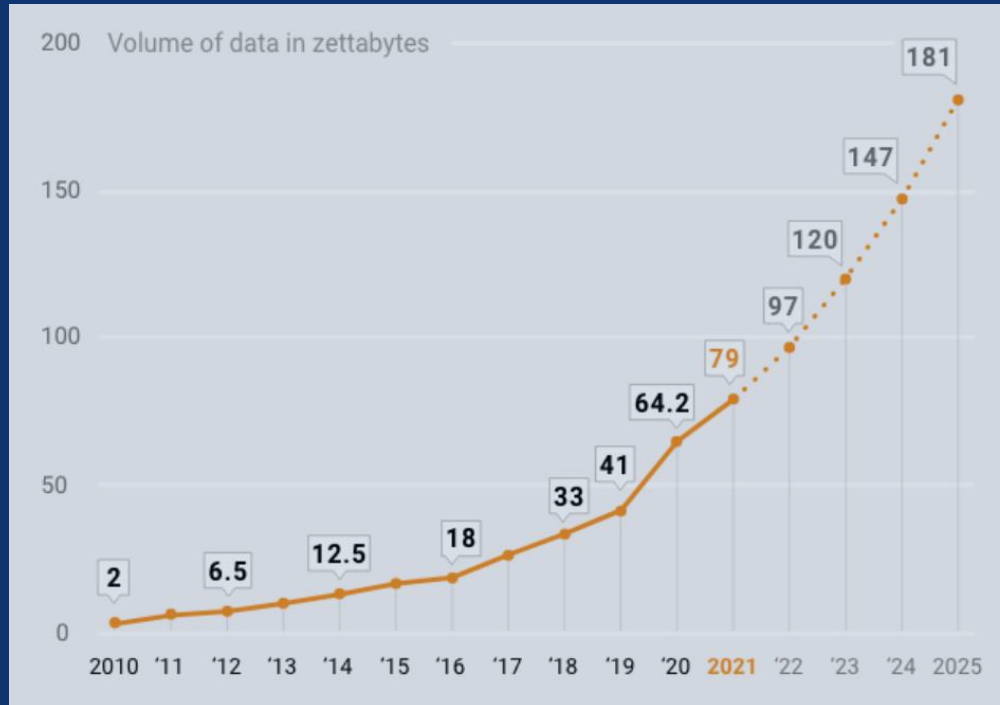
Jovan Njegić, [jovannj@diplomacy.edu](mailto:jovannj@diplomacy.edu)

Anja Đajić, [anjadj@diplomacy.edu](mailto:anjadj@diplomacy.edu)

# Data today



# Data today



# BUT.. WHAT IS ZETTABYTE?

Megabyte

1,000,000 bytes

Gigabyte

1,000,000,000 bytes

Terabyte

1,000,000,000,000 bytes

Petabyte

1,000,000,000,000,000 bytes

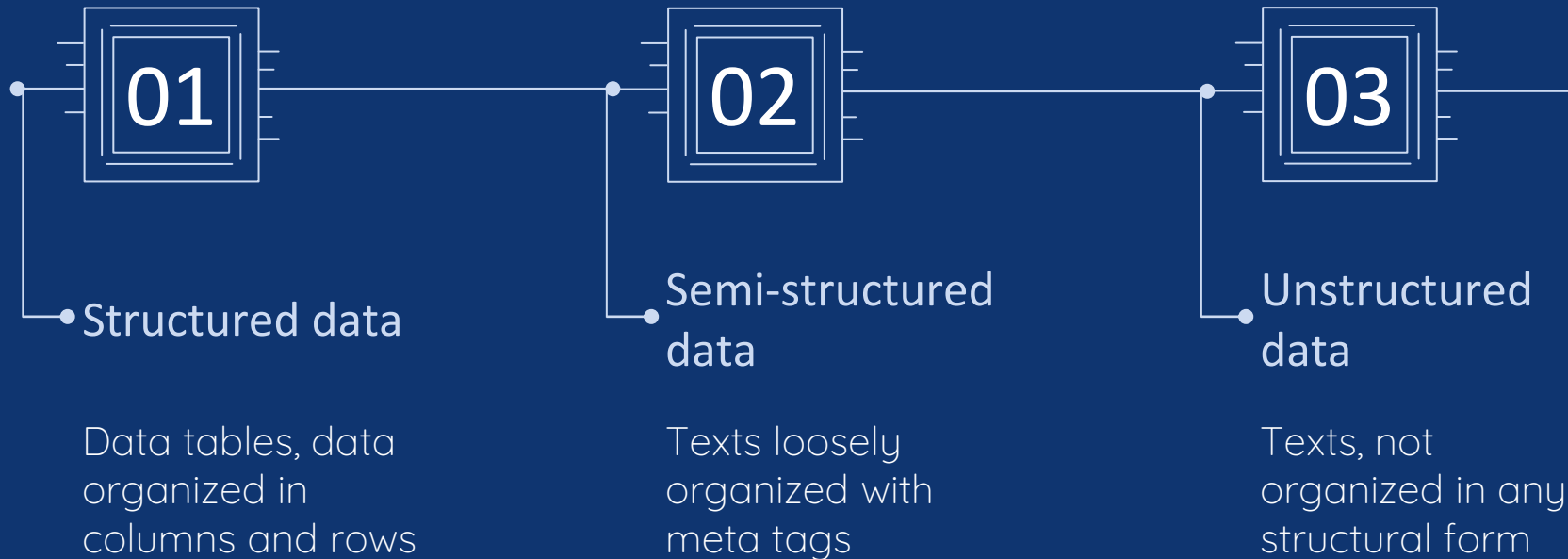
Exabyte

1,000,000,000,000,000,000 bytes

Zettabyte

1,000,000,000,000,000,000,000 bytes

# TYPES OF DATA

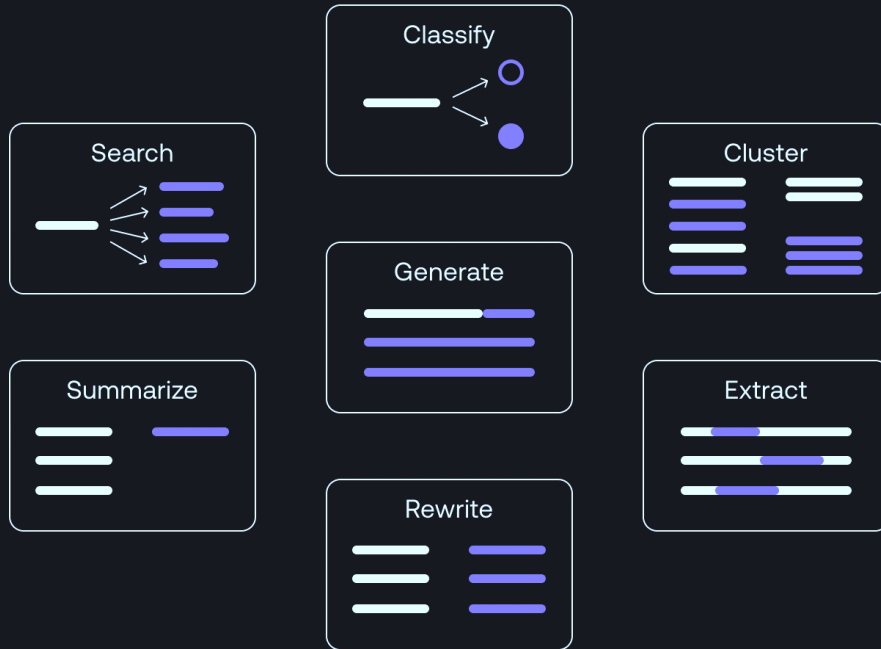




# What is NLP?

A branch of  
**ARTIFICIAL INTELLIGENCE**  
that focuses on the  
**INTERACTION**  
between  
**COMPUTERS**  
and  
**HUMAN LANGUAGE**

# NLP tasks





## Unsupervised Pre-training

Input (features)

a

robot

must



GPT-3  
(under training)



Output (Prediction)

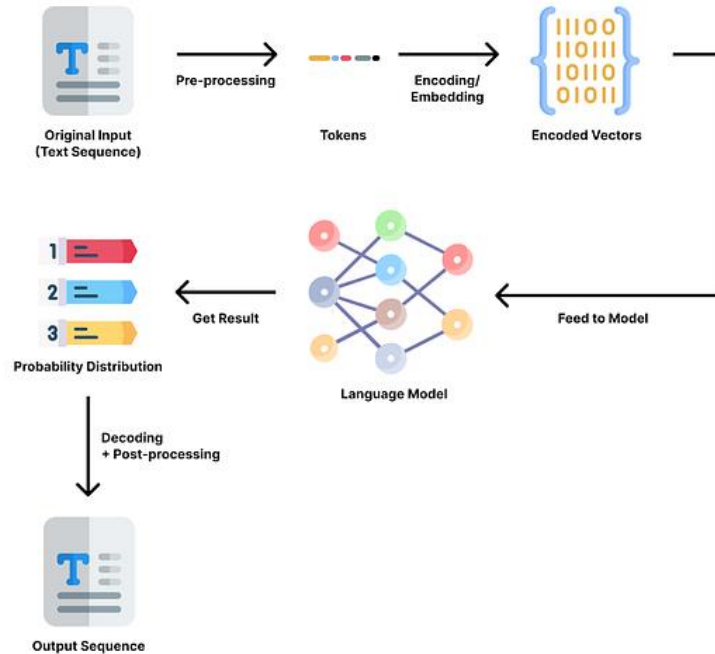
Correct output (label):

obey

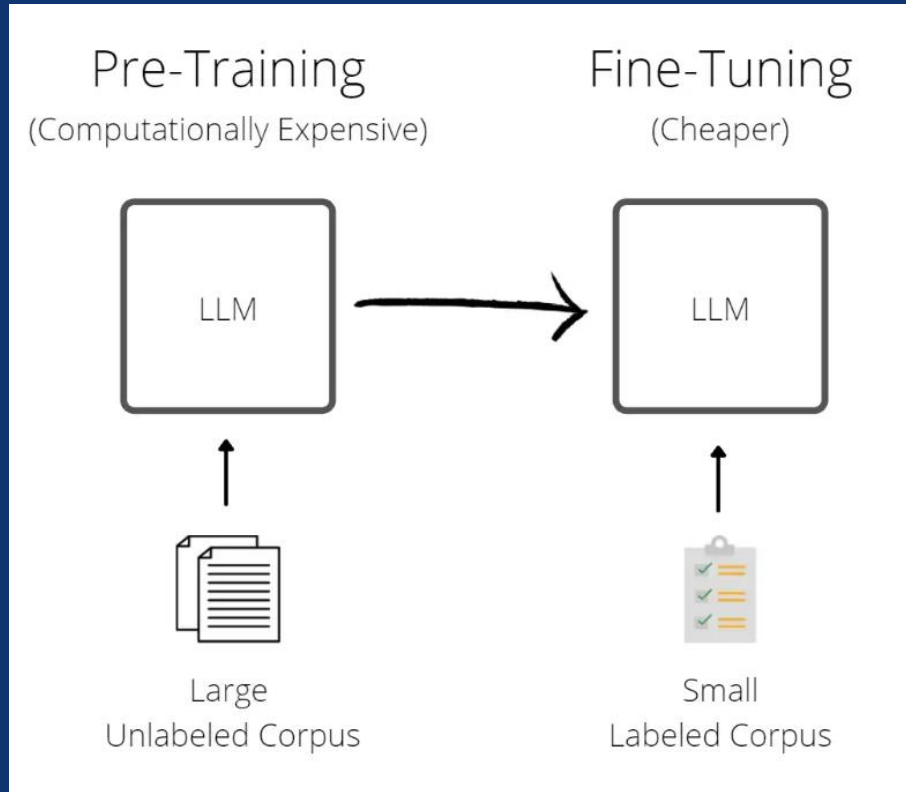
# Large Language Models



# Large Language Models



# Large Language Models



# Large Language Models

## LLM Market Map

### INFRASTRUCTURE

Inflection

 OpenAI

 Hugging Face

Adept Google

 Meta co:here

Tencent

 百度

AI21 labs



### APPLICATIONS

#### Search

YOU

 Twelve Labs

 Hebbia

ZIR AI

 Constructor.io

 DASHWORKS

#### Synthesis

 Mintlify

veezoo

 ZEBRIUM

DELV

artifct

 CopyMonkey

UNSCRAMBL

Nabla

#### Generation

 tavus

 copysmith

 textio

WRITER

 Jasper

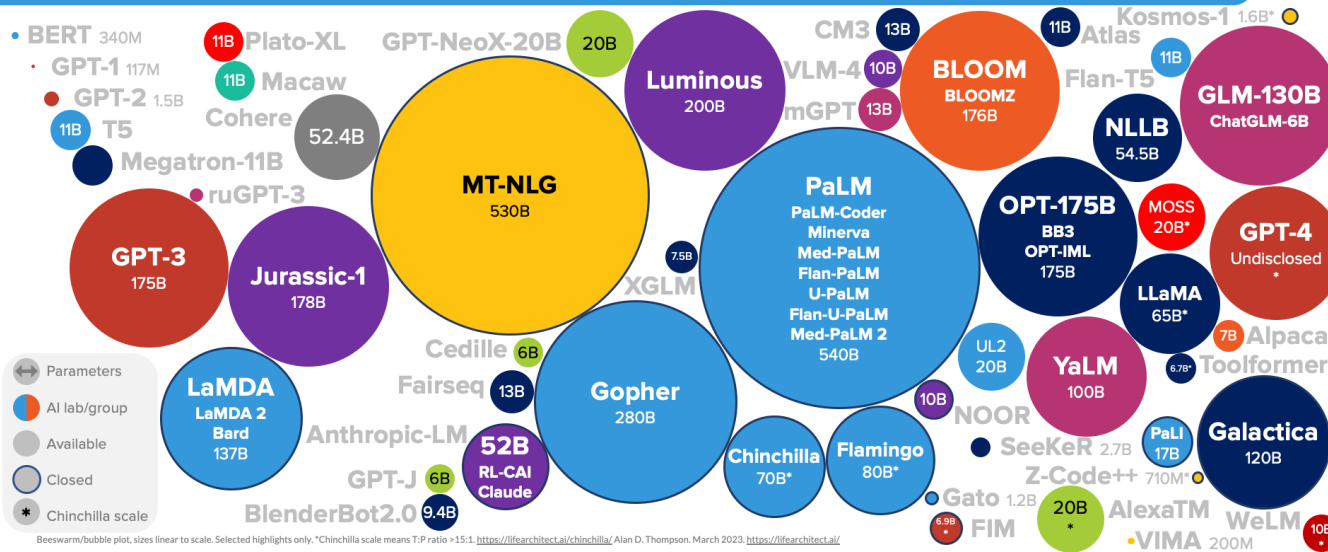
 GitHub Copilot



copy.ai  
anyword

# Large Language Models

## LANGUAGE MODEL SIZES TO MAR/2023



# Reinforcement Learning with Human Feedback

Step 1

**Collect demonstration data and train a supervised policy.**

A prompt is sampled from our prompt dataset.



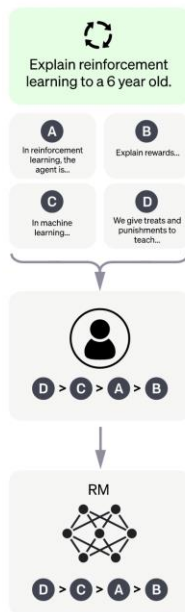
A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

**Collect comparison data and train a reward model.**

A prompt and several model outputs are sampled.



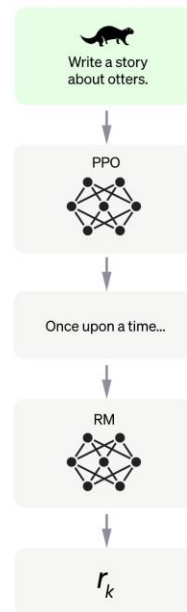
A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

Step 3

**Optimize a policy against the reward model using the PPO reinforcement learning algorithm.**

A new prompt is sampled from the dataset.



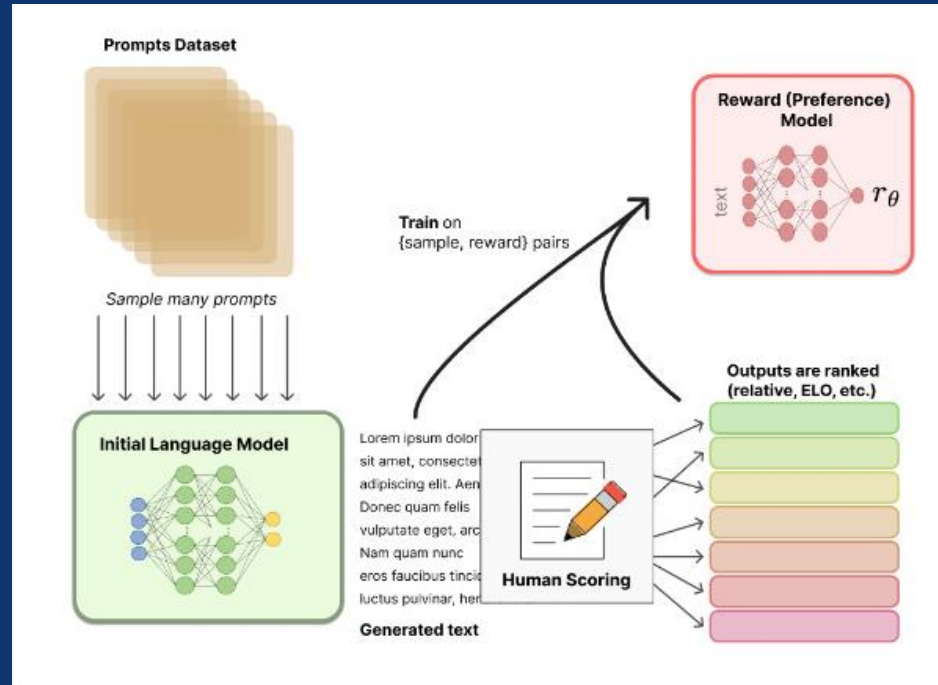
The PPO model is initialized from the supervised policy.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

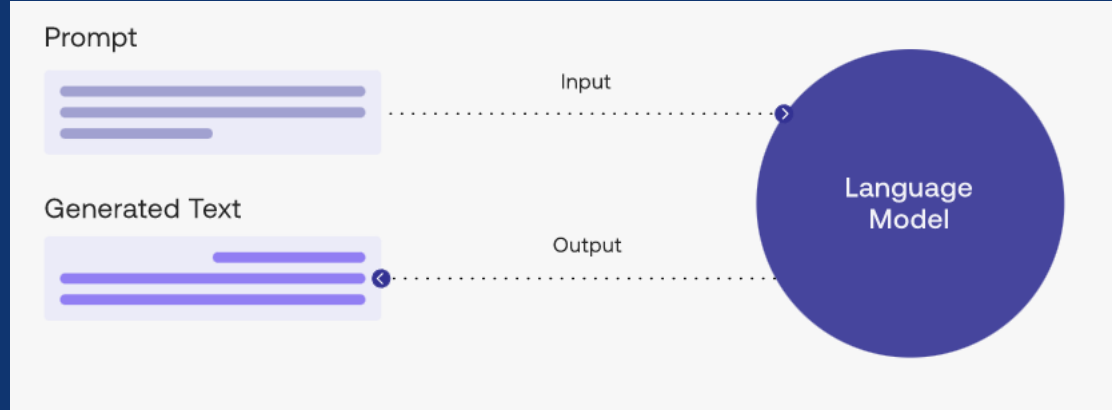
# Reinforcement Learning with Human Feedback





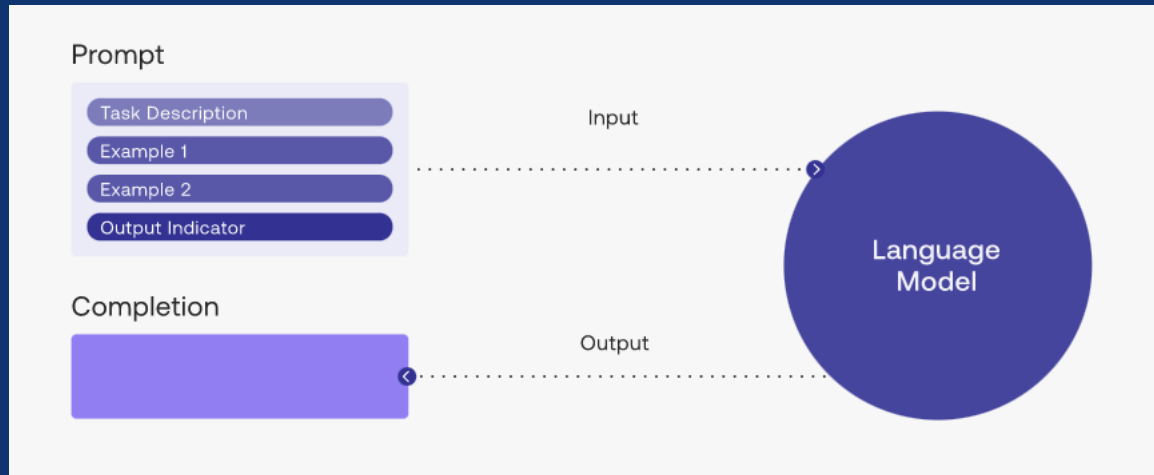
# Prompting Large Language Models

- Prompting refers to the process of providing input to an LLM in order to get a specific type of output
- The choice of prompt can **greatly affect** the **quality** of the LLM's output, and careful consideration should be given to the prompt design

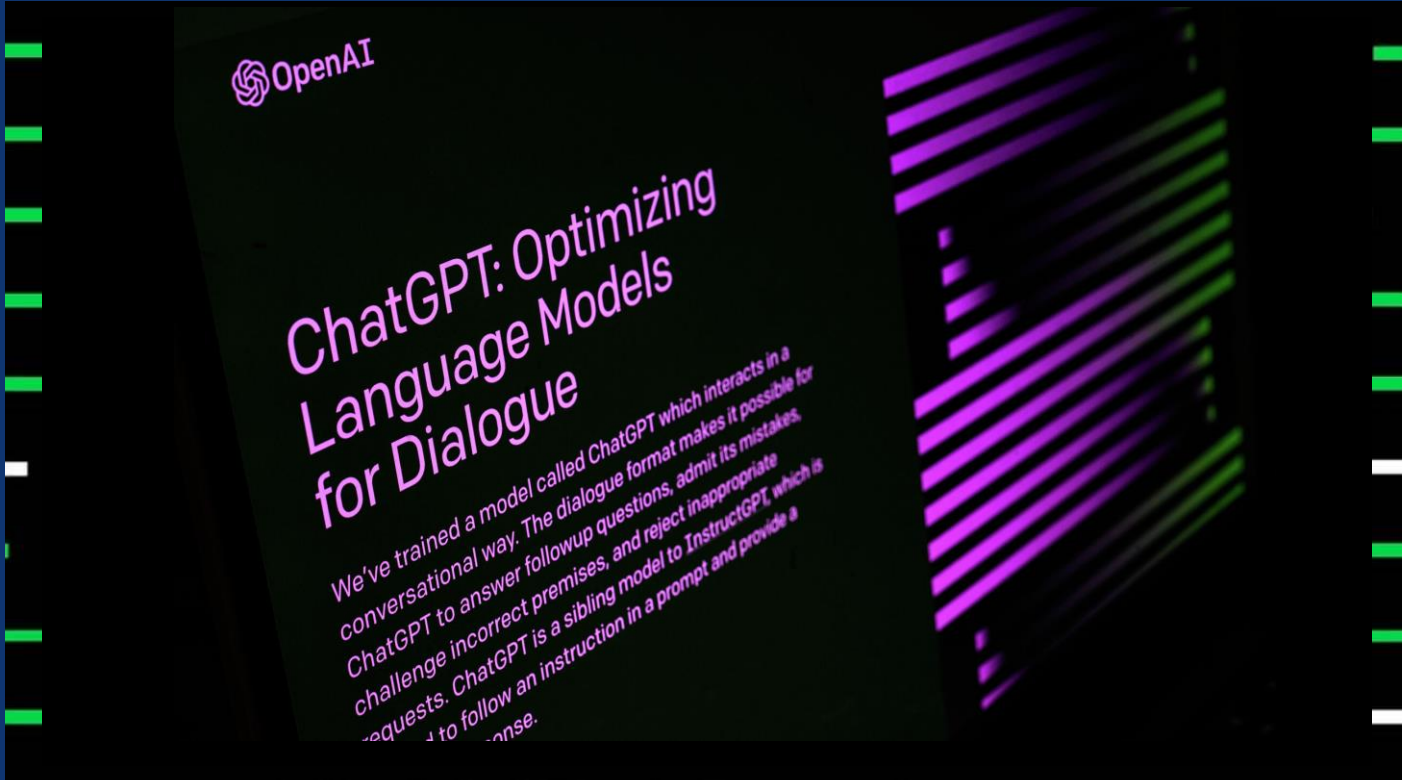


# Prompting Large Language Models

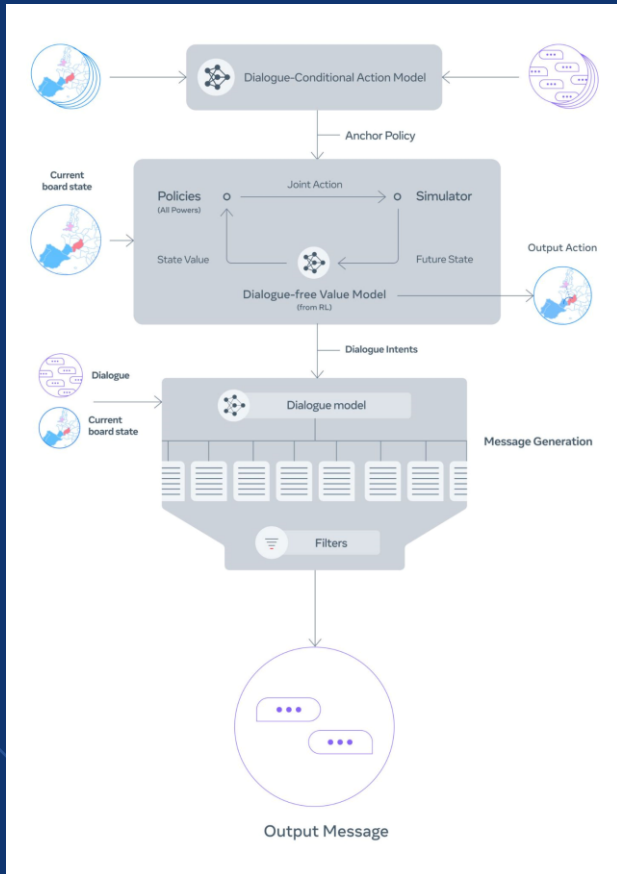
- A prompt guides the model to generate useful output



# GPT3 vs ChatGPT vs GPT4?



# Cicero



Autumn 1905 Movement, enter your orders.

18 5h remaining for Autumn 1905 Movement

ENG: 1 FRA: 7 ITA: 4 GER: 9 AUS: 2 TUR: 8 RUS: 3

AUTUMN 1905

PRESS INFO ORDERS MY GAMES HELP

Gunboat 101-12

Pl: 107 Autumn 1905 Movement  
Classic, No messaging, Anonymous players, Draw Size Scoring, Hidden draw votes, Anonymous, Gunboat (no press)  
18 12h / phase

POWER	STATUS	UNITS	CENTERS	BET	DELAYS
ENGLAND	🔒	1	1	101	1/1
FRANCE	🔒	7	7	101	1/1
ITALY	🔒	4	4	101	1/1
GERMANY	🔒	9	9	101	1/1
AUSTRIA	🔒	2	2	101	1/1
TURKEY	🔒	8	8	101	1/1
RUSSIA	🔒	3	3	101	1/1

# OTHER APPLICATIONS

1. Data Sandbox
2. How can flags help us understand AI?
3. Countries – Companies comparison



# Thank you!

Reach us at  
[ai@diplomacy.edu](mailto:ai@diplomacy.edu)

