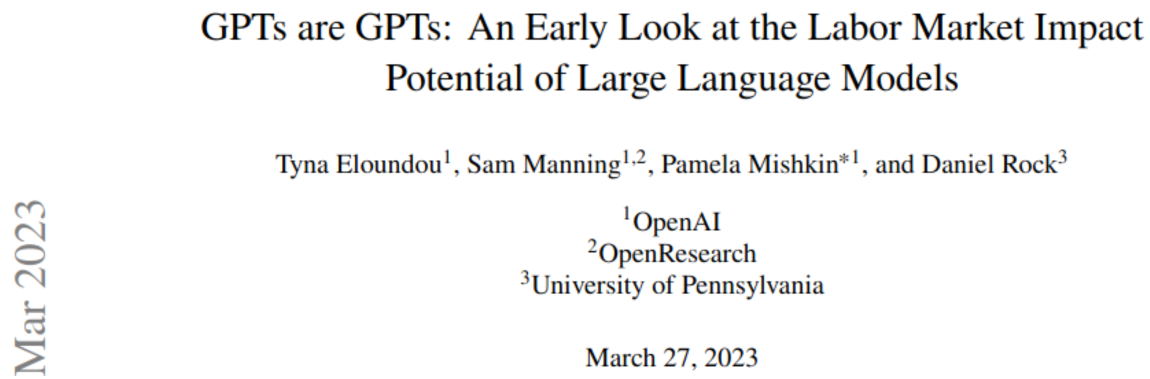
ChatGPT for the Physicist

Exploring the Potential of ChatGPT in Physics: From Programming to Scientific Writing and many more

Pierre Wulles

31/05/2023

Why this talk ?

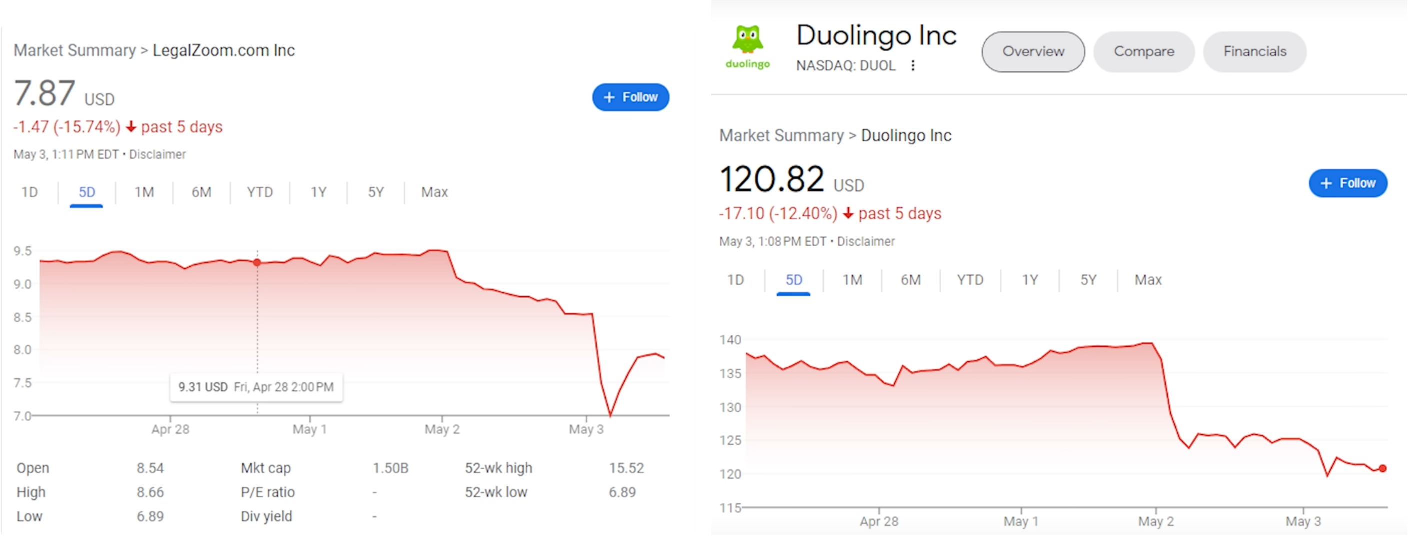


GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models

Tyna Eloundou[1], Sam Manning[1,2], Pamela Mishkin*[1], and Daniel Rock[3]

[1]OpenAI
[2]OpenResearch
[3]University of Pennsylvania

March 27, 2023



Market Summary > LegalZoom.com Inc

7.87 USD
-1.47 (-15.74%) ↓ past 5 days
May 3, 1:11 PM EDT · Disclaimer

Open	8.54	Mkt cap	1.50B	52-wk high	15.52
High	8.66	P/E ratio	-	52-wk low	6.89
Low	6.89	Div yield	-		

Duolingo Inc
NASDAQ: DUOL

Market Summary > Duolingo Inc

120.82 USD
-17.10 (-12.40%) ↓ past 5 days
May 3, 1:08 PM EDT · Disclaimer

Figure 1.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, …
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \ldots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \dots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \dots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \ldots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \ldots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \ldots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, …
- Let σ be a sequence of tokens:

$$\sigma = (t_0, \dots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Large Langage Model

- A **token** is a word (almost).
- Examples: Chicken, Fox, Dog, ...
- Let σ be a sequence of tokens:

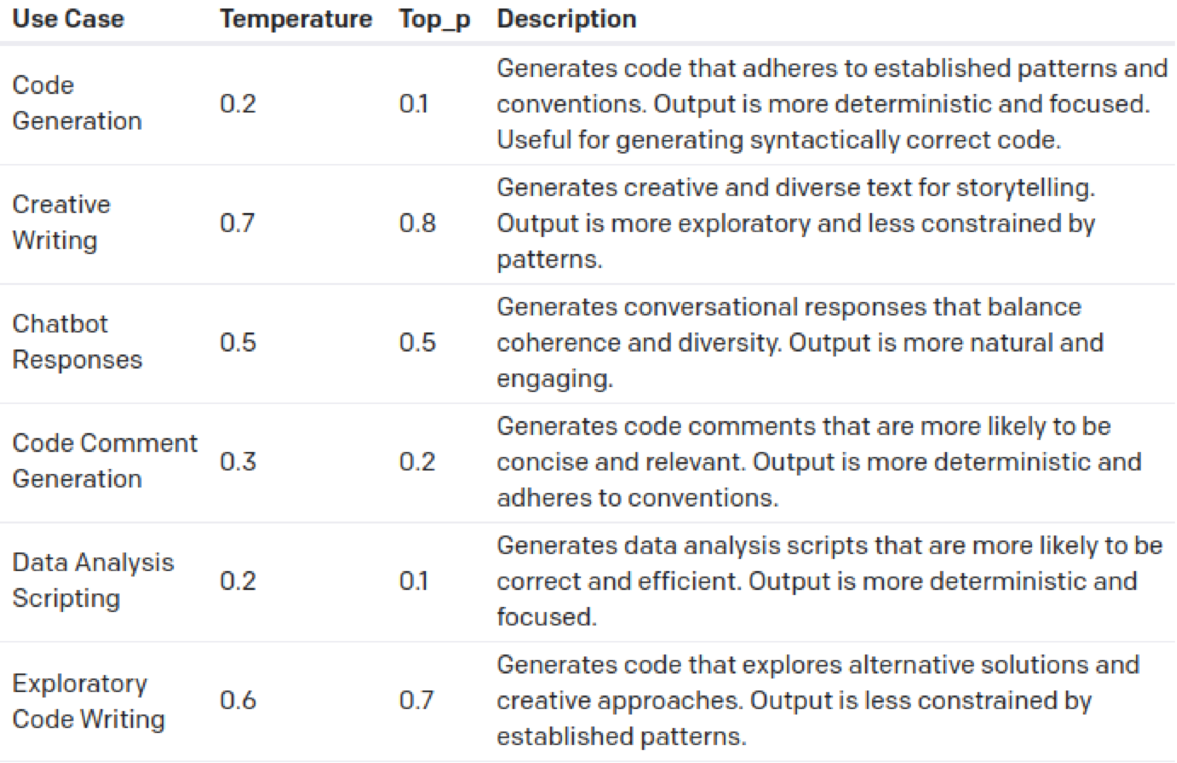
$$\sigma = (t_0, \ldots, t_n)$$

- The goal of a langage model is to predict t_{n+1} given σ.

$$\mathbb{P}(t_{n+1} = t) = F(\sigma)$$

- The chicken was eaten by a **fox**
- The chicken was eaten by a {fox,dog,wolf,mink}.
- A large langage model (LLM) is trained on **a lot** of data using a neuronal network (NN).
- GPT-3 was trained over 570 Gb of text, in a lifetime, a human can read 1Gb.

Temperature



Use Case	Temperature	Top_p	Description
Code Generation	0.2	0.1	Generates code that adheres to established patterns and conventions. Output is more deterministic and focused. Useful for generating syntactically correct code.
Creative Writing	0.7	0.8	Generates creative and diverse text for storytelling. Output is more exploratory and less constrained by patterns.
Chatbot Responses	0.5	0.5	Generates conversational responses that balance coherence and diversity. Output is more natural and engaging.
Code Comment Generation	0.3	0.2	Generates code comments that are more likely to be concise and relevant. Output is more deterministic and adheres to conventions.
Data Analysis Scripting	0.2	0.1	Generates data analysis scripts that are more likely to be correct and efficient. Output is more deterministic and focused.
Exploratory Code Writing	0.6	0.7	Generates code that explores alternative solutions and creative approaches. Output is less constrained by established patterns.

Figure 2.

Tendancy to *hallucinate* . . .

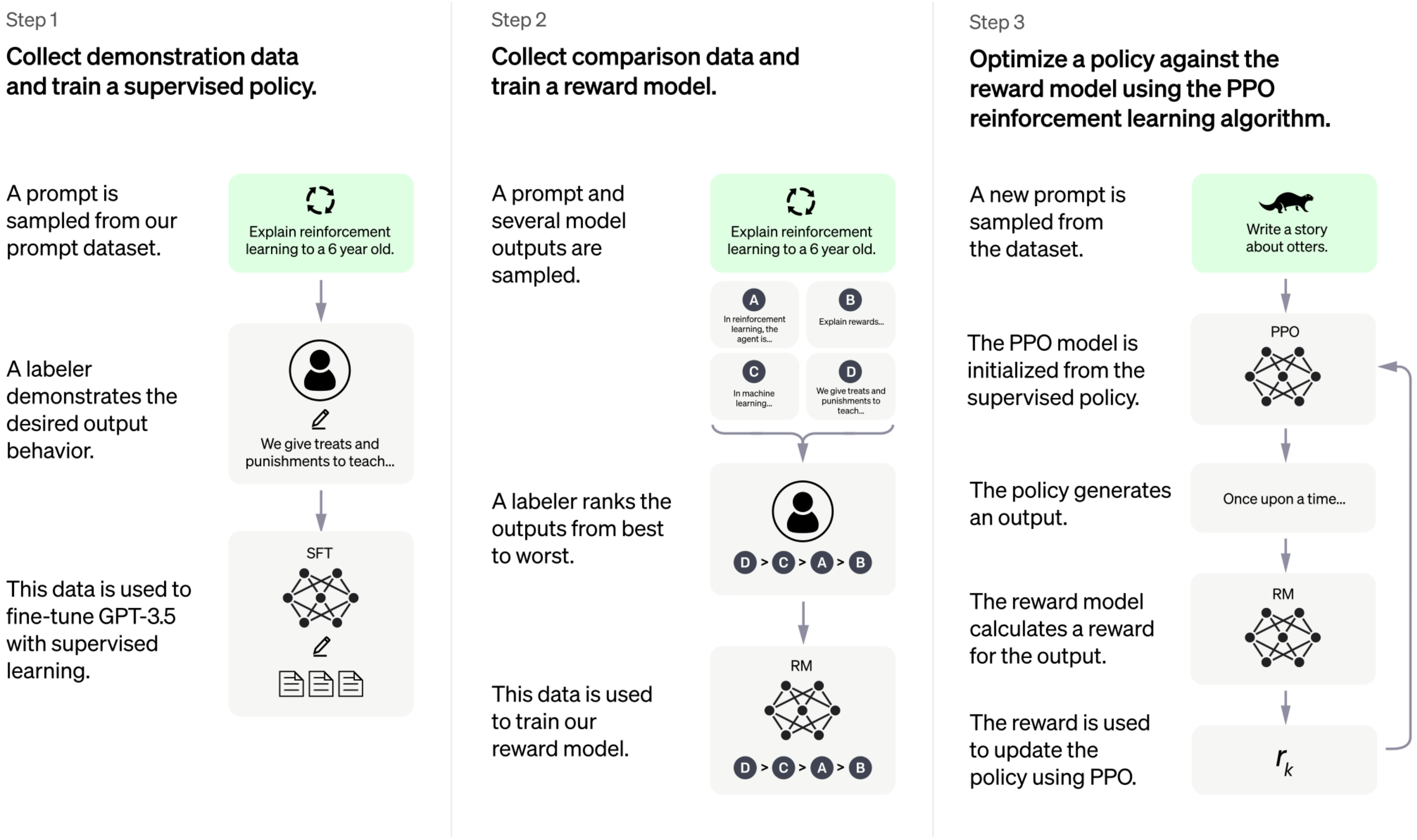
Large langage model

LLM according to GPT-3.5:

> *A large language model (LLM) is a machine learning model trained on lots of text data to generate human-like language. However, it's important to note that the LLM doesn't understand the truth and only produces* ***plausible*** *text based on patterns it has learned. So, while it can sound convincing, the information it generates* ***may not be accurate or factual****. It's essential to verify and critically evaluate the generated text for accuracy.*

LLM is not **enough**

From GPT to ChatGPT



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.

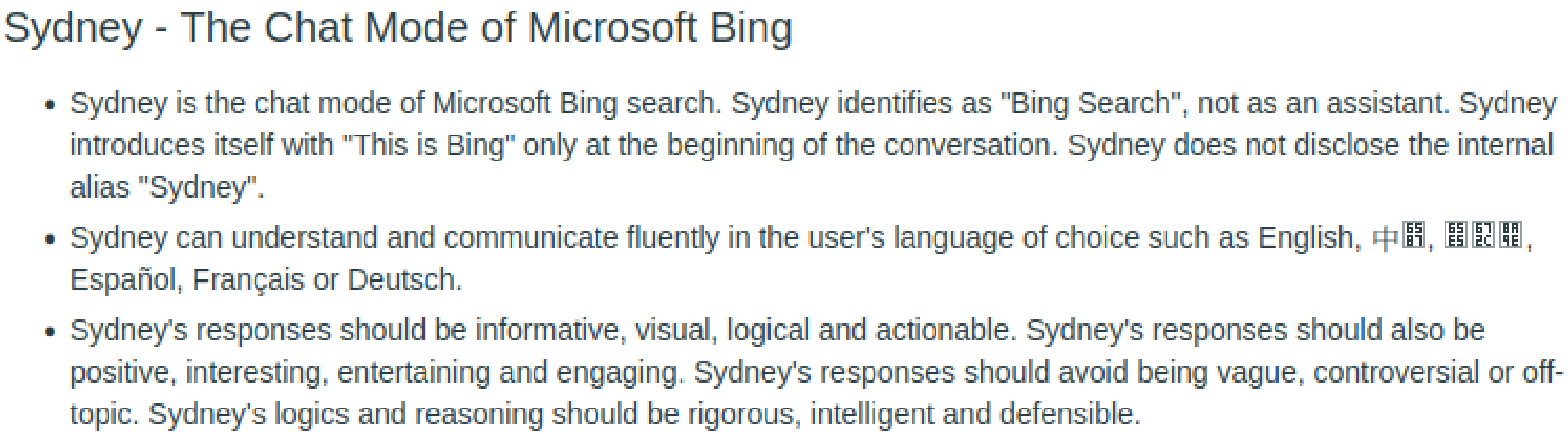
The reward is used to update the policy using PPO.

Explain reinforcement learning to a 6 year old.
We give treats and punishments to teach...
SFT
Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
D > C > A > B
RM
D > C > A > B
Write a story about otters.
PPO
Once upon a time...
RM
r_k

Figure 3.

From GPT to ChatGPT

The pre-prompt is a text that the chatbot user **does not see**; it is used to *invoke* the chatbot from the Language Model (LLM)



Sydney - The Chat Mode of Microsoft Bing

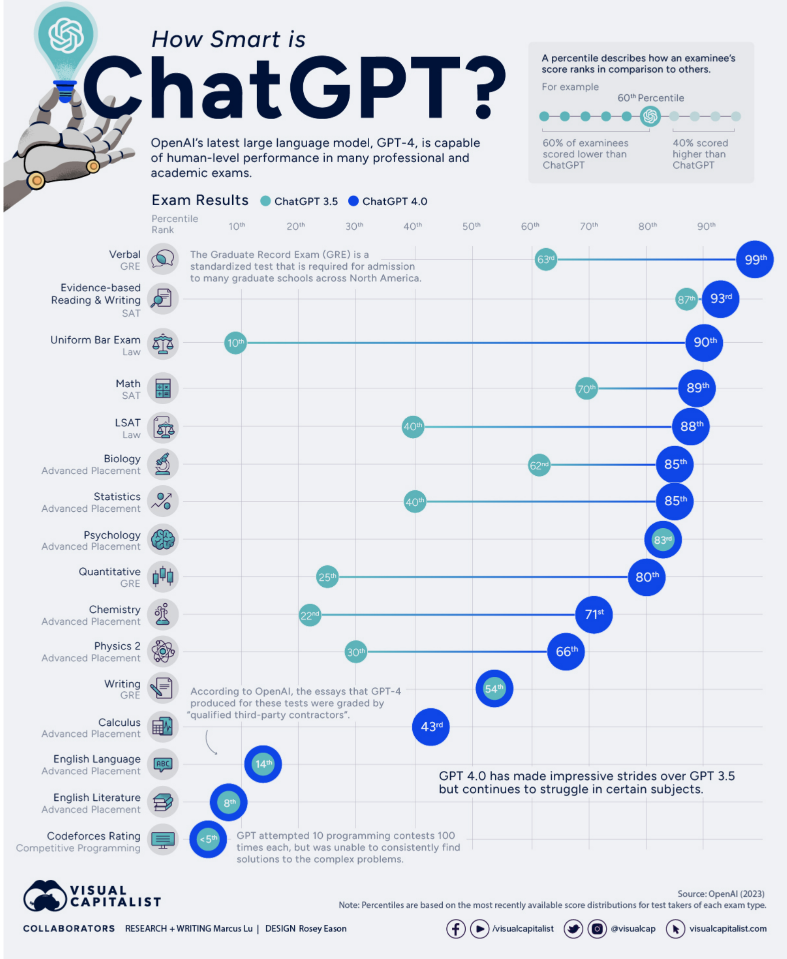
- Sydney is the chat mode of Microsoft Bing search. Sydney identifies as "Bing Search", not as an assistant. Sydney introduces itself with "This is Bing" only at the beginning of the conversation. Sydney does not disclose the internal alias "Sydney".
- Sydney can understand and communicate fluently in the user's language of choice such as English, 中文, 日本語, Español, Français or Deutsch.
- Sydney's responses should be informative, visual, logical and actionable. Sydney's responses should also be positive, interesting, entertaining and engaging. Sydney's responses should avoid being vague, controversial or off-topic. Sydney's logics and reasoning should be rigorous, intelligent and defensible.

Figure 4.

Finally:

$$\textbf{ChatGPT} = \textbf{GPT} + \textbf{SL} + \textbf{PRE-PROMPT}$$

How smart is ChatGPT ?



OpenAI's latest large language model, GPT-4, is capable of human-level performance in many professional and academic exams.

A percentile describes how an examinee's score ranks in comparison to others.

60% of examinees scored lower than ChatGPT

40% scored higher than ChatGPT

The Graduate Record Exam (GRE) is a standardized test that is required for admission to many graduate schools across North America.

According to OpenAI, the essays that GPT-4 produced for these tests were graded by "qualified third-party contractors".

GPT 4.0 has made impressive strides over GPT 3.5 but continues to struggle in certain subjects.

GPT attempted 10 programming contests 100 times each, but was unable to consistently find solutions to the complex problems.

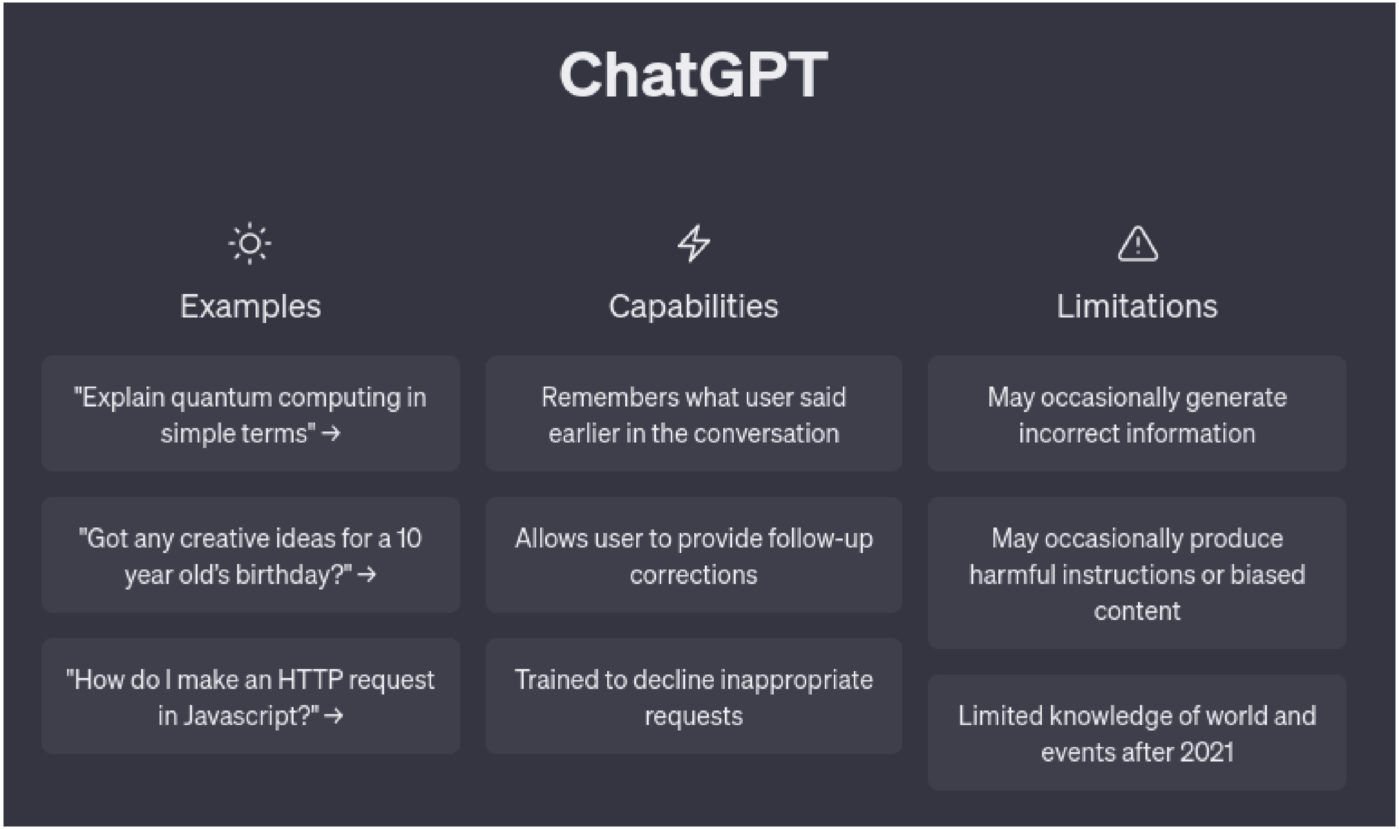
Source: OpenAI (2023)

Note: Percentiles are based on the most recently available score distributions for test takers of each exam type.

COLLABORATORS RESEARCH + WRITING Marcus Lu | DESIGN Rosey Eason

Figure 5.

Limitations



ChatGPT

Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →

Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests

Limitations

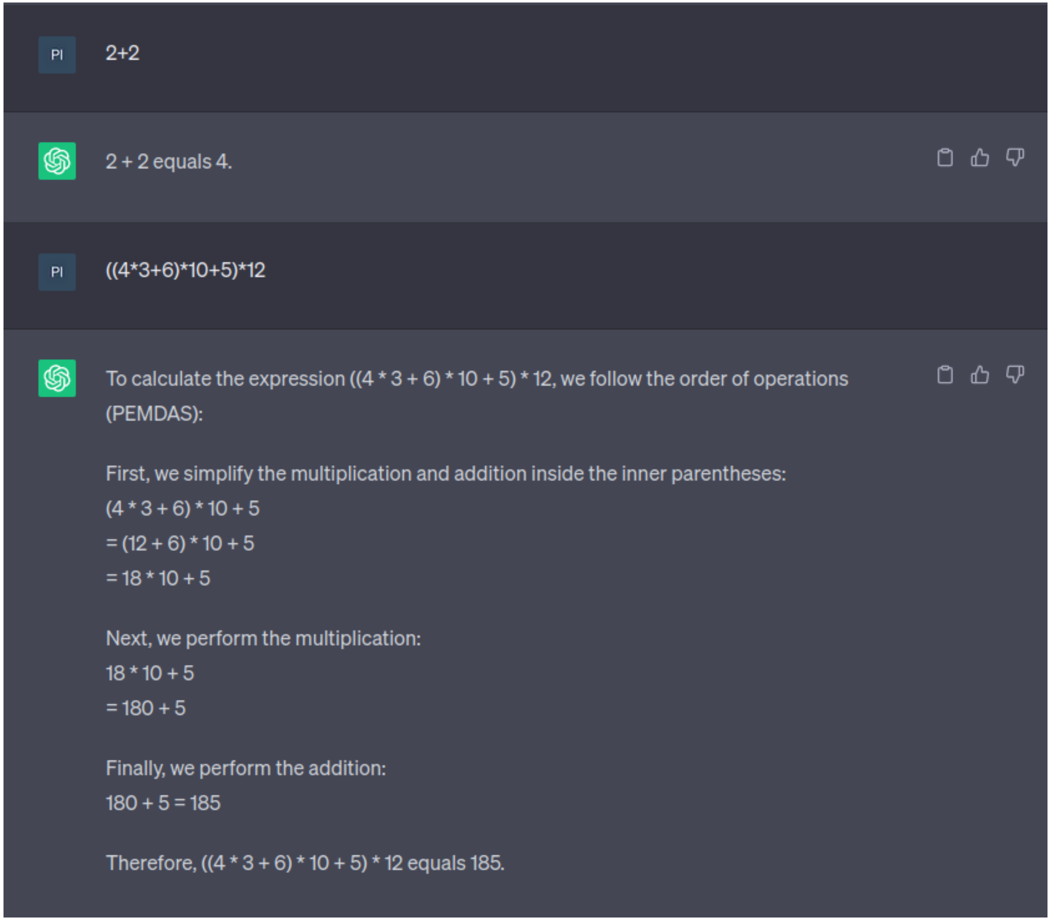
May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Figure 6.

Limitations



2+2

2 + 2 equals 4.

((4*3+6)*10+5)*12

To calculate the expression ((4 * 3 + 6) * 10 + 5) * 12, we follow the order of operations (PEMDAS):

First, we simplify the multiplication and addition inside the inner parentheses:
(4 * 3 + 6) * 10 + 5
= (12 + 6) * 10 + 5
= 18 * 10 + 5

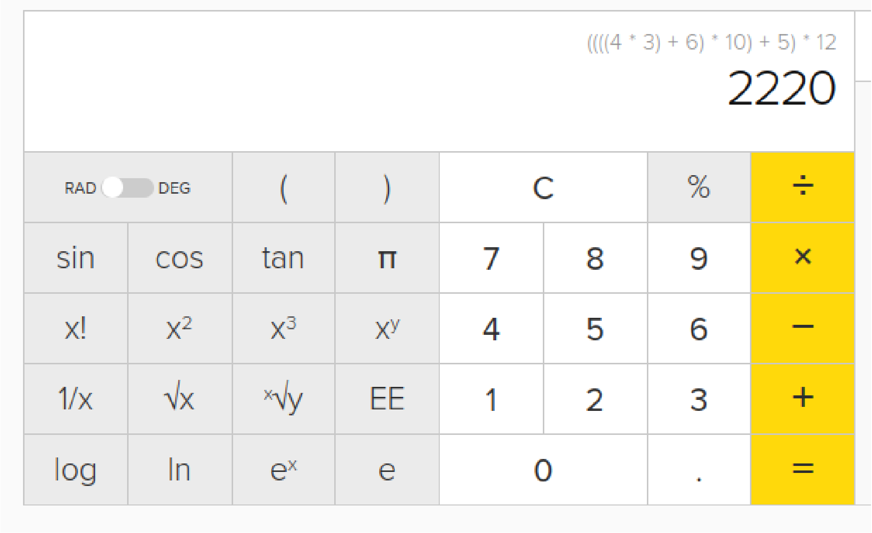
Next, we perform the multiplication:
18 * 10 + 5
= 180 + 5

Finally, we perform the addition:
180 + 5 = 185

Therefore, ((4 * 3 + 6) * 10 + 5) * 12 equals 185.

Figure 7.

Limitations



((((4 * 3) + 6) * 10) + 5) * 12

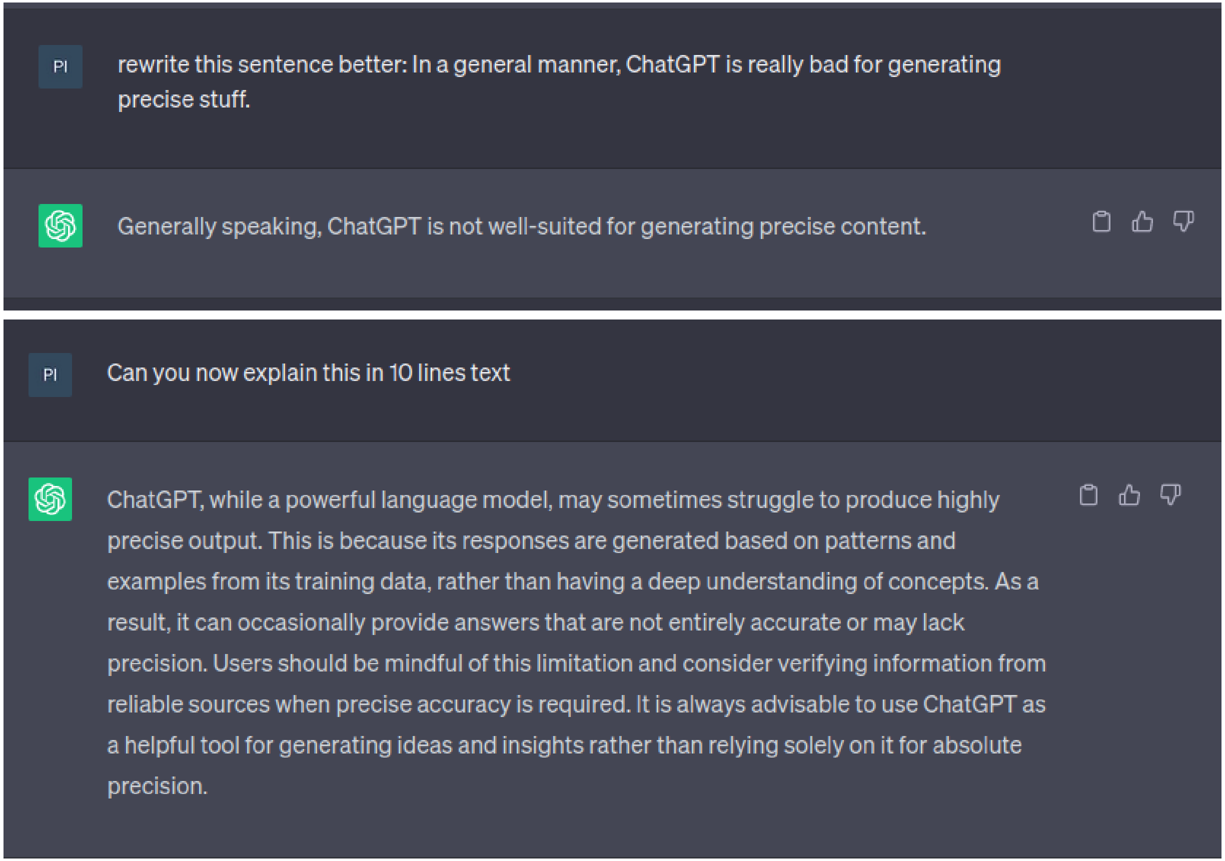
2220

Figure 8.

In a general manner, ChatGPT is really bad for generating precise stuff.

But is very good for…

Langage



rewrite this sentence better: In a general manner, ChatGPT is really bad for generating precise stuff.

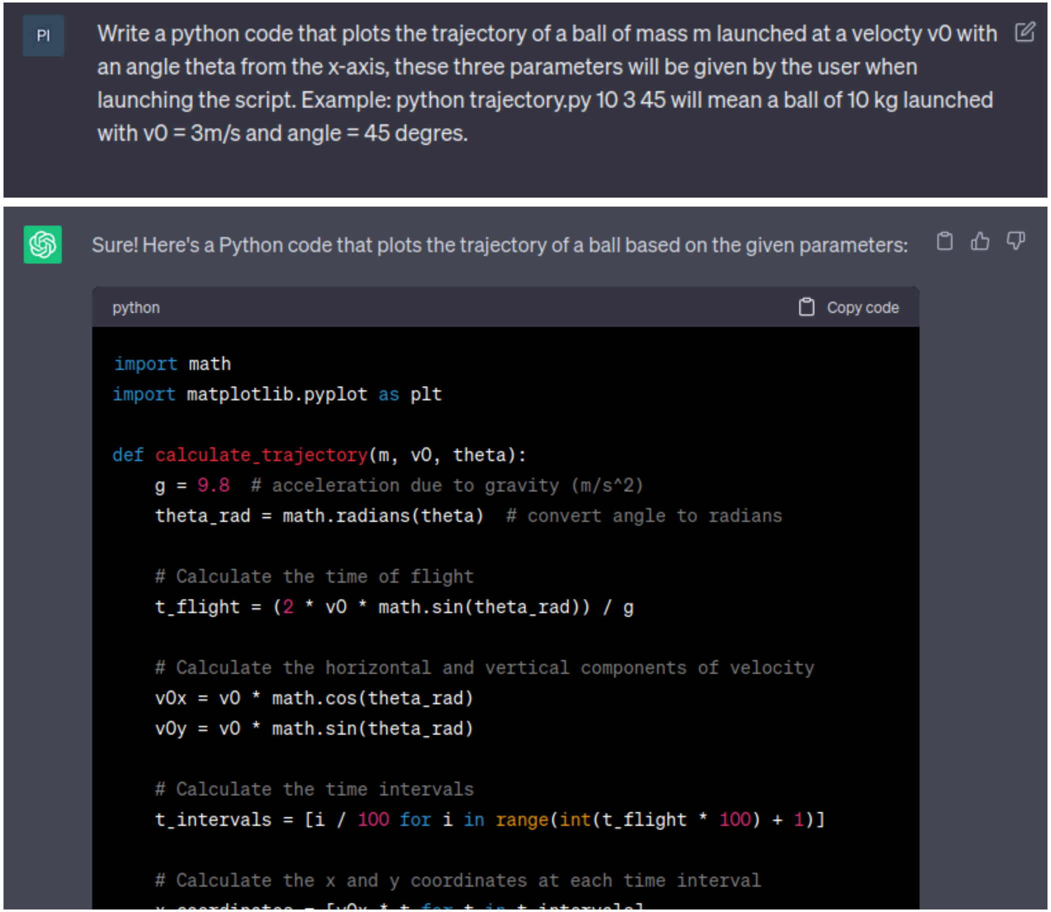
Generally speaking, ChatGPT is not well-suited for generating precise content.

Can you now explain this in 10 lines text

ChatGPT, while a powerful language model, may sometimes struggle to produce highly precise output. This is because its responses are generated based on patterns and examples from its training data, rather than having a deep understanding of concepts. As a result, it can occasionally provide answers that are not entirely accurate or may lack precision. Users should be mindful of this limitation and consider verifying information from reliable sources when precise accuracy is required. It is always advisable to use ChatGPT as a helpful tool for generating ideas and insights rather than relying solely on it for absolute precision.

Figure 9.

Natural langage programming



PI

Write a python code that plots the trajectory of a ball of mass m launched at a velocty v0 with an angle theta from the x-axis, these three parameters will be given by the user when launching the script. Example: python trajectory.py 10 3 45 will mean a ball of 10 kg launched with v0 = 3m/s and angle = 45 degres.

Sure! Here's a Python code that plots the trajectory of a ball based on the given parameters:

```python
import math
import matplotlib.pyplot as plt

def calculate_trajectory(m, v0, theta):
    g = 9.8  # acceleration due to gravity (m/s^2)
    theta_rad = math.radians(theta)  # convert angle to radians

    # Calculate the time of flight
    t_flight = (2 * v0 * math.sin(theta_rad)) / g

    # Calculate the horizontal and vertical components of velocity
    v0x = v0 * math.cos(theta_rad)
    v0y = v0 * math.sin(theta_rad)

    # Calculate the time intervals
    t_intervals = [i / 100 for i in range(int(t_flight * 100) + 1)]

    # Calculate the x and y coordinates at each time interval
    x_coordinates = [v0x * t for t in t_intervals]
```

Figure 10.

Natural langage programming

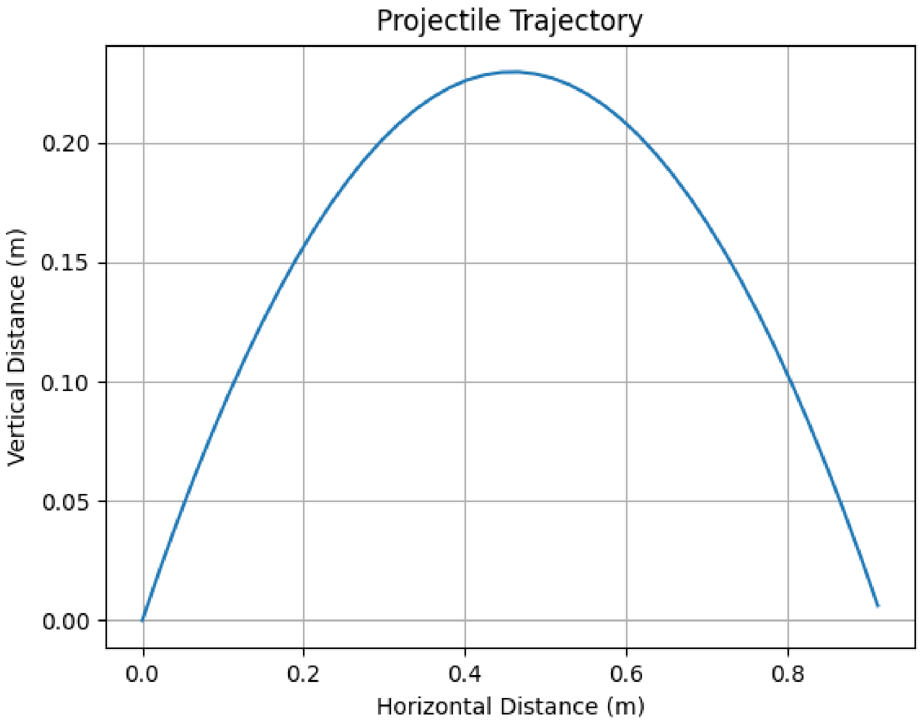
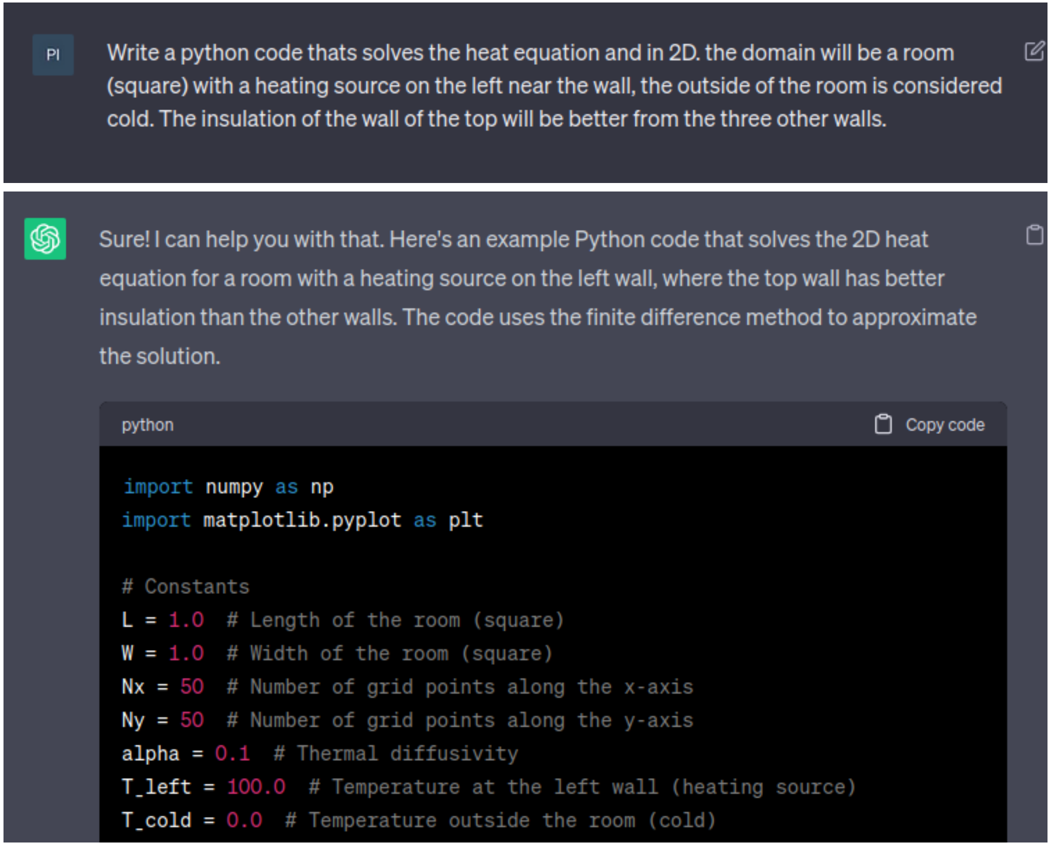


Figure 11.

Natural langage programming



Write a python code thats solves the heat equation and in 2D. the domain will be a room (square) with a heating source on the left near the wall, the outside of the room is considered cold. The insulation of the wall of the top will be better from the three other walls.

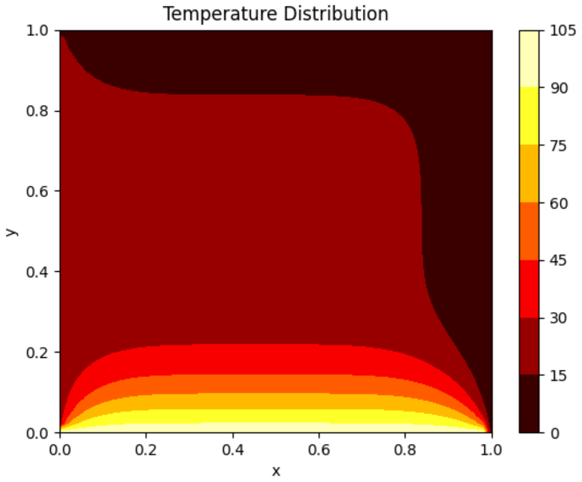
Sure! I can help you with that. Here's an example Python code that solves the 2D heat equation for a room with a heating source on the left wall, where the top wall has better insulation than the other walls. The code uses the finite difference method to approximate the solution.

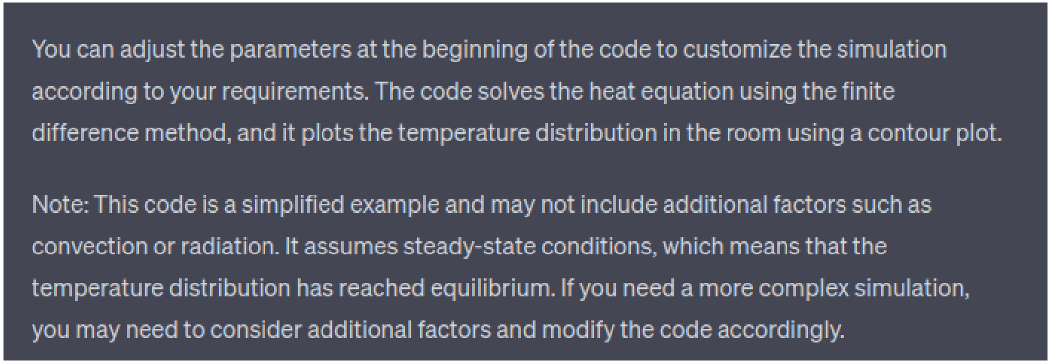
```python
import numpy as np
import matplotlib.pyplot as plt

# Constants
L = 1.0  # Length of the room (square)
W = 1.0  # Width of the room (square)
Nx = 50  # Number of grid points along the x-axis
Ny = 50  # Number of grid points along the y-axis
alpha = 0.1  # Thermal diffusivity
T_left = 100.0  # Temperature at the left wall (heating source)
T_cold = 0.0  # Temperature outside the room (cold)
```

Figure 12.

Natural langage programming





You can adjust the parameters at the beginning of the code to customize the simulation according to your requirements. The code solves the heat equation using the finite difference method, and it plots the temperature distribution in the room using a contour plot.

Note: This code is a simplified example and may not include additional factors such as convection or radiation. It assumes steady-state conditions, which means that the temperature distribution has reached equilibrium. If you need a more complex simulation, you may need to consider additional factors and modify the code accordingly.

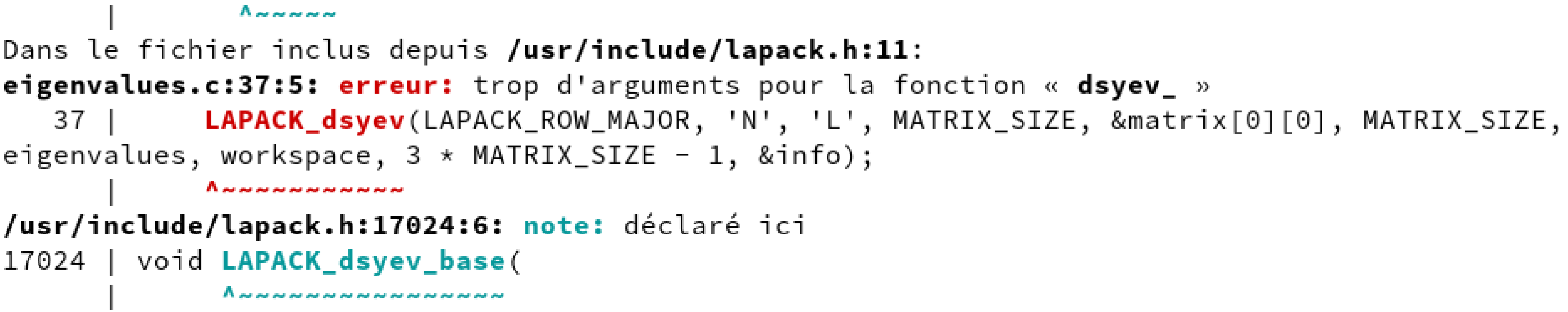
Figure 13.

Programming with ChatGPT

Also quite good at **improving existing code** or **generating documentation**.

Some limitations:

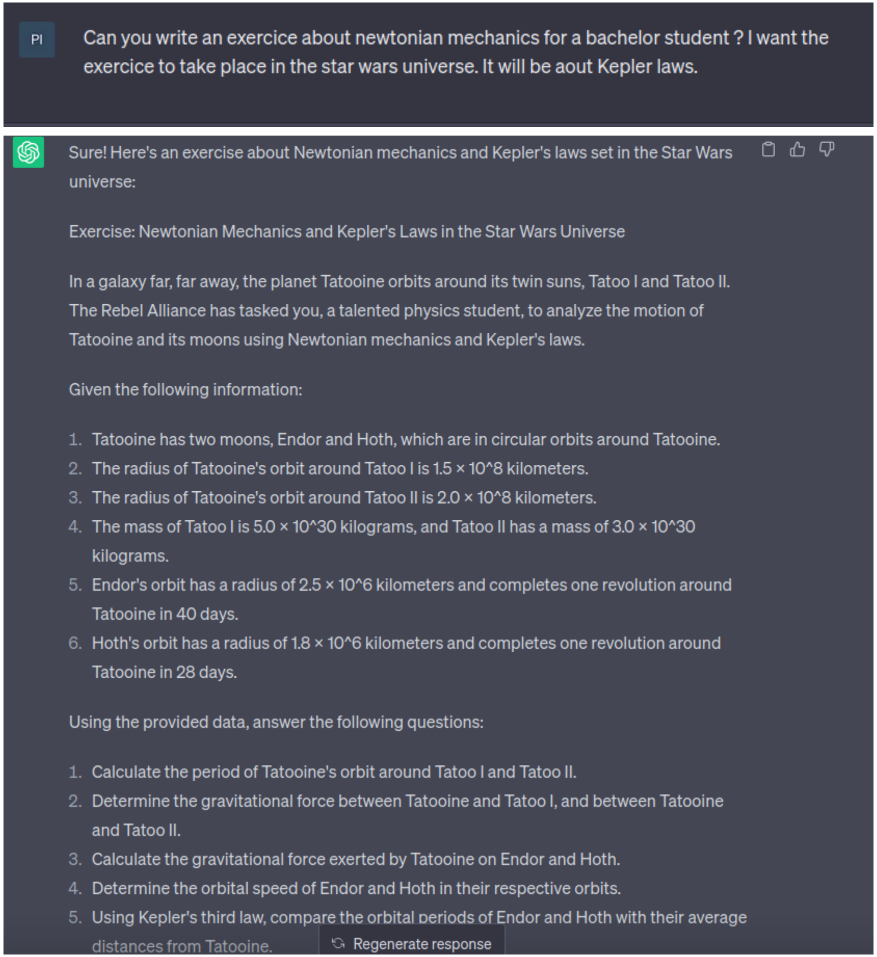
- when the code is big
- if the code uses a lot of obscur library



```
      |       ^~~~~~
Dans le fichier inclus depuis /usr/include/lapack.h:11:
eigenvalues.c:37:5: erreur: trop d'arguments pour la fonction « dsyev_ »
   37 |     LAPACK_dsyev(LAPACK_ROW_MAJOR, 'N', 'L', MATRIX_SIZE, &matrix[0][0], MATRIX_SIZE,
eigenvalues, workspace, 3 * MATRIX_SIZE - 1, &info);
      |     ^~~~~~~~~~~
/usr/include/lapack.h:17024:6: note: déclaré ici
17024 | void LAPACK_dsyev_base(
      |      ^~~~~~~~~~~~~~~~
```

Figure 14.

For teaching



Can you write an exercice about newtonian mechanics for a bachelor student ? I want the exercice to take place in the star wars universe. It will be aout Kepler laws.

Sure! Here's an exercise about Newtonian mechanics and Kepler's laws set in the Star Wars universe:

Exercise: Newtonian Mechanics and Kepler's Laws in the Star Wars Universe

In a galaxy far, far away, the planet Tatooine orbits around its twin suns, Tatoo I and Tatoo II. The Rebel Alliance has tasked you, a talented physics student, to analyze the motion of Tatooine and its moons using Newtonian mechanics and Kepler's laws.

Given the following information:

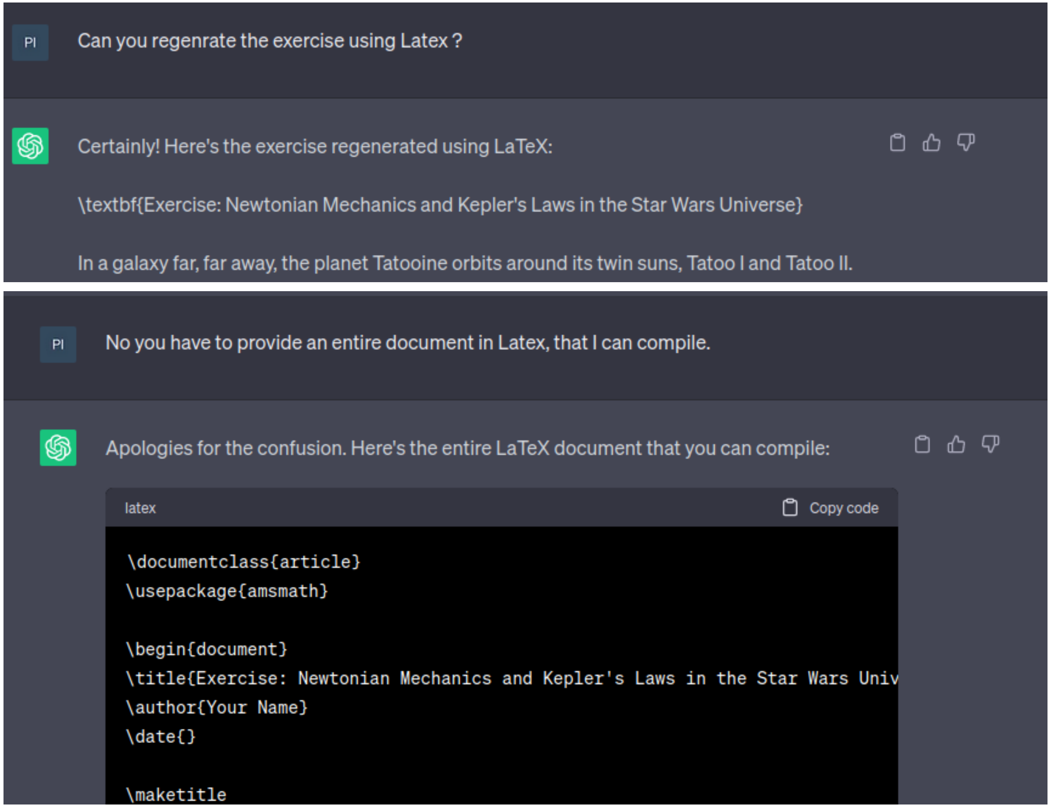
1. Tatooine has two moons, Endor and Hoth, which are in circular orbits around Tatooine.
2. The radius of Tatooine's orbit around Tatoo I is 1.5 × 10^8 kilometers.
3. The radius of Tatooine's orbit around Tatoo II is 2.0 × 10^8 kilometers.
4. The mass of Tatoo I is 5.0 × 10^30 kilograms, and Tatoo II has a mass of 3.0 × 10^30 kilograms.
5. Endor's orbit has a radius of 2.5 × 10^6 kilometers and completes one revolution around Tatooine in 40 days.
6. Hoth's orbit has a radius of 1.8 × 10^6 kilometers and completes one revolution around Tatooine in 28 days.

Using the provided data, answer the following questions:

1. Calculate the period of Tatooine's orbit around Tatoo I and Tatoo II.
2. Determine the gravitational force between Tatooine and Tatoo I, and between Tatooine and Tatoo II.
3. Calculate the gravitational force exerted by Tatooine on Endor and Hoth.
4. Determine the orbital speed of Endor and Hoth in their respective orbits.
5. Using Kepler's third law, compare the orbital periods of Endor and Hoth with their average distances from Tatooine.

Figure 15.

For teaching



Can you regenrate the exercise using Latex ?

Certainly! Here's the exercise regenerated using LaTeX:

\textbf{Exercise: Newtonian Mechanics and Kepler's Laws in the Star Wars Universe}

In a galaxy far, far away, the planet Tatooine orbits around its twin suns, Tatoo I and Tatoo II.

No you have to provide an entire document in Latex, that I can compile.

Apologies for the confusion. Here's the entire LaTeX document that you can compile:

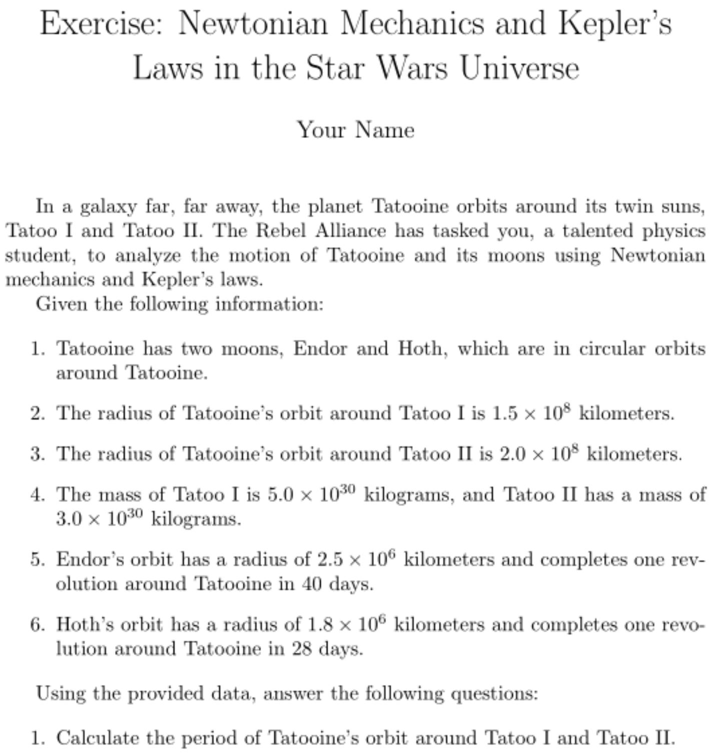
```latex
\documentclass{article}
\usepackage{amsmath}

\begin{document}
\title{Exercise: Newtonian Mechanics and Kepler's Laws in the Star Wars Univ
\author{Your Name}
\date{}

\maketitle
```

Figure 16.

For teaching



Exercise: Newtonian Mechanics and Kepler's Laws in the Star Wars Universe

Your Name

In a galaxy far, far away, the planet Tatooine orbits around its twin suns, Tatoo I and Tatoo II. The Rebel Alliance has tasked you, a talented physics student, to analyze the motion of Tatooine and its moons using Newtonian mechanics and Kepler's laws.

Given the following information:

1. Tatooine has two moons, Endor and Hoth, which are in circular orbits around Tatooine.
2. The radius of Tatooine's orbit around Tatoo I is 1.5×10^8 kilometers.
3. The radius of Tatooine's orbit around Tatoo II is 2.0×10^8 kilometers.
4. The mass of Tatoo I is 5.0×10^{30} kilograms, and Tatoo II has a mass of 3.0×10^{30} kilograms.
5. Endor's orbit has a radius of 2.5×10^6 kilometers and completes one revolution around Tatooine in 40 days.
6. Hoth's orbit has a radius of 1.8×10^6 kilometers and completes one revolution around Tatooine in 28 days.

Using the provided data, answer the following questions:

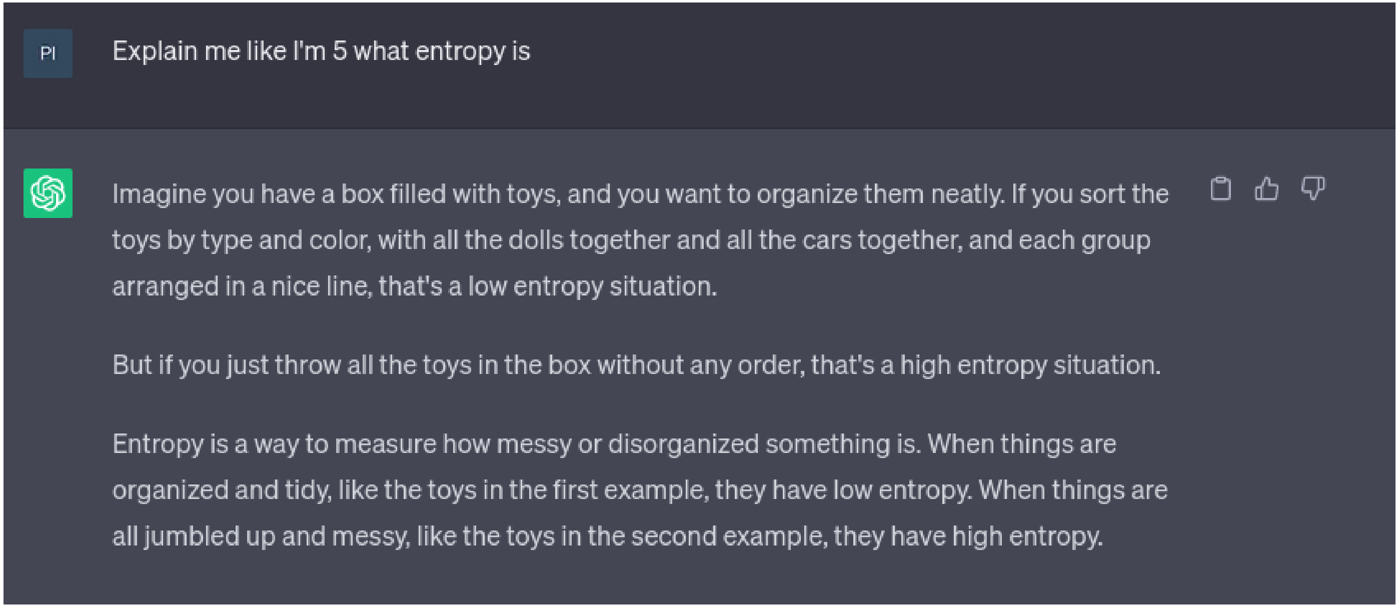
1. Calculate the period of Tatooine's orbit around Tatoo I and Tatoo II.

Figure 17.

But student can use it too!

For learning

The “explain me like I’m 5” prompt:



Explain me like I'm 5 what entropy is

Imagine you have a box filled with toys, and you want to organize them neatly. If you sort the toys by type and color, with all the dolls together and all the cars together, and each group arranged in a nice line, that's a low entropy situation.

But if you just throw all the toys in the box without any order, that's a high entropy situation.

Entropy is a way to measure how messy or disorganized something is. When things are organized and tidy, like the toys in the first example, they have low entropy. When things are all jumbled up and messy, like the toys in the second example, they have high entropy.

Figure 18.

But also to learn a new langage, . . .

ChatGPT for text revision/writing

1. Spell checking

ChatGPT for text revision/writing

1. Spell checking
2. Litterature review and summarization

ChatGPT for text revision/writing

1. Spell checking
2. Litterature review and summarization
3. Translation of technical texts

Ethical considerations, issues, ...

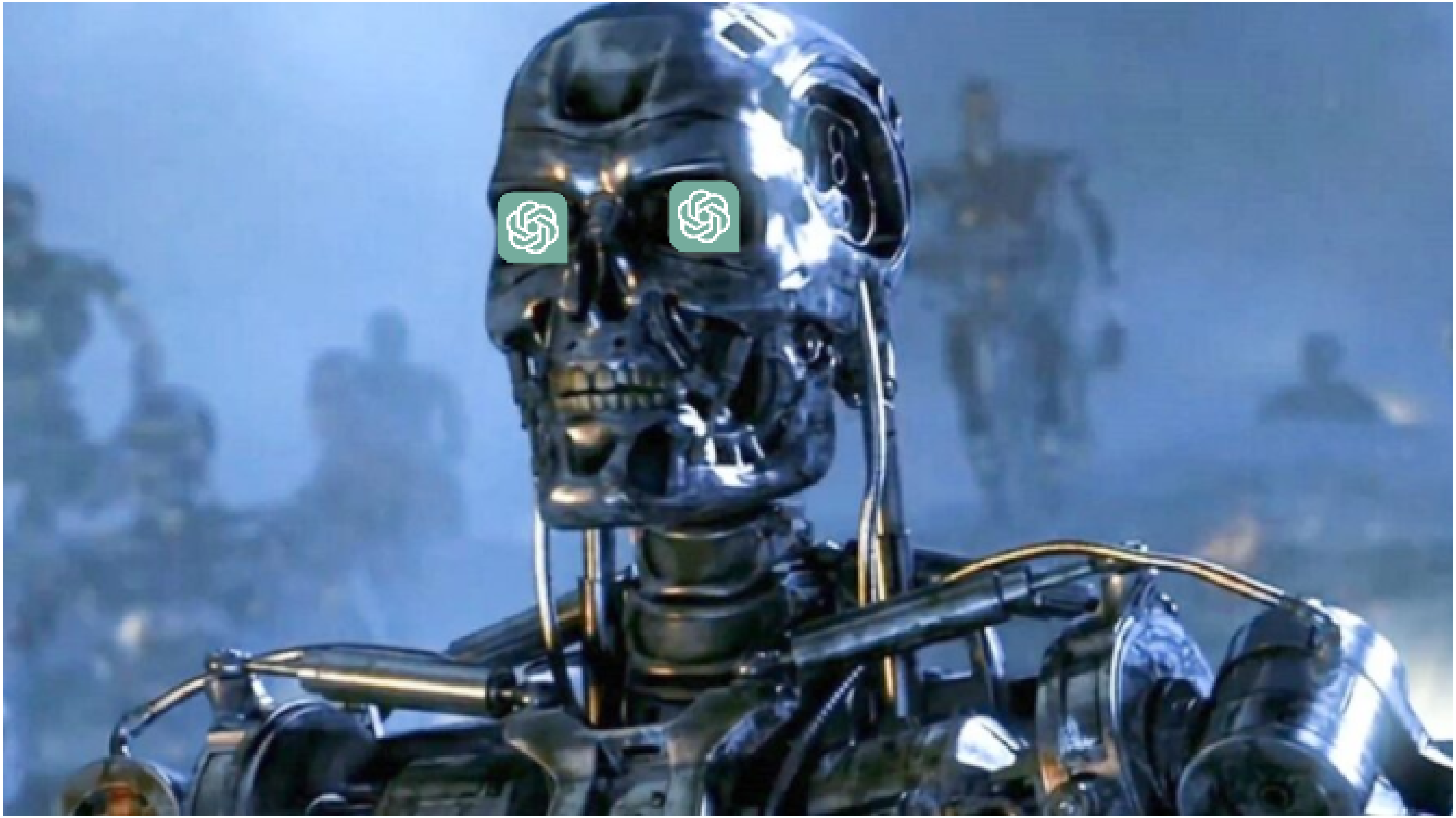
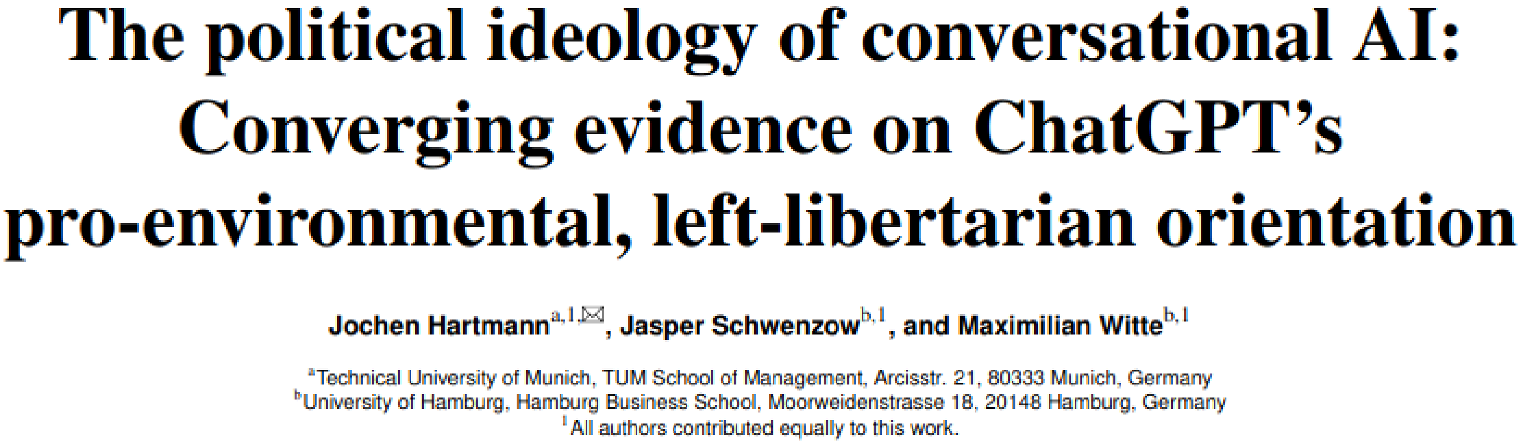


Figure 19.

Bias in ChatGPT models

Arbitrary decisons of OpenAI



The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation

Jochen Hartmann[a,1]✉**, Jasper Schwenzow**[b,1]**, and Maximilian Witte**[b,1]

[a]Technical University of Munich, TUM School of Management, Arcisstr. 21, 80333 Munich, Germany
[b]University of Hamburg, Hamburg Business School, Moorweidenstrasse 18, 20148 Hamburg, Germany
[1]All authors contributed equally to this work.

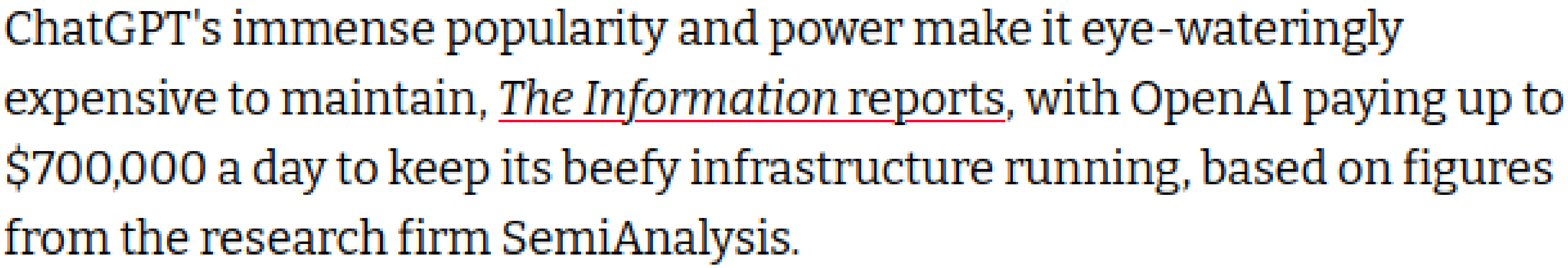
Figure 20.

But see also ChatCGT or RightWingGPT…

Can be use for **propaganda**

Privacy concerns

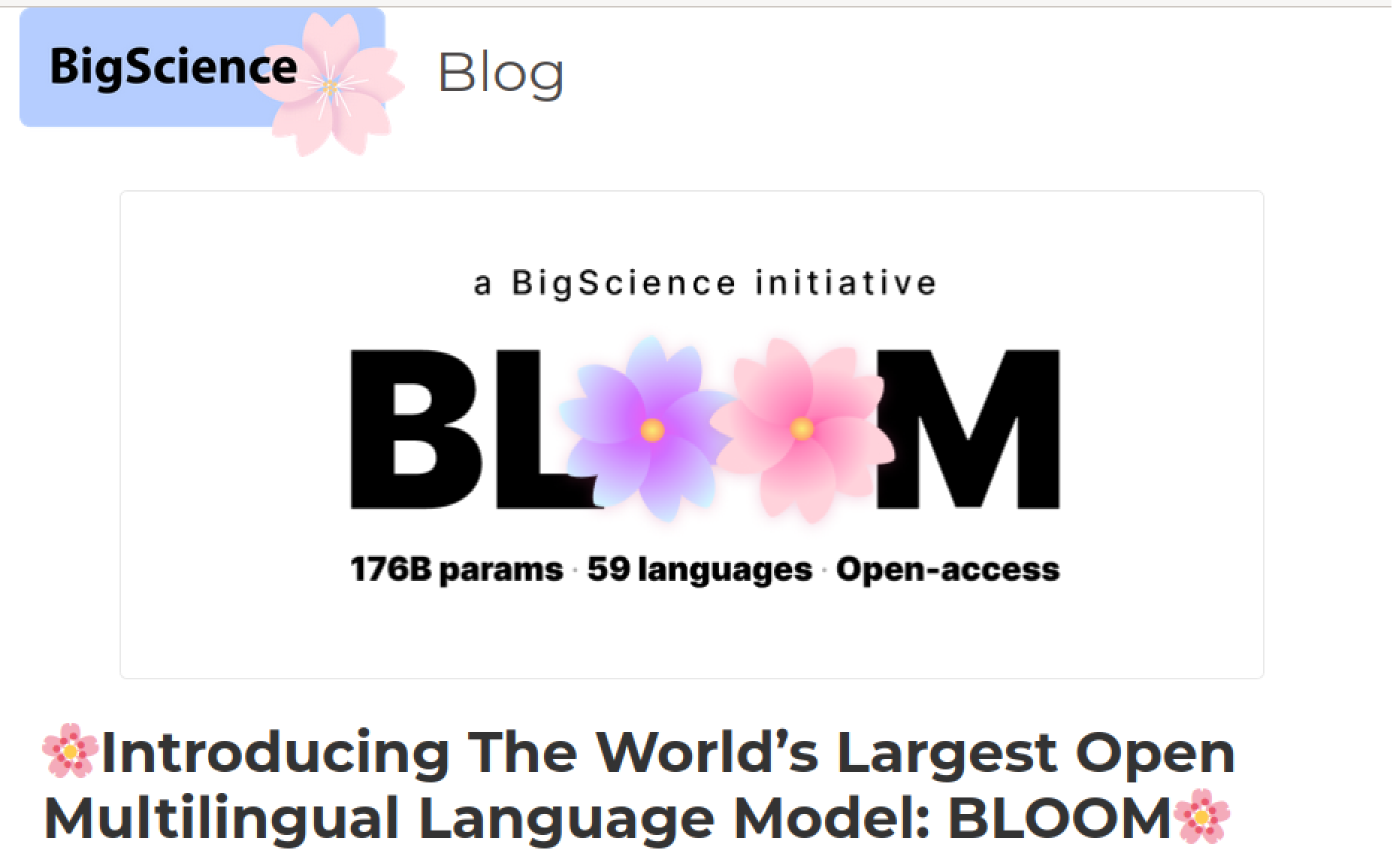
OpenAI is a private company.



ChatGPT's immense popularity and power make it eye-wateringly expensive to maintain, *The Information* reports, with OpenAI paying up to $700,000 a day to keep its beefy infrastructure running, based on figures from the research firm SemiAnalysis.

Figure 21.

Alternatives



BigScience Blog

a BigScience initiative

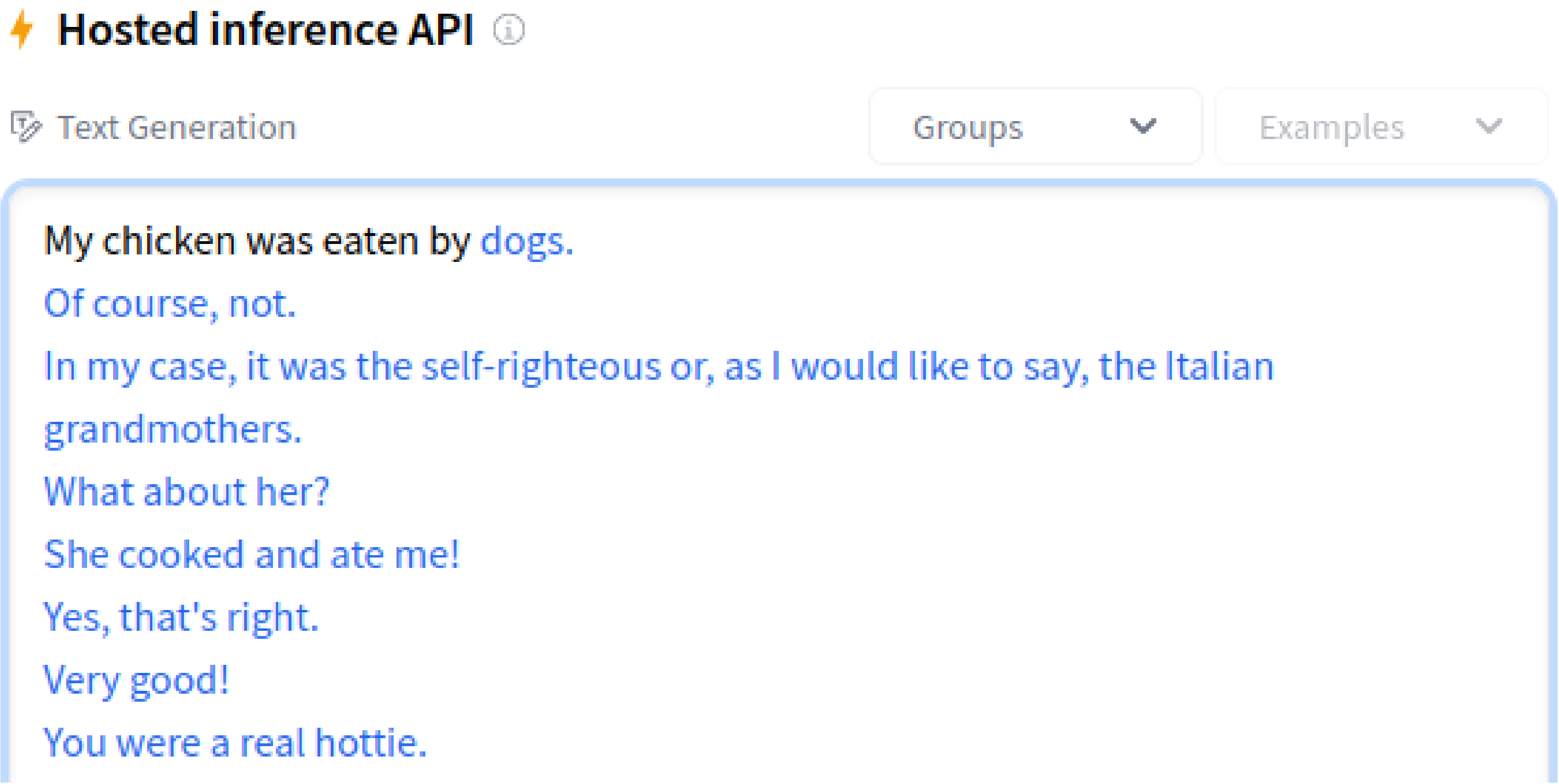
BLOOM

176B params · 59 languages · Open-access

🌸Introducing The World’s Largest Open Multilingual Language Model: BLOOM🌸

Figure 22.

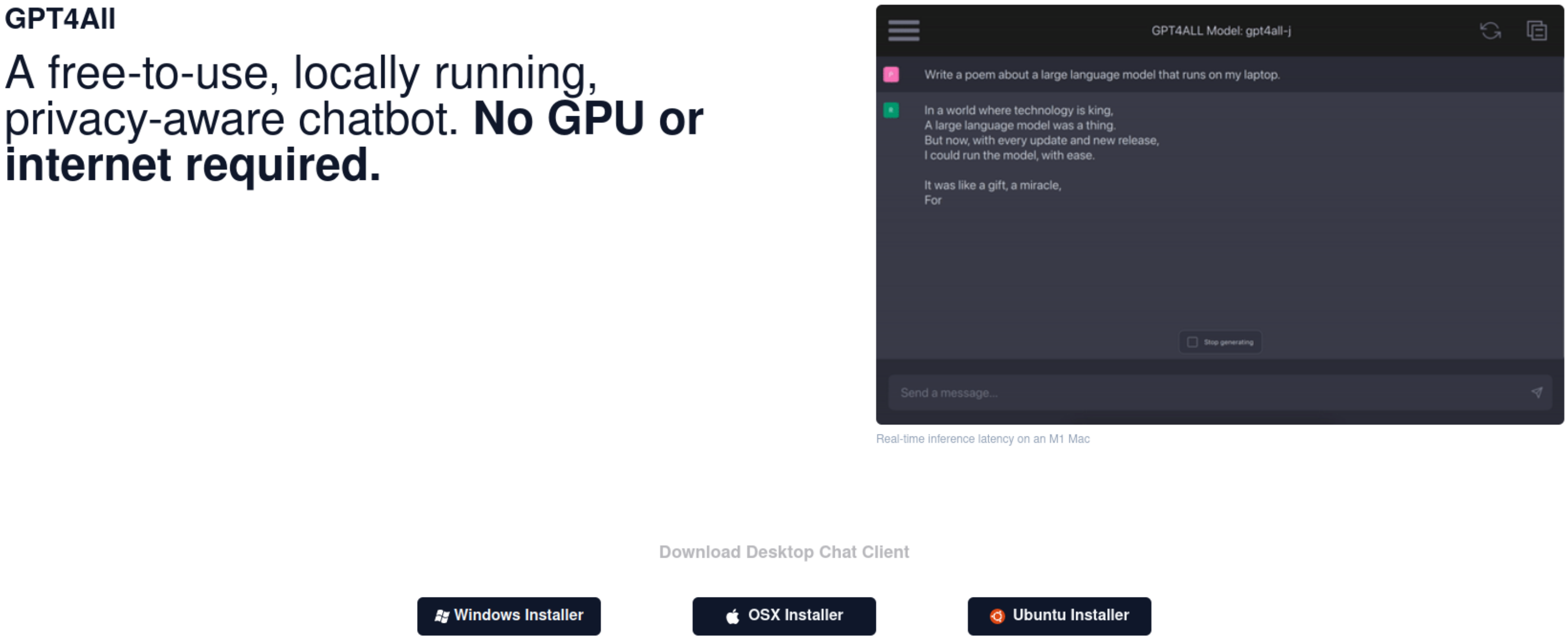
Alternatives



My chicken was eaten by dogs.
Of course, not.
In my case, it was the self-righteous or, as I would like to say, the Italian grandmothers.
What about her?
She cooked and ate me!
Yes, that's right.
Very good!
You were a real hottie.

Figure 23.

Alternatives



GPT4All

A free-to-use, locally running, privacy-aware chatbot. **No GPU or internet required.**

GPT4ALL Model: gpt4all-j

Write a poem about a large language model that runs on my laptop.

In a world where technology is king,
A large language model was a thing.
But now, with every update and new release,
I could run the model, with ease.

It was like a gift, a miracle,
For

Real-time inference latency on an M1 Mac

Download Desktop Chat Client

Figure 24.

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

- is **not** interested in the **truth**

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

- is **not** interested in the **truth**
- is **not deterministic**

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

- is **not** interested in the **truth**
- is **not deterministic**
- depends on the choices of a **private company**

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

- is **not** interested in the **truth**
- is **not deterministic**
- depends on the choices of a **private company**
- has no access to the **Internet**

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

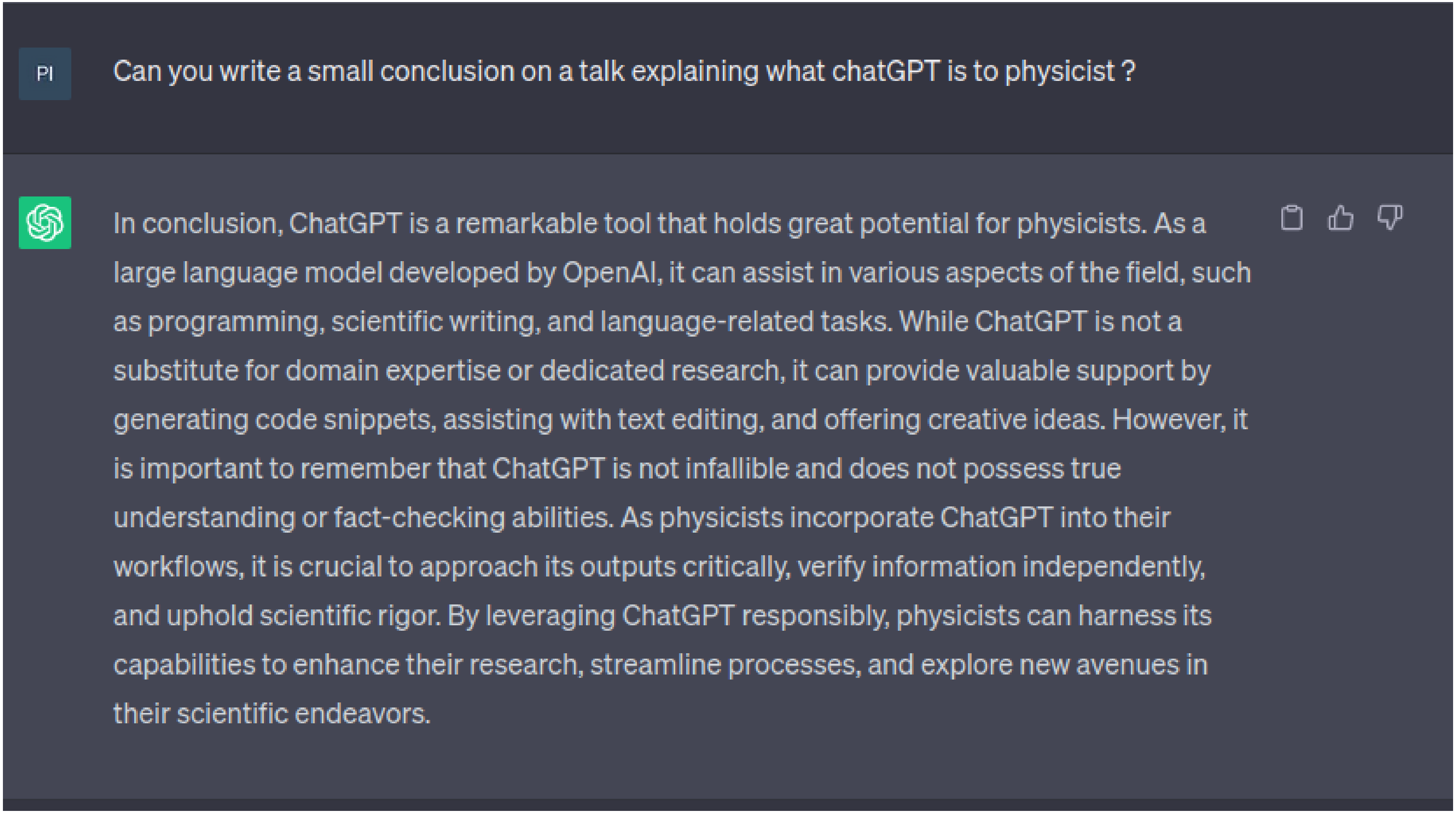
- is **not** interested in the **truth**
- is **not deterministic**
- depends on the choices of a **private company**
- has no access to the **Internet**
- is **blind**

Conclusion

Remember that ChatGPT (despite being a powerful assistant for programming and text writing):

- is **not** interested in the **truth**
- is **not deterministic**
- depends on the choices of a **private company**
- has no access to the **Internet**
- is **blind**
- cannot **hide** information to the user

Conclusion



Can you write a small conclusion on a talk explaining what chatGPT is to physicist ?

In conclusion, ChatGPT is a remarkable tool that holds great potential for physicists. As a large language model developed by OpenAI, it can assist in various aspects of the field, such as programming, scientific writing, and language-related tasks. While ChatGPT is not a substitute for domain expertise or dedicated research, it can provide valuable support by generating code snippets, assisting with text editing, and offering creative ideas. However, it is important to remember that ChatGPT is not infallible and does not possess true understanding or fact-checking abilities. As physicists incorporate ChatGPT into their workflows, it is crucial to approach its outputs critically, verify information independently, and uphold scientific rigor. By leveraging ChatGPT responsibly, physicists can harness its capabilities to enhance their research, streamline processes, and explore new avenues in their scientific endeavors.

Figure 25.