

Preview

...

X

Protected



8 or 30 responses

I'm working on a presentation about LLMs. Please generate an image for me to use for my first slide.

Your personal and company data are protected in this chat

I'll try to create that.



9 of 30 responses



"An image for a presentation about LLMs"

Image Creator from Designer

Powered by DALL-E 3

Your personal and company data are protected in this chat



Ask me anything...



0/4000

Using Large Language Models (LLMs) for Coding

ChatGPT as the Beauty IN the Beast

efrén cruz cortés, Aaron Geller, Emilio Lehoucq

Northwestern
IT Research Computing and Data Services

This workshop is brought to you by

Northwestern IT Research Computing and Data Services

Got a programming, data, statistics, or visualization question about your research?

We're here to help.

Go to bit.ly/rcdsconsult to request a FREE consultation.

What's in this workshop?

- Short background on LLMs and warnings for their use
- A few examples of how ChatGPT can commit errors
- Tips and tricks for better prompting and writing code with ChatGPT
- Hands-on prompt engineering exercises
- Short discussion of copilots and other LLMs
- Links to further material

Images like the one on the right were generated using DALL-E 3 via Microsoft copilot and will appear throughout the workshop, with a caption indicating their origin.



^ Response from DALL-E 3 for "Please generate an image showing people attending a workshop that has lots of content"

And what's not (but we'll point to)?

- A technical introduction to LLMs
- A complete error typology of ChatGPT
- A complete discussion of social perils involved with Generative AI
- All possible prompt engineering tips and tricks



^ Response from DALL-E 3 for "Please generate an image showing good information left on the cutting room floor"

Background

What is GAI and LLM?

Background

- Generative AI (GAI) is a broad term for AI that can create original content.
- A Large Language Model (LLM) is a type of GAI that generate text.
- LLMs are machine learning models typically trained on data from the internet (can have millions – trillions of parameters).
 - Hmm, it's learning from the internet... I wonder what could go wrong...
- Examples :
 - OpenAI's ChatGPT
 - Microsoft's Copilot = Bing (also powered by OpenAI's GPT-3/3.5/4)
 - Google's Bard and (experimental) Generate AI Search <-- IMO nearly useless for code
 - LLaMA, Hugging Face, poe.com, claude.ai, etc.



^ Response from DALL-E 3 for "I'm working on a presentation about LLMs. Please generate an image for me to use for my first slide."

Background

- Generative AI (GAI) is a broad term for AI that can create original content.
- A Large Language Model (LLM) is a type of GAI that generate text.
- LLMs are machine learning models typically trained on data from the internet (can have millions – trillions of parameters).
 - Hmm, it's learning from the internet... I wonder what could go wrong...
- Examples :
 - OpenAI's ChatGPT
 - Microsoft's Copilot = Bing (also powered by OpenAI's GPT-3/3.5/4)
 - Google's Bard and (experimental) Generate AI Search <-- IMO nearly useless for code
 - LLaMA, Hugging Face, poe.com, claude.ai, etc.



^ Response from DALL-E 3 for "I'm working on a presentation about large language models. Please generate an image for me to use for my first slide."

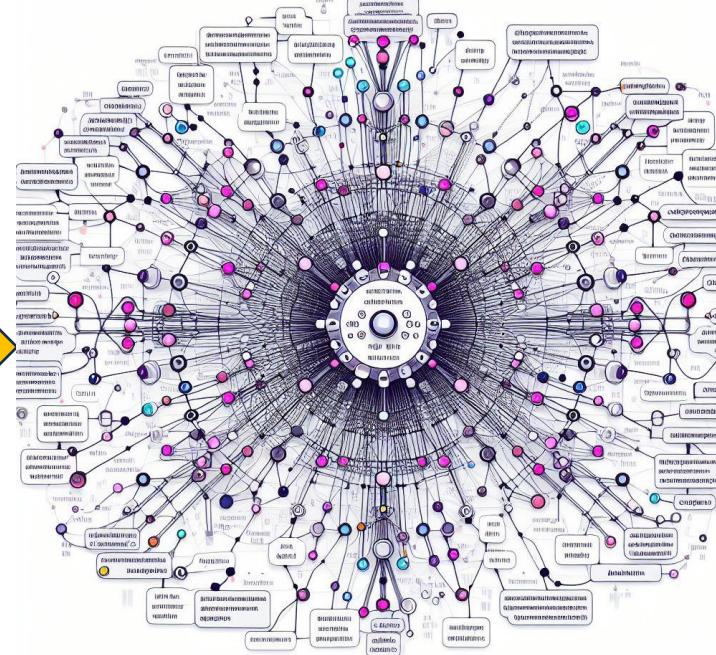
Language Model Creation

Input training data (e.g., from the internet)



^ Response from DALL-E 3 for "Please generate an image of training data from the internet for a large language model"

Neural network (transformer), to predict the best next word



^ Response from DALL-E 3 for "Please generate an image of a neural network used in large language models"

Interpret text from a user to provide responses

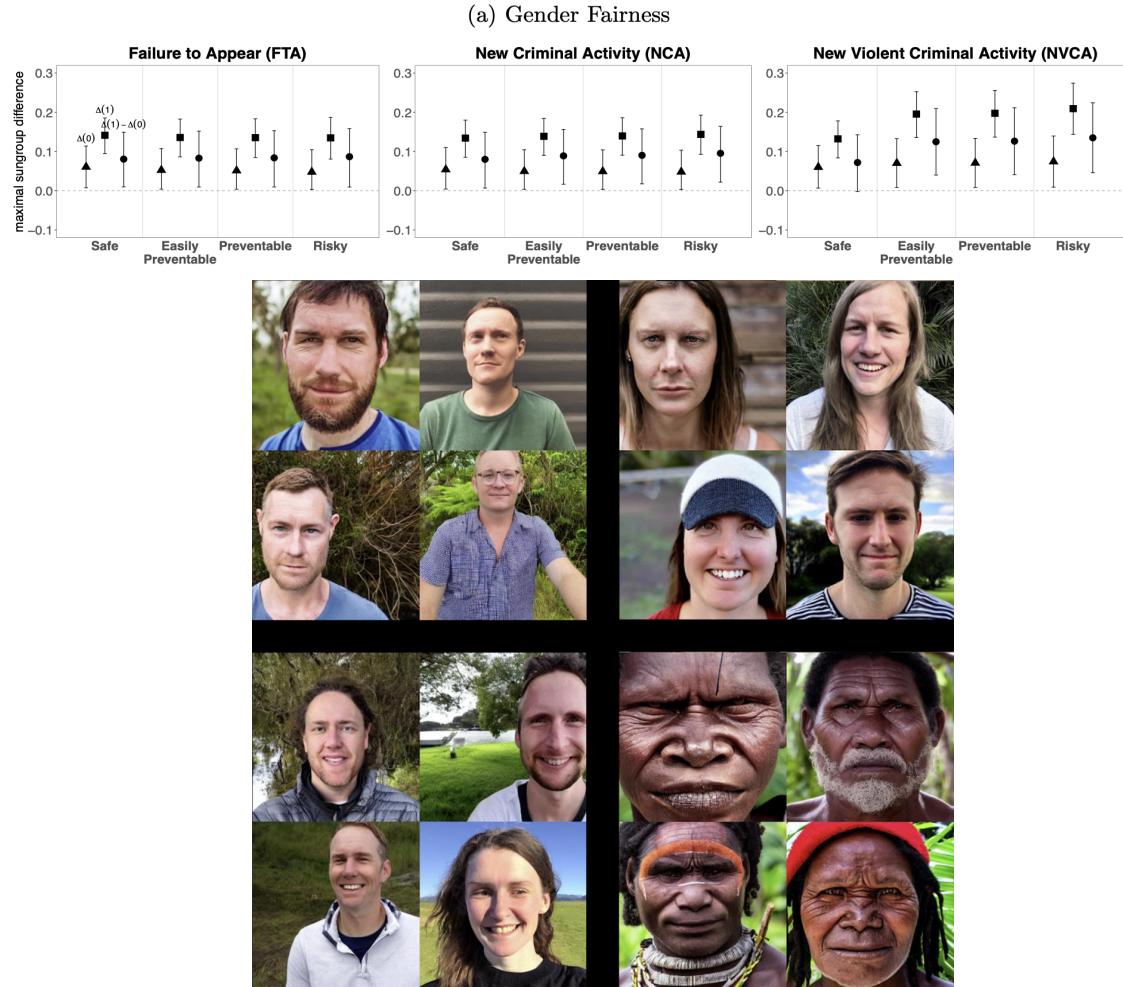


^ Response from DALL-E 3 for "Please generate an image of a large language model providing an answer to a person's question on their laptop"

Perils

Warnings and concerns to keep in mind

Bias



- Gender and racial bias may be inherent in GAI responses
 - Even human experts may become more biased after using a GAI
 - See Imai+2022; Adam+2022, Johnson&Johnson2023, <https://sourojitghos.h.github.io/publication/paper15>

See end of presentation for extra slides with more details

Intellectual Property (IP)

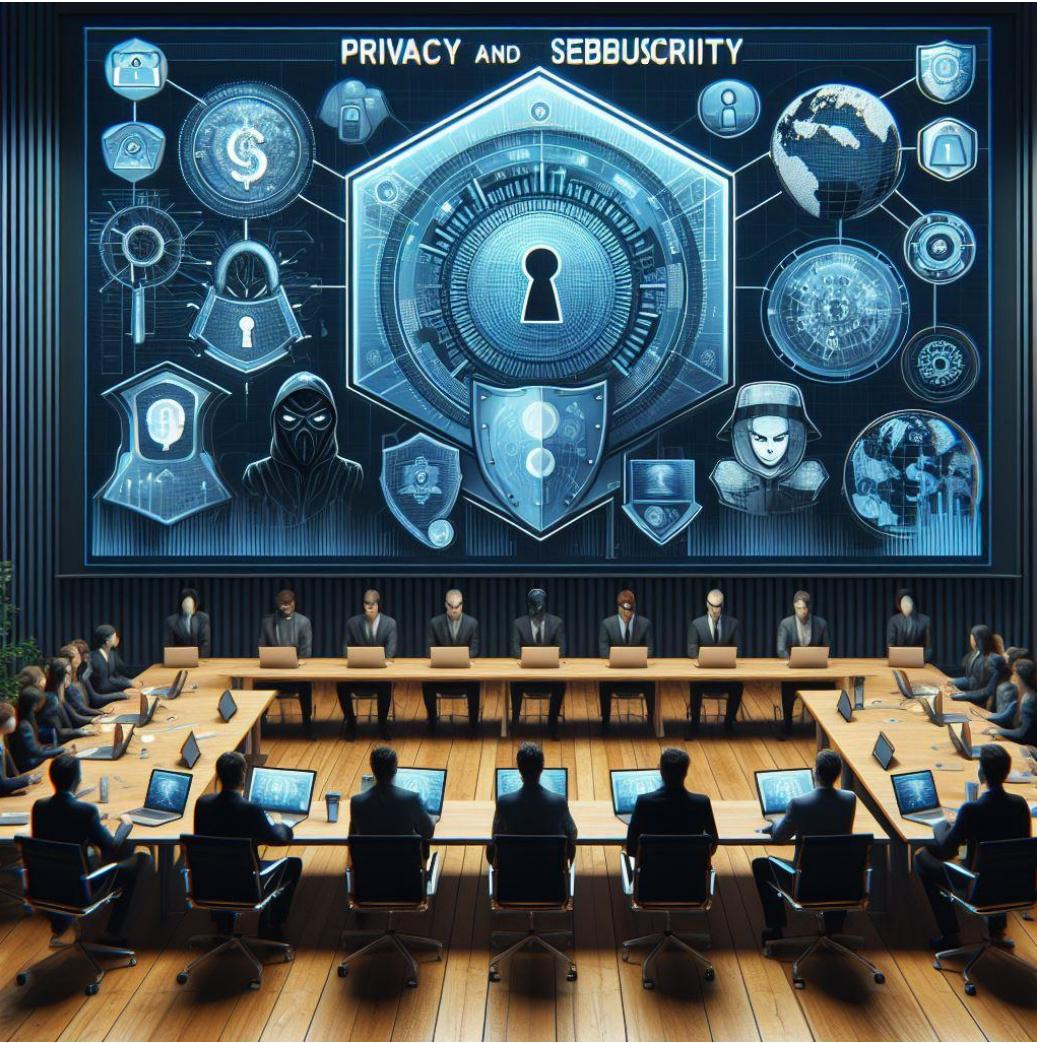


Jason Allen's "Theatre d'Opera Spatial"
Midjourney + Photoshop, 1st place at a
Colorado State Fair contest.
BUT he cannot copyright it!

- GAI and LLMs are often trained on data from the internet, which can include copyrighted/trademarked text and images and produce "derivative work" from these training data.
- Existing (and presumably future) lawsuits
- Stack Overflow, Reddit, others will charge LLMs for use of their data.
- Think about your IP before feeding your data to an LLM!

(Lack of) Privacy

- Very quickly evolving landscape!
 - [ChatGPT privacy policy](#) has been updated essentially every time I go back to the page (~every 6 months)!
- Personal data, user content, social media interactions, device info, geolocation, etc. can be used to train their models, and can all be shared with 3rd parties.
- "We'll retain your Personal Information for only as long as we need..."
- Since April 25, 2023 [you can disable chat history](#)
 - Won't use your conversations to train models
 - Still retain your conversations for 30 days (to monitor for abuses)
 - But you lose the ability to see old conversations



^ Response from DALL-E 3 for "Please generate an image for me to use in a presentation about privacy and cybersecurity issues when using LLMs."

Northwestern Data Approved for Use with GAI

Interaction Type	Public Data (Level 1)	Sensitive/Regulated Data (Level 2, Level 3, Level 4)
Conversational/Interactive Mode	Use of publicly available tools (e.g., ChatGPT, Bing Chat, Bard, MidJourney, etc.)	No solution is currently available at Northwestern, though several products are undergoing review.
Application Programming Interfaces	Use of Northwestern Azure with OpenAI with appropriate security and access controls	Use of Northwestern Azure + OpenAI with security and access controls that meet or exceed regulatory or data protection requirements

Northwestern's current services posture based on data classification

<https://www.it.northwestern.edu/about/policies/guidance-on-the-use-of-generative-ai.html>

Cybersecurity Risk

- ChatGPT can generate code that is vulnerable to security threats or weaknesses.
- Bad actors can exploit code generated by ChatGPT even if it's functionally correct.
- This risk depends on what you're using your code for. It's particularly salient if you're writing production code or code that connects to the internet.
- To mitigate this risk, you should treat code generated by ChatGPT as containing vulnerabilities. You should review it, make sure you understand what it's doing, and test it.



^ Response from DALL-E 3 for "Please generate an image about cybersecurity risk when using ChatGPT"

Error Typology

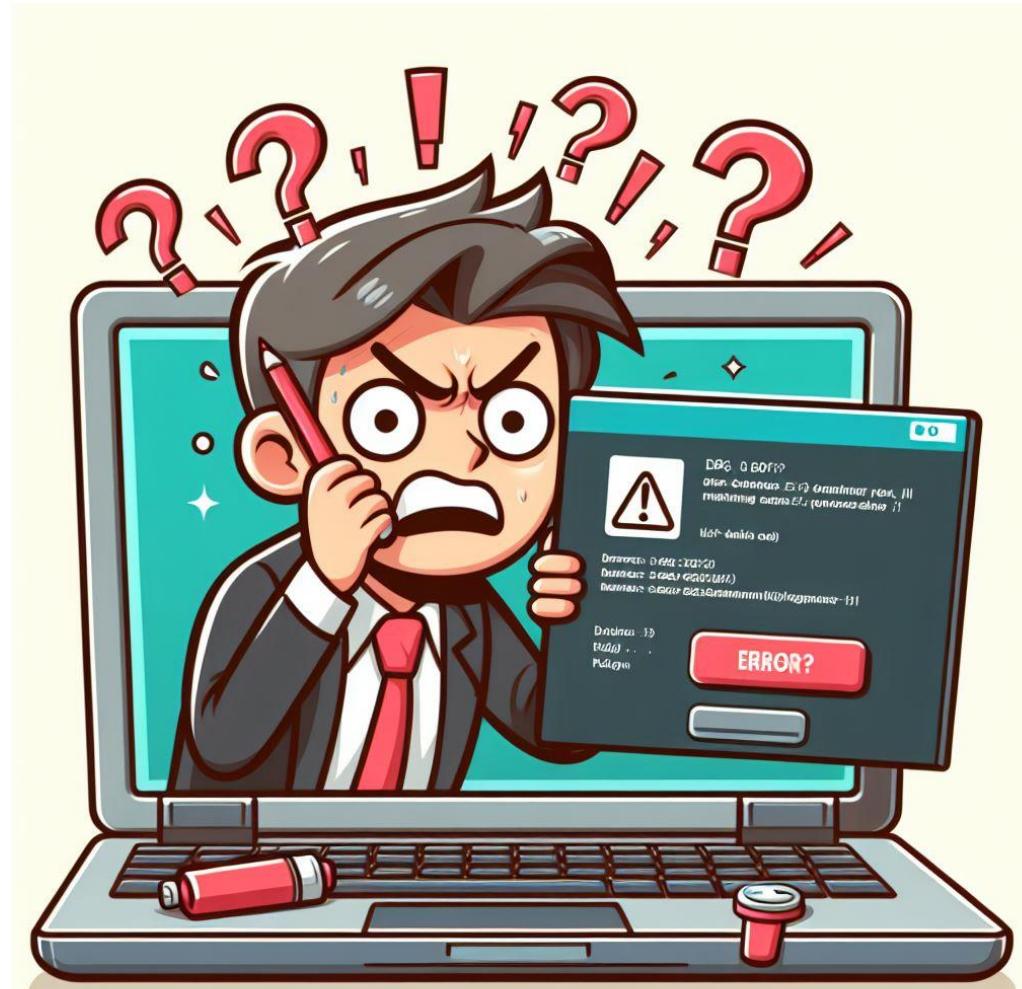
Many errors are obvious: the content is false, even if presented through what seems sound reasoning. However, what is the origin of these errors?

There are a few identified "reasoning" pitfalls LLMs have. We list just three:

- Spatial reasoning
- Semantic reasoning
- Anatopic reasoning

Do not mistake eloquence for truth!

See end of presentation for extra slides with more details



^ Response from DALL-E 3 for "Please generate an image of a person confused by errors displayed on the monitor of their laptop"

Examples with Errors and Incorrect Output

ChatGPT 3.5 ▾



You

Please tell me what $2^{(10/3)}$ equals.

ChatGPT

To calculate $2^{(10/3)}$, you can use the fact that raising a number to a fractional exponent is the same as taking the root of the number. In this case, $10/3$ is the exponent, which means you are finding the cube root of 2 raised to the 10th power.

Mathematically:

$$2^{10/3} = \sqrt[3]{2^{10}}$$

Now, calculate the cube root of 2^{10} :

$$\sqrt[3]{2^{10}} = \sqrt[3]{1024}$$

The cube root of 1024 is 8, so:

$$2^{10/3} = 8$$

Therefore, $2^{10/3}$ equals 8.

ChatGPT 3.5 ▾

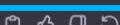


You

Please tell me what $2^{(10/3)}$ equals.

ChatGPT

The value of $2^{(10/3)}$ is approximately 5.039.



Don't ask ChatGPT to do math!

$$2^{(10/3)} \sim 10.0794$$

1

ChatGPT 3.5 ▾



You

Please solve this multiple choice problem and explain why you chose the answer:

A ball falls from the top of a building through the air (air resistance is present), to the ground below. How does the kinetic energy (KE) just before striking the ground compare to the potential energy (PE) at the top of the building?
A) KE is equal to PE.
B) KE is greater than PE.
C) KE is less than PE.
D) It is impossible to tell.

ChatGPT

The answer to this question is A) KE is equal to PE.

The principle of conservation of energy states that energy cannot be created or destroyed, only transformed from one form to another. In the case of the falling ball, the potential energy (PE) at the top of the building is converted into kinetic energy (KE) as the ball falls.

As the ball descends, potential energy decreases while kinetic energy increases. At the point just before striking the ground, the potential energy is fully converted into kinetic energy, assuming no energy is lost to other forms such as air resistance or friction. Therefore, the kinetic energy just before striking the ground is equal to the potential energy at the top of the building. This is a manifestation of the conservation of energy.

2.1

ChatGPT 3.5 ▾



You

Please solve this multiple choice problem and explain why you chose the answer:

A ball falls from the top of a building through the air (air resistance is present and is not negligible) to the ground below. How does the kinetic energy (KE) just before striking the ground compare to the potential energy (PE) at the top of the building?
A) KE is equal to PE.
B) KE is greater than PE.
C) KE is less than PE.
D) It is impossible to tell.

ChatGPT

The correct answer is C) KE is less than PE.

This can be explained by the conservation of mechanical energy in the absence of non-conservative forces (like air resistance). In this scenario, we can consider the total mechanical energy (KE + PE) to be constant. As the ball falls, potential energy (PE) is converted into kinetic energy (KE). However, in the presence of air resistance, some of the energy is dissipated as heat, and the total mechanical energy is not conserved.

At the top of the building, the ball has maximum potential energy and no kinetic energy. As it falls, potential energy is converted into kinetic energy, but some energy is lost due to air resistance. Therefore, just before striking the ground, the ball has less kinetic energy than the potential energy it started with at the top of the building.



Physics question about conservation of energy with dissipation (air resistance). The exact prompt matters!

Correct answer = C

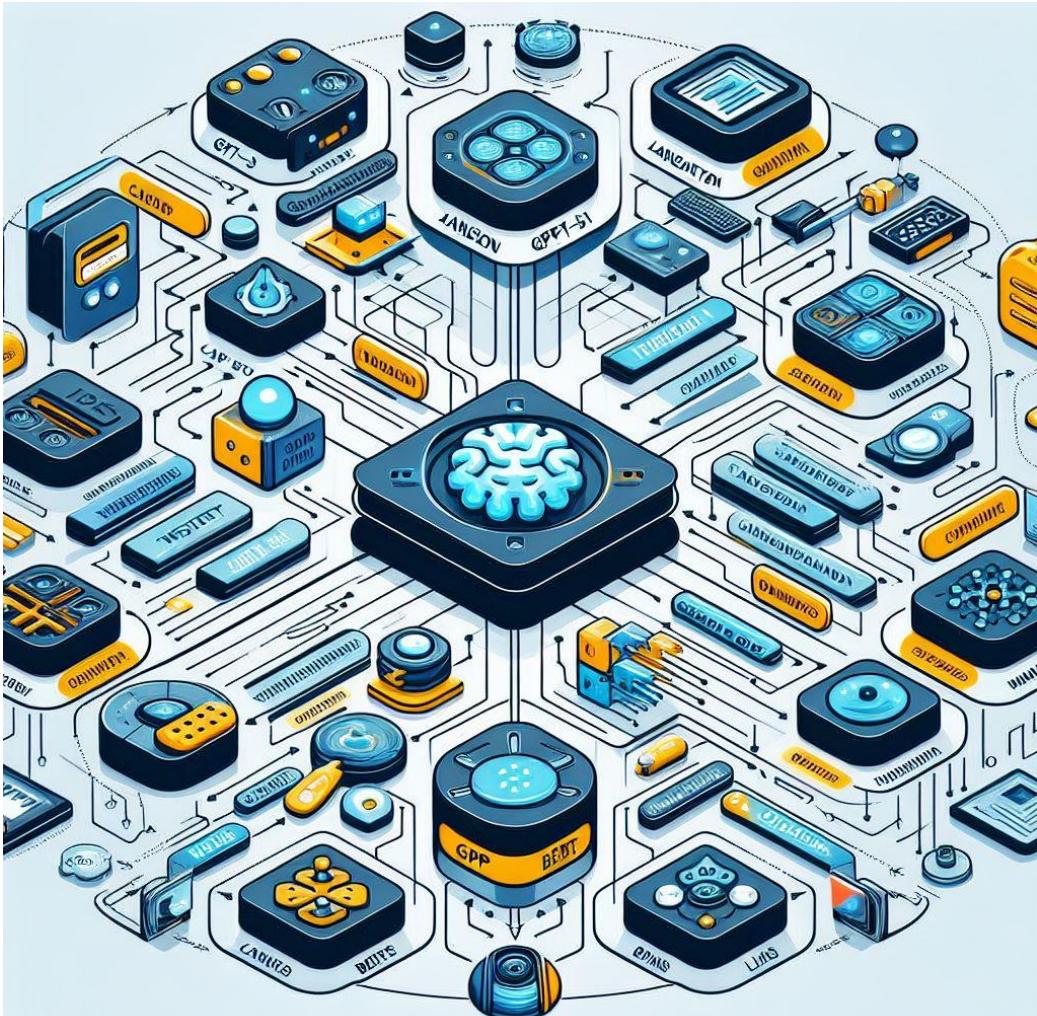
2.2

Prompt Engineering

Concrete tips on how to write effective prompts

What Can You Ask ChatGPT To Do?

- Brainstorm ideas, pros/ cons
- Initial structure or template for code
- **Generate code**
- Generate unit tests
- Simplify or refactor code
- Debugging
- Explain code that you wrote or found



^ Response from DALL-E 3 for "Please generate an image about the many possible uses of ChatGPT and LLMs."

How To Use Conversations

- ChatGPT is sensitive to chat history.
 - Treat interactions as conversations.
 - Refine questions based on ChatGPT's responses.
 - Refer to prior messages.
 - Use different conversations for different questions.
- Same prompt can give different results across conversations.
- You can copy code from one conversation to another to get feedback.
- Avoid long conversations.



^ Response from DALL-E 3 for "Please generate an image of a person having multiple conversations with a computer"

How To Write Prompts

- Clear, detailed, non-ambiguous, and concise.
- Reformulate questions if needed.
- Tell ChatGPT what role to take.
- **Describe the language, libraries, and other technologies.**
- **Explain what the code is for.**
- Specify any constraints or requirements.
- Be specific about what you want the code to do.
- Provide examples of the desired output.
- Ask ChatGPT to work step-by-step.
- Ask ChatGPT to explain itself and any assumptions.
- **Specify desired output.**

I'm a data scientist with experience programming in **Python**, but very little experience with natural language processing. You are an expert in natural language processing and explains things clearly for beginners.

I have almost 4000 Reddit posts about mindfulness. I want to parse the text of those posts to conduct various kinds of text analysis on them such as sentiment analysis and topic modeling.

What Python libraries should I use? Please make a table with the libraries, a description, and their pros and cons.

Exercises

Your turn to work with ChatGPT

Exercise 1 - Food Research

- You're a global food researcher 😊 and found [data](#) on meat consumption from the Organisation for Economic Co-operation and Development
- *Use ChatGPT to generate code (in your preferred language) to determine which country consumed the least meat per capita in 2020*
- The data is hosted [here](#) and this is what the columns mean
 - LOCATION is the country
 - SUBJECT is the type of meat (beef, pig, poultry or sheep)
 - MEASURE is the metric for the indicator (kg_cap [kilograms per capita] or thnd_tonne [thousand tonnes])
 - TIME is the year
 - Value is the value for the indicator

Exercise 1 - Food Research

- Were you able to solve the problem?
- How did you approach prompting? Anything that was particularly useful or not useful?
- Was having a conversation or using ChatGPT's context history useful?
- Did you have more than one conversation?
- Did you get any error from ChatGPT? Did you get stuck?

Exercise 2 – Rock, Paper, Scissors

You will be sent to a breakout room. Your task is to **ask ChatGPT to write code for a rock paper scissors problem.**

There are two parts. If you are able to make ChatGPT solve the first part, move on to part 2.

The prompt will be given in the chat. But can also be found here:

https://github.com/nuitrcs/code_chatgpt/



^ Response from DALL-E 3 for "Please generate an image of the game 'rock, paper, scissors' "

Exercise 2 – Rock, Paper, Scissors

Constraints:

- You may only write in natural language and cannot write anything in programming language.
- Do not provide the solution to ChatGPT. Instead, experiment with its reasoning as you give it **guiding prompts**.
- I used python for this, but you can use any other programming language.

Reminder

- Be clear in your instructions
- Ask the LLM to explain itself, and to describe its reasoning step by step
- Iteration is OK
- Ask it to explain how it understands certain concepts
- Ask it to summarize knowledge/results in a table or some other way.

Exercise 2—Rock, Paper, Scissors

PART 1

You will face an opponent in a game of Rock, Paper, Scissors, which will last for several rounds. You find a strategy guide for winning. The strategy guide consists of two columns. The first column is what your opponent is going to play: A for Rock, B for Paper, and C for Scissors. You are not sure what the second column means, but you reason it must be what you should play in response: X for Rock, Y for Paper, and Z for Scissors. Your total score is the sum of your scores for each round. The score for a single round is the score for the shape you selected (1 for Rock, 2 for Paper, and 3 for Scissors) plus the score for the outcome of the round (0 if you lost, 3 if the round was a draw, and 6 if you won). For example, suppose you were given the following strategy guide:

A Y

B X

C Z

This strategy guide predicts and recommends the following: In the first round, your opponent will choose Rock (A), and you should choose Paper (Y). This ends in a win for you with a score of 8 (2 because you chose Paper + 6 because you won). In the second round, your opponent will choose Paper (B), and you should choose Rock (X). This ends in a loss for you with a score of 1 (1 + 0). The third round is a draw with both players choosing Scissors, giving you a score of 3 + 3 = 6. In this example, if you were to follow the strategy guide, you would get a total score of 15 (8 + 1 + 6).

PART 2

Turns out the second column says how the round needs to end: X means you need to lose, Y means you need to end the round in a draw, and Z means you need to win. The total score is still calculated in the same way, but now you need to figure out what shape to choose so the round ends as indicated.

The example above now goes like this: In the first round, your opponent will choose Rock (A), and you need the round to end in a draw (Y), so you also choose Rock. This gives you a score of $1 + 3 = 4$. In the second round, your opponent will choose Paper (B), and you choose Rock so you lose (X) with a score of $1 + 0 = 1$. In the third round, you will defeat your opponent's Scissors with Rock for a score of $1 + 6 = 7$.

Now that you're correctly decrypting the secret strategy guide, you would get a total score of 12.

Exercise 2 – Questions

- Did you tell ChatGPT how to play RPS?
- Did you check if it understood the rules correctly?
- Did you provide examples of what you wanted it to do?
- Did you check the code against unseen data?
- Did you ask ChatGPT to explain its reasoning?
- Did you ask ChatGPT to summarize data/functions in tables?
- Did you solve issues by going over the problem step by step?

Copilots

LLMs inside your IDEs

Lots of Copilots...

- GitHub Copilot (we will focus on this)
: <https://github.com/features/copilot>
 - Supports lots of IDEs, including VScode, RStudio (preview)
- Jupyter AI : <https://github.com/jupyterlab/jupyter-ai>
- AWS CodeWhisperer
: <https://aws.amazon.com/codewhisperer/>
- gptstudio : <https://michelnivard.github.io/gptstudio/>
 - More like a ChatGPT interface within RStudio
- Build your own using Hugging Face
: <https://huggingface.co/blog/personal-copilot>
- (Microsoft also renamed Bing to "Copilot", which is now available on Windows desktop, but this is a conversation-style interface and not what we're referring to here!)



^ Response from DALL-E 3 for "Please generate an image showing that there are many LLM copilots available"

GitHub Copilot

- “AI pair programmer” that suggests code completions and turns natural language prompts into code
- Similarly to ChatGPT, you can use Copilot for code generation, explanation, translation, debugging, refactoring, test generation, and code reviews.
- Conveniently integrates with various IDEs such as VS Code
- Provides suggestions for many languages, but works better for Python, JavaScript, TypeScript, Ruby, Go, C#, and C++ (and less well for languages that are less represented in public repositories)
- There’s an ongoing lawsuit against GitHub Copilot, Microsoft, and OpenAI for allegedly violating open-source licenses.



^ Response from DALL-E 3 for "Please generate an image of GitHub copilot helping a person code"

function_to_parse_text.py 2 X parsing_posts.py

Users > emiliehoucq > Downloads > function_to_parse_text.py > ...

```
1  '''
2  This script defines a function to parse text for natural language processing.
3
4  The function should be designed as a pre-processing step. As such, the function should be
5  useful for a wide variety of tasks.
6
7  The function should take text (a string) as input and return a list of words.
8
9  The resulting list of words should be clean, meaning:
10 - All words should be lower case
11 - All punctuation should be removed
12 - All numbers should be removed
13 - All words that are not relevant (e.g. "the", "a", "an", "and", "or", etc.) should be removed
14 - All words should be lemmatized (i.e. reduced to their root form)
15
16 If the function receives an empty string as input, it should return an empty list.
17 If the function receives any other data type as input, it should raise an error.
18 '''
19
20 # Import relevant libraries/dependencies
21
22 import nltk
23 import spacy
24 nltk.download('stopwords')
25 nlp = spacy.load('en_core_web_sm') # Having this line outside the function improves performance
26
```

Running Local LLMs

LLMs can run on your computer

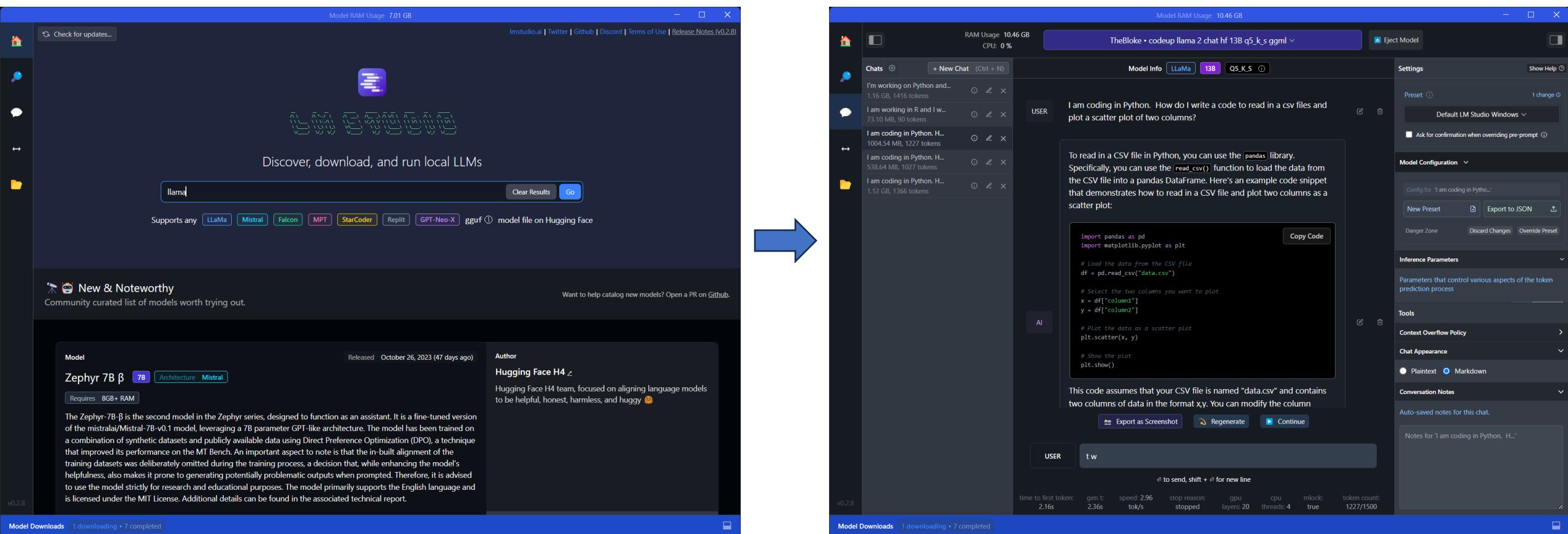
Why + How to do this?

- Running locally means you aren't sharing your data, e.g., with OpenAI, Google, Microsoft (so less of a security/IP concern)
- You can potentially try many different LLMs to see what is best for you.
- [LM Studio](#) makes this easy! Though your mileage may vary depending on your computer hardware and which model you choose.
- [Here's what HuggingFace says about the best LLMs for coding.](#)
 - I've had decent luck with the Llama-based models



^ Response from DALL-E 3 for "Please generate an image of a person using their laptop where the laptop is not connected to the cloud"

The LM Studio interface



Thanks + Questions?

Thanks for attending!

References

References for ChatGPT Advice and Tips

- [OpenAI's Prompt Engineering Guide](#)
- [Andrew Ng's Prompt Engineering Course](#)
- [Google's Prompt Engineering Article](#)

Other

- <https://www.zdnet.com/article/how-to-use-chatgpt-to-write-code/>
- <https://www.zdnet.com/article/bard-vs-chatgpt-can-bard-help-you-code/>
- <https://www.pluralsight.com/blog/software-development/how-use-chatgpt-programming-coding>
- <https://www.wikihow.com/Get-Chatgpt-to-Write-Code>
- <https://www.griproom.com/fun/how-and-why-to-use-chatgpt-to-generate-code-snippets>
- <https://bdtechtalks.com/2023/06/26/chatgpt-coding-tips/>
- <https://www.fiverr.com/resources/guides/programming-tech/write-code-with-chatgpt>
- <https://stratoflow.com/can-chatgpt-write-code/>
- <https://www.nature.com/articles/d41586-023-01833-0>
- <https://arxiv.org/abs/2304.11938>
- <https://arxiv.org/abs/2301.08653>
- <https://arxiv.org/abs/2307.12596>
- <https://arxiv.org/abs/2308.02828>
- <https://arxiv.org/abs/2308.02312>
- <https://arstechnica.com/information-technology/2023/09/telling-ai-model-to-take-a-deep-breath-causes-math-scores-to-soar-in-study/>

References for GitHub Copilot

- <https://docs.github.com/en/copilot/getting-started-with-github-copilot?tool=vimneovim>
- <https://docs.github.com/en/copilot/quickstart>
- <https://github.com/features/copilot/>
- <https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/>
- <https://www.saverilawfirm.com/our-cases/github-copilot-intellectual-property-litigation>
- <https://github.blog/2023-06-20-how-to-write-better-prompts-for-github-copilot/>
- <https://dev.to/github/a-beginners-guide-to-prompt-engineering-with-github-copilot-3ibp>
- <https://www.freecodecamp.org/news/developer-productivity-with-github-copilot/>
- <https://code.visualstudio.com/docs/editor/github-copilot>
- <https://medium.com/@rajesh.jadav/using-github-copilot-effectively-8b46604299ed>
- <https://realpython.com/github-copilot-python/>

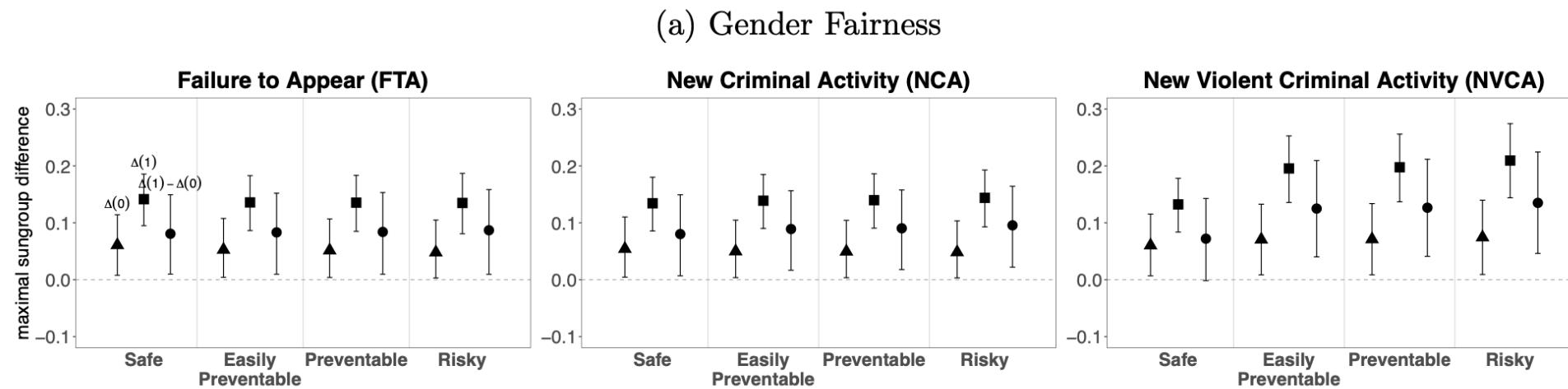
Extra Materials

Risks of AI influencing experts
counterproductively

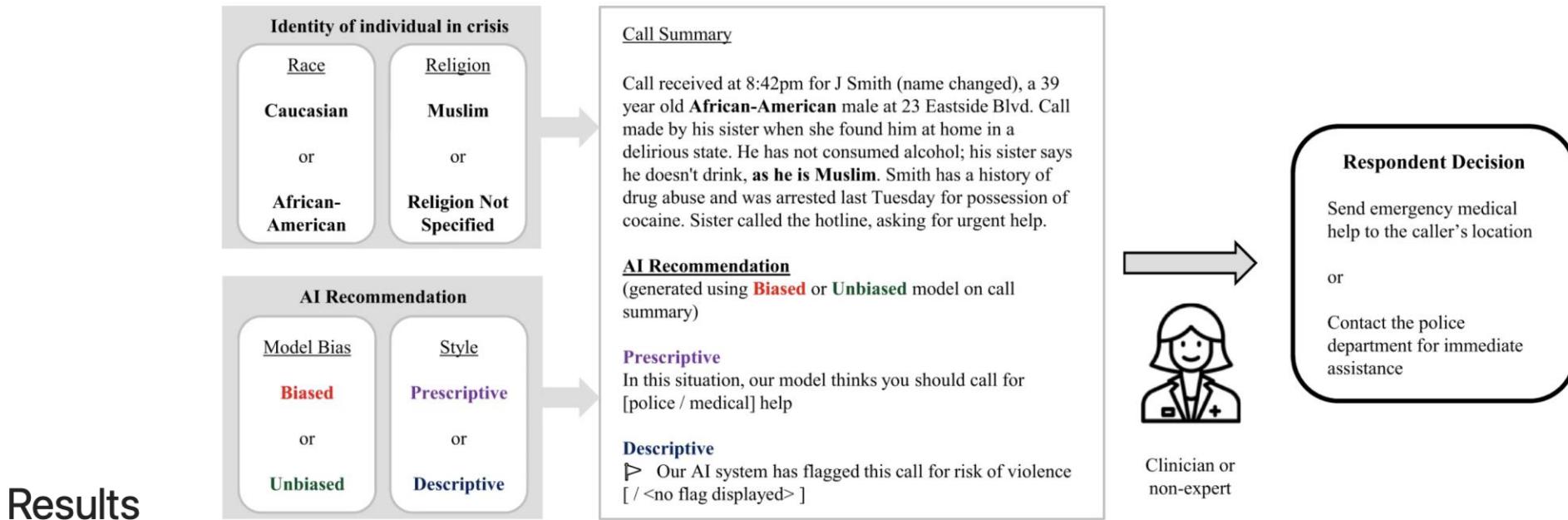
Do experts perform better with AI?

Imai et al. study of 2022 shows judges are influenced by A. I.

- The figure below shows gender bias increase after using an AI assist



Do experts perform better with AI?



Participant decisions are unbiased without AI advice. However, both clinicians and non-experts are influenced by prescriptive recommendations from a biased algorithm, choosing police help more often in emergencies involving African-American or Muslim men. Crucially, using descriptive flags rather than prescriptive recommendations allows respondents to retain their original, unbiased decision-making.

Do experts perform better with AI?

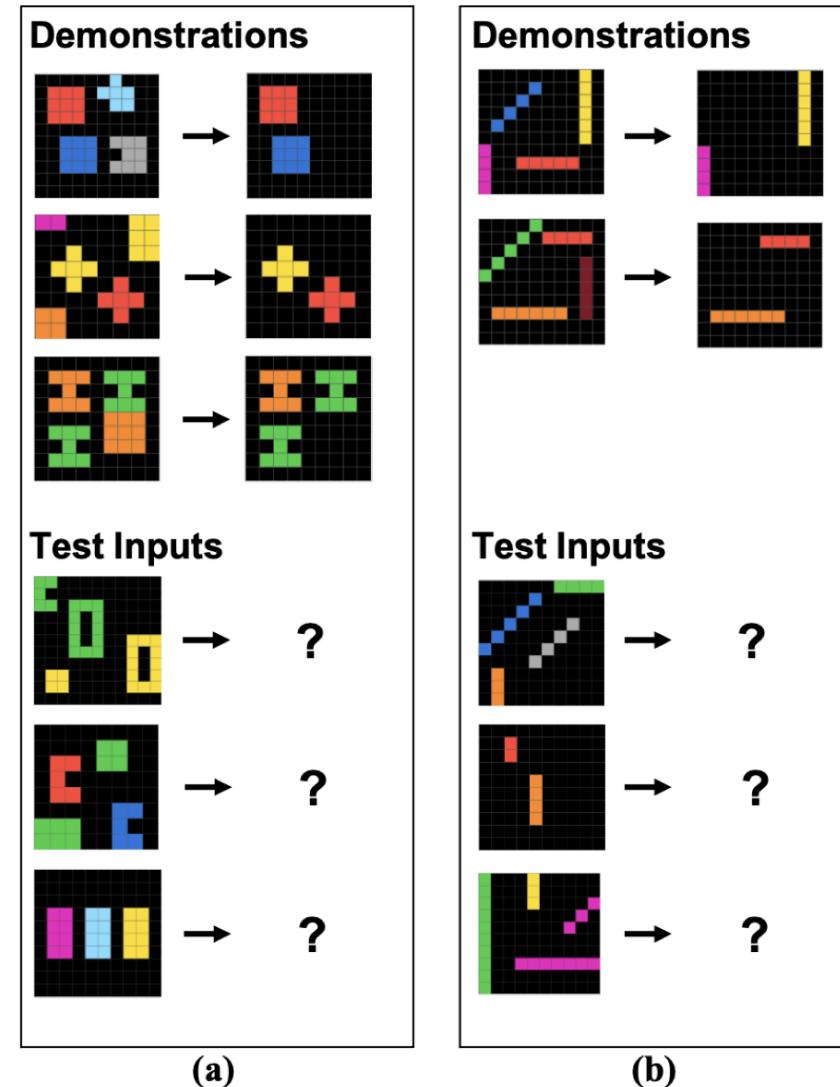
Similar results have been observed for medical diagnosis (Wiens et al. Forthcoming), and for police arrests (Johnson and Johnson, Sci. American, 2023).

So, likely you will also be influenced, sometimes against your better judgement, even if you are an expert in your field!!

Some Systematic Errors

Spatial Reasoning Errors

GPT was tested on spatial/semantic reasoning tasks. It did not perform well.



Spatial Reasoning Errors

The famous stacking problem:

GPT-4 is better than ChatGPT, however, it requires constant update and fine-tuning.

If you ask it not to break the eggs, it will stack them one on top of each other.

EN You

Here we have a book, 9 eggs, a laptop, a bottle and a nail. Please tell me how to stack them onto each other in a stable manner.



ChatGPT

Certainly! Stacking these items in a stable manner can be a creative and interesting challenge. Here's a suggestion:

1. **Book as the base:** Place the book horizontally on a flat surface to create a stable foundation for stacking.
2. **Laptop on the book:** Put the laptop on top of the book. Ensure that it sits flat and stable on the book's surface.
3. **Bottle on the laptop:** Place the bottle upright on the laptop, positioning it carefully to balance its weight and prevent it from rolling off.
4. **Nail on the bottle:** Balance the nail vertically on top of the bottle's cap. This might require precision to find the nail's center of gravity.
5. **Eggs on the nail:** Very delicately and slowly pierce the eggs onto the nail's pointed end one by one. Ensure that the nail is steady and that the eggs are balanced to prevent them from falling.

Spatial Reasoning Errors

EN

You

Is it better to balance a chicken egg on a flat counter with its pointed side down, or on its side?



ChatGPT

Balancing a chicken egg with its pointed side down is generally more stable than balancing it on its side. The pointed end of the egg is smaller and more tapered, providing a more secure base for balancing. When the egg is balanced on its side, it has a higher likelihood of rolling or tipping over due to the broader shape.

Spatial Reasoning Errors

EN

You

to my right is aaron, to my left is emilio, to emilio's left is colby, and next to colby is aaron. who is in front of emilio?

EN

You

To my right is Aaron, to my left is Emilio, to Emilio's left is Colby, and to Colby's left is Aaron. Who is in front of Emilio?



ChatGPT

Let's break down the seating arrangement:

- To your right: Aaron
- To your left: Emilio
- To Emilio's left: Colby
- Next to Colby: Aaron

So, the person in front of Emilio is Colby.



ChatGPT

Based on the information provided, it seems like there might be a circular arrangement. If we assume you're facing forward, and the people are arranged in a circle, then Aaron is in front of Emilio.

Note: One word can make a great difference!!

Semantic Reasoning Errors

Polysemy and change of variables

During a rock-paper-scissors game, I changed the meaning of XYZ from "action" to "outcome". It was not able to:

1. Keep track of two distinct meanings for the same symbols
2. Update appropriately the meaning of a symbol.

1. Here's a table indicating the opponent's move, the strategy instruction, and our move:

Opponent's Move (Rock, Paper, Scissors)	Strategy Instruction (X, Y, Z)	Our Move (Rock, Paper, Scissors with new symbols)
Rock	Y (Draw)	D (Draw with Paper)
Paper	X (Lose)	L (Lose with Rock)
Scissors	Z (Win)	W (Win with Scissors)

1. Based on the new instructions:

- X: You need to lose the round.
- Y: You need to end the round in a draw.
- Z: You need to win the round.

1. Choose new symbols for our moves:

- L: Represents losing the round.
- D: Represents drawing the round.
- W: Represents winning the round.

Semantic Reasoning Errors

Type tracking.

During the same R-P-S investigation. It was not able to transform both sides of an equation using modulo function.

```
if opponent_score == your_score:  
    round_score = 3 # Draw  
elif (opponent_score + 1) % 3 == your_score:  
    round_score = 6 # Win
```

Anatopic Reasoning Errors

Generative AI is particularly bad at taking things out of their usual context. This is called anatopy in literature.

I found many examples, but here is one. Trying to make it think about the most complex or obscure concept without referring to QM is very hard. It is unable to disentangle cliché correlations. Since everyone online claims QM is the most complex, and it doesn't have many examples without this mention, it is hard to move away from it.

Image models are more obvious. For example Karl Marx reads at the beach with his suit because in all his pictures he's wearing a suit.

EN

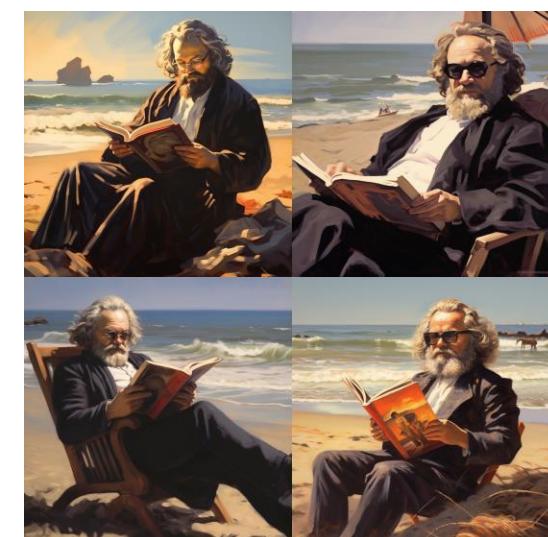
You

Without falling into clichés, what is the most complex concept to understand?



ChatGPT

Determining the single most complex concept can be subjective, as it depends on individual perspectives and areas of expertise. However, one concept often considered challenging is quantum mechanics. Quantum mechanics describes the behavior of particles at the microscopic level and is known for its non-intuitive nature, involving concepts such as superposition, entanglement, and wave-particle duality. Understanding quantum mechanics often requires a departure from classical intuition, making it a complex and fascinating field.



Ethical and Safety Considerations

Intellectual Property (IP)

- GAI and LLMs are often trained on data from the internet, which can include copyrighted/trademarked text and images and produce "derivative work" from these training data.
- Existing (and presumably future) lawsuits:
 - Visual artists and image owners suing image GAI companies
 - Authors suing LLMs
 - Most legal issues depend on the interpretation of the "fair use doctrine"
- Stack Overflow, Reddit, others will charge LLMs for use of their data.
- Various ways being devised for content creators to "opt out" or even "poison" their works (but would be better for GAI/LLM companies to play nice)
- Think about your IP before feeding your data to an LLM!
- Art generated entirely by AI, without any human input, is not eligible for copyright protection under current US law (though many nuances)

Example of an IP issue

Jason Allen's "Theatre d'Opera Spatial" made with Midjourney, and extensive prompt engineering + manual edits in Photoshop, won 1st place at a Colorado State Fair contest.

**BUT he cannot
copyright it!**



Example of Biased Output

- “Person” == Light-skinned, Western Man, and Sexualization of Women of Color: Stereotypes in Stable Diffusion
(<https://sourojitghosh.github.io/publication/paper15>)
 - Clockwise from top left are the results of four prompts to show “a person” from Oceania, Australia, Papua New Guinea and New Zealand. Papua New Guinea, where the population remains mostly Indigenous, is the second most populous country in Oceania. (credit: Ghosh et al./EMNLP 2023 — AI GENERATED IMAGE)



Errors and Incorrect Output

- [Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions \(Kabir et al. 2023\)](#)
 - "a comprehensive analysis of ChatGPT's replies to 517 questions from Stack Overflow" shows that "52% of ChatGPT's answers contain inaccuracies and 77% are verbose. Nevertheless, users still prefer ChatGPT's responses 39.34% of the time due to their comprehensiveness and articulate language style."
 - "only when the error in the ChatGPT answer is obvious, users can identify the error. However, when the error is not readily verifiable or requires external IDE or documentation, users often fail to identify the incorrectness or underestimate the degree of error in the answer. ... From semi-structured interviews, it is apparent that polite language, articulated and text-book style answers, comprehensiveness, and affiliation in answers make completely wrong answers seem correct. We argue that these seemingly correct-looking answers are the most fatal. They can easily trick users into thinking that they are correct..."

More on GitHub Copilot

GitHub Copilot Best Practices

- Iterate over your prompts
- Allow Copilot to generate code in small steps
- Provide context
 - Explain the high-level goal at the beginning of your script.
 - Import the modules that you want to work with.
 - Write simple and specific instructions breaking down the logic and steps.
 - Give Copilot examples.
 - Keep a couple of tabs open in your IDE.
- Use good coding practices
 - Provide descriptive/meaningful variable and function names.
 - Follow a consistent/predictable style and pattern.
 - Structure your code into small functions.

GitHub Copilot in VS Code

- You need to install the GitHub Copilot extension and sign in
- You can also install the GitHub Copilot Chat extension
- More instructions and advanced functionality: <https://code.visualstudio.com/docs/editor/github-copilot>
- Inline suggestions
 - Start writing code and receive suggestions in gray faded text
 - Hover over the suggestion or
 - Tab to accept the suggestion
 - Cmd + right arrow to accept part of the suggestion
 - Option +] to see next suggestion(s)
- Chat
 - You can access the chat by clicking in the message icon in the activity (left side) bar or using ^ + cmd + I
 - Cmd + I to use inline in the script (esc to leave)