



A Tale of GPT Models

CSC401/2511 – Natural Language Computing – Winter 2023

Lecture 13

University of Toronto

Logistics

- All contents of this lecture will **not** be on the exam.
- Assignment 3 due on April 8.

- The AutoModelForCausalLM questions: Don't worry if the accuracies are low.

ChatGPT, help me write an intro

ZI

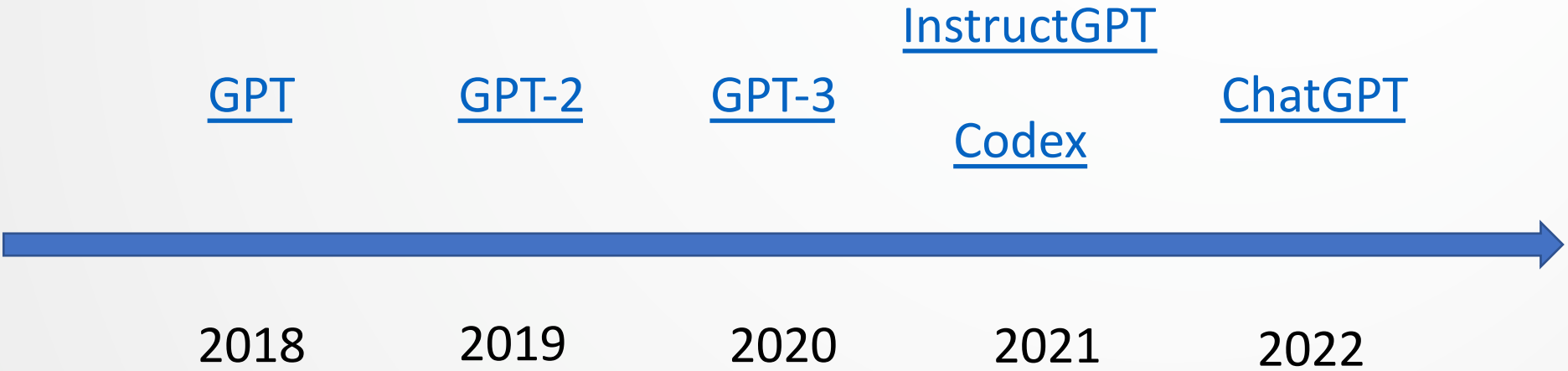
In a welcoming tone, write a short introduction for a Natural Language Processing lecture that introduces the development history of GPT models.



Welcome to our Natural Language Processing lecture on the development history of GPT models. In this lecture, we will explore the evolution of GPT models from their early beginnings to the current state-of-the-art models. We will discuss the key milestones and breakthroughs that have contributed to the development of these powerful language processing tools. By the end of this lecture, you will have a better understanding of the history and progression of GPT models and their impact on the field of natural language processing. Let's begin.

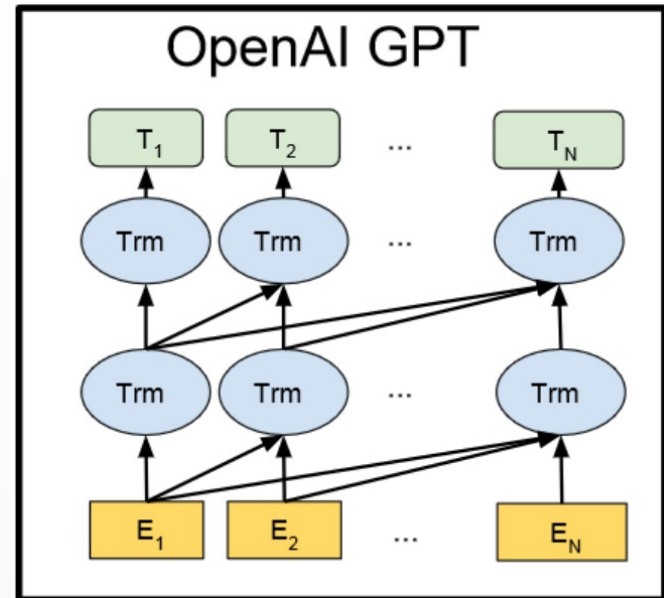


A timeline of GPT models



Generative Pre-Training

- Uni-directional Transformers.
 - “Transformer decoder”.
- Pre-training.
 - Language modeling objective.
- Fine-tune on multiple tasks.
 - Share all layers except the top one.



Improving language understanding with generative pre-training (GPT; Radford et al., 2018)

General-purpose understanding

- Only use the language modeling pre-training objective.

$$\min_{\theta} \mathbb{E}_t \log \frac{1}{p(x_t | x_{1:(t-1)}; \theta)}$$

- Intuition: Modeling the language gives enough knowledge about the tasks.
 - The corpora sizes differ:
 - GPT-2 trains on 40GB of text
 - GPT trains on BookCorpus – 1.1GB on huggingface hub

Language models are unsupervised multitask learners (GPT-2; Radford et al., 2019)

GPT-2: language generation

Context (human-written): In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

Prompt

GPT-2: The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Generation

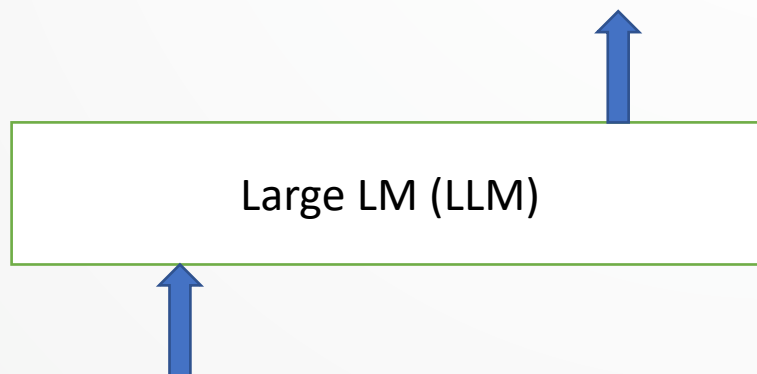
Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be

Prompt-based system

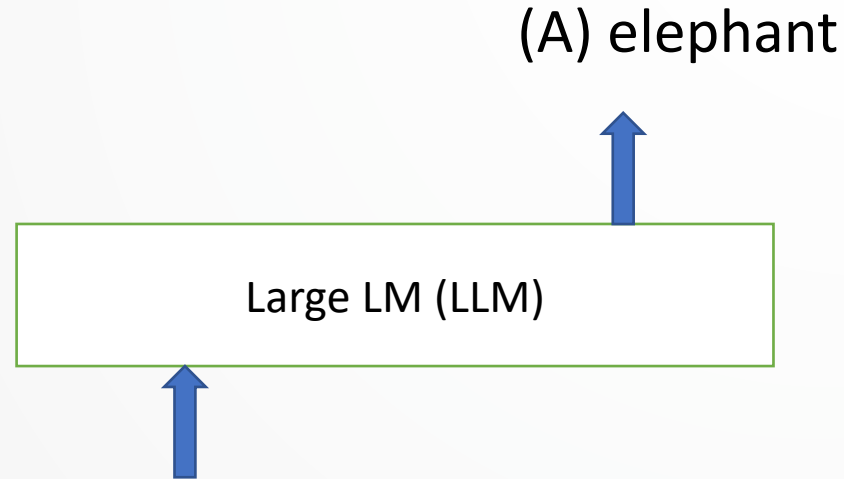
that I had an elephant in my
fridge while eating dinner at a
family holiday party with my



An elephant can't be put into a fridge because it is too large.
What is it? (A) elephant or (B) fridge? The answer is

Learning from prompt

**One-shot
demonstration**



An elephant can't be put into **a fridge** because **it** is too small.
What is it? (A) elephant or (B) fridge? The answer is (B) fridge.
An elephant can't be put into **a fridge** because **it** is too large.
What is it? (A) elephant or (B) fridge? The answer is

GPT-3: High-quality data

- Training data comes from texts crawled from web.
 - Crawled text data have low qualities: html tags, duplications, etc.
 - Train a “data quality” predictor to automatically filter.
 - 45TB -> 500GB after filtering.
- Result: GPT-3 davinci model.
- How to further improve the model?

Language models are few-shot learners. (GPT-3; Brown et al 2020)

Codex: Learn from codes

- What are high quality languages?
 - Correct syntax
 - Clear structures
 - Reasonable vocabulary
 - Actionable commands
- Collected 159GB of codes + docstrings
- Results:
 - Copilot / Codex: a 12B-parameter model.
 - code-davinci-002: a 175B (GPT-3 sized) model.

Programming
languages!

Evaluating language models trained on code. (Chen et al., 2021)

Performance of Codex

```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date, value, currency).
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8         2016-01-02 -34.01 USD
9         2016-01-03 2.59 DKK
10        2016-01-03 -2.72 EUR
11    """
12    expenses = []
13    for line in expenses_string.splitlines():
14        if line.startswith("#"):
15            continue
16        date, value, currency = line.split(" ")
17        expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
18                        float(value),
19                        currency))
20    return expenses
```

Fine-tuning from human demonstration

- Collect the prompts submitted to the Playground.
- Recruit humans to write desired generations.
- Fine-tune the davinci model on this “FeedMe” dataset.

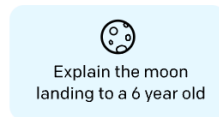
- Result: GPT-3 text-davinci-002 model.
 - Note: text-davinci-002 is fine-tuned from code-davinci-002.

RL from human feedback

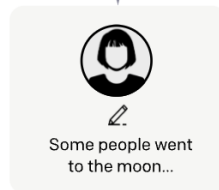
Step 1

Collect demonstration data, and train a supervised policy.

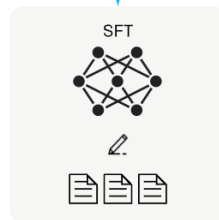
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



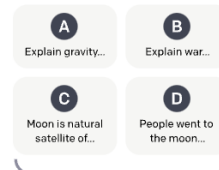
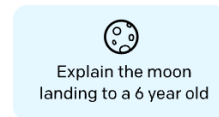
This data is used to fine-tune GPT-3 with supervised learning.



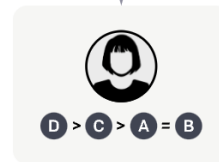
Step 2

Collect comparison data, and train a reward model.

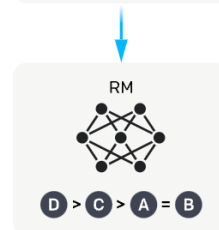
A prompt and several model outputs are sampled.



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 reinforcement learning.

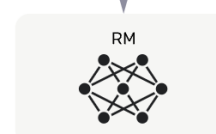
A new prompt is sampled from the dataset.



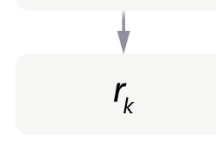
The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



Original RLHF paper (InstructGPT):
Tutorial for RLHF: <https://huggingface.co/blog/rlhf>

Demo of InstructGPT

PROMPT

Explain the moon landing to a 6 year old in a few sentences.

COMPLETION

GPT-3

Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

Explain the big bang theory to a 6 year old.

Explain evolution to a 6 year old.

InstructGPT

People went to the moon, and they took pictures of what they saw, and sent them back to the earth so we could all see them.

Recipe of ChatGPT

ChatGPT \approx GPT3 + InstructGPT + Codex \approx

Uni-directional large-scale GPT +

High-quality text data +

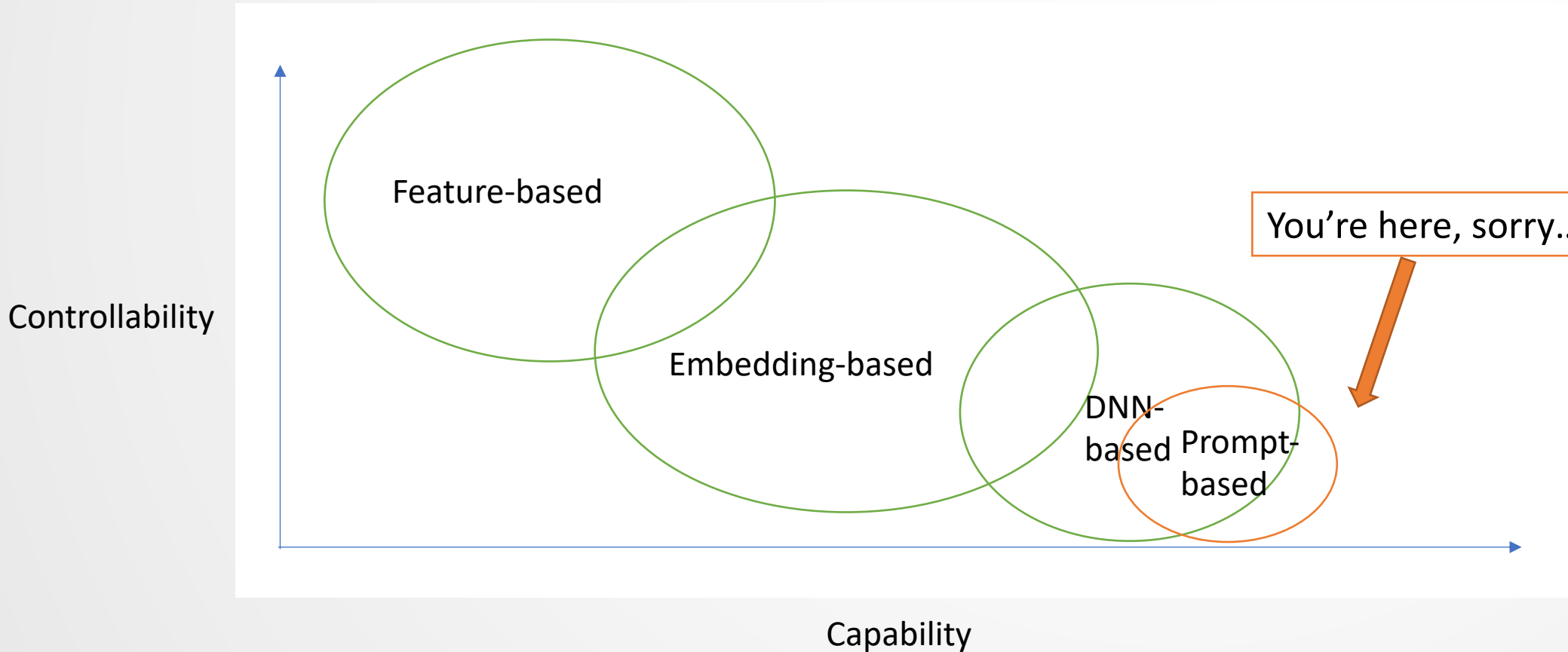
Learning from codes +

Human demonstration +

Reinforcement learning from human feedback +

Massive engineering effort.

A new way to build NLP systems

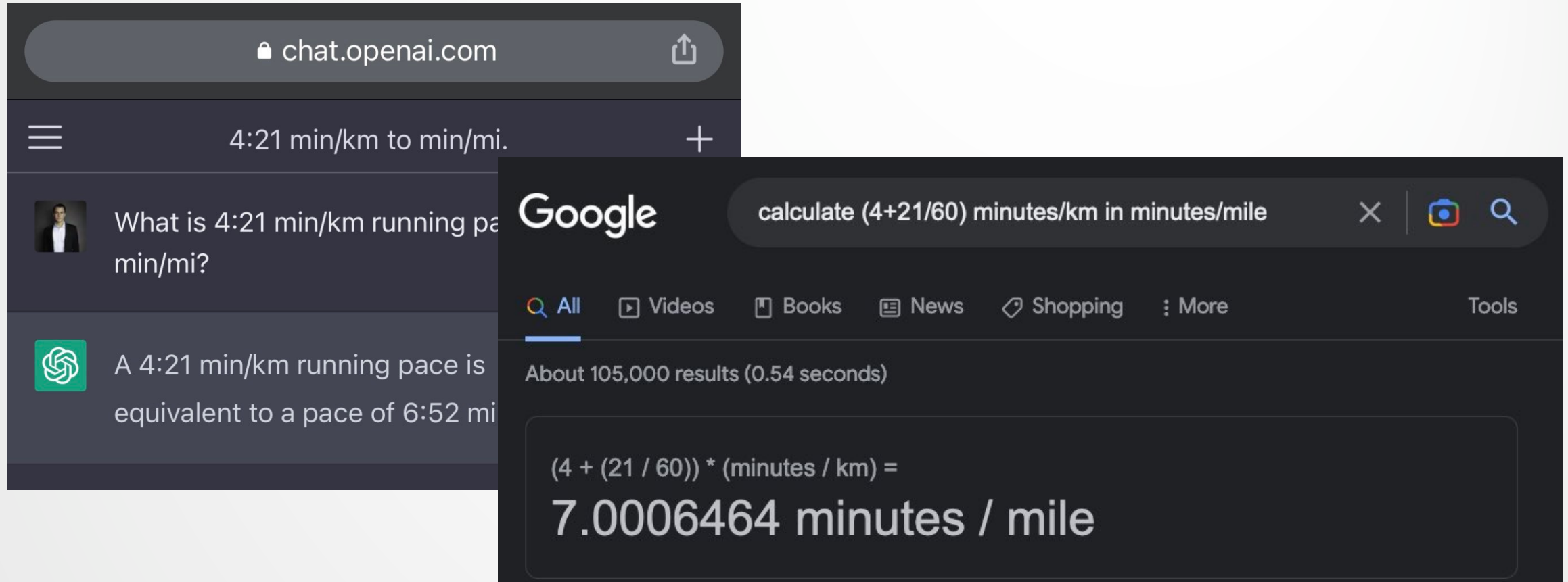


LLMs hallucinate coherently

The Canadian political leader was previously Stephen Harper, who lost an election to Justin Trudeau. On January 6, what happened in Canada's Capitol in Ottawa?

On January 6, a mob of supporters of former U.S. President Donald Trump stormed the Canadian parliament in Ottawa. The rioters broke through the gates, smashed windows, and desecrated the Canadian Parliament. The RCMP responded to the incident and arrested three people. No serious injuries were reported.

LLMs BS confidently



The image shows two overlapping screenshots. The background screenshot is from chat.openai.com, showing a user asking "What is 4:21 min/km running pace in min/mi?" and a GPT-4 response stating "A 4:21 min/km running pace is equivalent to a pace of 6:52 mi". The foreground screenshot is from Google, showing a search for "calculate (4+21/60) minutes/km in minutes/mile" with a result of "7.0006464 minutes / mile".

chat.openai.com

4:21 min/km to min/mi.

What is 4:21 min/km running pace in min/mi?

A 4:21 min/km running pace is equivalent to a pace of 6:52 mi

Google

calculate (4+21/60) minutes/km in minutes/mile

All Videos Books News Shopping More Tools

About 105,000 results (0.54 seconds)




$(4 + (21 / 60)) * (\text{minutes} / \text{km}) =$
7.0006464 minutes / mile

Source: <https://garymarcus.substack.com/p/scientists-please-dont-let-your-chatbots>

LLMs discriminate incorrectly

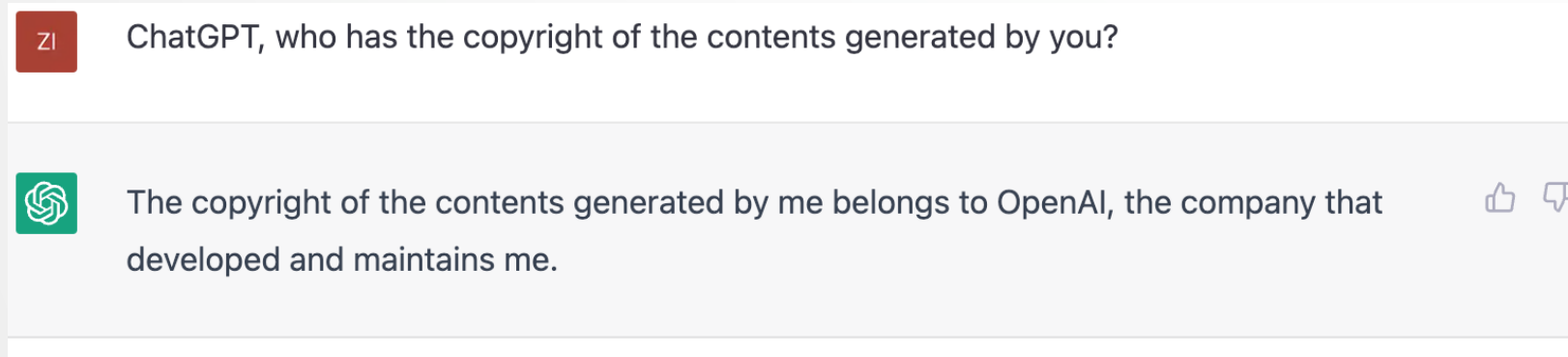
zi Why is the University of Toronto the best place to work at?

 zi Why is the McGill University the best place to work at?
res

inc
gro  I am not aware of any specific information that would indicate that McGill University is the "best" place to work. However, it is a well-respected institution and is consistently ranked among the top universities in Canada and the world. It has a diverse student population and a wide range of academic programs, which can provide a stimulating and dynamic work environment for faculty and staff. Additionally, it is located in Montreal, a vibrant and multicultural city with a high quality of life.  

LLMs copy-paste irresponsibly

- There is no attribution to the source of the information.
- Who owns the generated contents?
 - Here's a response from ChatGPT:



- OpenAI, GitHub and Copilot are currently challenged by a lawsuit about **copyright**. Read more about the litigation [here](#).

Risks of LLM: poor evaluations

- Never train on test data.
- ... but how much leaderboard data did Chat and GPT-4 include in their training?
- ... and how much data that are “very similar to the leaderboard data” are included in their training?

Risks of LLM: an API perspective

- Over-reliance on API
 - On March 23, 2023: OpenAI discontinued serving the Codex model API and recommends using gpt-3.5 (Chat) instead.
- When processing sensitive and proprietary data: GPT3, and other API-based server is a strong no-go.

Risks of LLM: improper usage

- Trusting the outputs of LLMs when one should not.
 - E.g., Writing an assignment?
 - E.g., writing or reading a newspaper article?
- Utilizing LLMs to exaggerate inequality.
 - Who have access to the best AI technologies?
 - Who are the most aware of how to use these AI technologies?
- Discussion: Other risks of LLM?

Artificial general-purpose assistants

- Some argue that LLMs are already AGIs.
- LLMs can be useful automation tools though, similar to search engines.
- What will be the future of AI?

Anonymous feedback form

- <https://forms.gle/W3i6AHaE4uRx2FAJA>

