



# COMP 336 I Natural Language Processing

Lecture 15: LLM prompting, in-context learning,  
scaling laws

Spring 2025

# Announcements


- Final exam is scheduled at 2:30 pm - 4:30 pm on May 20 @Rm 167 Main Bldg.
- [#assignment-2](#) due next week!
  - Join [#assignment-2](#) Slack channel for discussion

# Latest AI news

**Google** @Google · Mar 26  
Introducing **Gemini 2.5**, our most intelligent AI model.

Our first release, an experimental version of **2.5 Pro**, unlocks state-of-the-art performance in math and science. 🔥

Learn more 📖



A black rectangular area containing the Gemini 2.5 logo, which consists of a blue four-pointed starburst icon to the left of the text "Gemini 2.5" in white. A small "GIF" icon is visible in the bottom left corner of the image area.

214 559 3.4K 282K

**OpenAI** @OpenAI · Mar 26  
4o image generation has arrived.

It's beginning to roll out today in ChatGPT and Sora to all Plus, Pro, Team, and Free users.



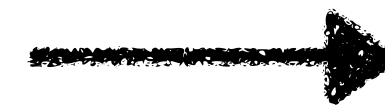
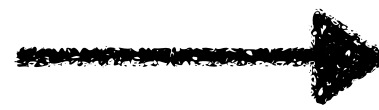
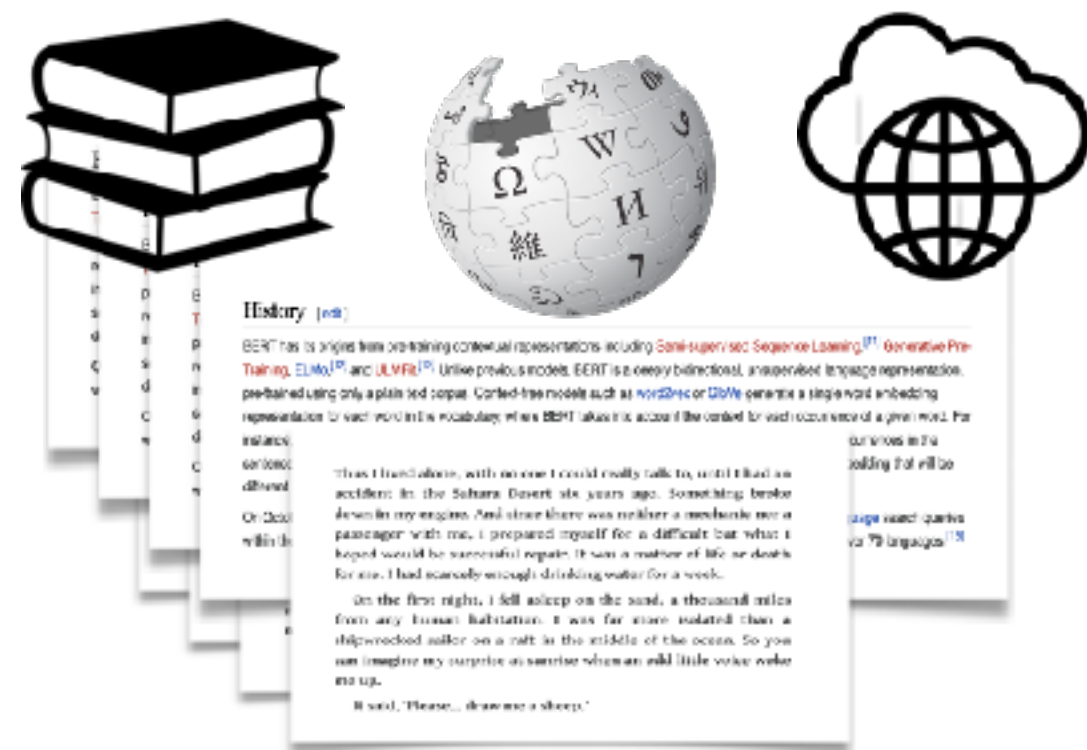
A hand holding a brown, textured boarding pass card. The card has the following text: "BOARDING PASS" with an airplane icon, "INTRODUCING 4o IMAGE GENERATION NOW IN CHATGPT AND SORA", "MARCH 25 11AM PDT", and "As you can tell by now it's very good at text. Look at all this accurate text!". The background is a blurred airport terminal.

743 2.7K 20K 4.2M

GPT4o image generation

# Pretraining: training objectives?

- During pretraining, we have a large text corpus (**no task labels**)
  - **Key question: what labels or objectives used to train the vanilla Transformers?**



**Training labels/objectives?**

Pretraining Transformers

# Pretraining objectives



BERT (Encoder-only)  
Devlin et al., 2018

The cabs \_\_\_ the same rates as those \_\_\_ by horse-drawn cabs and were \_\_\_ quite popular, \_\_\_ the Prince of Wales (the \_\_\_ King Edward VII) travelled in \_\_\_. The cabs quickly \_\_\_ known as "hummingbirds" for \_\_\_ noise made by their motors and their distinctive black and \_\_\_ livery. Passengers \_\_\_ the interior fittings were \_\_\_ when compared to \_\_\_ cabs but there \_\_\_ some complaints \_\_\_ the \_\_\_ lighting made them too \_\_\_ to those outside \_\_\_.

charged, used, initially, even, future, became, the, yellow, reported, that, luxurious, horse-drawn, were that, internal, conspicuous, cab

Masked token prediction



T5 (Encoder-decoder)  
Raffel et al., 2019

Original text  
Thank you ~~for inviting~~ me to your party ~~last~~ week.

Inputs  
Thank you <X> me to your party <Y> week.

Targets  
<X> for inviting <Y> last <Z>

Denosing span-mask prediction



GPT - 4  
Decoder-only

Text: Second Law of Robotics: A robot must obey the orders given it by human beings

Generated training examples

Example #	Input (features)	Correct output (labels)
1	Second law of robotics :	a
2	Second law of robotics : a	robot
3	Second law of robotics : a robot	must
...		

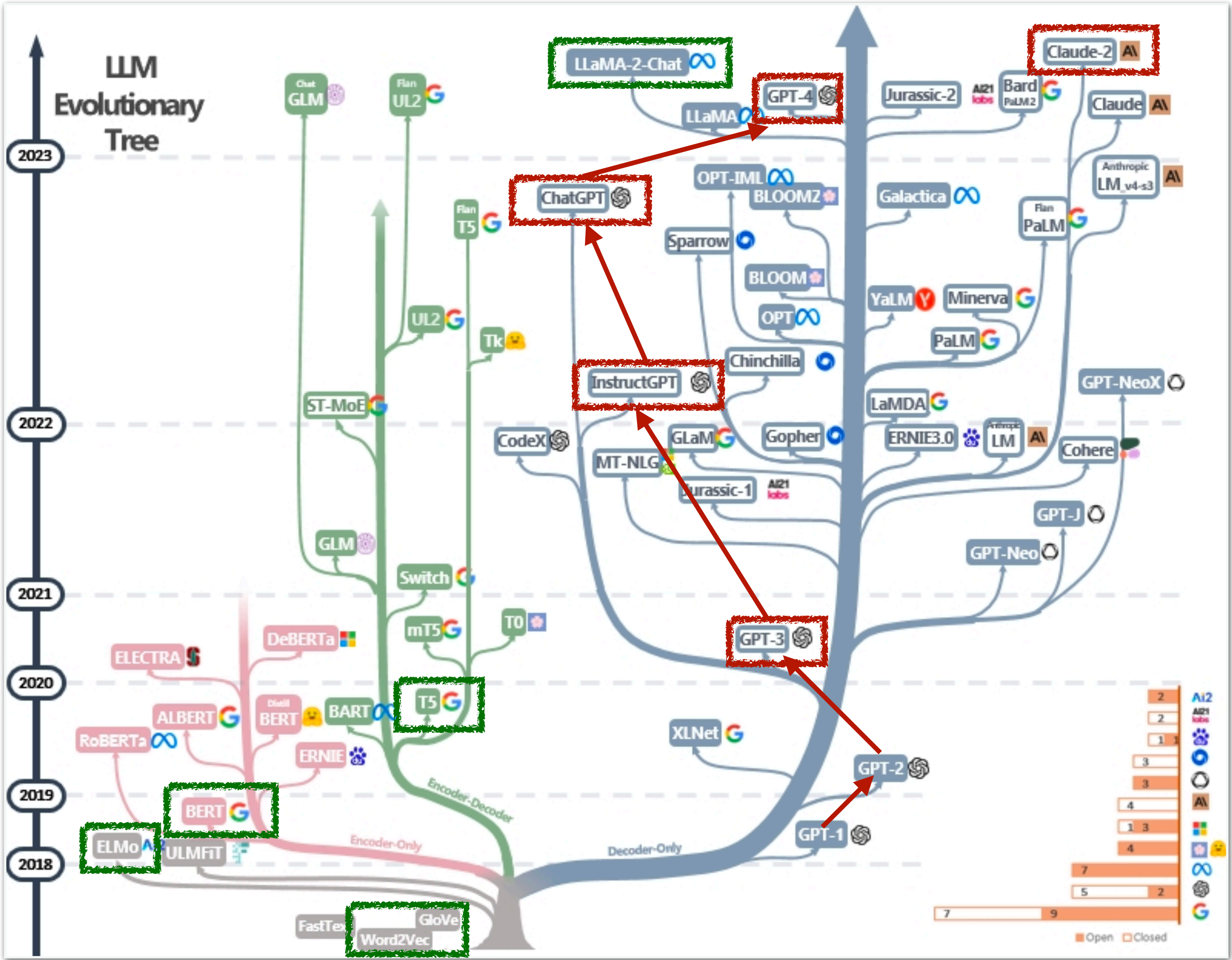
Next token prediction

# Evolution tree of pretrained LMs

~200 billion

Model size  
(# of parameters)  
~1000 times larger

~300 million



Open-sourced

Close-sourced

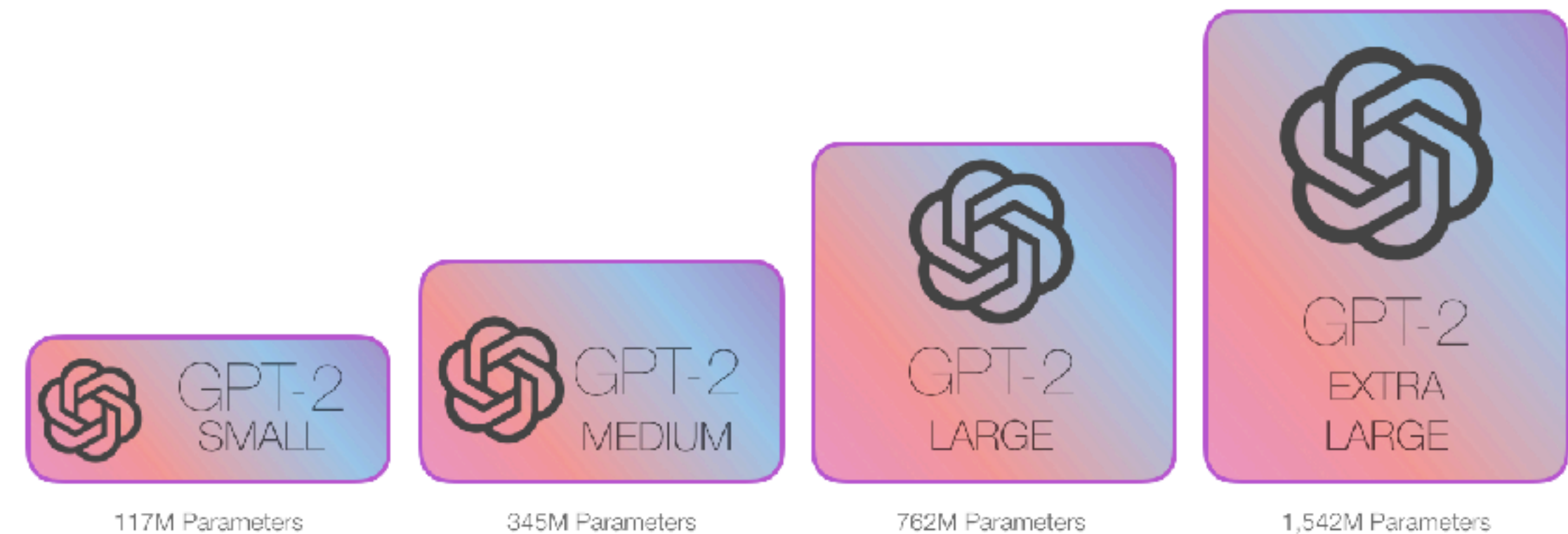
# From GPT-1 to GPT-2 to GPT-3

- All decoder-only Transformer-based language models
- Model size ↑, training corpora ↑

GPT-2



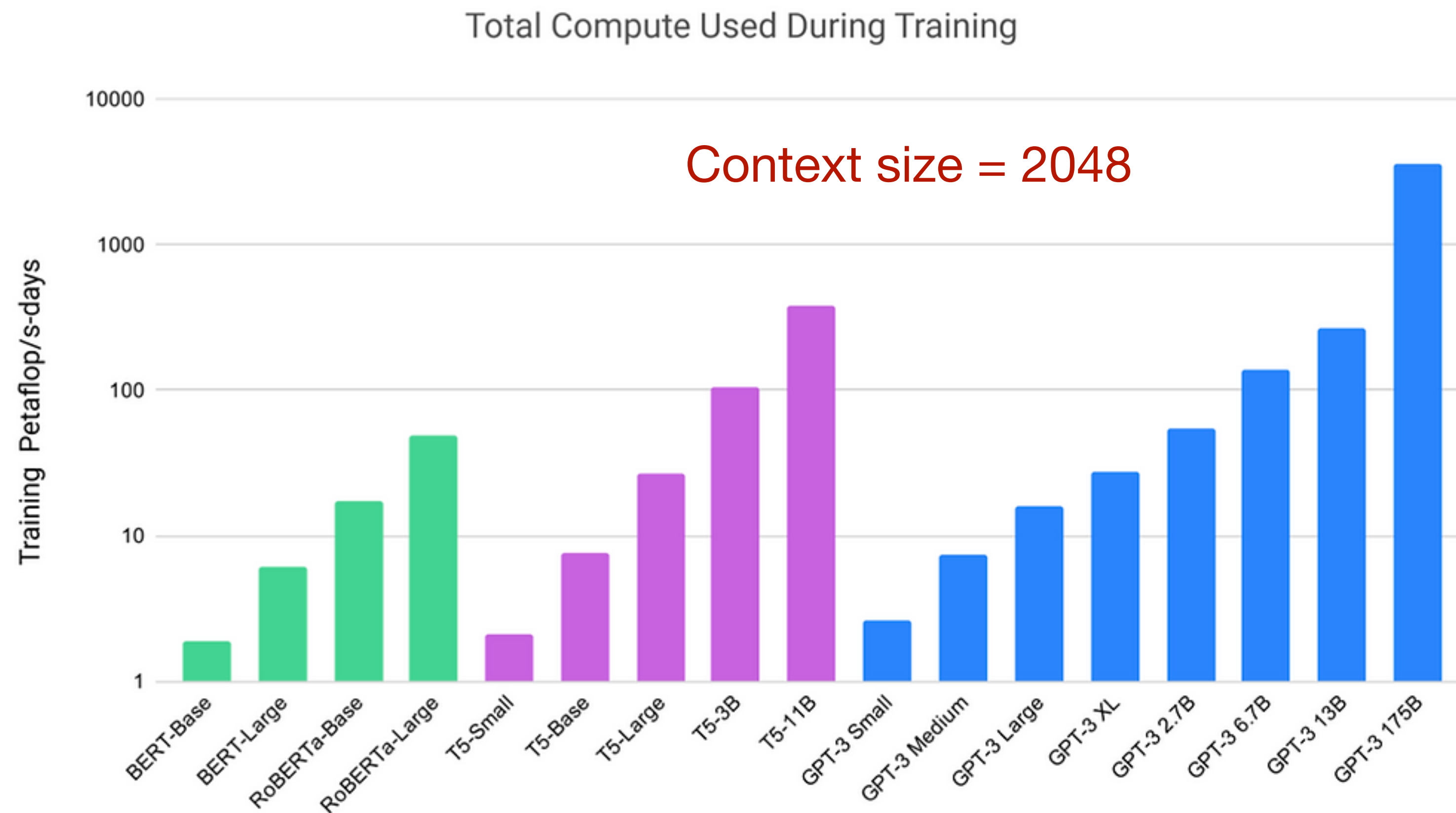
Context size = 1024



.. trained on 40Gb of Internet text ..

# GPT-3: language models are few-shot learners

- GPT-2 → GPT-3: 1.5B → **175B** (# of parameters), ~14B → **300B** (# of tokens)



Training computation is measured using floating-point operations or “FLOP”.

One FLOP represents a single arithmetic operation involving floating-point numbers, such as addition, subtraction, multiplication, or division.



# Before GPT3: Modern learning paradigm

- **Pre-training + supervised training/fine-tuning**

- First train Transformer using a lot of general text using unsupervised learning. This is called **pretraining**.
- Then train the pretrained Transformer for a specific task using supervised learning. This is called **finetuning**.



# Paradigm shift since GPT-3

- Before GPT-3, **Pre-training + supervised training/fine-tuning** is the default way of doing learning in models like BERT/T5/GPT-2
  - SST-2 has 67k examples, SQuAD has 88k (passage, answer, question) triples
- Fine-tuning requires computing the gradient and applying a parameter update on every example (or every K examples in a mini-batch)
- However, this is very expensive for the 175B GPT-3 model

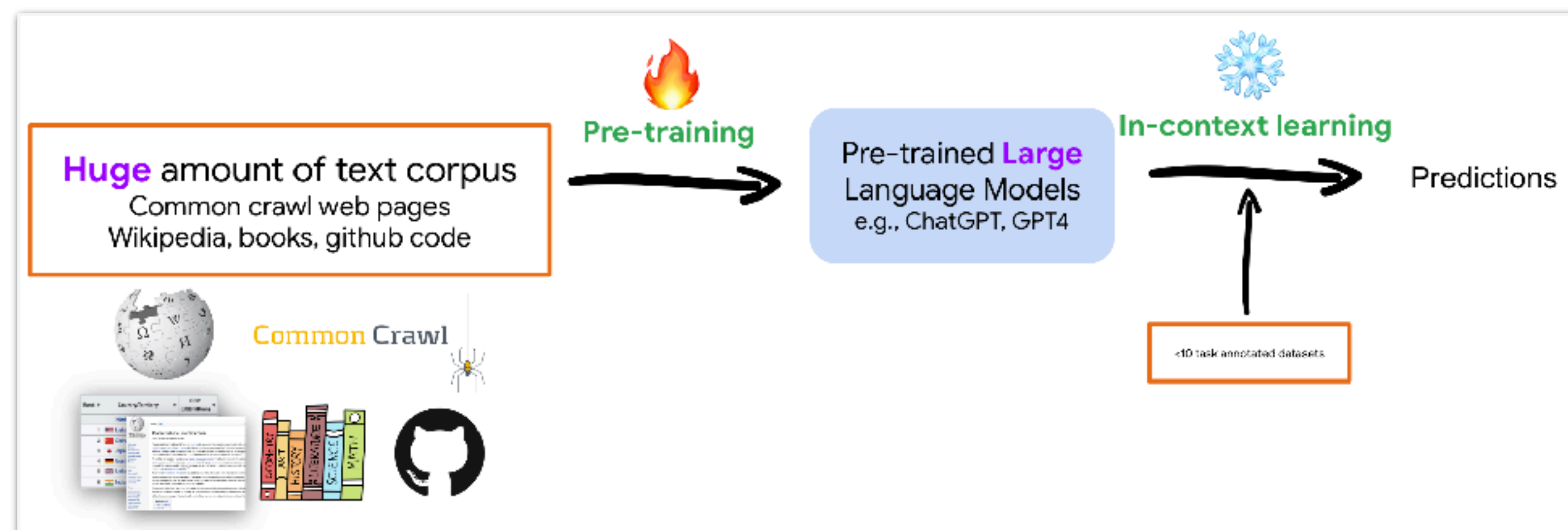
## Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.



# Latest learning paradigm shift since GPT-3

- **Pre-training + prompting/in-context learning (no training this step)**
  - First train a **large (>7~175B)** Transformer using a lot of general text using unsupervised learning. This is called **large** language model **pretraining**.
  - Then **directly use** the pretrained large Transformer (**no further finetuning/training**) for any different task given only a natural language description of the task or a few task (x, y) examples. This is called **prompting/in-context learning**.

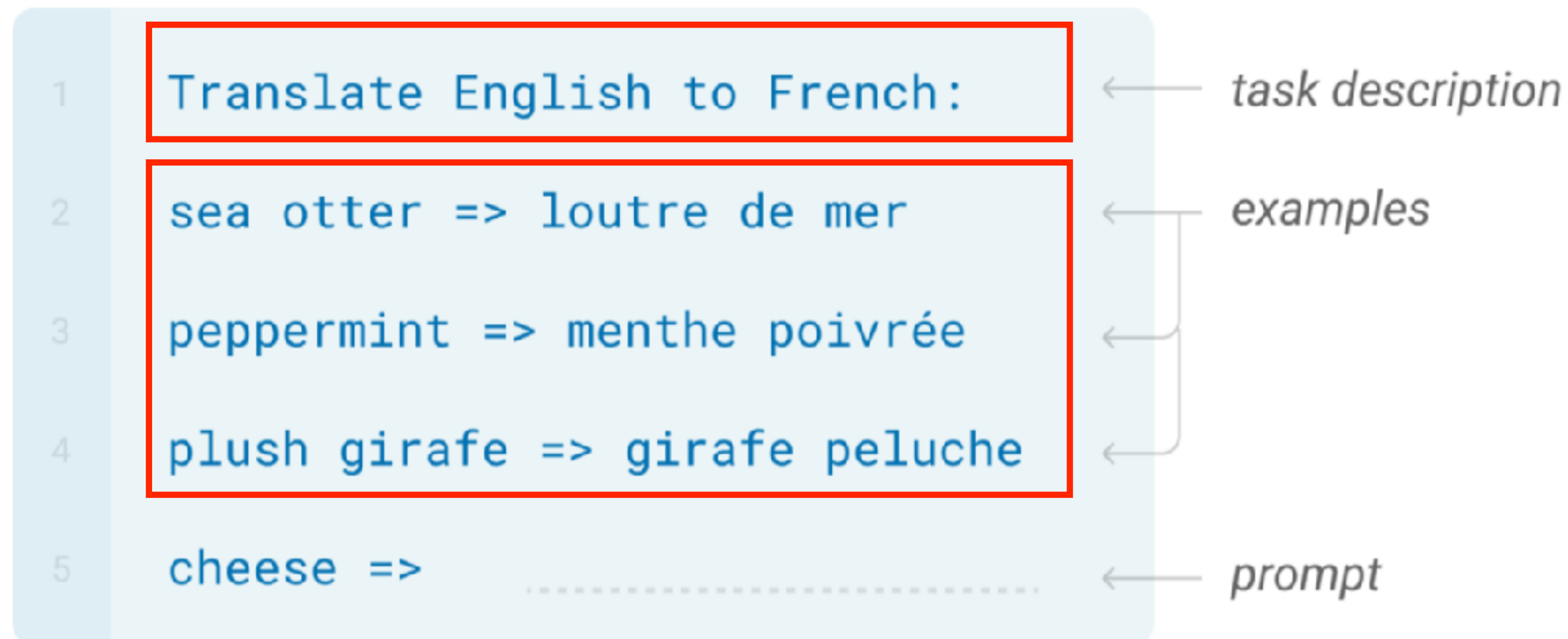


# GPT-3: few-shot in-context learning

- GPT-3 proposes an alternative: **in-context learning**

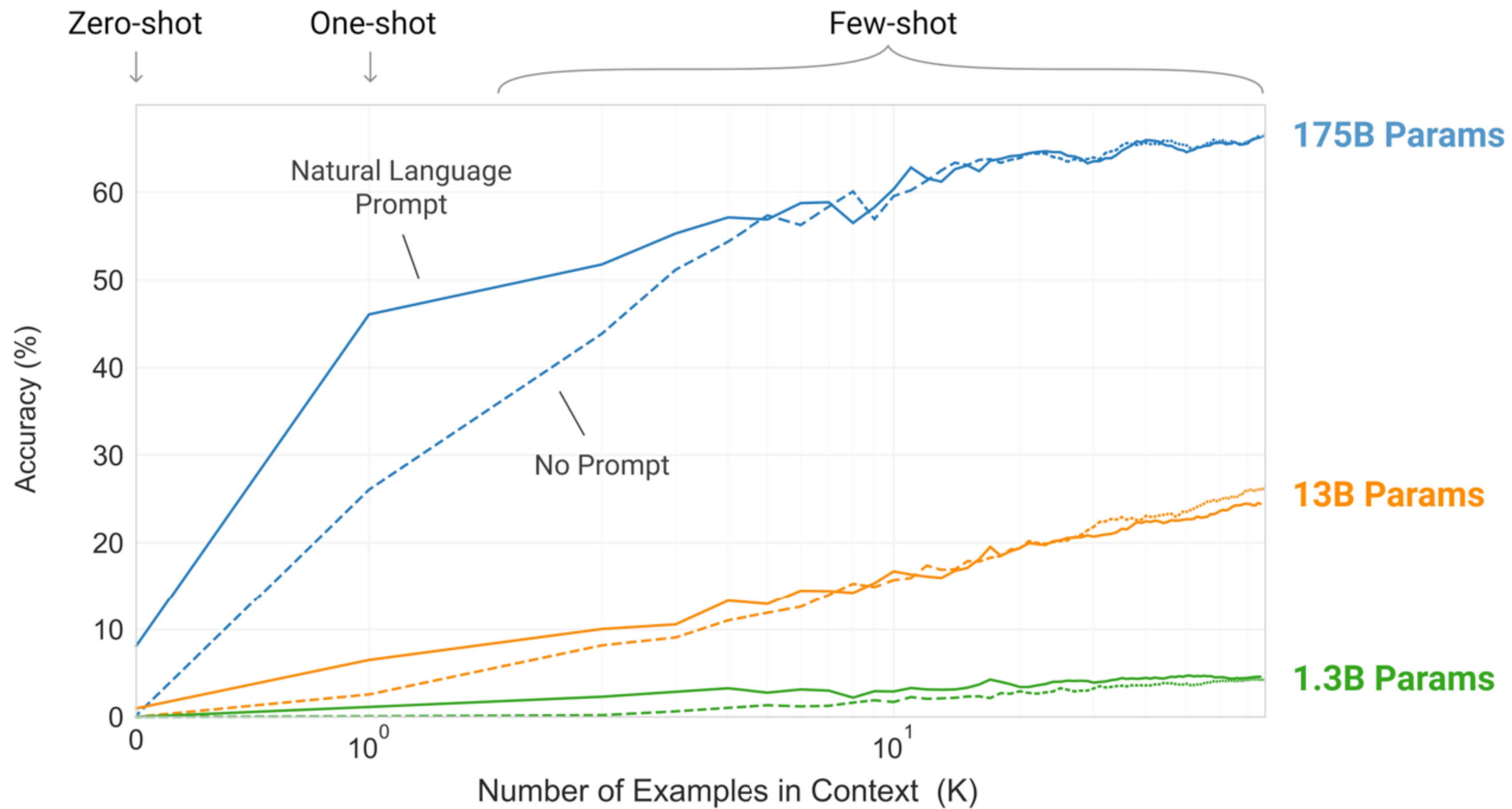
## Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.



- This is just a forward pass, **no gradient update at all!**
- You only need to feed a small number of examples (e.g., 32)  
(On the other hand, you can't feed many examples at once too as it is bounded by context size)

# GPT-3's scaling laws in performance



# Chain-of-thought (CoT) prompting

## Standard Prompting

### Input

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: 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?

### Model Output

A: The answer is 27. ❌

## Chain of Thought Prompting

### Input

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?

### Model Output

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. ✅

# Why in-context learning with LLMs?

- Amazing zero/few-shot performance
  - Save a lot of annotation! 🎉
- Easy to use without training
  - Just talk to them! 👍
- One model for many NLP applications 😊
  - No need to annotate and fine-tune for different tasks

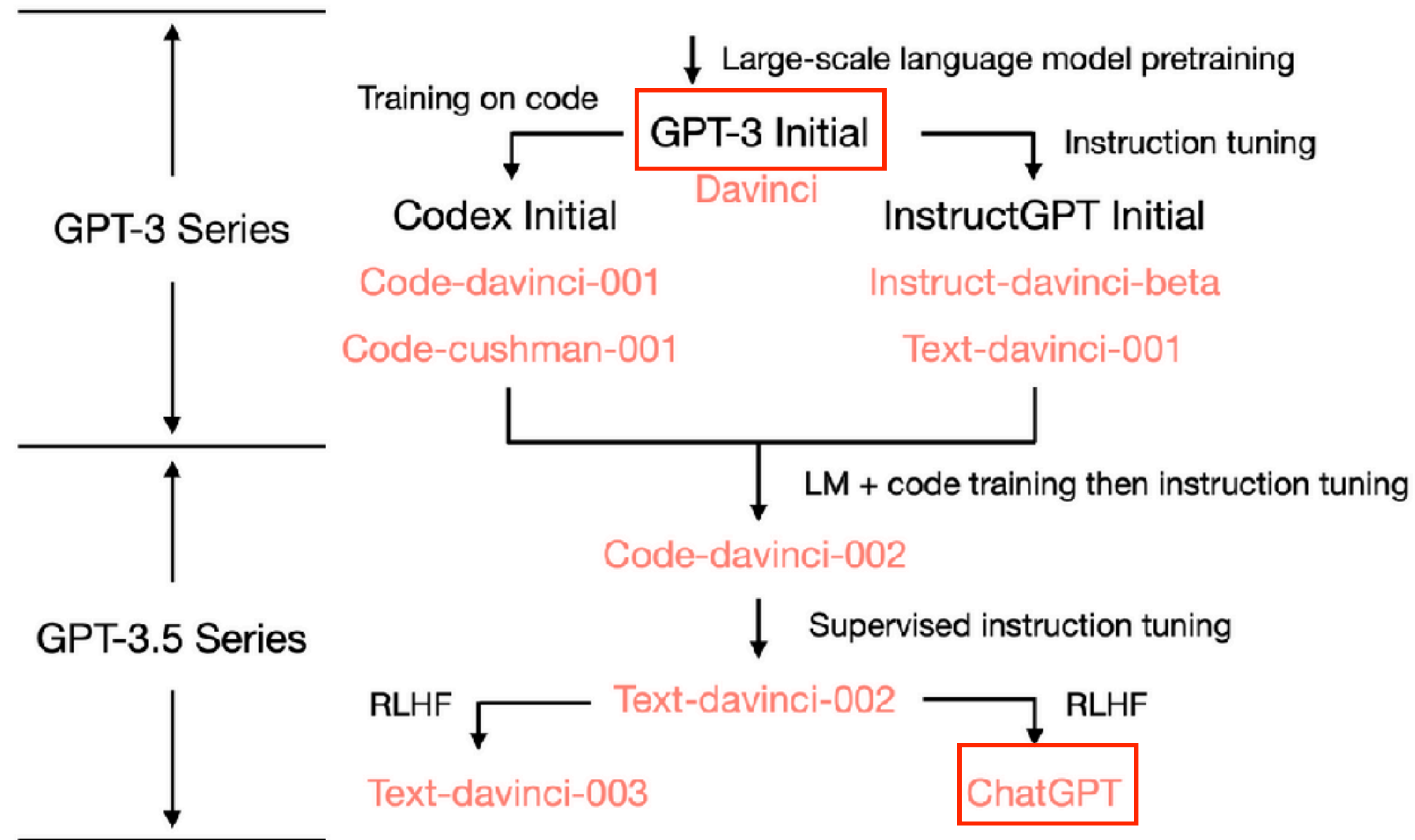
But, again, they are sensitive to prompts! Need to design a good prompt or train a good example retriever! 😂

# What happened after GPT-3?

(Is model size ↑, training corpora ↑ the only way to go?)



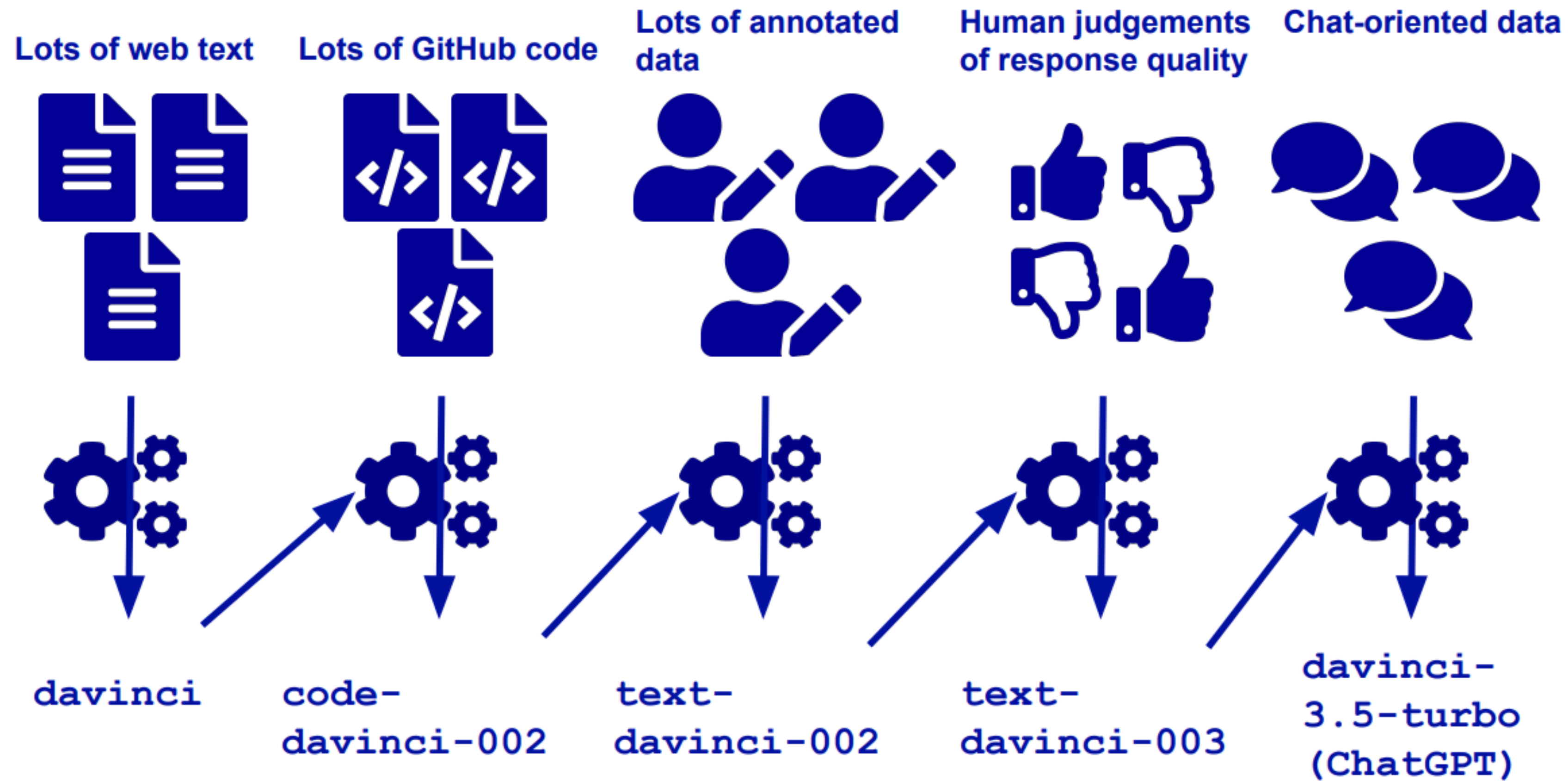
# How was ChatGPT developed?



## What's new?

- Training on code
- **Supervised instruction tuning**
- **RLHF = Reinforcement learning from human feedback**

# How was ChatGPT developed?



(Slide credit: Graham Neubig)

# InstructGPT: Supervised instruction tuning + RLHF

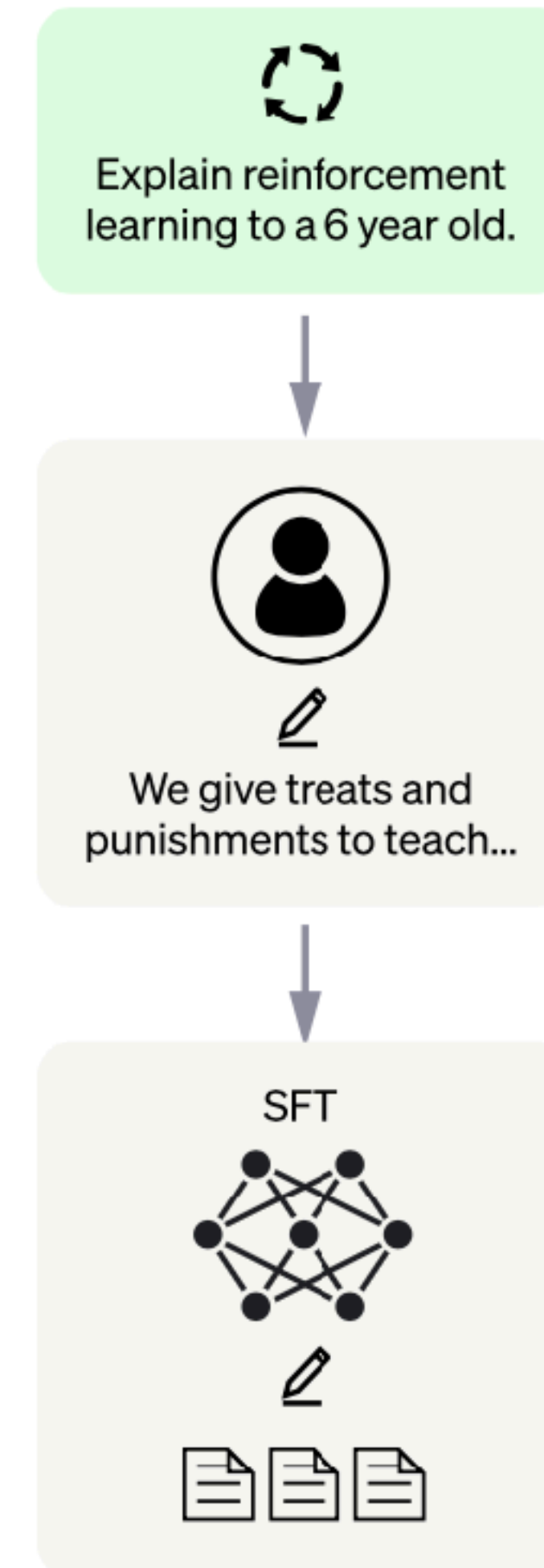
Step 1

**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.



# Supervised instruction tuning

Use-case	Prompt	Use-case	(%)	Number of Prompts		
Brainstorming	List five ideas for how to regain enthusiasm for my career	Generation	45.6%	SFT Data		
Generation	Write a short story where a bear goes to the beach, makes friends with a seal, and then returns home.	Open QA	12.4%	split	source	size
Rewrite	This is the summary of a Broadway play: "" { summary } "" This is the outline of the commercial for that play: ""	Brainstorming	11.2%	train	labeler	11,295
		Chat	8.4%	train	customer	1,430
		Rewrite	6.6%	valid	labeler	1,550
		Summarization	4.2%	valid	customer	103
		Classification	3.5%			
		Other	3.5%			
		Closed QA	2.6%			
		Extract	1.9%			

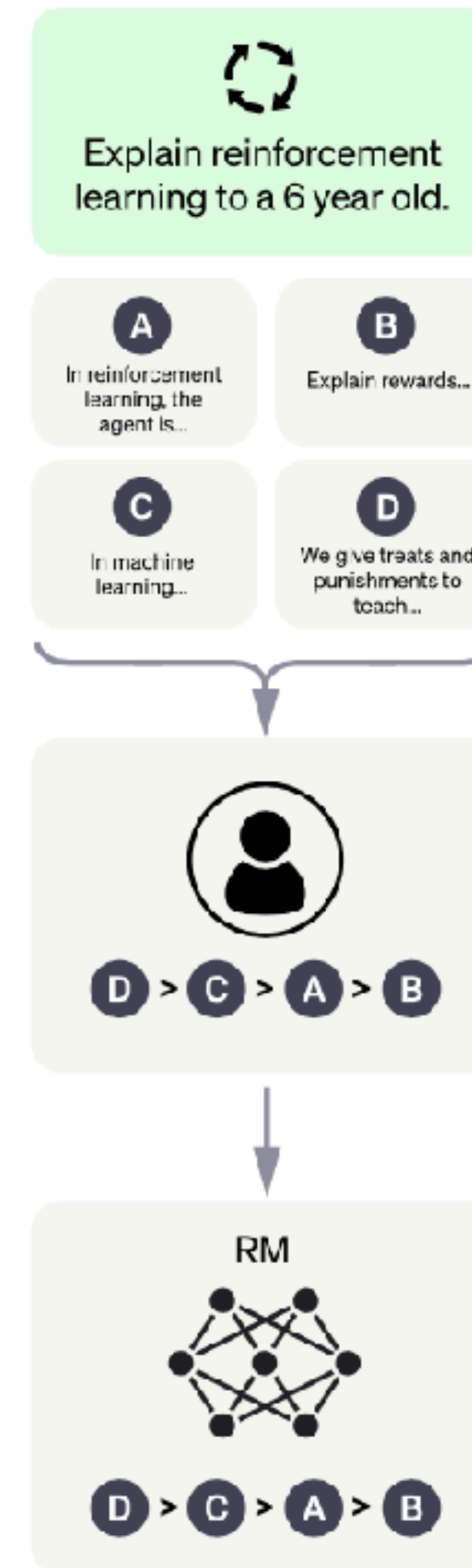
SFT data: only ~13k (not public)

# InstructGPT: Supervised instruction tuning + RLHF

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.

# InstructGPT: Supervised instruction tuning + RLHF

Step 3

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

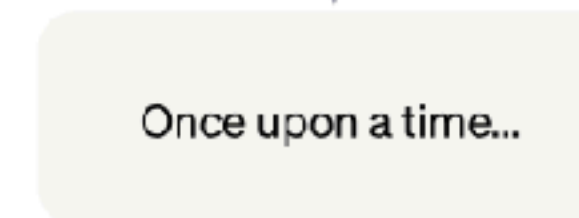
A new prompt is sampled from the dataset.



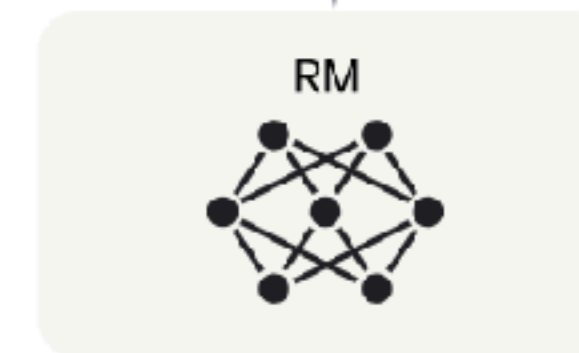
The PPO model is initialized from the supervised policy.



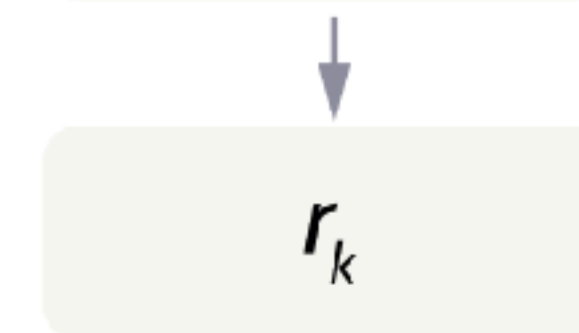
The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



# ChatGPT = InstructGPT + dialogue data

## Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

“We trained this model using Reinforcement Learning from Human Feedback (RLHF), **using the same methods as InstructGPT**, but with slight differences in the data collection setup. We trained an initial model using supervised fine-tuning: human AI trainers provided conversations in which they played both sides—the user and an AI assistant. We gave the trainers access to model-written suggestions to help them compose their responses. **We mixed this new dialogue dataset with the InstructGPT dataset**, which we transformed into a dialogue format.”

Human feedback data is the key!

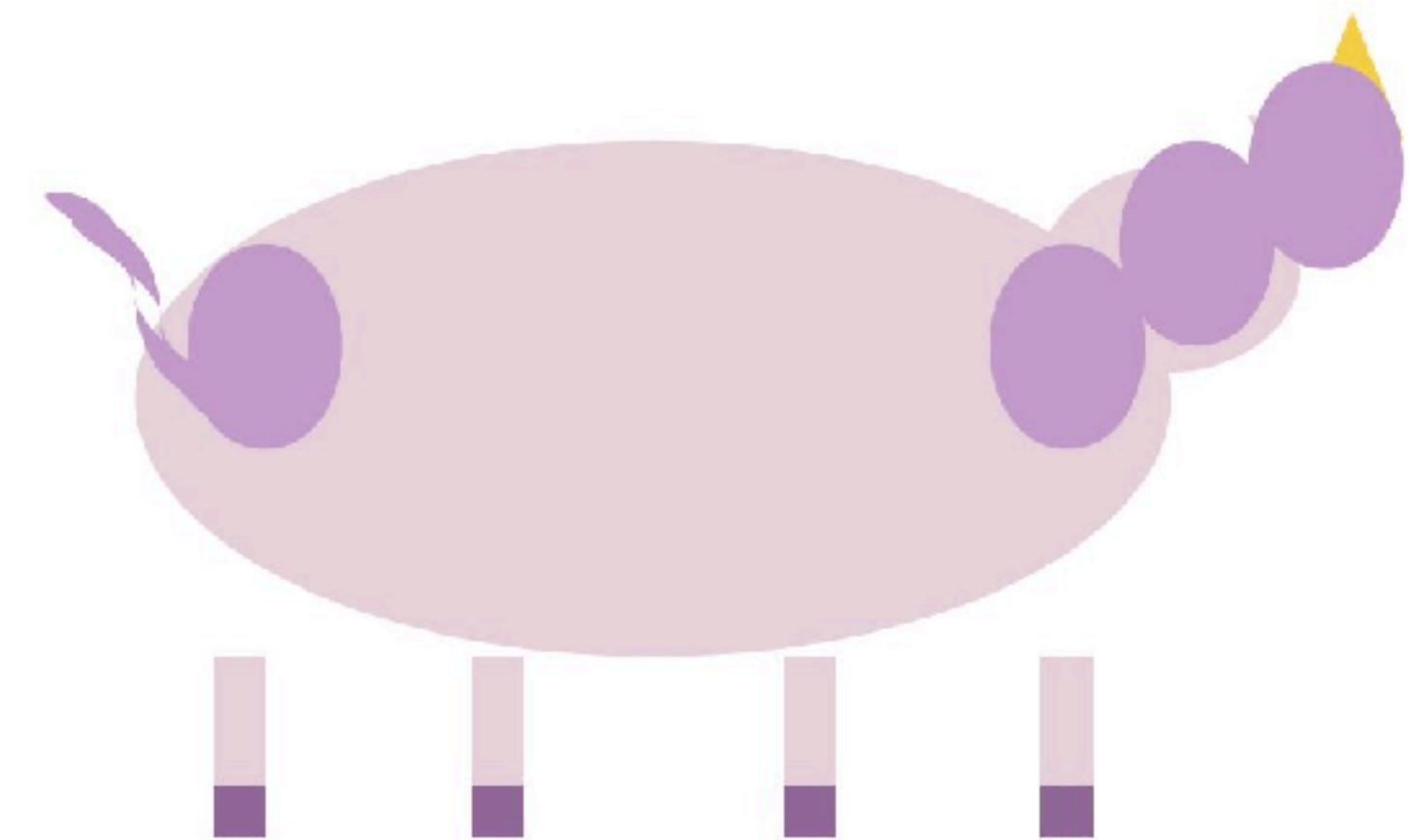
<https://openai.com/blog/chatgpt>

# GPT-4

- GPT-4: a **multi-modal** model capable of processing image and text inputs and producing text outputs.
- Model size and training details unknown
- Can process up to **32k** context size

**Prompt:** Draw a unicorn in TiKZ.

**GPT-4:** [Produces  $\text{\LaTeX}$  compiling to following picture.]

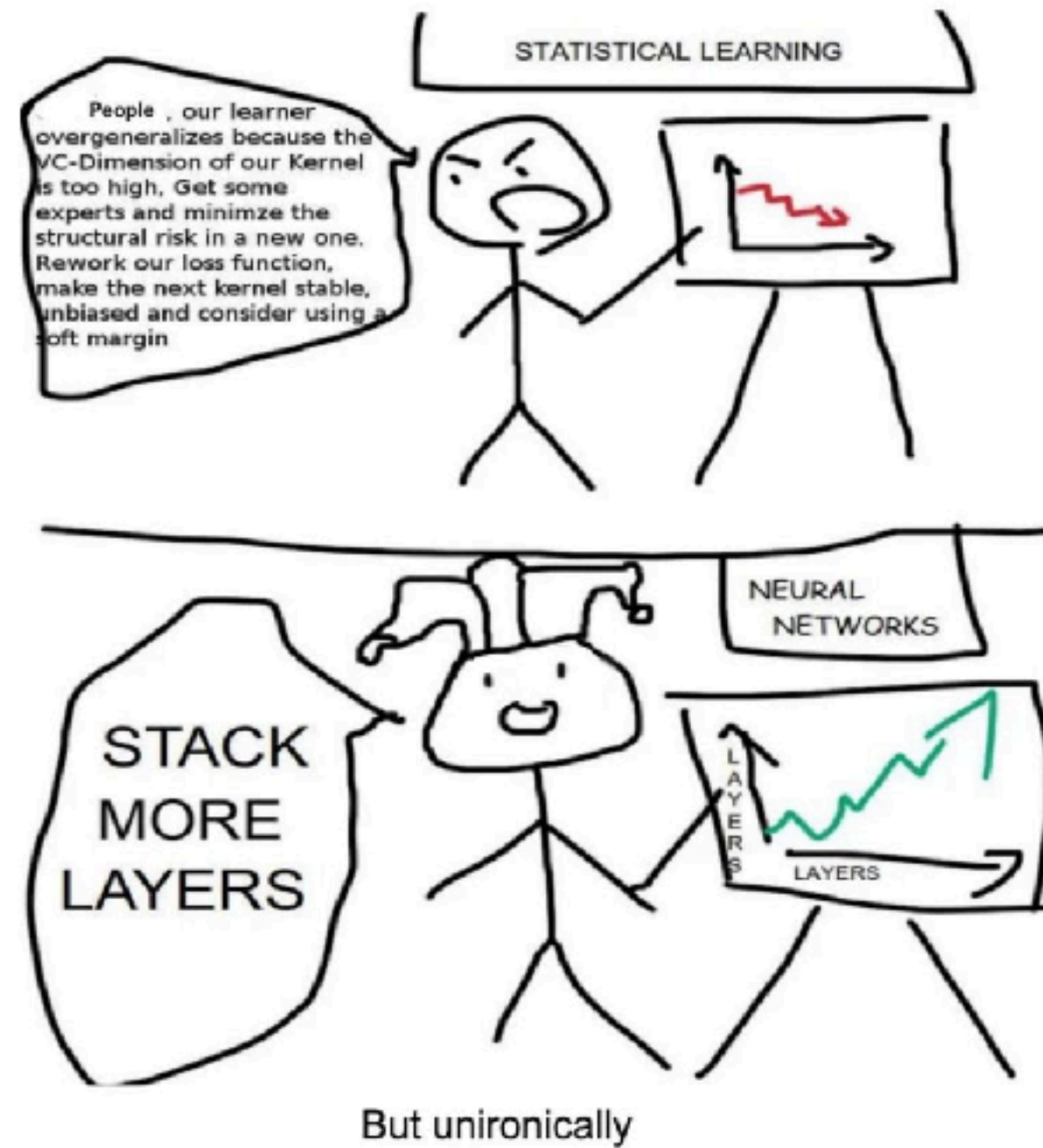




# GPT-4

## GPT-4 visual input example, Moar Layers:

User Can you explain why this is funny. Think about it step-by-step.



GPT-4 The comic is satirizing the difference in approaches to improving model performance between statistical learning and neural networks.

In statistical learning, the character is shown to be concerned with overfitting and suggests a series of complex and technical solutions, such as minimizing structural risk, reworking the loss function, and using a soft margin.

In contrast, the neural networks character simply suggests adding more layers to the model. This is often seen as a common solution to improving

## GPT-4 visual input example, Extreme Ironing:

User What is unusual about this image?



Source: <https://www.barnorama.com/wp-content/uploads/2016/12/03-Confusing-Pictures.jpg>

GPT-4 The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.