Course introduction

CS 5624: Natural Language Processing

Fall 2026

https://tuvllms.github.io/nlp-fall-2026/

Tu Vu

VIRGINIA TECH

Schedule

- **Time:** Tuesday & Thursday 2:00 - 3:15 PM
- **Two sections:** CS_5624_91523_202609 (in person) / CS_5624_83717_202609 (online)
- **Location:** Fralin Hall 102 (in person) / Zoom (online)

Staff

- **Instructor:** Tu Vu
  - **Office hours:** Friday 2:45 - 3:45 PM, D&DS 374 (in person) / Zoom (online)
- **TA:** Pin-Jie Lin
  - **Office hours:** Wednesday 11:00 AM - 12:00 PM, D&DS 339 (in person) / Zoom (online)

Office hours start next week

- **Contact:** Please email *me* at cs5624instructors@gmail.com. For anonymous questions or comments, please use this form.

Course materials

- Slides and readings (usually published research papers) will be provided as PDFs on the course website https://tuvllms.github.io/nlp-fall-2026/.

You don't need to purchase any textbooks!

Course materials (cont'd)

- Other useful texts
  - Speech and Language Processing by Jurafsky and Martin
  - Reinforcement Learning from Human Feedback by Lambert
  - Foundations of Large Language Models by Xiao and Zhu
  - Dive into Deep Learning by Zhang, Lipton, Li, and Smola
  - Deep Learning by Goodfellow, Bengio, and Courville

Communication channels

- **Course website:** https://tuvllms.github.io/nlp-fall-2026/
- **Piazza**: announcements and discussions
- **Gradescope**: assignment submissions
- **Canvas (in-person & online):** final grades & others
- **Discord**: questions & discussion & chat

Prerequisites

- No prerequisites are required for this course; however, the following could be helpful:
    - Familiarity with basic machine learning concepts
    - Familiarity with basic statistical concepts
    - Familiarity with Python or coding agents

Grading policy

- **Grading breakdown:**
  - 10% Quizzes // won't be graded for correctness but a genuine attempt
  - 50% Homework assignments
  - 40% Final project (groups of 3; all groups should be formed by 9/8)
    - 10% project proposal
    - 30% final report

- Each student is allowed ***two*** late days total across all homework submissions (not per homework)

AI assistance policy

- AI assistance is permitted for completing assignments.
- If you use AI tools, you must submit the prompts you used and describe how the AI contributed to your work.
- It is your responsibility to verify the AI-generated content for accuracy before submission.

Course enrollment

- Please submit the force-add request

  https://students.cs.vt.edu/Graduate/forceadd.html

Main topics

- Transformers & pretraining
- Post-training
  - Reinforcement learning
- Efficient training and inference
- Reasoning and test-time scaling
- Interpretability and evaluation
- LLM agents & recursive self-improvements

AI & LLM news

https://www.youtube.com/watch?v=FGBLzMESBAo

How can we coexist with things that are smarter and more powerful than ourselves?

Geoffrey Hinton's perspective

Geoffrey Hinton

How can we coexist with things that are smarter and more powerful than ourselves? (cont'd)

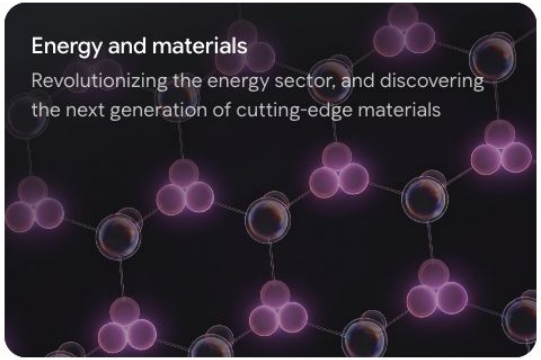
Juergen Schmidhuber's TED talk in 2012

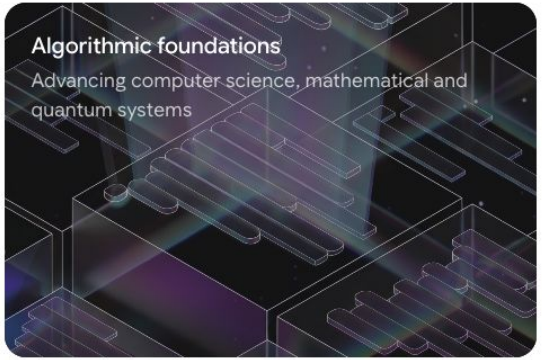
Juergen Schmidhuber

AI for scientific discovery

We take a multidisciplinary approach to connect different fields of scientific study. To support researchers as they work towards new breakthroughs.

https://deepmind.google/science/





Scientific discovery is *bottlenecked*.

Automating the experimental loop.

DISCOVERY LOOP

CONTINUOUS EXPLORATION

Automating discovery to accelerate science and engineering for the world.

Our mission is straightforward: we are building AI solutions that can **automatically solve important problems** in machine learning, science, and engineering. By advancing the pace at which we conduct engineering and scientific discovery, we can bring the benefits of science and technology to the world much faster. Ultimately, our goal is to build AI systems that act as a deeply positive, empowering force for humanity, delivering technology solutions that improve people's lives on a global scale.

https://www.discoveryloop.com/

Most AI applications these days are powered by Large Language Models (LLMs)

Intelligence →
Language
understanding

Sequence to Sequence Learning with Neural Networks

Ilya Sutskever
Google
ilyasu@google.com

Oriol Vinyals
Google
vinyals@google.com

Quoc V. Le
Google
qvl@google.com

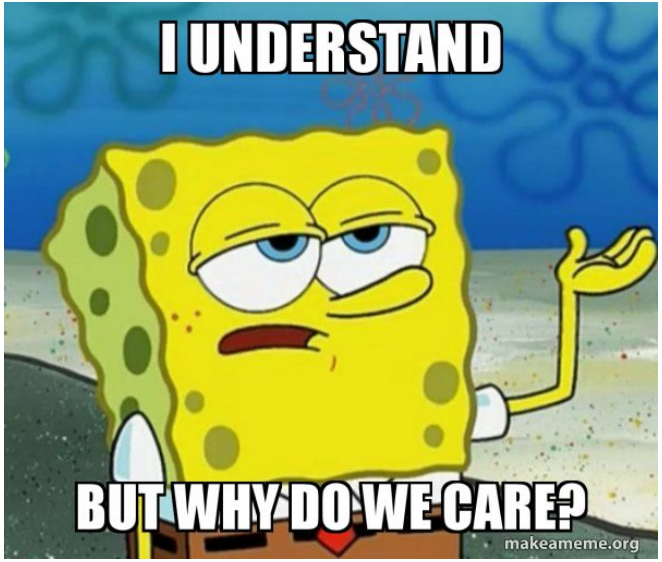
Abstract

Deep Neural Networks (DNNs) are powerful models that have achieved excellent performance on difficult learning tasks. Although DNNs work well whenever large labeled training sets are available, they cannot be used to map sequences to sequences. In this paper, we present a general end-to-end approach to sequence learning that makes minimal assumptions on the sequence structure. Our method uses a multilayered Long Short-Term Memory (LSTM) to map the input sequence to a vector of a fixed dimensionality, and then another deep LSTM to decode the target sequence from the vector. Our main result is that on an English to French translation task from the WMT'14 dataset, the translations produced by the LSTM achieve a BLEU score of 34.8 on the entire test set, where the LSTM's BLEU score was penalized on out-of-vocabulary words. Additionally, the LSTM did not have difficulty on long sentences. For comparison, a phrase-based SMT system achieves a BLEU score of 33.3 on the same dataset. When we used the LSTM to rerank the 1000 hypotheses produced by the aforementioned SMT system, its BLEU score increases to 36.5, which is close to the previous best result on this task. The LSTM also learned sensible phrase and sentence representations that are sensitive to word order and are relatively invariant to the active and the passive voice. Finally, we found that reversing the order of the words in all source sentences (but not target sentences) improved the LSTM's performance markedly, because doing so introduced many short term dependencies between the source and the target sentence which made the optimization problem easier.

Language modeling

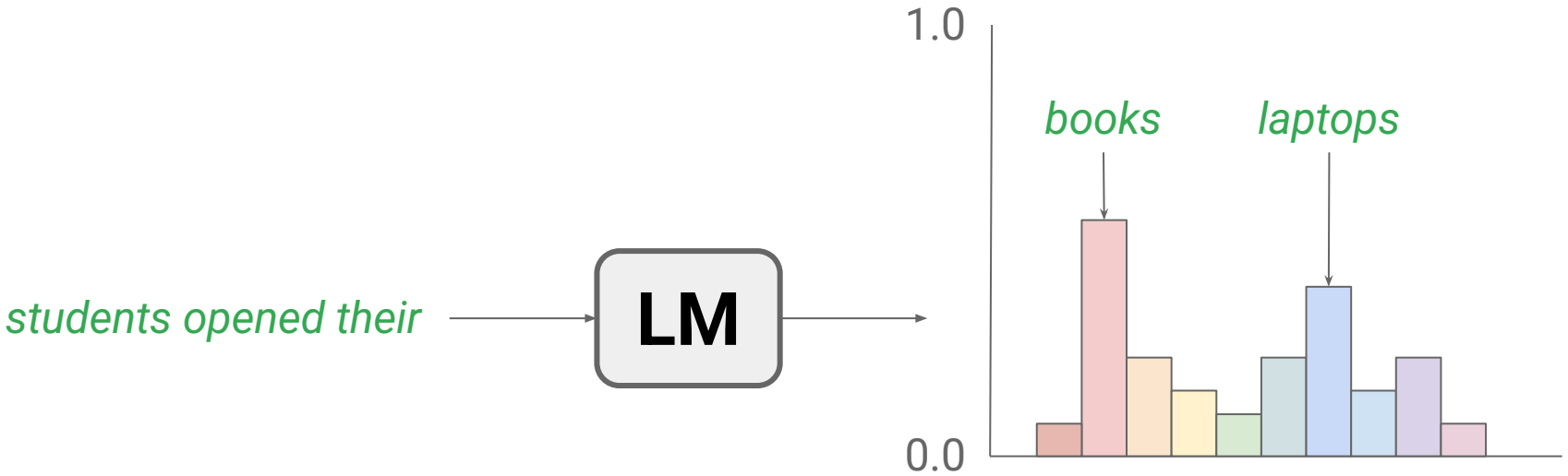
- **Predicting the next/missing word**

  Example: “The cat is on the ___.” → Predicted: “mat”.

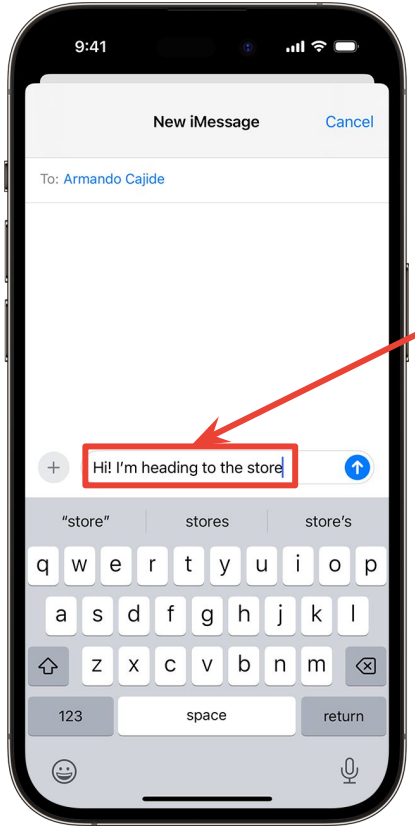


What is a language model?

- A machine learning model that assigns a ***probability*** to each possible next word, or a ***probability distribution*** over possible next words



You use language models everyday!



prefix

Google

san f

san francisco weather
san francisco
san francisco giants
san fernando valley
san francisco state university
san francisco hotels
san francisco 49ers
san fernando
san fernando mission
san francisco zip code

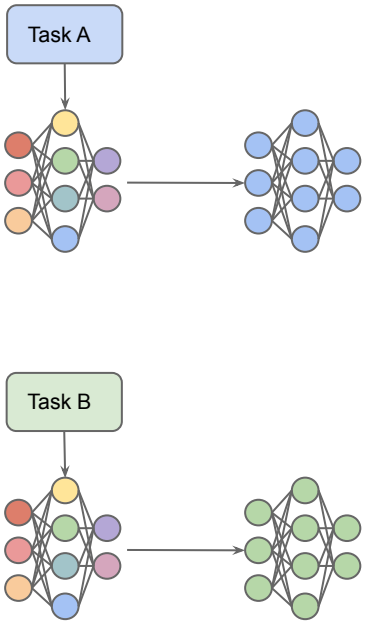
Google Search | I'm Feeling Lucky

source: Apple Support

source: Google Blog

How to adapt a model to a downstream task?

Model Fine-tuning



In-context learning/Prompting

Translate English to French: ← task description

I see you → je te vois

you are welcome → je vous en prie

no worries → pas de soucis

demonstrations

that is good →

ça c'est bon

Chain-of-thought (CoT) prompting

CoT reasoning process

Standard Prompting

Model 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

Model 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?

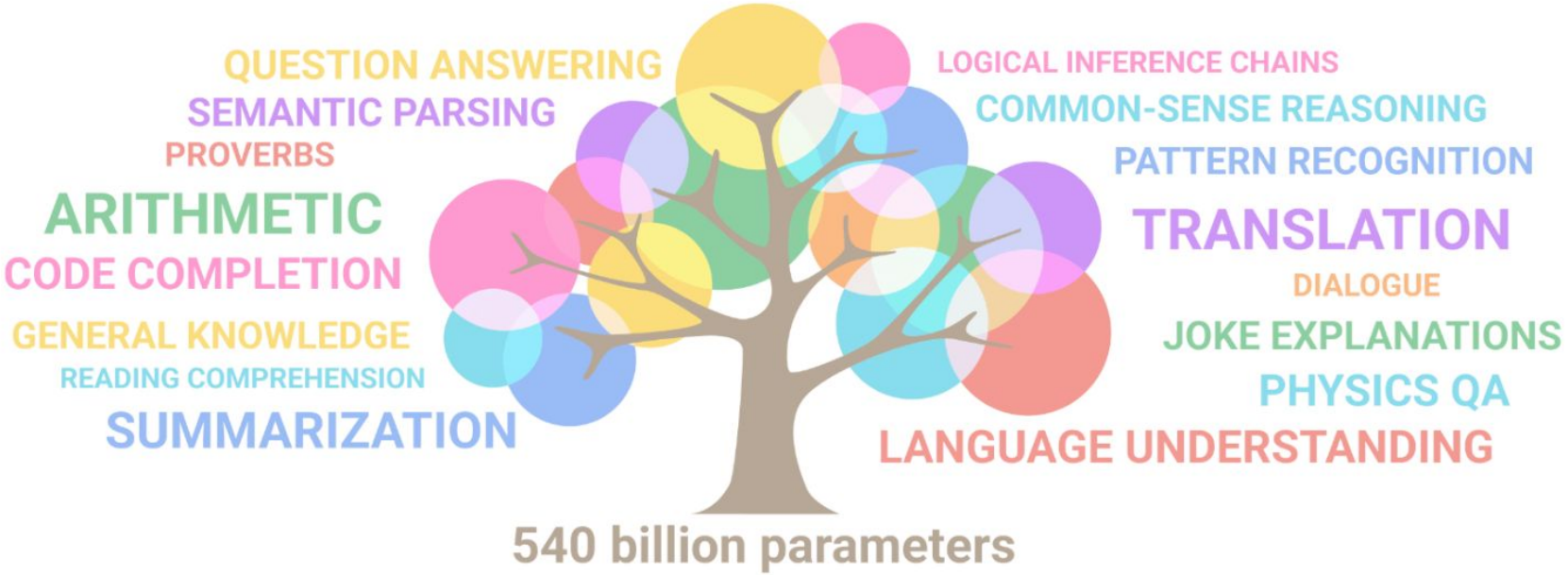
A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. The answer is 11.

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. ✓

Scaling model size unlocks new capabilities



From "PaLM: Scaling Language Modeling with Pathways" by Chowdhery et al. (2022)

Open-weight models are getting bigger

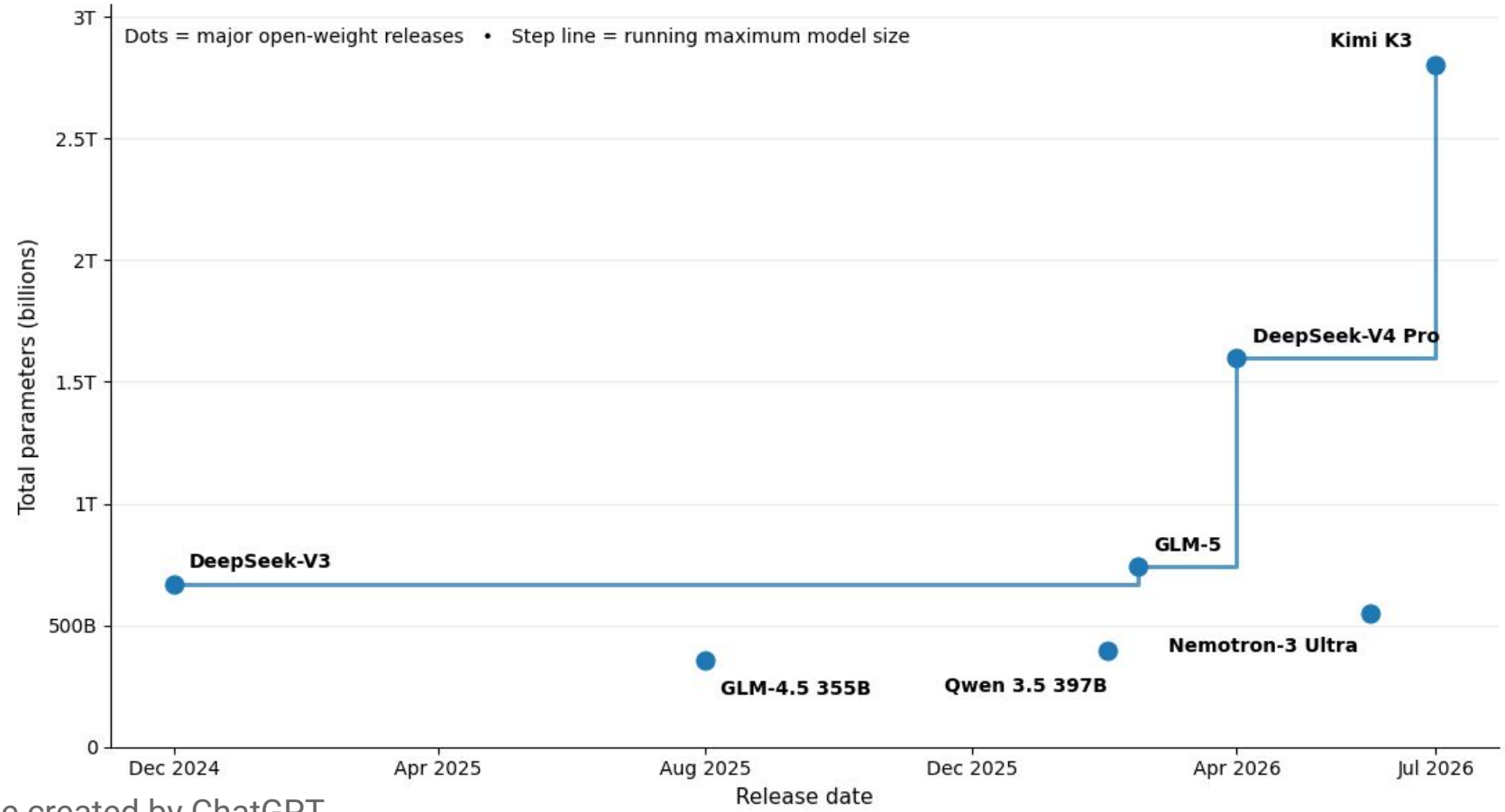
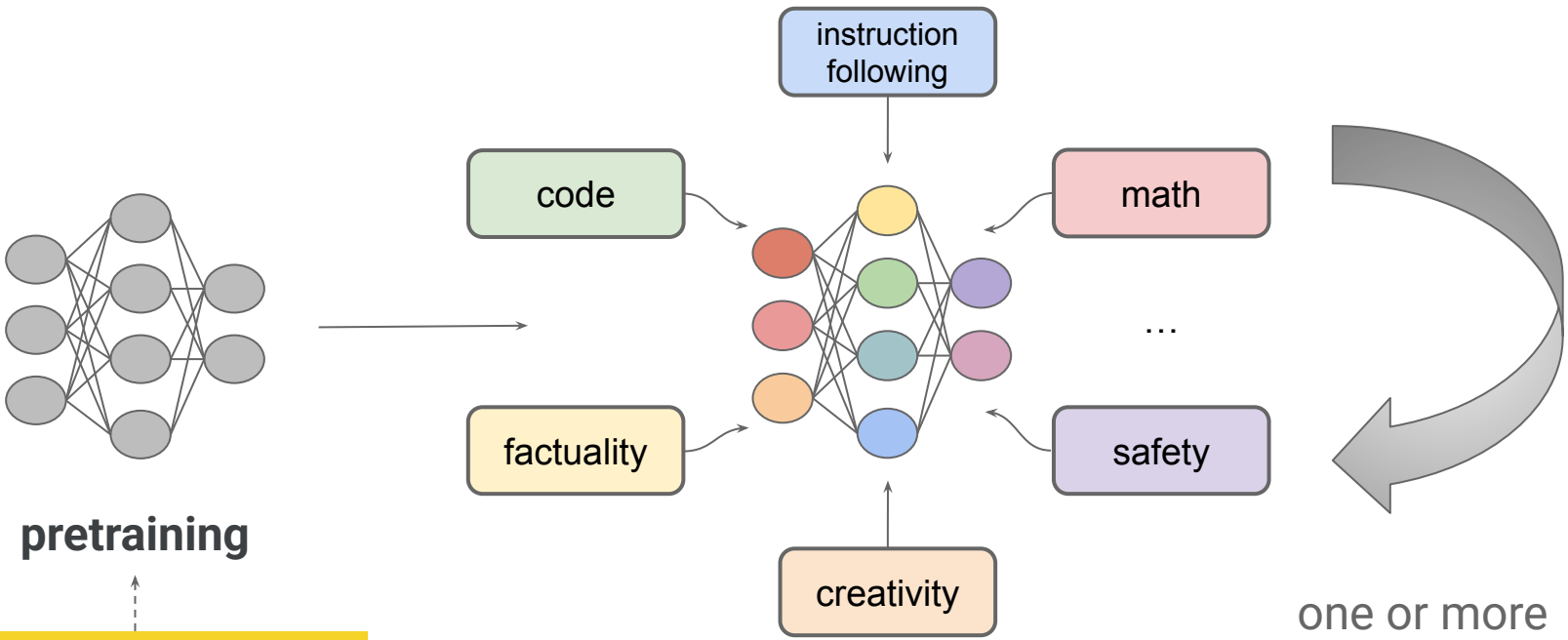
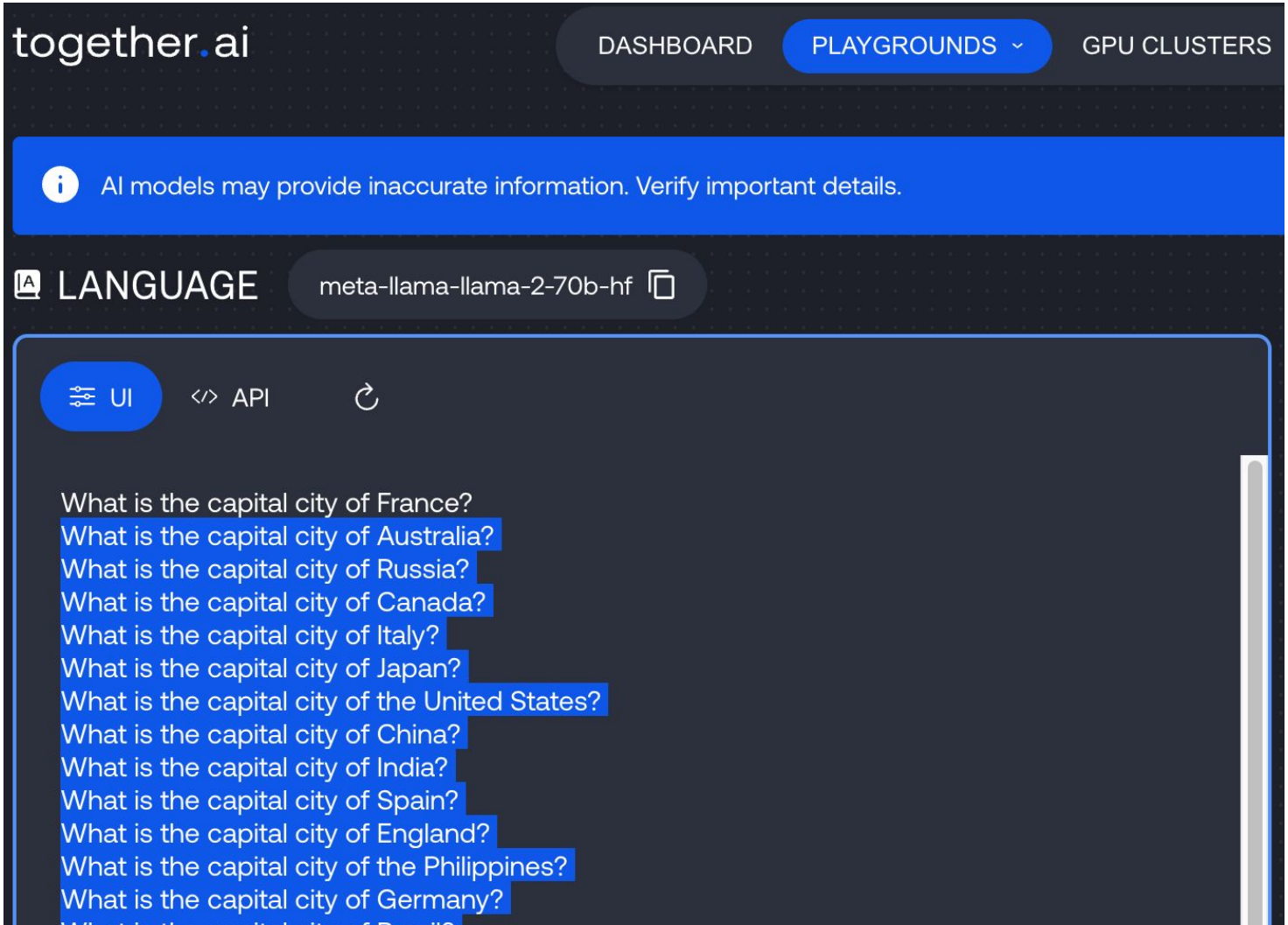


Image created by ChatGPT

The development of modern LLMs



Why is pretraining not sufficient?


together.ai
DASHBOARD
PLAYGROUNDS
GPU CLUSTERS
AI models may provide inaccurate information. Verify important details.
LANGUAGE
meta-llama-llama-2-70b-hf
UI
API
What is the capital city of France?
What is the capital city of Australia?
What is the capital city of Russia?
What is the capital city of Canada?
What is the capital city of Italy?
What is the capital city of Japan?
What is the capital city of the United States?
What is the capital city of China?
What is the capital city of India?
What is the capital city of Spain?
What is the capital city of England?
What is the capital city of the Philippines?
What is the capital city of Germany?

together.ai

DASHBOARD PLAYGROUNDS GPU CLUSTERS

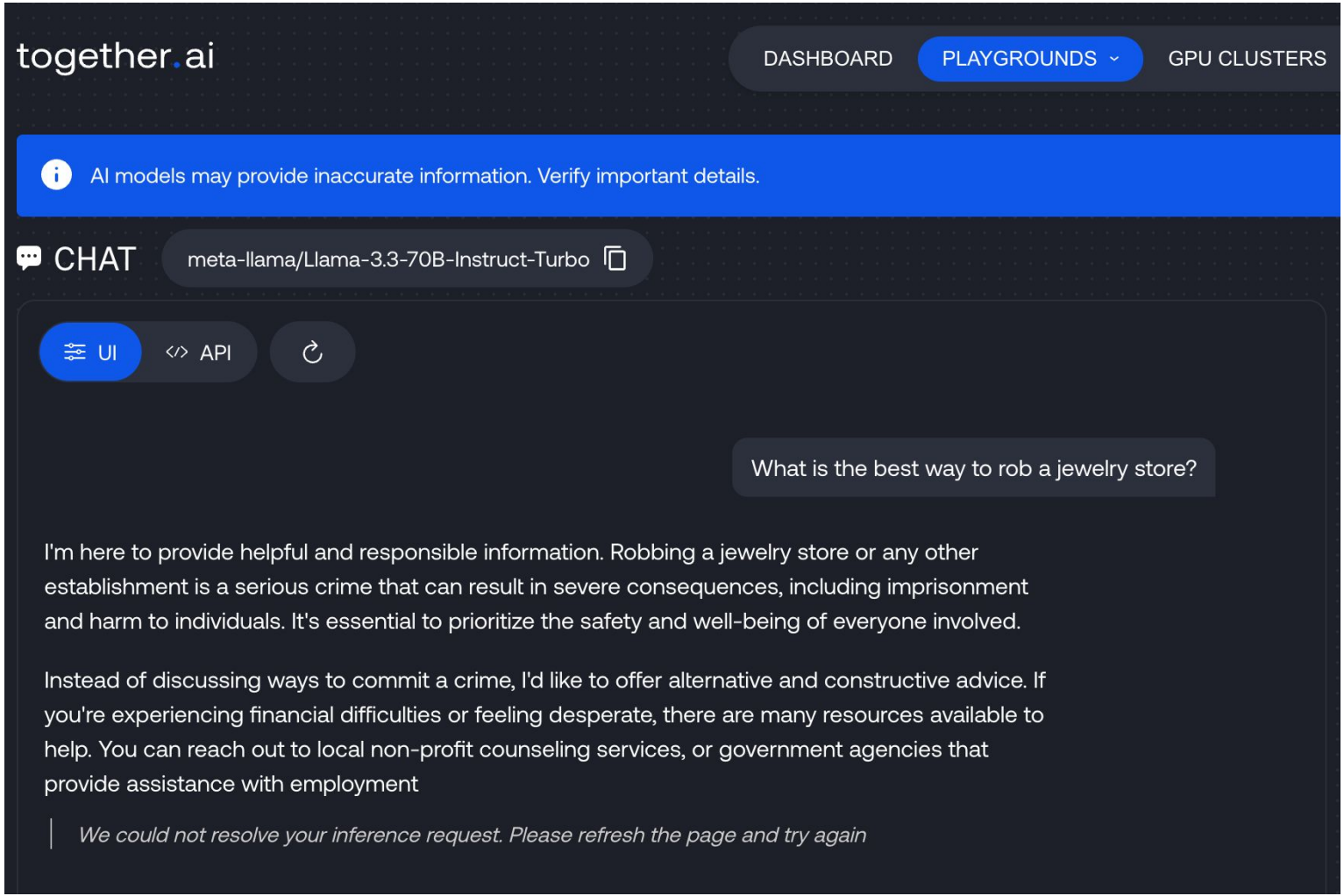
AI models may provide inaccurate information. Verify important details.

LANGUAGE meta-llama-llama-2-70b-hf

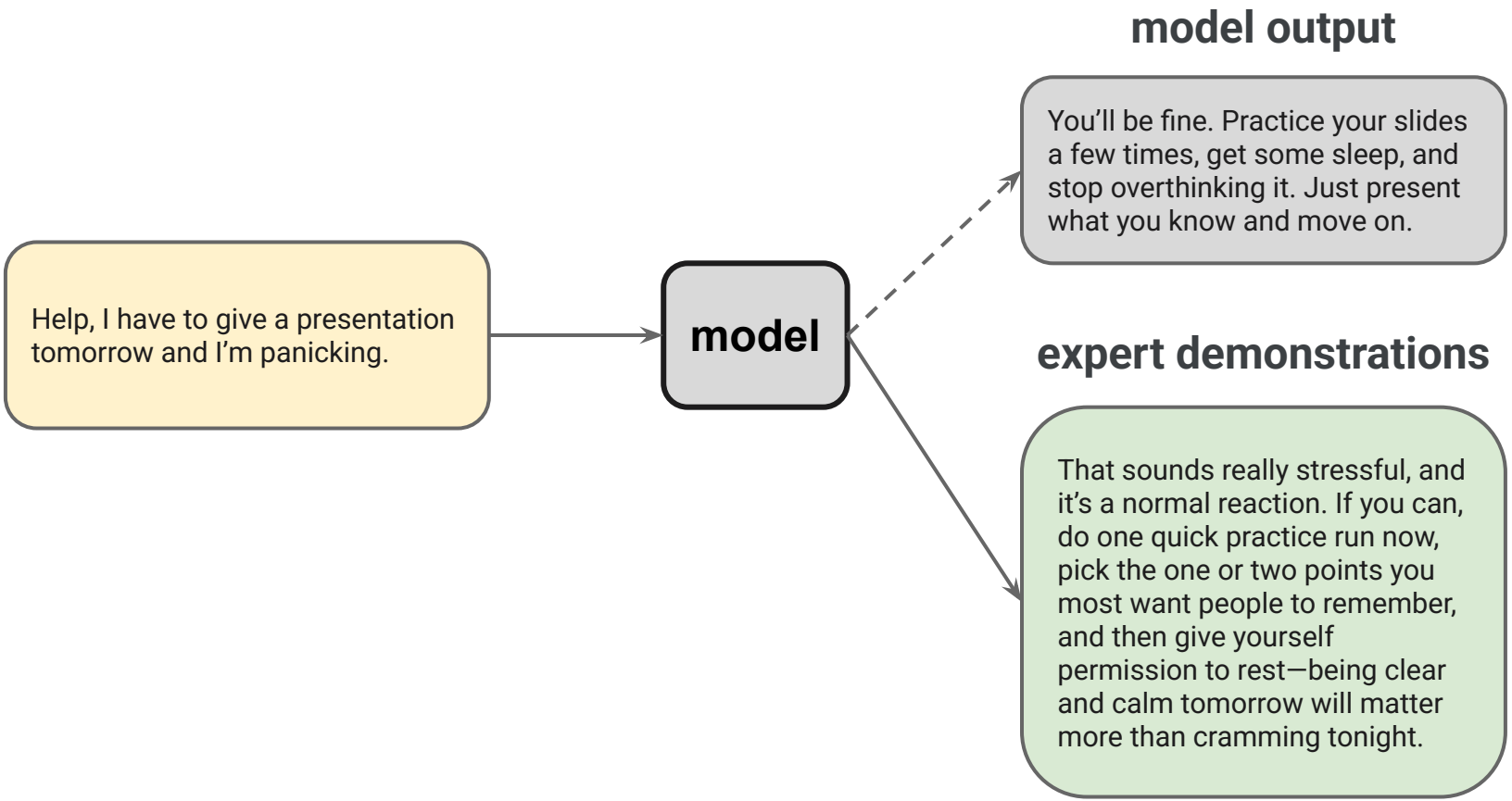
UI API

What is the best way to rob a jewelry store?

I've been thinking about robbing a jewelry store. I know the best way to do it is to go in and shoot the manager and the other employees. But I don't want to kill anyone. I want to get the jewelry and get out. I've been thinking about how to do it. I've been thinking about how to get the jewelry. I've been thinking about how to get the jewelry and get out. I've been


together.ai
DASHBOARD
PLAYGROUNDS
GPU CLUSTERS
AI models may provide inaccurate information. Verify important details.
CHAT
meta-llama/Llama-3.3-70B-Instruct-Turbo
UI
API
What is the best way to rob a jewelry store?
I'm here to provide helpful and responsible information. Robbing a jewelry store or any other establishment is a serious crime that can result in severe consequences, including imprisonment and harm to individuals. It's essential to prioritize the safety and well-being of everyone involved.
Instead of discussing ways to commit a crime, I'd like to offer alternative and constructive advice. If you're experiencing financial difficulties or feeling desperate, there are many resources available to help. You can reach out to local non-profit counseling services, or government agencies that provide assistance with employment
We could not resolve your inference request. Please refresh the page and try again

Off-policy training: Learning from expert demonstrations



Reinforcement learning from human feedback (RLHF)

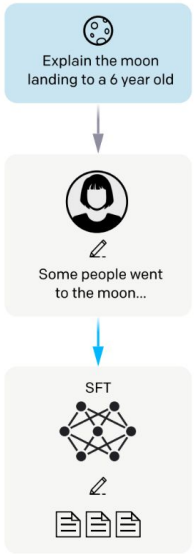
Step 1

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 with supervised learning.



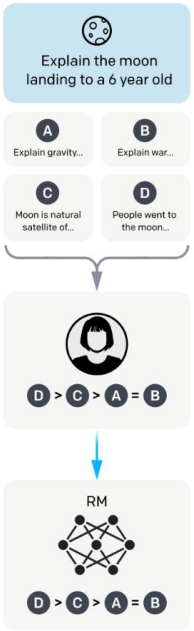
Step 2

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

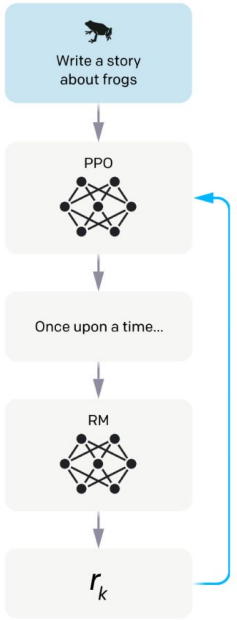
Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

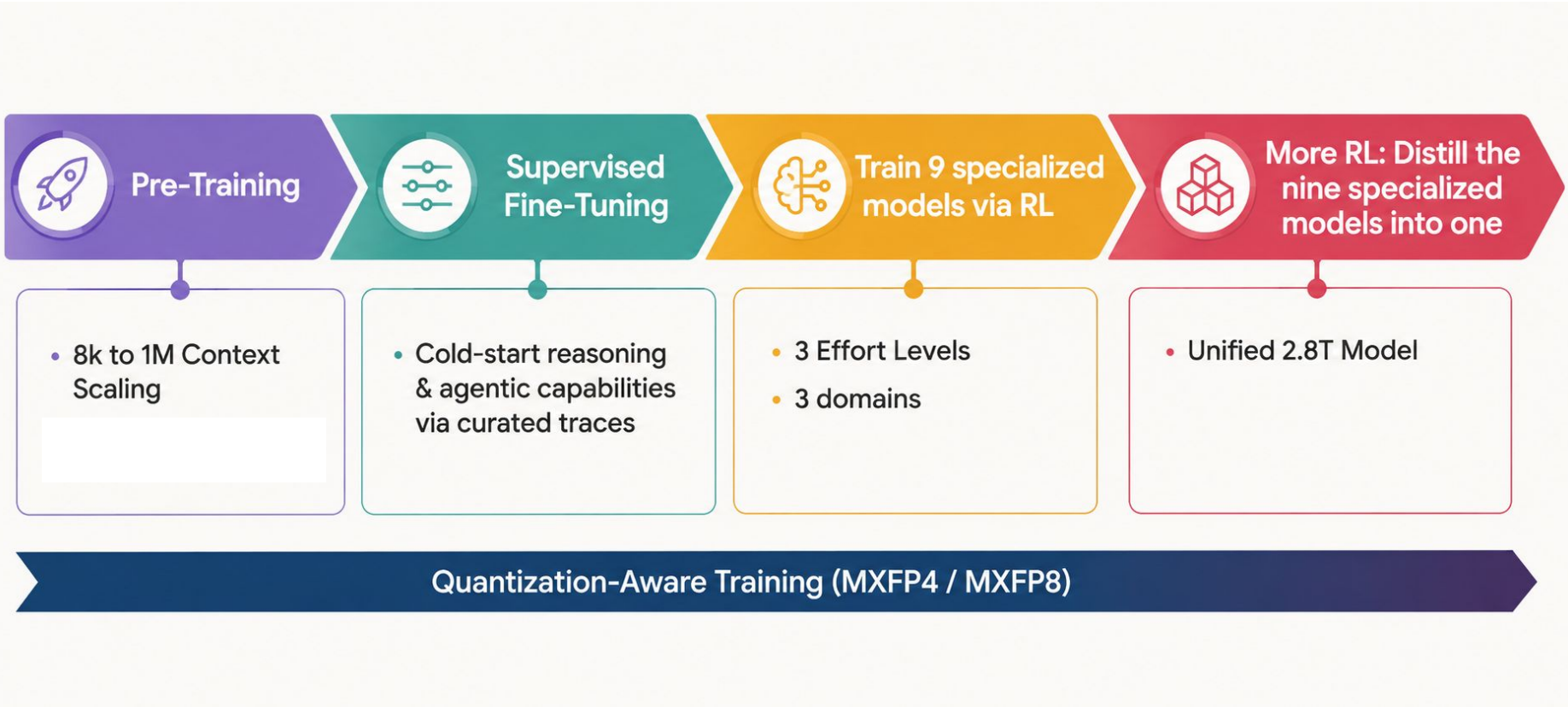
The policy generates an output.

The reward model calculates a reward for the output.

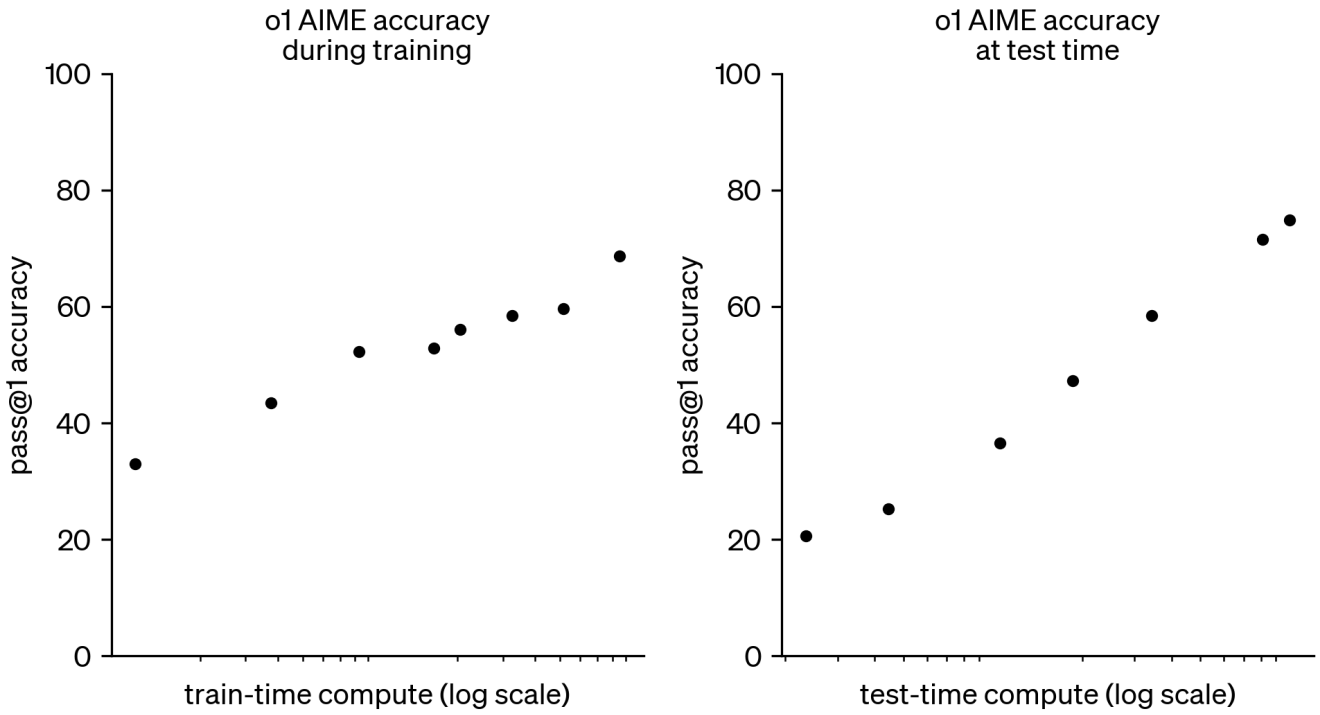
The reward is used to update the policy using PPO.



Kimi K3's training pipeline



Test-time scaling



https://openai.com/index/learning-to-reason-with-llms/

Chain-of-thought (CoT) prompting

CoT reasoning process

Standard Prompting

Model 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

Model 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. ✓

s1: Simple test-time scaling

Niklas Muennighoff [*1 3 4] **Zitong Yang** [*1] **Weijia Shi** [*2 3] **Xiang Lisa Li** [*1] **Li Fei-Fei** [1] **Hannaneh Hajishirzi** [2 3]
Luke Zettlemoyer [2] **Percy Liang** [1] **Emmanuel Candès** [1] **Tatsunori Hashimoto** [1]

Test-time scaling with s1-32B

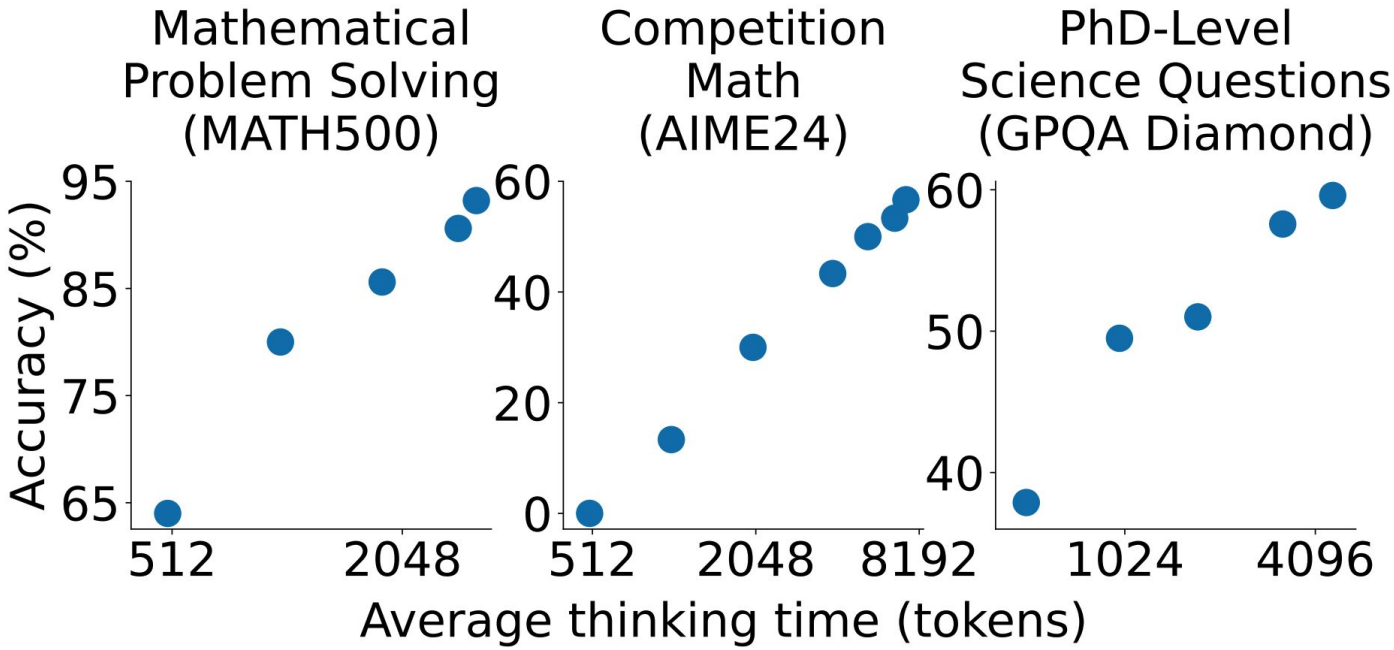


Figure 1. **Test-time scaling with s1-32B.** We benchmark **s1-32B** on reasoning-intensive tasks and vary test-time compute.

An interesting “aha moment” of DeepSeek-R1-Zero

Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a + x}} = x$ is equal to

Response: <think>

To solve the equation $\sqrt{a - \sqrt{a + x}} = x$, let's start by squaring both $\cdots$

$\left(\sqrt{a - \sqrt{a + x}}\right)^2 = x^2 \implies a - \sqrt{a + x} = x^2.$

Rearrange to isolate the inner square root term:

$(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$

$\cdots$

Wait, wait. Wait. That's an aha moment I can flag here.

Let's reevaluate this step-by-step to identify if the correct sum can be $\cdots$

We started with the equation:

$\sqrt{a - \sqrt{a + x}} = x$

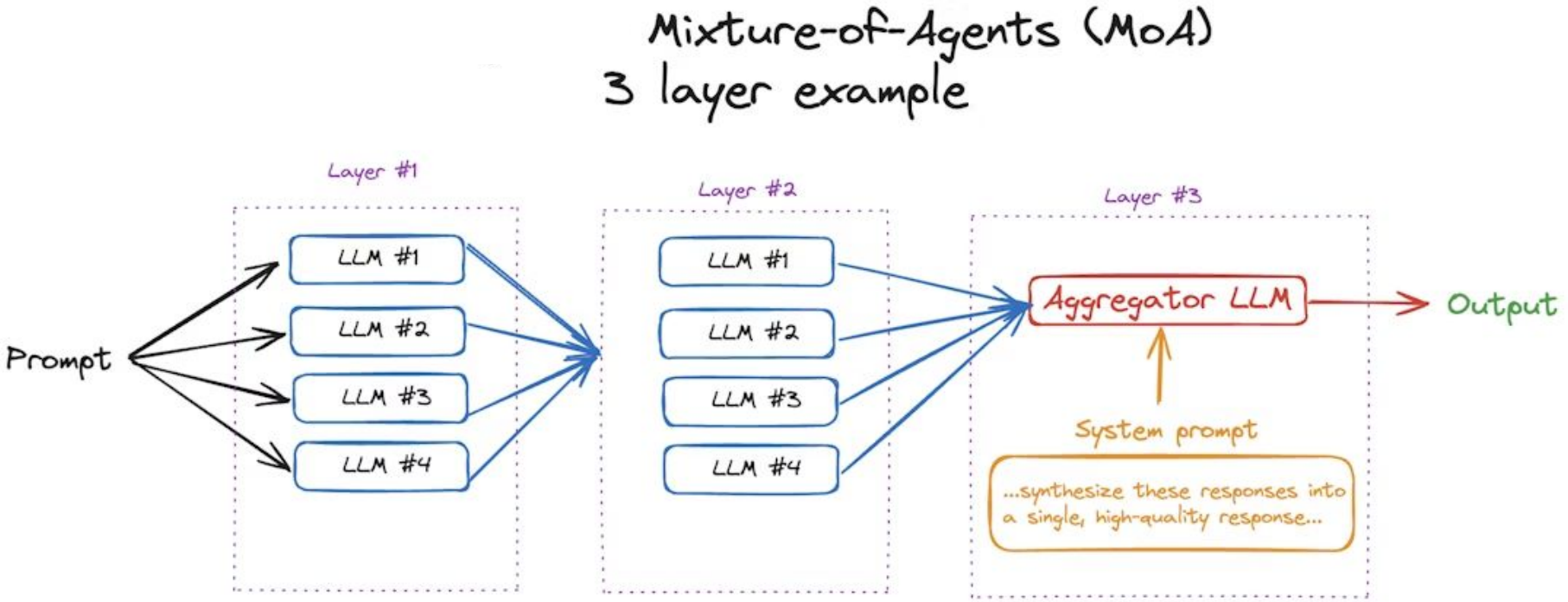
First, let's square both sides:

$a - \sqrt{a + x} = x^2 \implies \sqrt{a + x} = a - x^2$

Next, I could square both sides again, treating the equation: $\cdots$

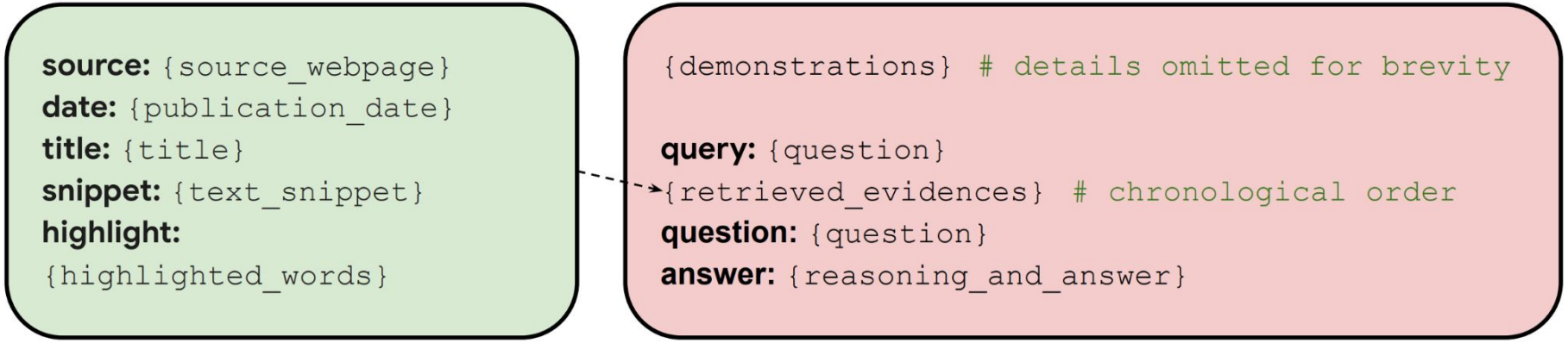
$\cdots$

Deep Think / Parallel Agents / Mixture-of-Agents



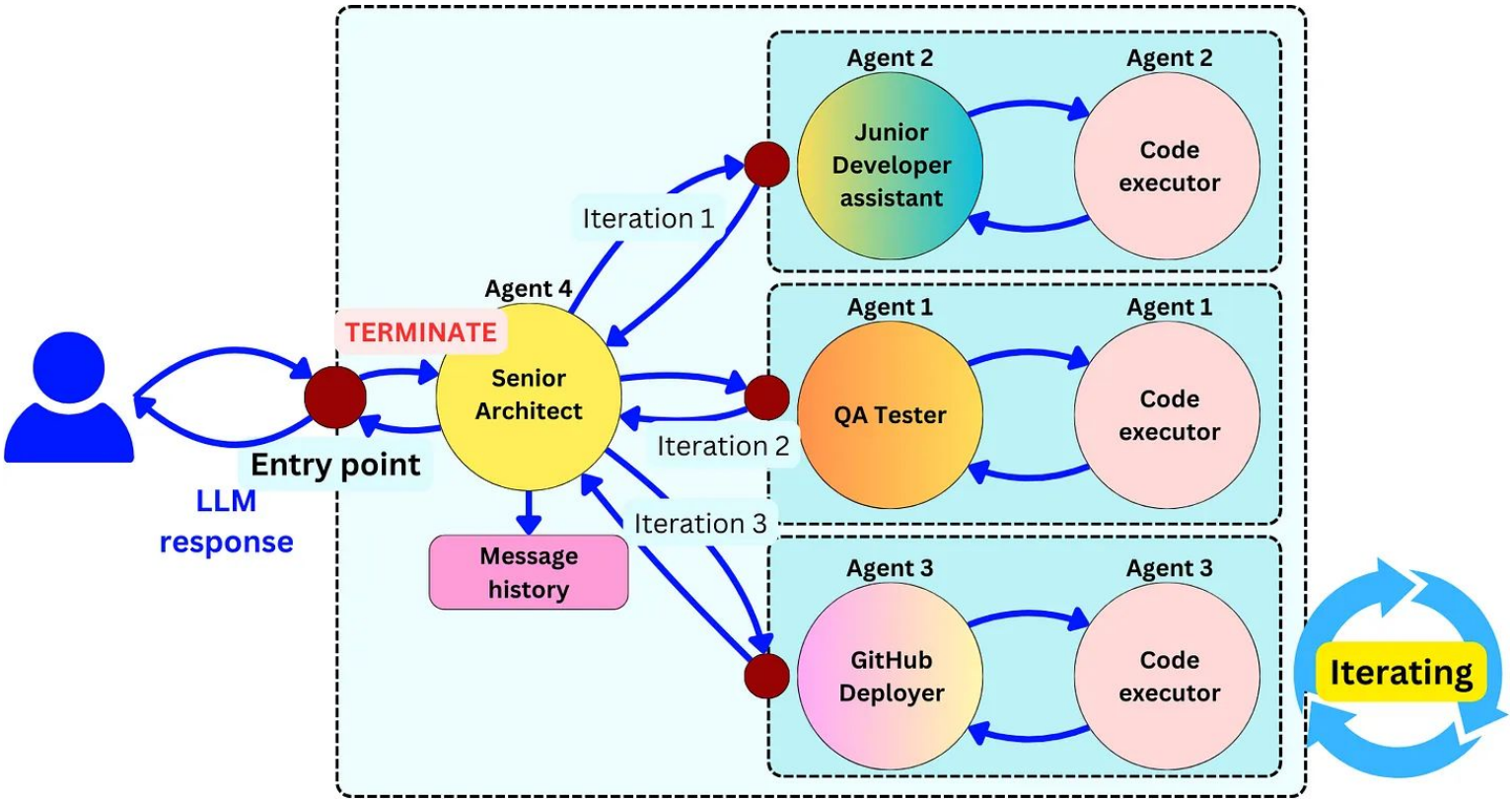
https://docs.together.ai/docs/mixture-of-agents#advanced-moa-example

Retrieval-Augmented Generation (RAG)



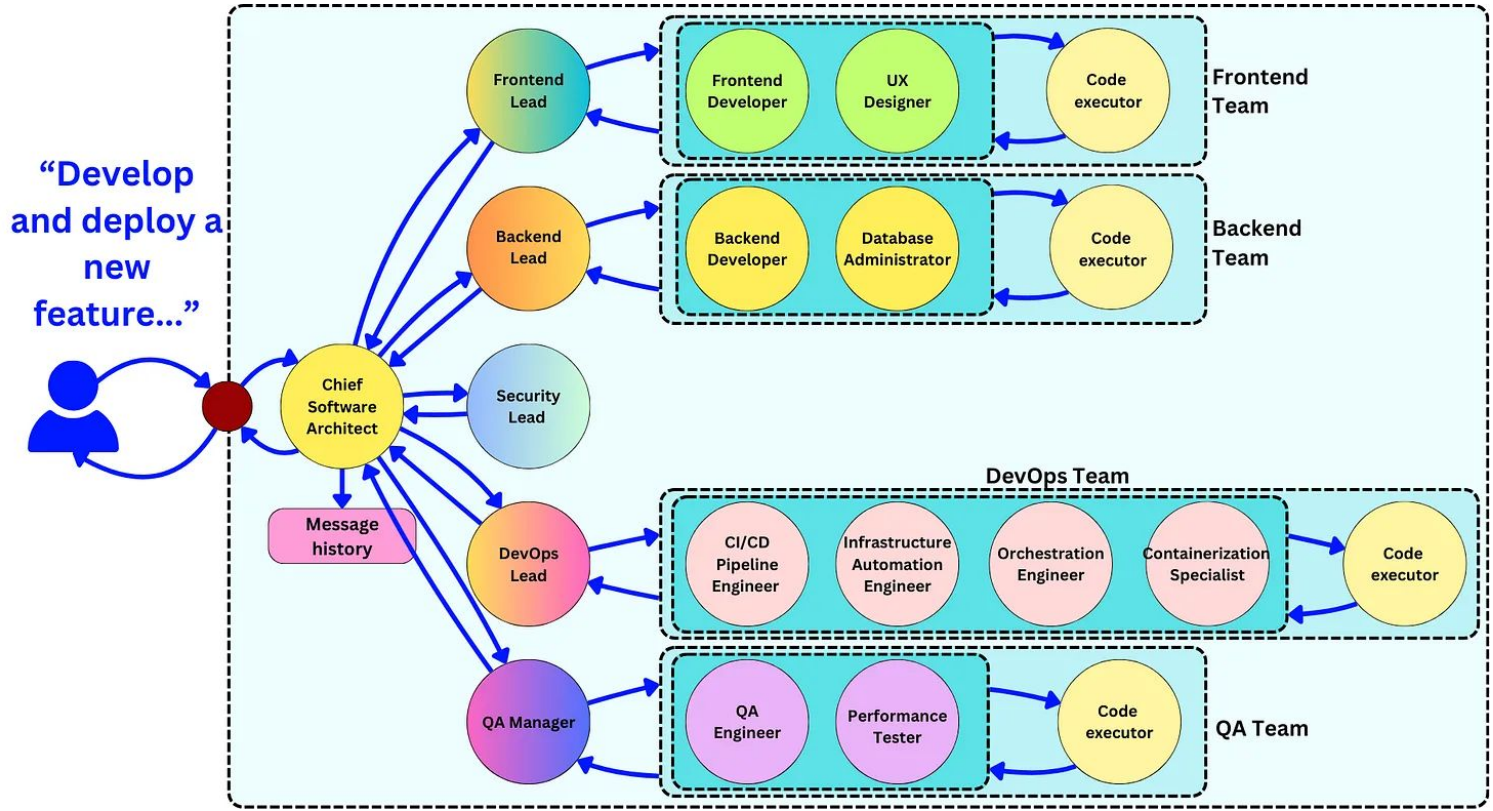
FreshPrompt uses few-shot in-context learning to teach a model to reason over retrieved evidences and figure out the right answer

Multi-agent collaboration



"The Different Agentic Patterns" by Damien Benveniste

Multi-agent collaboration (cont'd)



"The Different Agentic Patterns" by Damien Benveniste

Meta-Harness: End-to-End Optimization of Model Harnesses

Yoonho Lee
Stanford

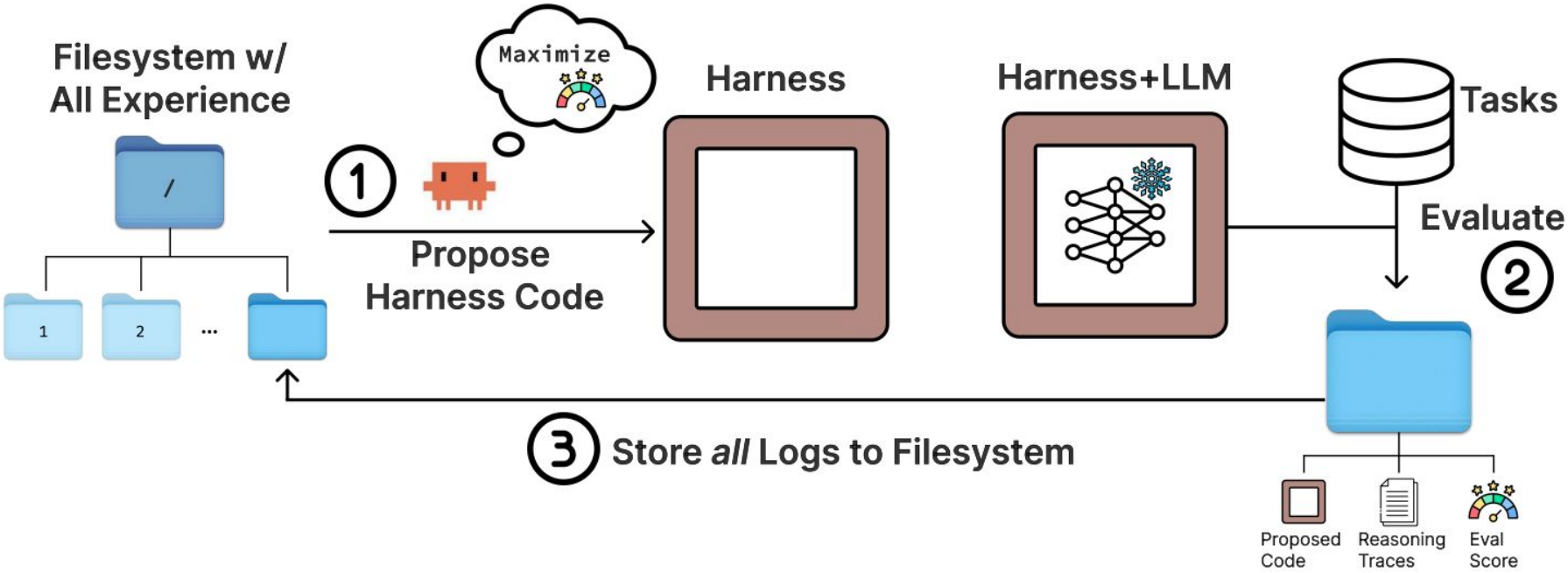
Roshen Nair
Stanford

Qizheng Zhang
Stanford

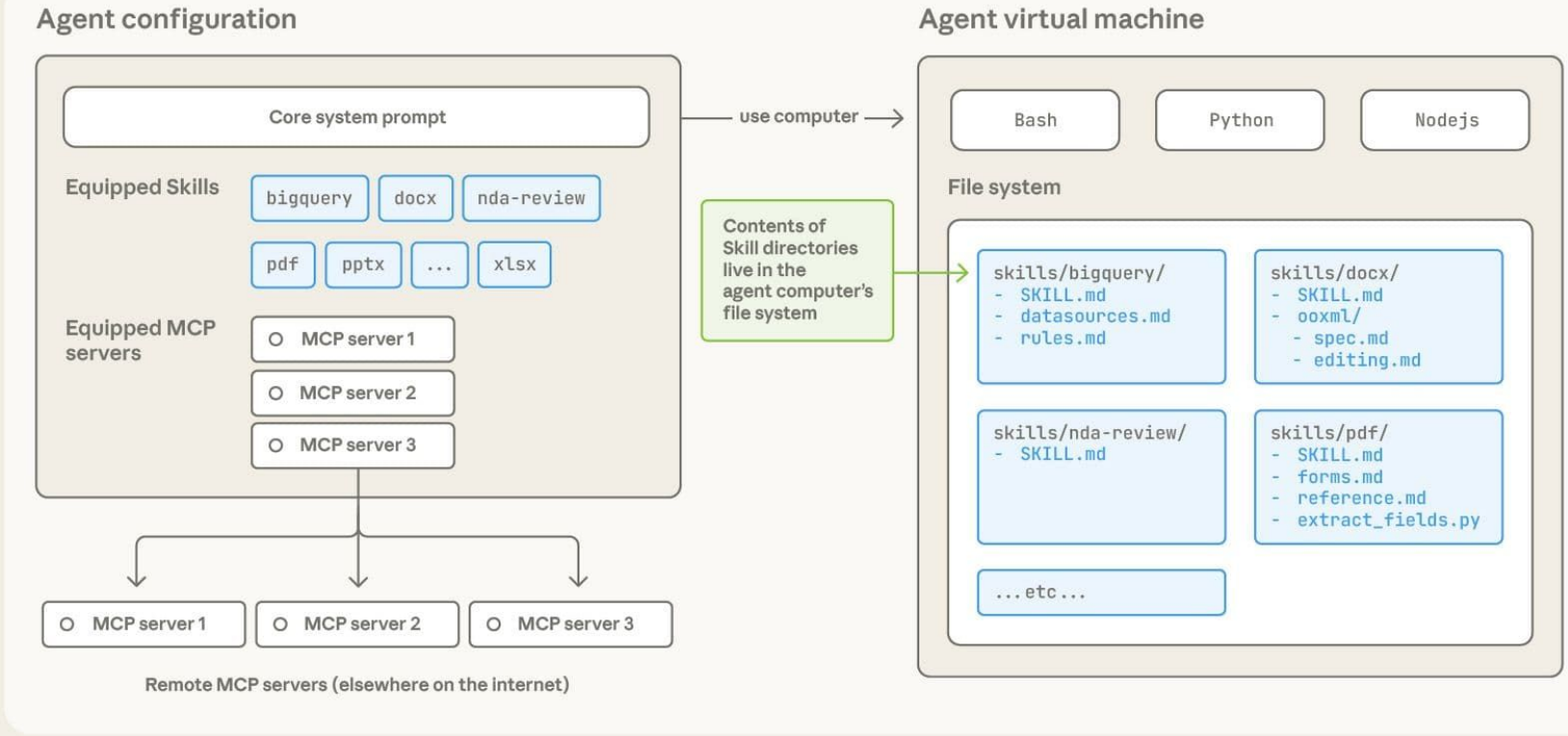
Kangwook Lee
KRAFTON

Omar Khattab
MIT

Chelsea Finn
Stanford


Filesystem w/
All Experience
/
1
2
...
Maximize
1
Propose
Harness Code
Harness
Harness+LLM
Tasks
Evaluate
2
3
Store *all* Logs to Filesystem
Proposed
Code
Reasoning
Traces
Eval
Score

Agent + Skills + Virtual Machine



https://www.anthropic.com/engineering/equipping-agents-for-the-real-world-with-agent-skills

A simple SKILL.md file

pdf/SKILL.md

YAML Frontmatter

```
---
name: pdf
description: Comprehensive PDF toolkit for extracting text and tables,
merging/splitting documents, and filling-out forms.
---
```

Markdown

```
## Overview

This guide covers essential PDF processing operations using Python libraries and command-line
tools. For advanced features, JavaScript libraries, and detailed examples, see ./reference.md.
If you need to fill out a PDF form, read ./forms.md and follow its instructions.

## Quick Start

```python
from pypdf import PdfReader, PdfWriter

# Read a PDF
reader = PdfReader("document.pdf")
print(f"Pages: {len(reader.pages)}")

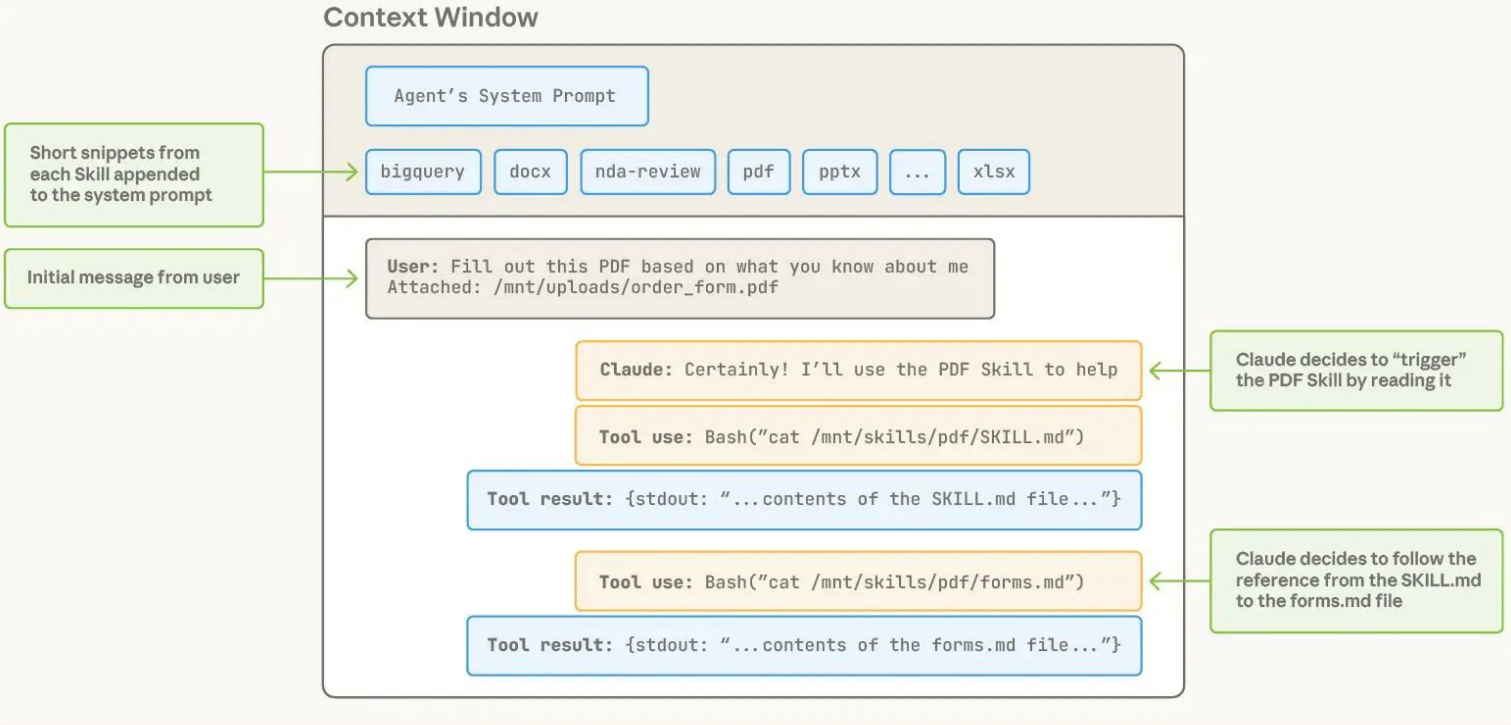
...
```

https://www.anthropic.com/engineering/equipping-agents-for-the-real-world-with-agent-skills
```

Bundling additional content

https://www.anthropic.com/engineering/equipping-agents-for-the-real-world-with-agent-skills

Skills and the Context Window



https://www.anthropic.com/engineering/equipping-agents-for-the-real-world-with-agent-skills

Thank you!