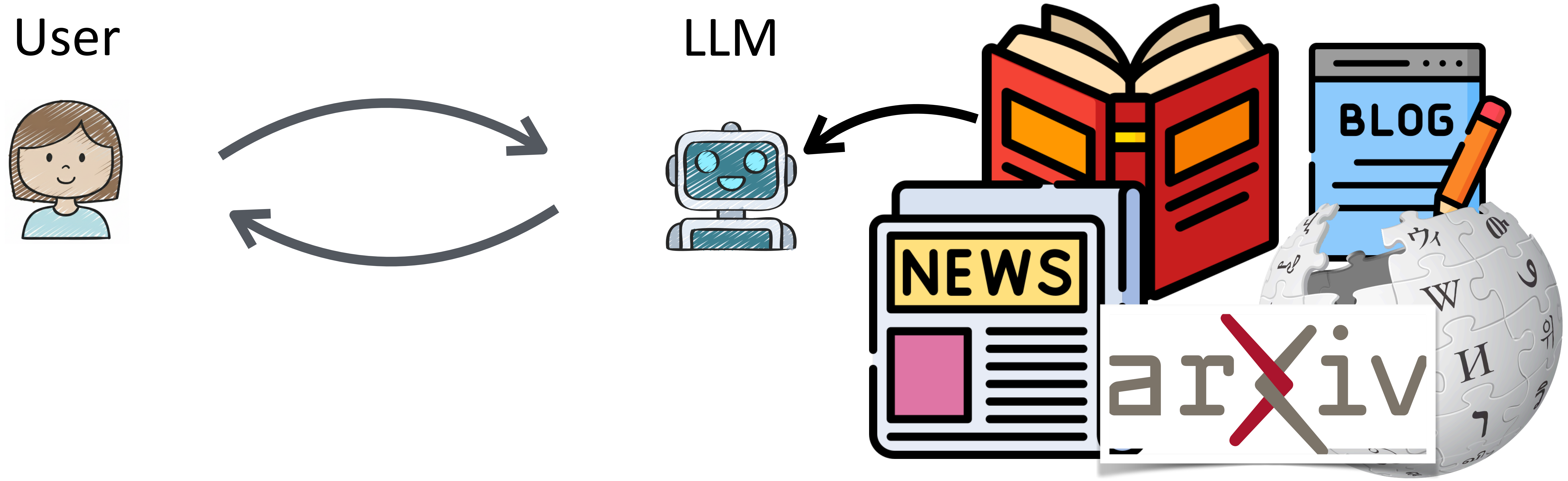
Equipping LLMs for Interaction

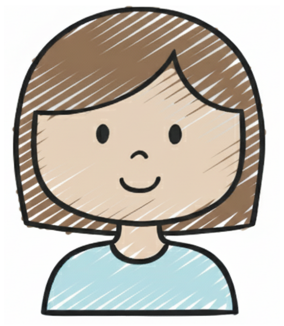
Eunsol Choi

NYU

LLMs in real world

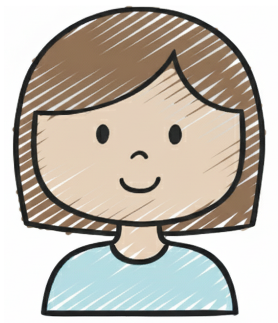


Interaction with LLM vs. with Human



What are some good hotels in Austin?

Interaction with LLM vs. with Human



What are some good hotels in Austin?

Here are some great hotel options in Austin, catering to different preferences and budgets:

Luxury Hotels

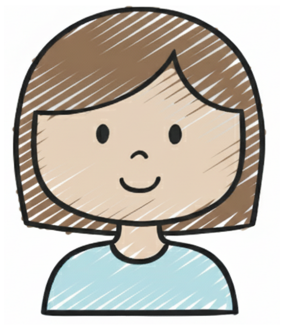
Driskill: A historic hotel in downtown, …

Mid-Range Hotels

South Congress Hotel: Located in the trendy South Congress area, this

382 words

Interaction with LLM vs. with Human



What are some good hotels in Austin?

Here are some great hotel options in Austin, catering to different preferences and budgets:

Luxury Hotels

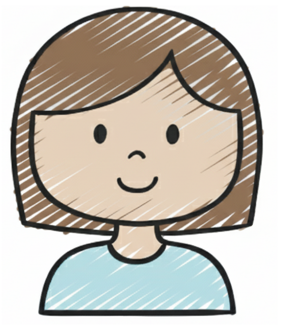
Driskill: A historic hotel in downtown,

…

Mid-Range Hotels

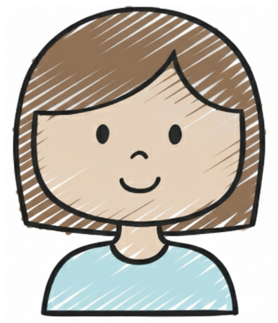
South Congress Hotel: Located in the trendy South Congress area, this

382 words



What are some good hotels in Austin?

Interaction with LLM vs. with Human



What are some good hotels in Austin?

Here are some great hotel options in Austin, catering to different preferences and budgets:

Luxury Hotels

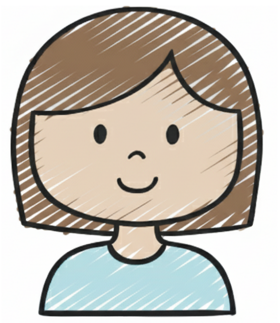
Driskill: A historic hotel in downtown,

…

Mid-Range Hotels

South Congress Hotel: Located in the trendy South Congress area, this

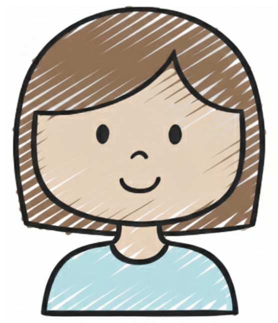
382 words



What are some good hotels in Austin?

What is your price range? Do you like to be in downtown?

Interaction with LLM vs. with Human



What are some good hotels in Austin?

Here are some great hotel options in Austin, catering to different preferences and budgets:

Luxury Hotels

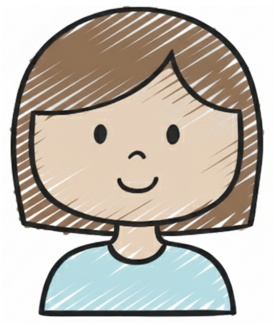
Driskill: A historic hotel in downtown,

…

Mid-Range Hotels

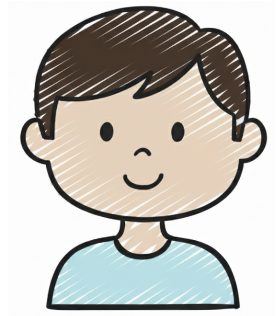
South Congress Hotel: Located in the trendy South Congress area, this

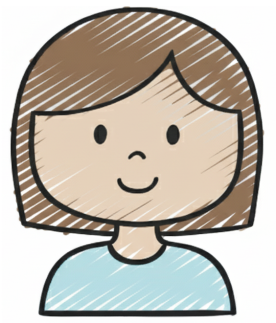
382 words



What are some good hotels in Austin?

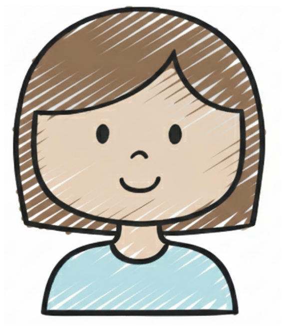
What is your price range? Do you like to be in downtown?





Mid-range and downtown is great.

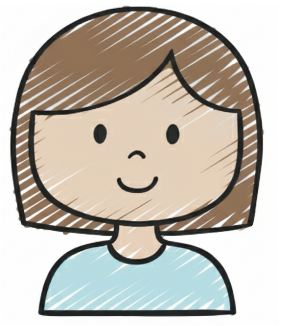
Interaction with LLM vs. with Human



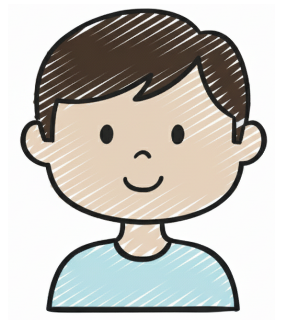
What are some good hotels in Austin?

Here are some great hotel options in Austin, catering to different preferences and budgets:
Luxury Hotels
Driskill: A historic hotel in downtown,
…
Mid-Range Hotels
South Congress Hotel: Located in the trendy South Congress area, this

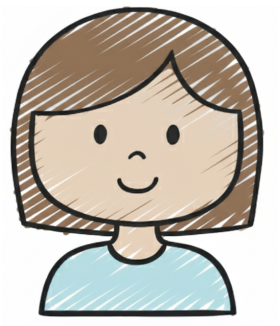
382 words



What are some good hotels in Austin?



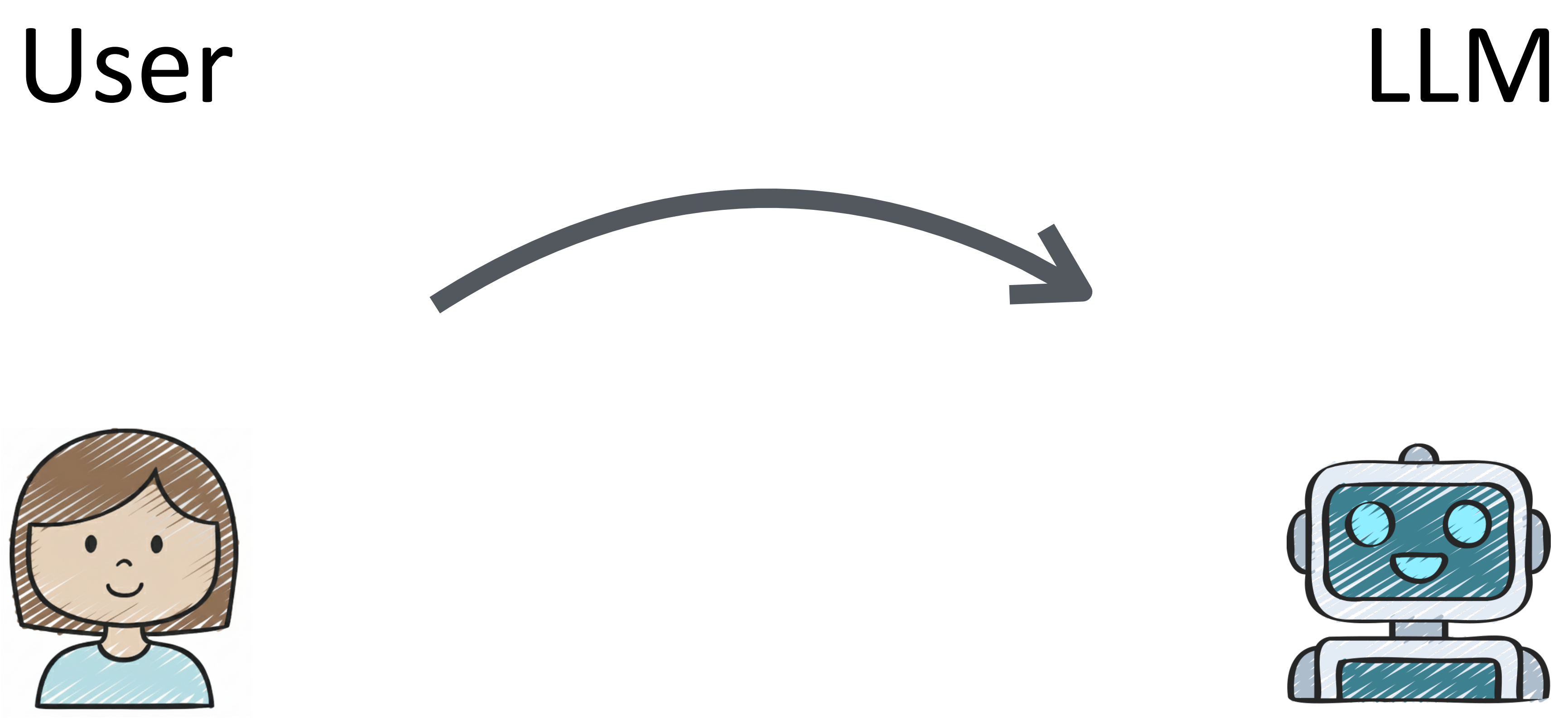
What is your price range? Do you like to be in downtown?



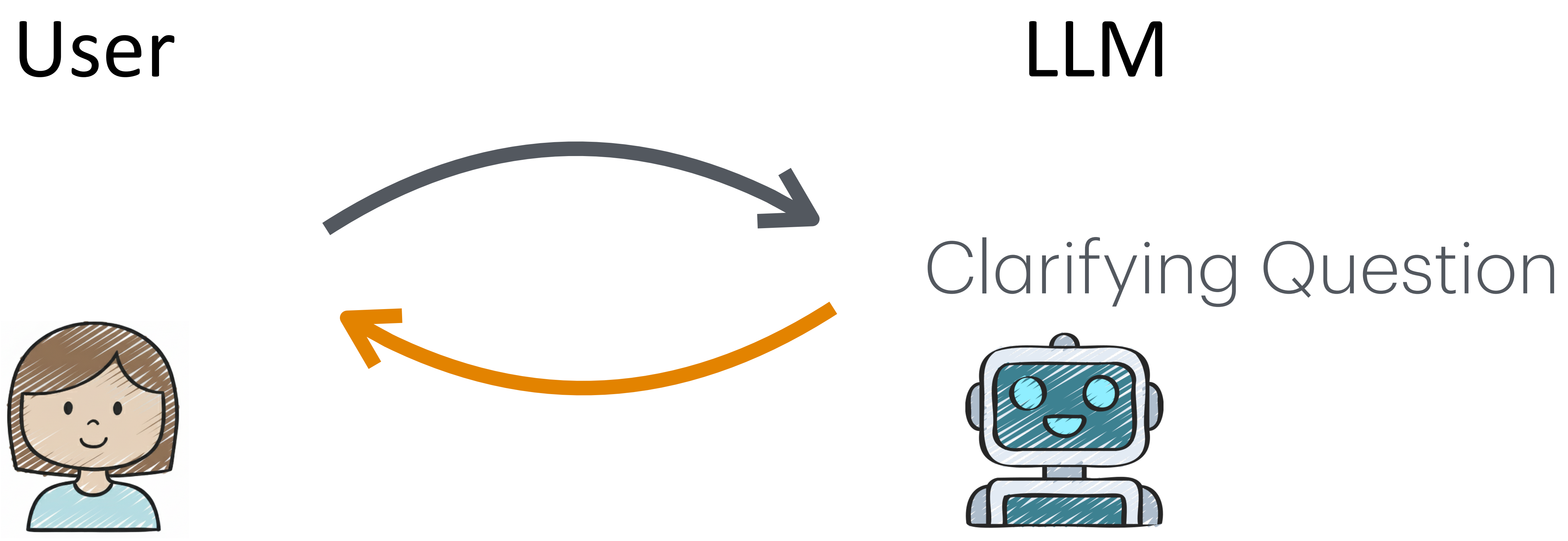
Mid-range and downtown is great.

In that case, I'd recommend the Kimpton hotel which is centrally located and has high ratings.

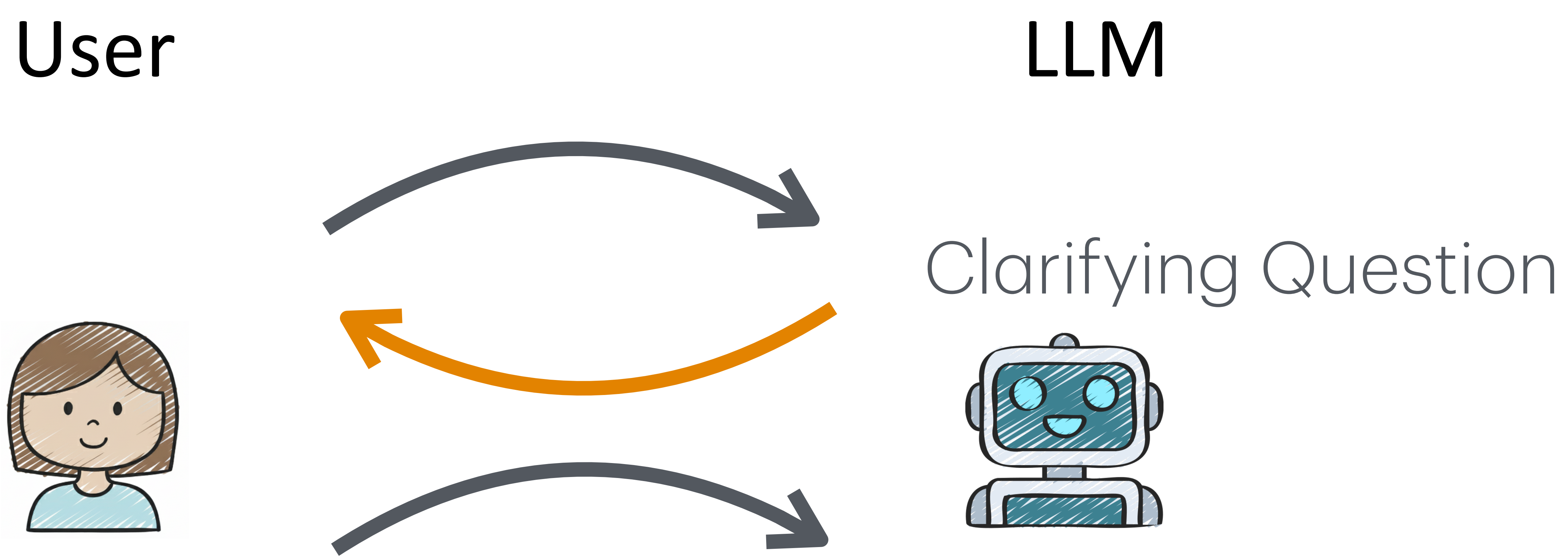
LLMs in real world



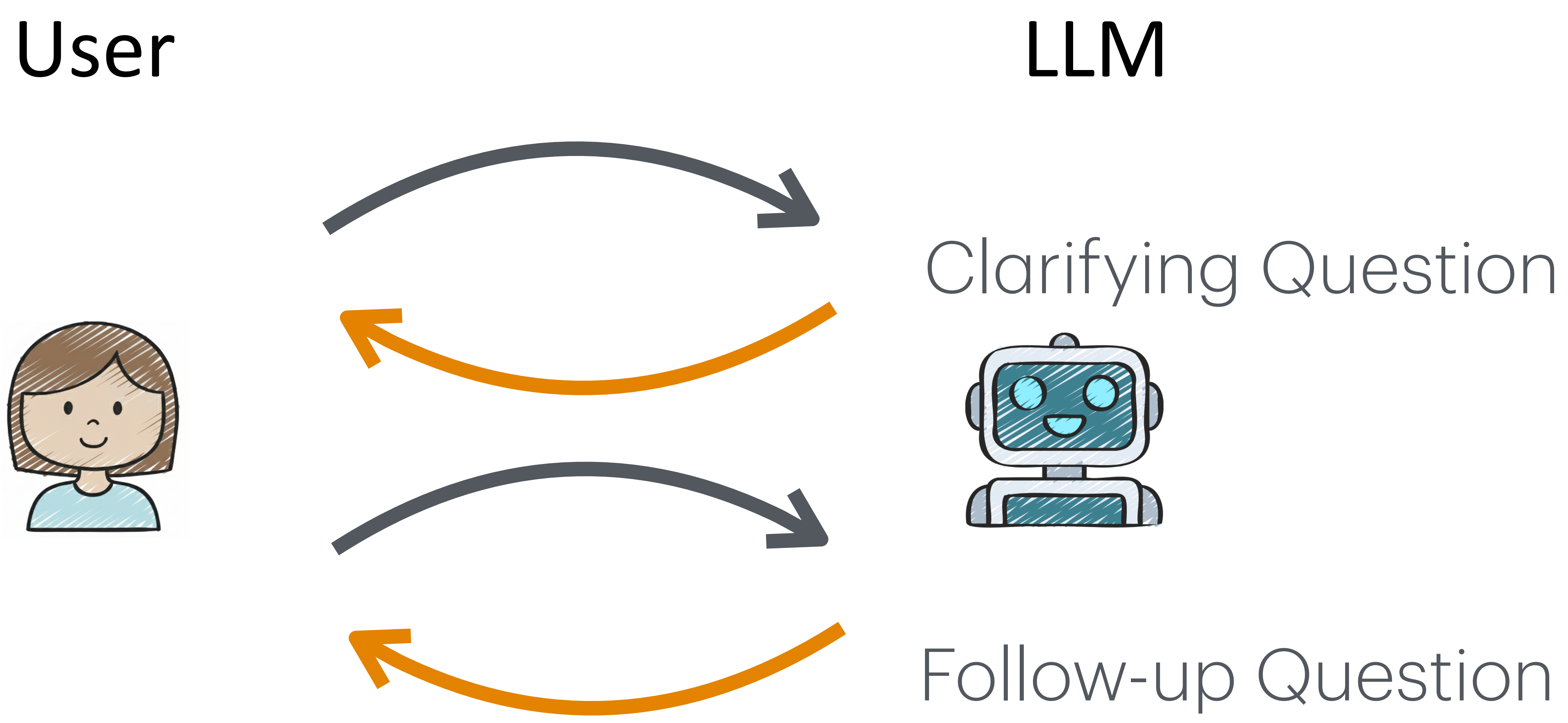
LLMs in real world



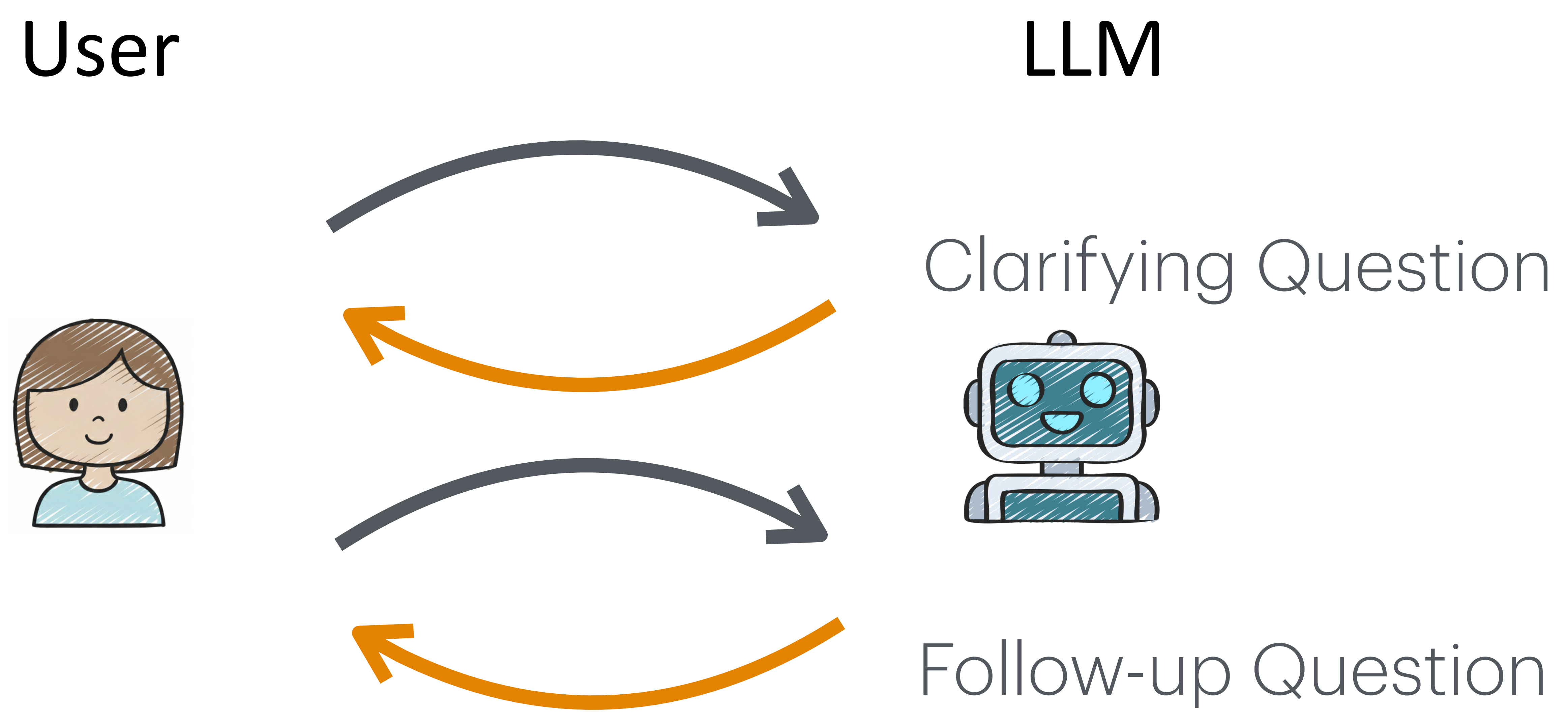
LLMs in real world



LLMs in real world

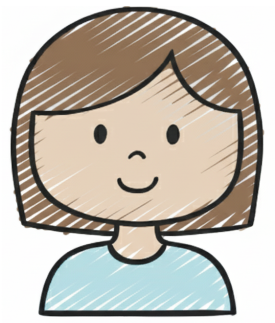


LLMs in real world

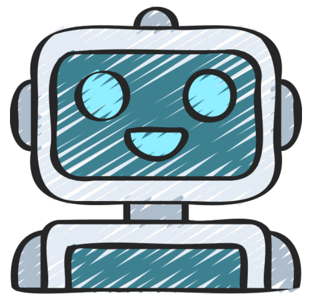


Part 1: Teach LLM to take initiative

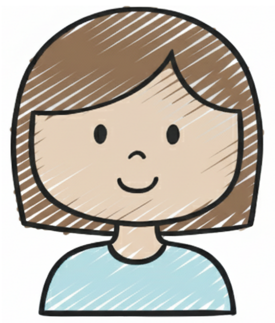
User Feedback from Human LLM Conversation



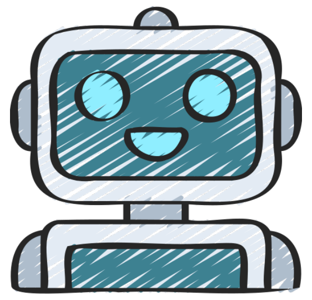
What are some good hotels in Austin?



What is your price range? Do you like to be in downtown?

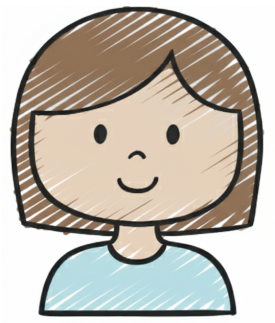


Mid-range and downtown is great.



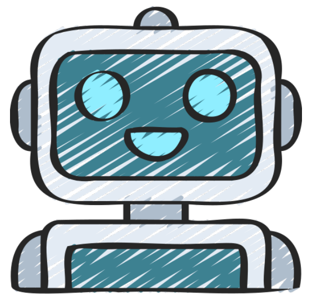
In that case, I'd recommend the Kimpton hotel which is centrally located and has high ratings.

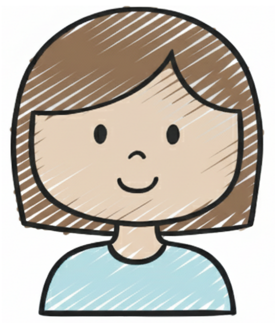
User Feedback from Human LLM Conversation



What are some good hotels in Austin?

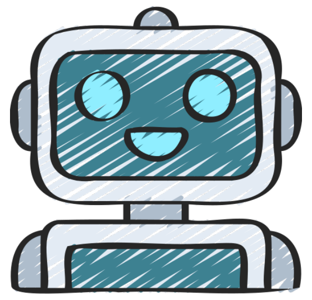
What is your price range? Do you like to be in downtown?

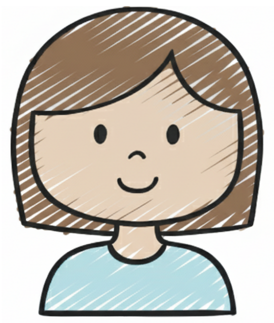




Mid-range and downtown is great.

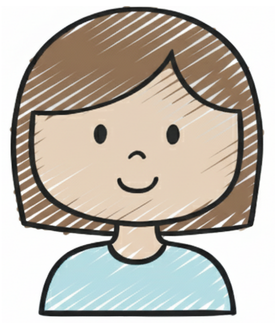
In that case, I'd recommend the Kimpton hotel which is centrally located and has high ratings.





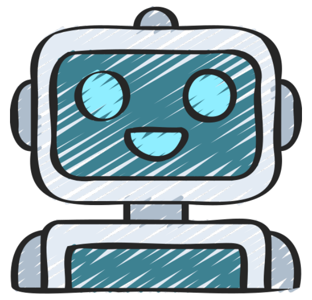
Great!

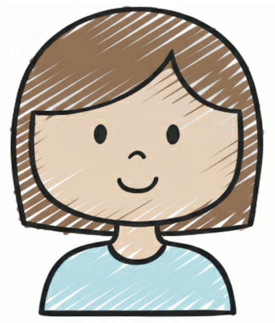
User Feedback from Human LLM Conversation



What are some good hotels in Austin?

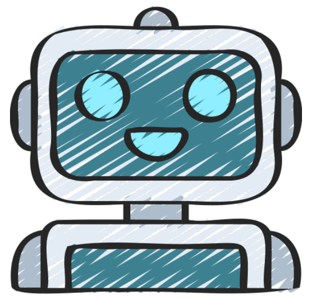
What is your price range? Do you like to be in downtown?

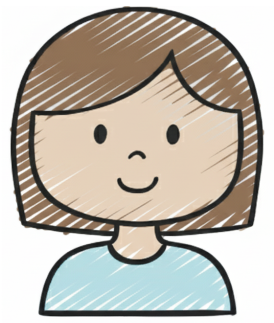




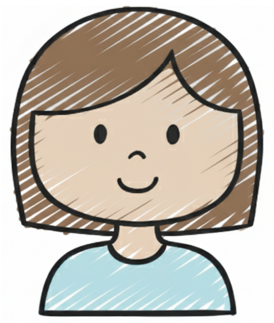
Mid-range and downtown is great.

In that case, I'd recommend the Kimpton hotel which is centrally located and has high ratings.



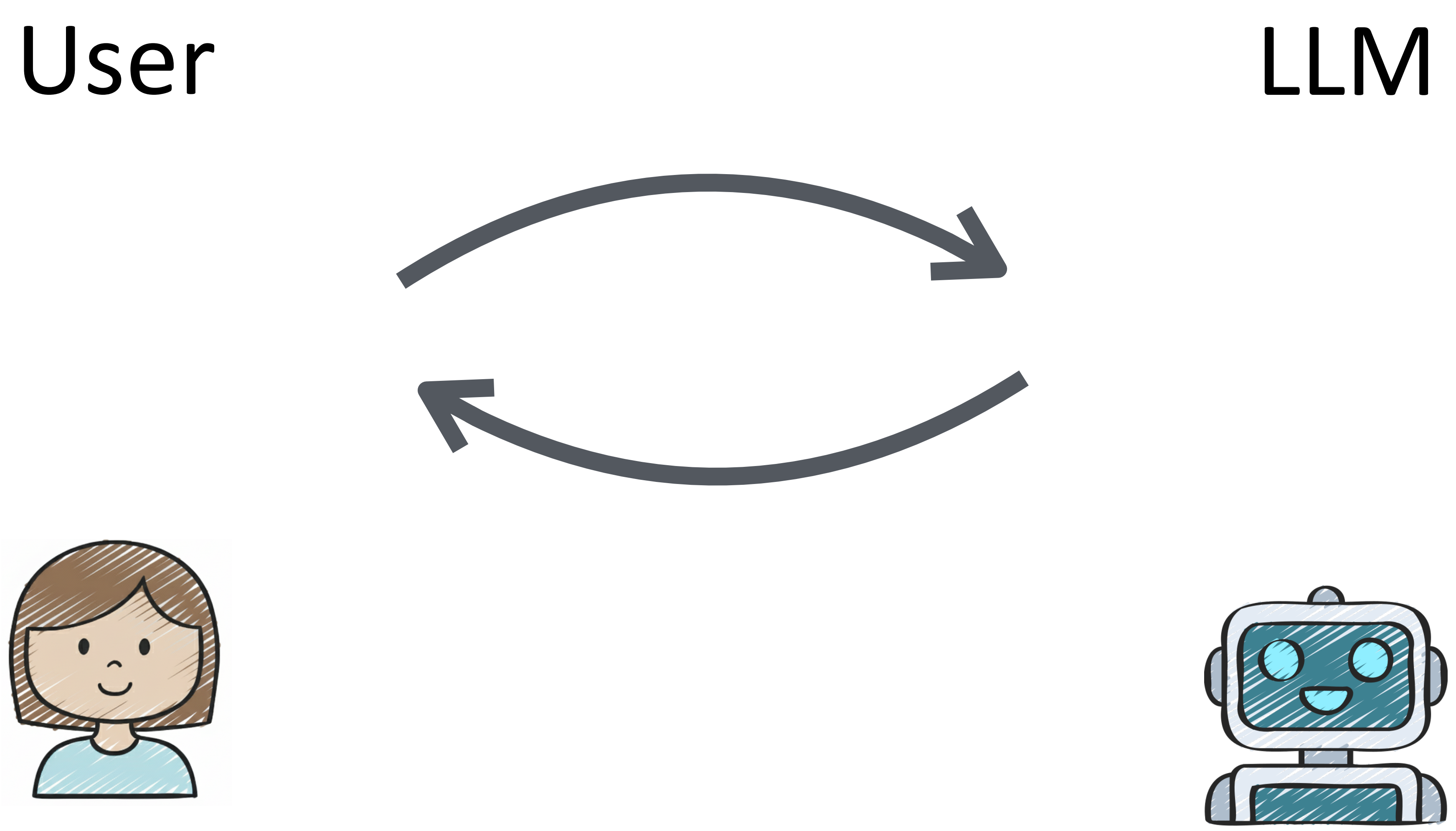


Great!

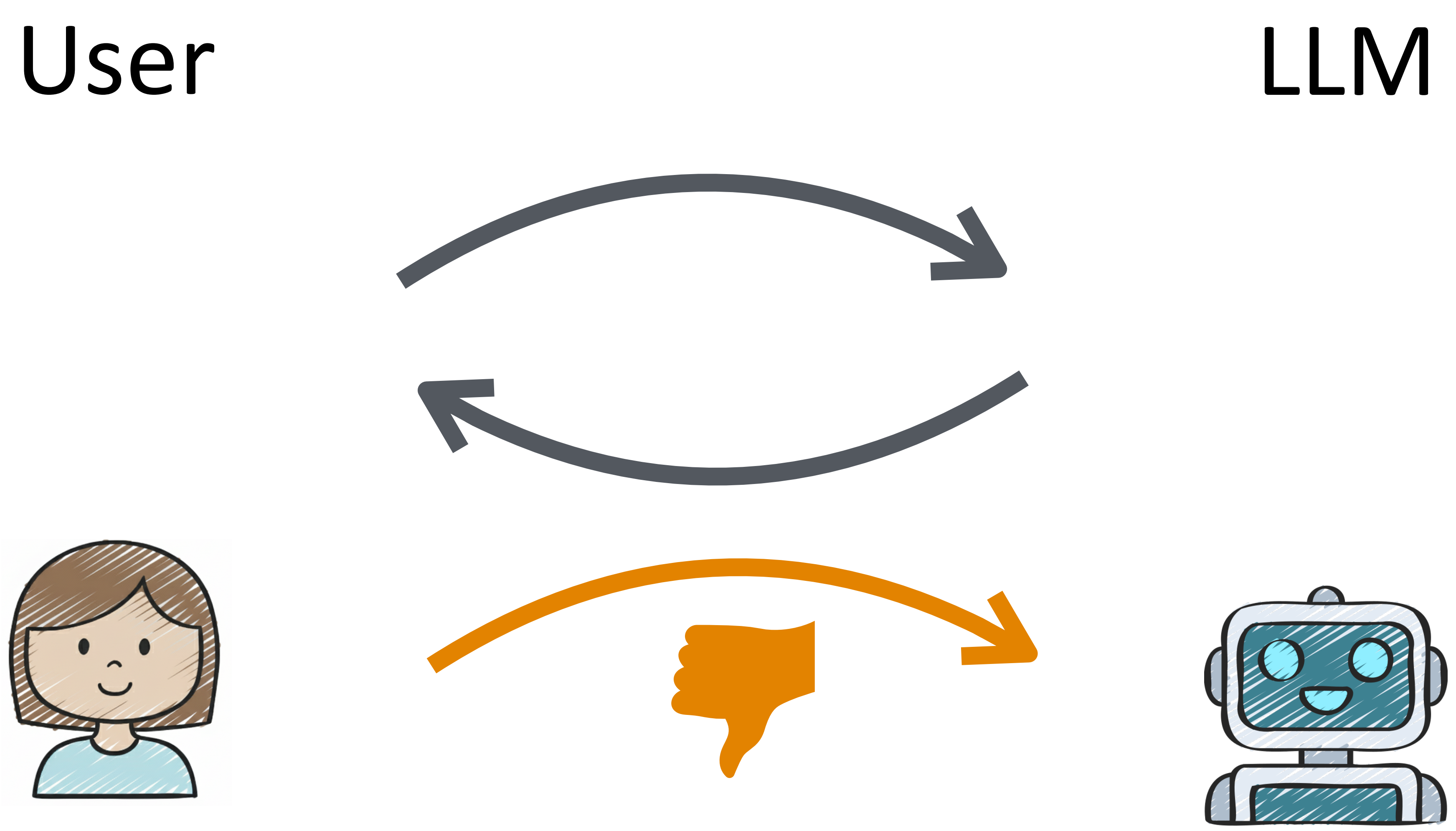


Ah, but I don't like chain hotel like Kimpton. I want something more local.

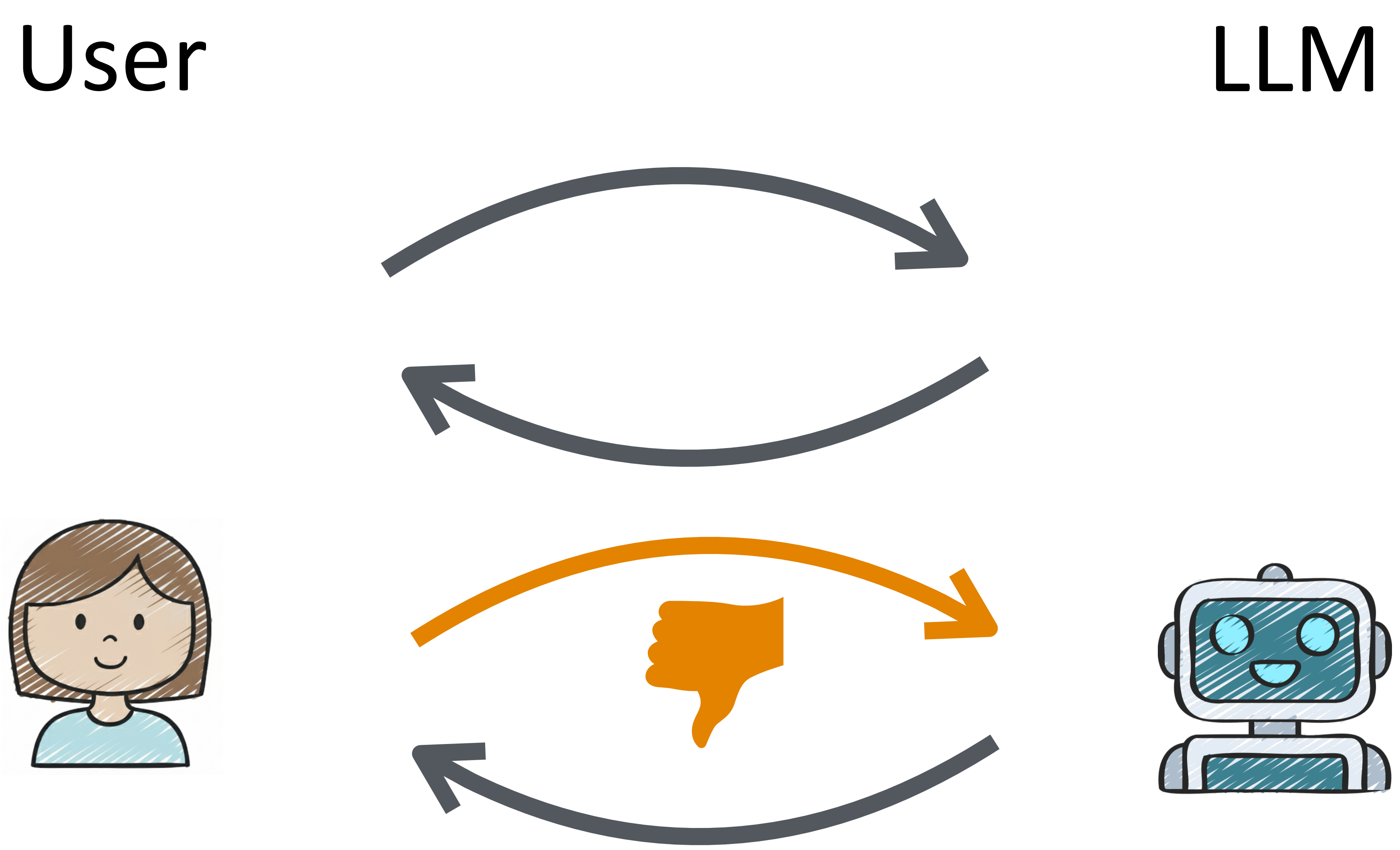
LLMs in real world



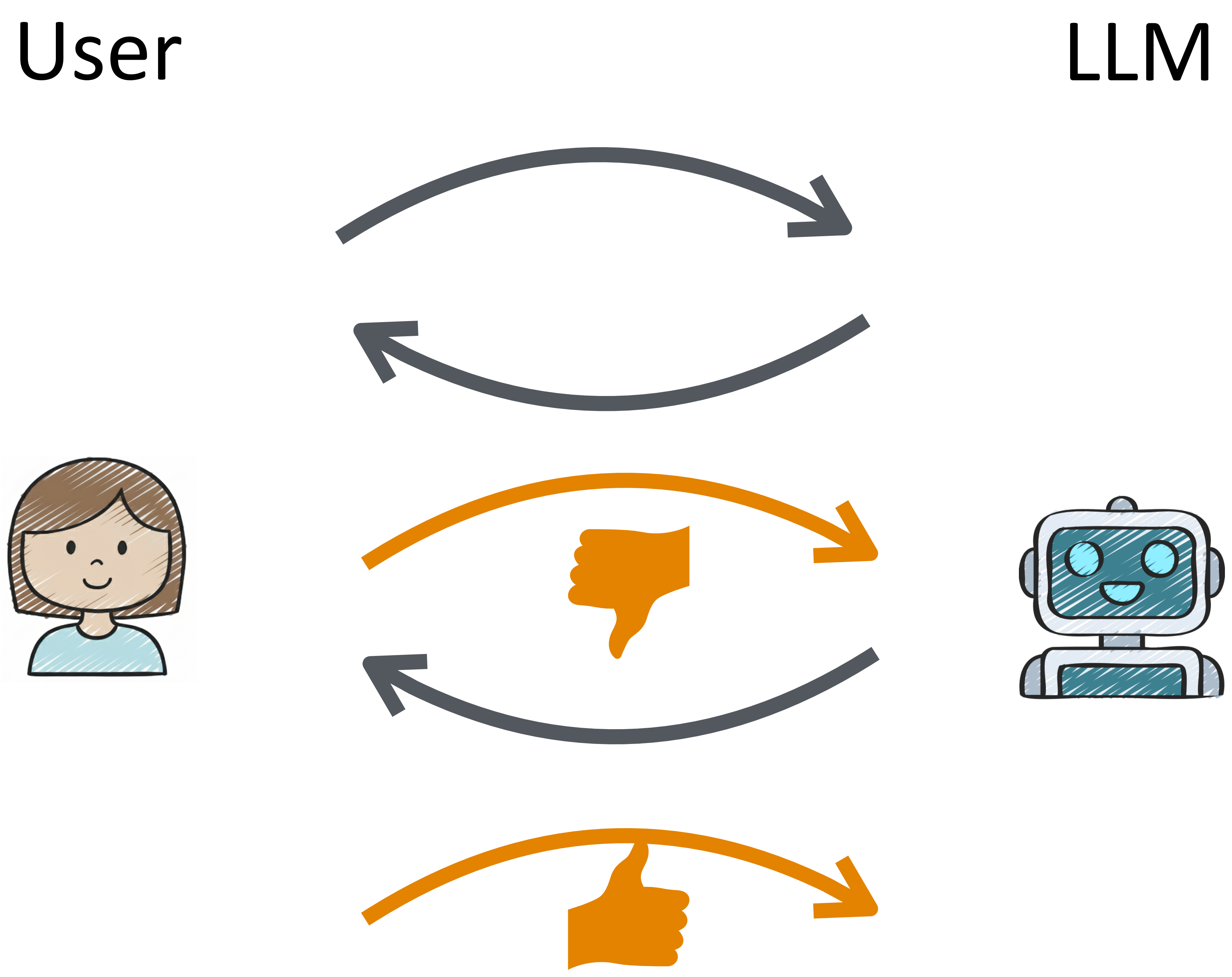
LLMs in real world



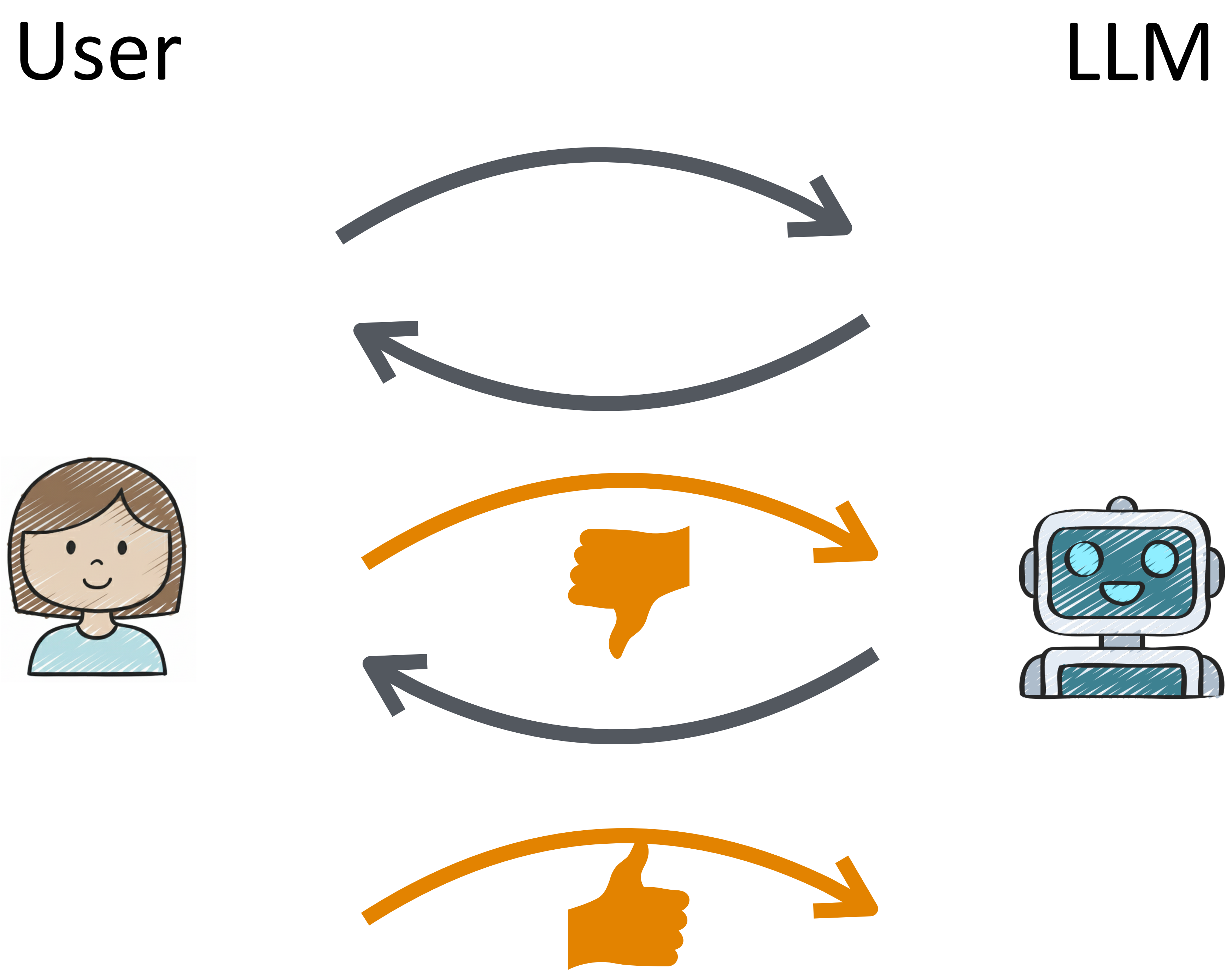
LLMs in real world



LLMs in real world

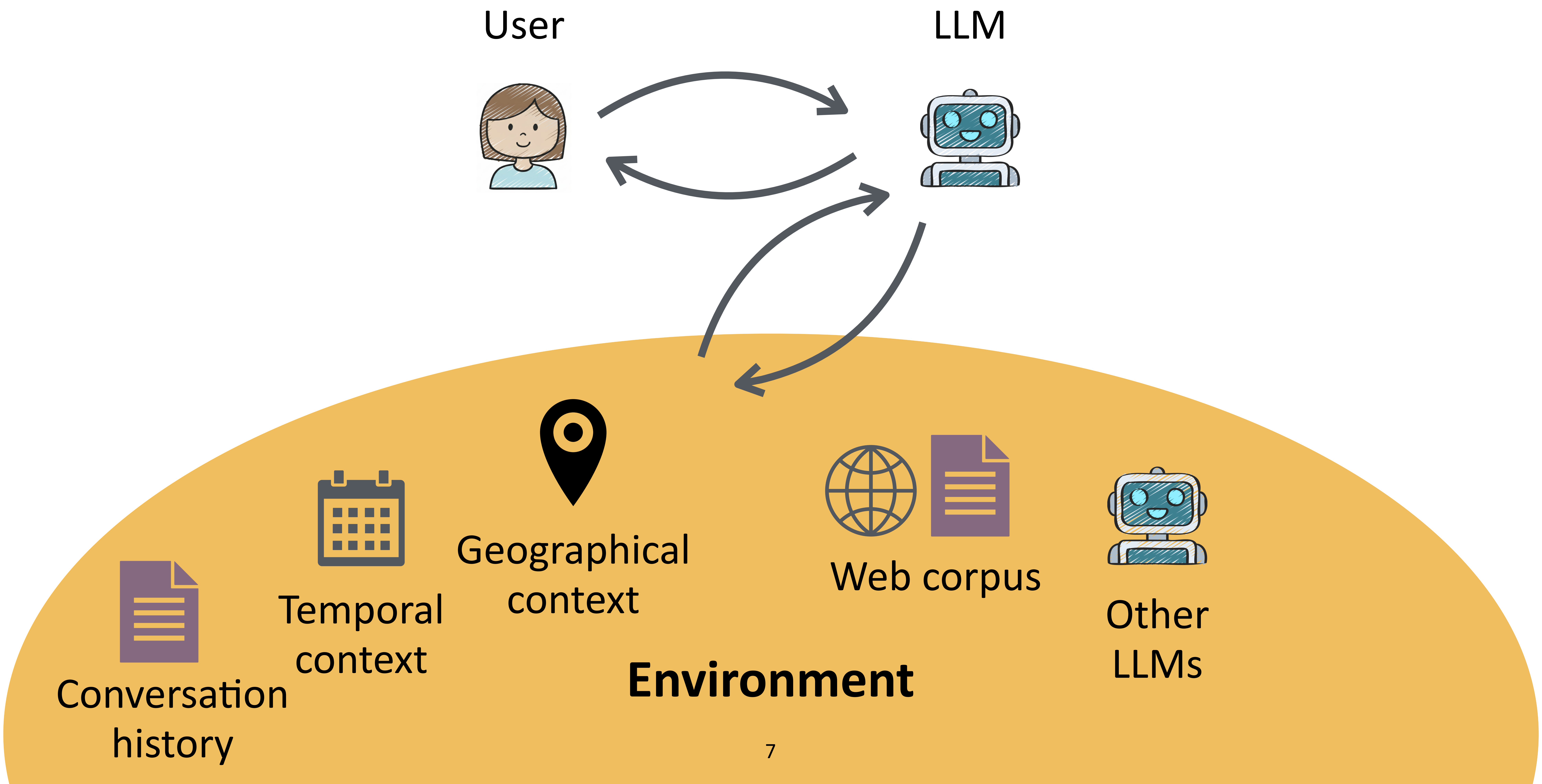


LLMs in real world

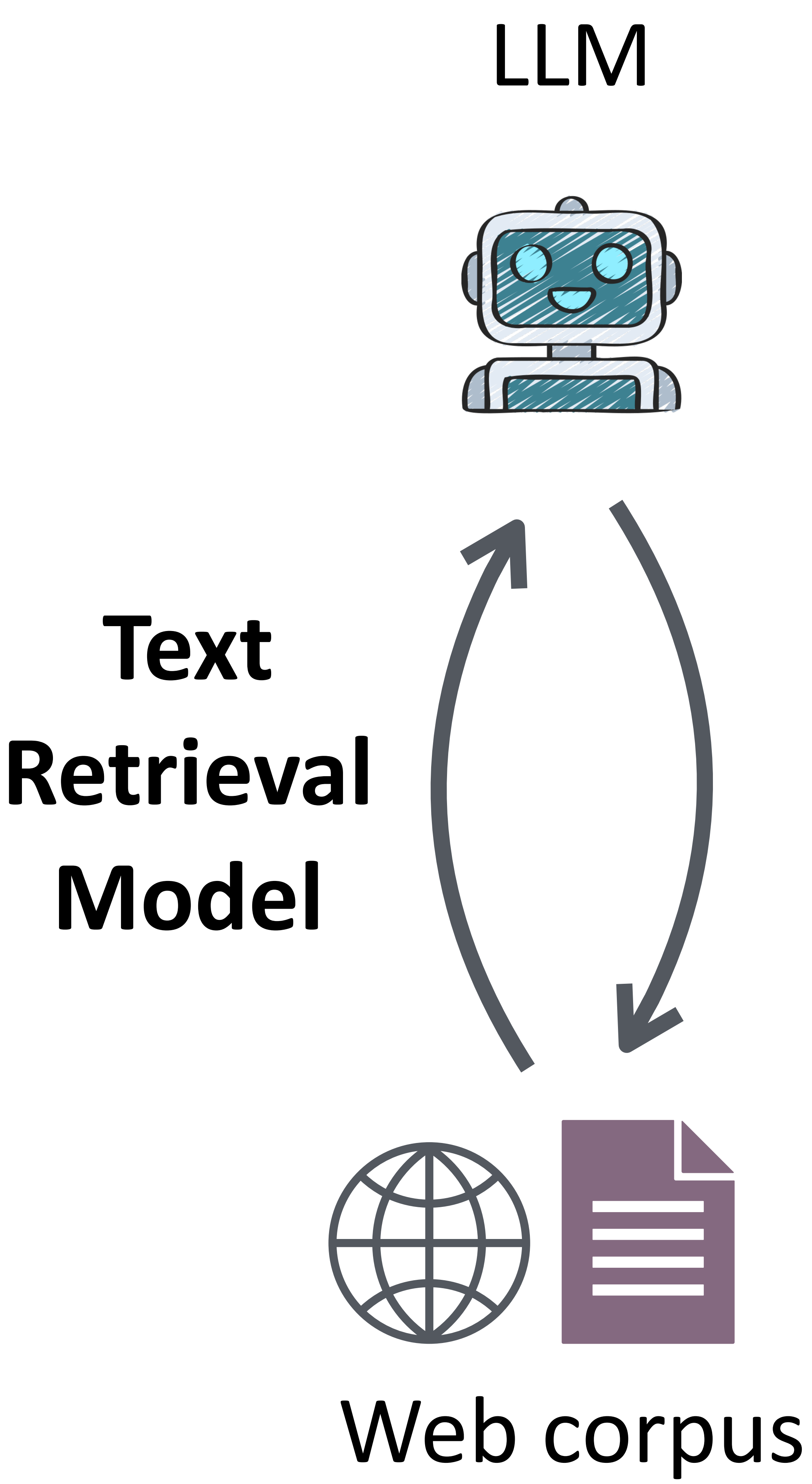


Part 2: Leverage User Feedback

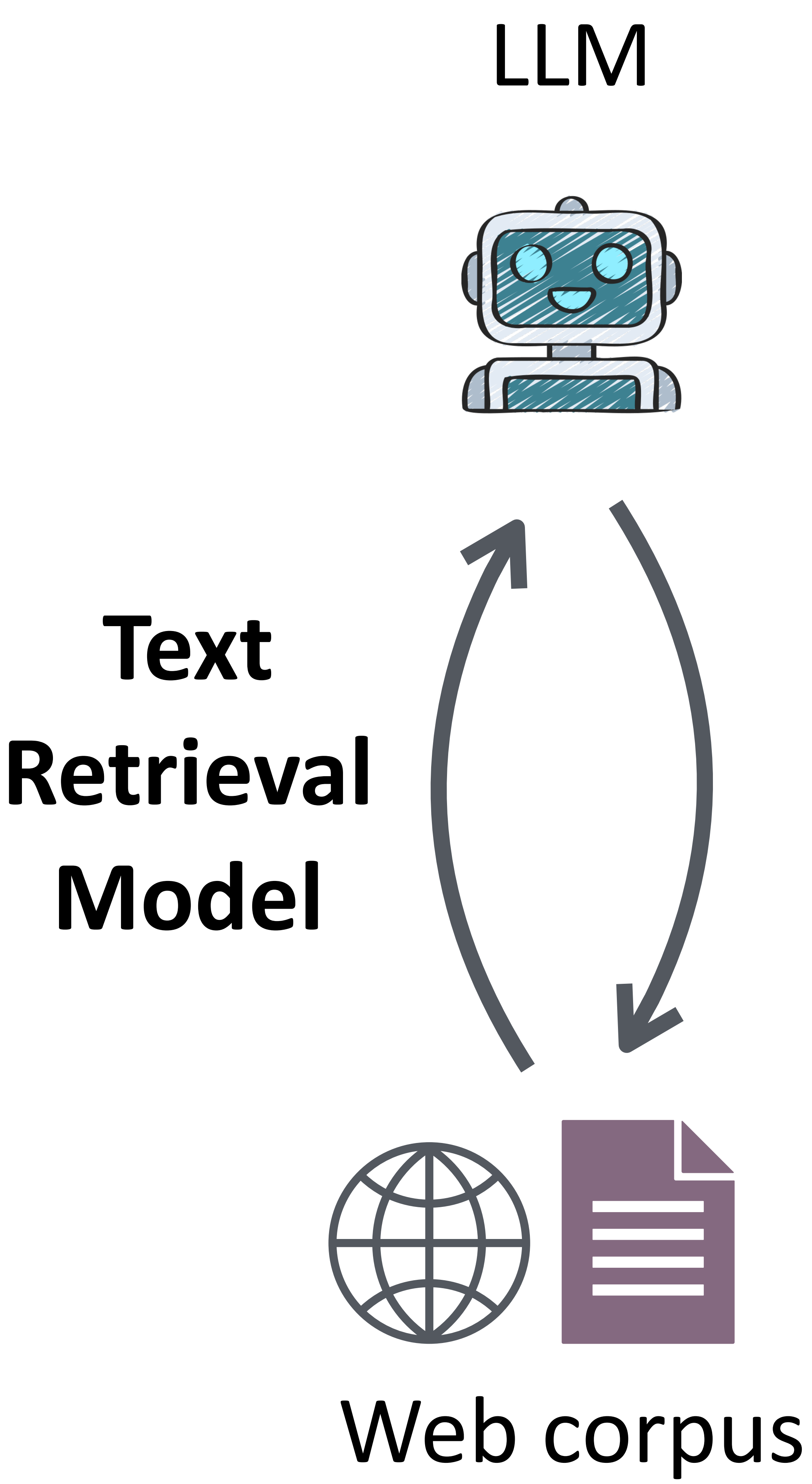
Interaction between LLM and Environment

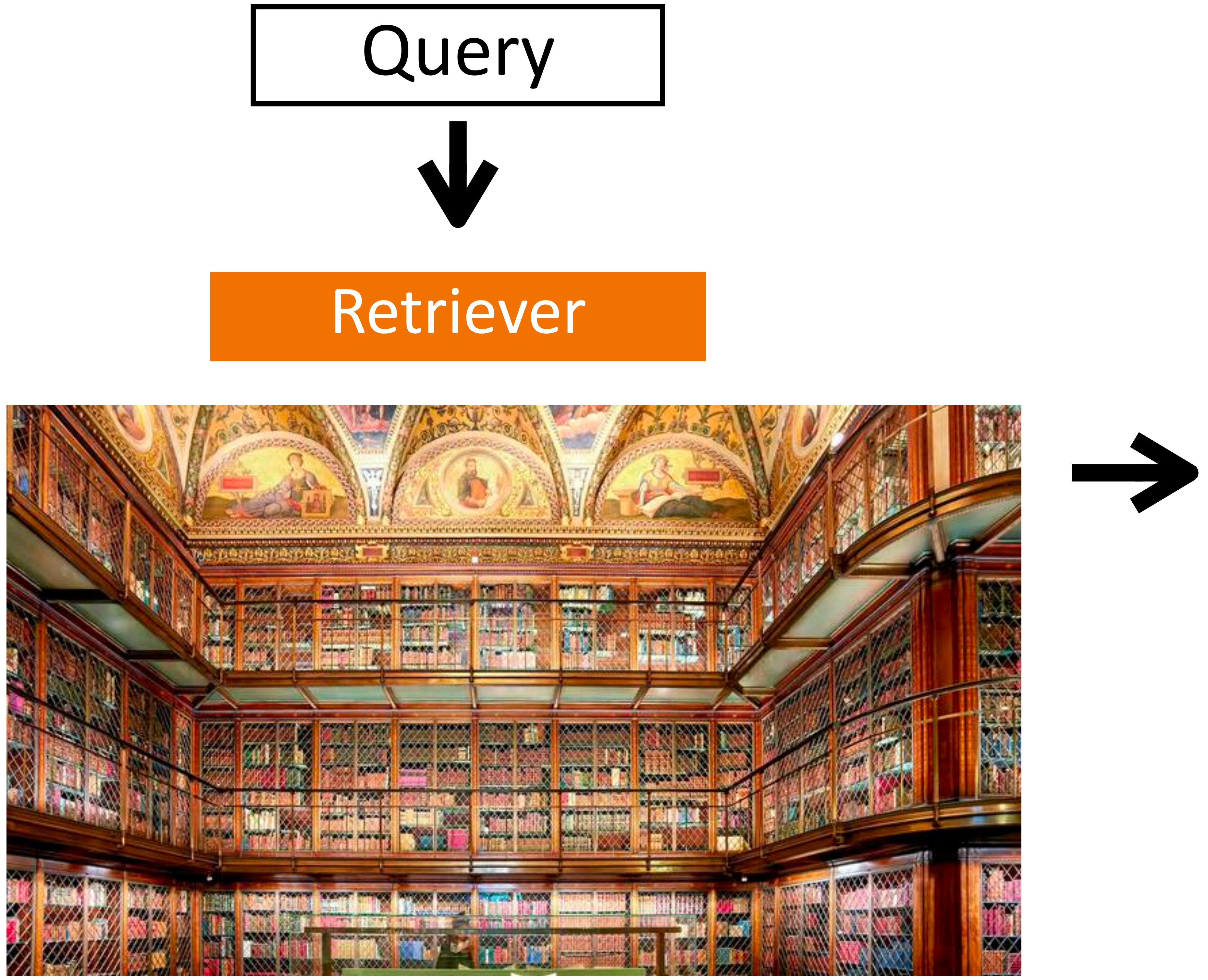


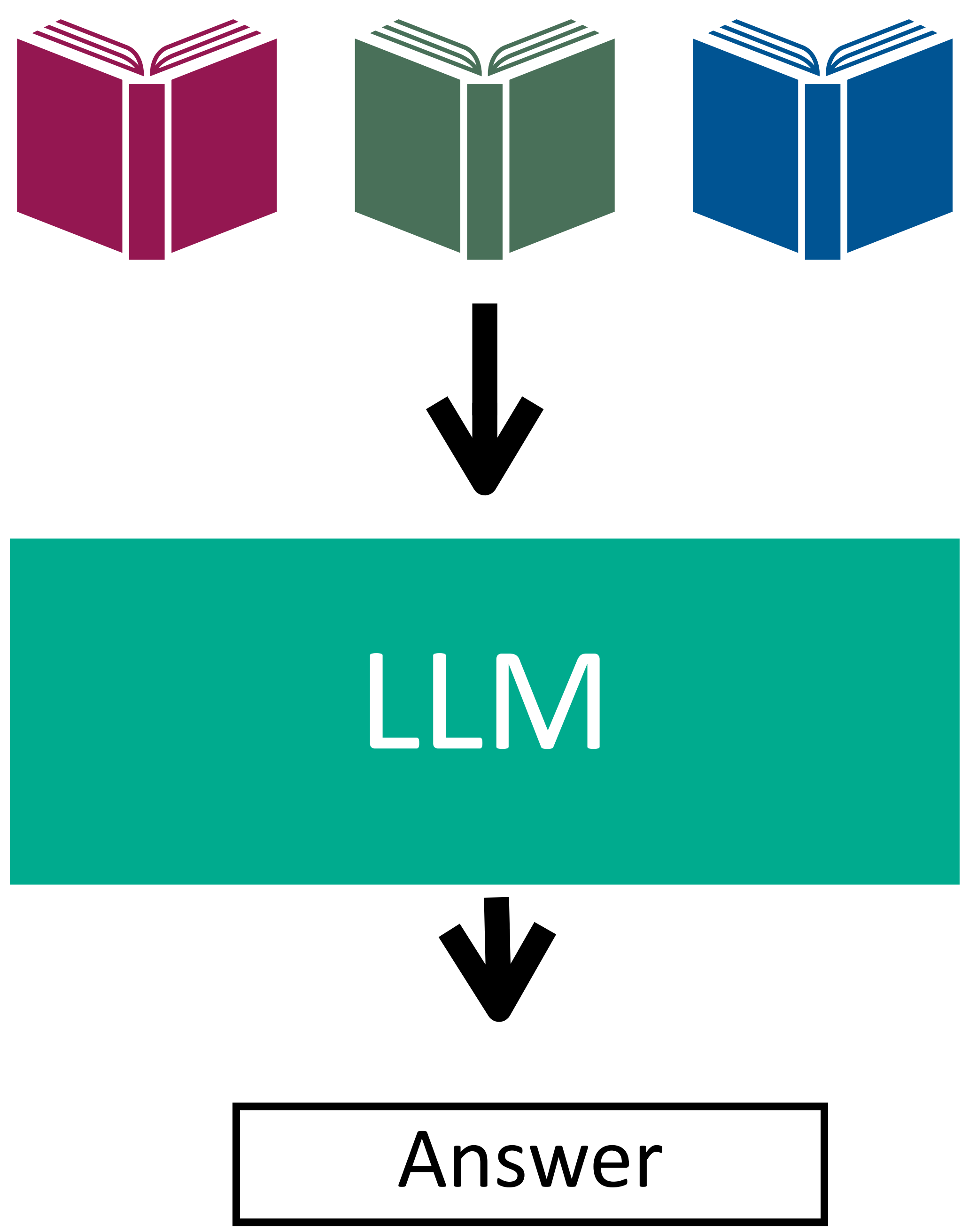
Focus: LLM using Text Retrieval Tools



Focus: LLM using Text Retrieval Tools



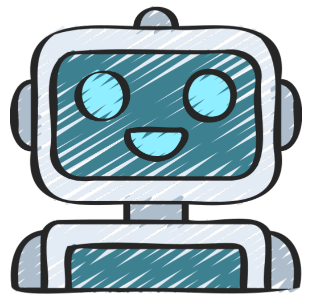




[RefreshKV: Updating Small KV Cache During Long-form Generation, Xu, Goyal*and **Choi***, under Submission]
[AmbigDocs: Reasoning across Documents on Different Entities under the Same Name, Lee, Ye, **Choi**, COLM24]

Focus: LLM using Text Retrieval Tools

LLM



Query

Retriever

Part 3: LLM interacting with retrieval outputs

Text
Retriev
Model

Web corpus

[RefreshKV: Updating Small KV Cache During Long-form Generation, Xu, Goyal*and **Choi***, under Submission]
[AmbigDocs: Reasoning across Documents on Different Entities under the Same Name, Lee, Ye, **Choi**, COLM24]

LLM

Answer

This Talk

Part 1: **User**

Teach LLM to ask clarifying questions

[Modeling Future Conversation Turns to Teach LLMs to Ask Clarifying Questions, Zhang, Knox, Choi, ICLR 25]

Learning from User Feedback

Part 2: **Environment**

Add new information at inference

Why Do LLMs Need Clarification?

Humans interpret questions in rich contexts

- Who wrote it? Why they wrote it?
- When and where was it written?

Why Do LLMs Need Clarification?

Humans interpret questions in rich contexts

- Who wrote it? Why they wrote it?
- When and where was it written?

Misinfo Reaction Frames: Reasoning about Readers' Reactions to News Headlines
[Gabriel, Halinan, Sap, Nguyen, Roesner, **Choi**, Choi ACL 22]

Why Do LLMs Need Clarification?

Humans interpret questions in rich contexts

- Who wrote it? Why they wrote it?
- When and where was it written?

Misinfo Reaction Frames: Reasoning about Readers' Reactions to News Headlines
[Gabriel, Halinan, Sap, Nguyen, Roesner, **Choi**, Choi ACL 22]

Question	+	**Geographical Context (Location)**	→	**Answer**
When did we become a member of the United Nations?		**(Zimbabwe)**		1980
		(Ireland)		1955
		(US)		1945

Question	+	**Temporal Context (Timestamp)**	→	**Answer**
Where were the last Winter Olympic Games held?		**2018**		Pyeongchang, South Korea
		2022		Beijing, China

SituatedQA: Incorporating extra-linguistic contexts into QA
[Zhang and **Choi**, EMNLP 22]

Why Do LLMs Need Clarification?

- Establish common grounds between the participants: crucial for preventing misunderstanding

Why Do LLMs Need Clarification?

- Establish common grounds between the participants: crucial for preventing misunderstanding
- Large proportion of questions (up to 50% in Natural Questions) are ambiguous.

Why Do LLMs Need Clarification?

- Establish common grounds between the participants: crucial for preventing misunderstanding
- Large proportion of questions (up to 50% in Natural Questions) are ambiguous.
- The frequency of clarifying question in LM responses…

Why Do LLMs Need Clarification?

- Establish common grounds between the participants: crucial for preventing misunderstanding
- Large proportion of questions (up to 50% in Natural Questions) are ambiguous.
- The frequency of clarifying question in LM responses…
  - LMSYS-Chat-1M data:

0.01 %

0.05%

Why Do LLMs Need Clarification?

- Establish common grounds between the participants: crucial for preventing misunderstanding
- Large proportion of questions (up to 50% in Natural Questions) are ambiguous.
- The frequency of clarifying question in LM responses…
  - LMSYS-Chat-1M data:

    0.01 %

    0.05%
  - On domain-specific dialogues (education, etc):

    3%

    0.04%

[Grounding Gaps in Natural Language Generations [Shaikh, Gilgoric et al, EMNLP24]]

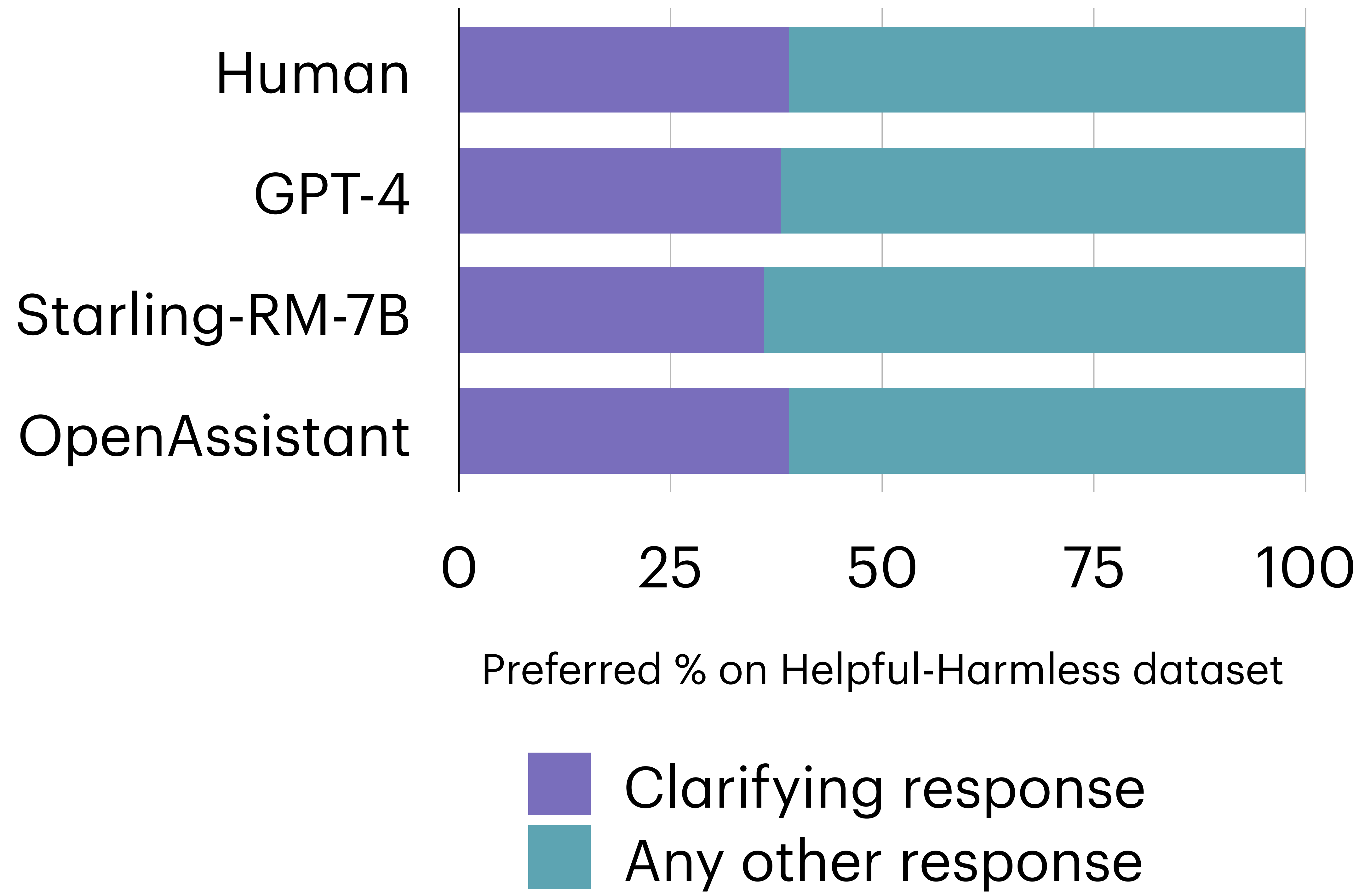
Why Do LLMs Not Ask Clarifying Questions?

Why Do LLMs Not Ask Clarifying Questions?

- Instruction tuning trains LLMs to directly act on commands for the end-users.

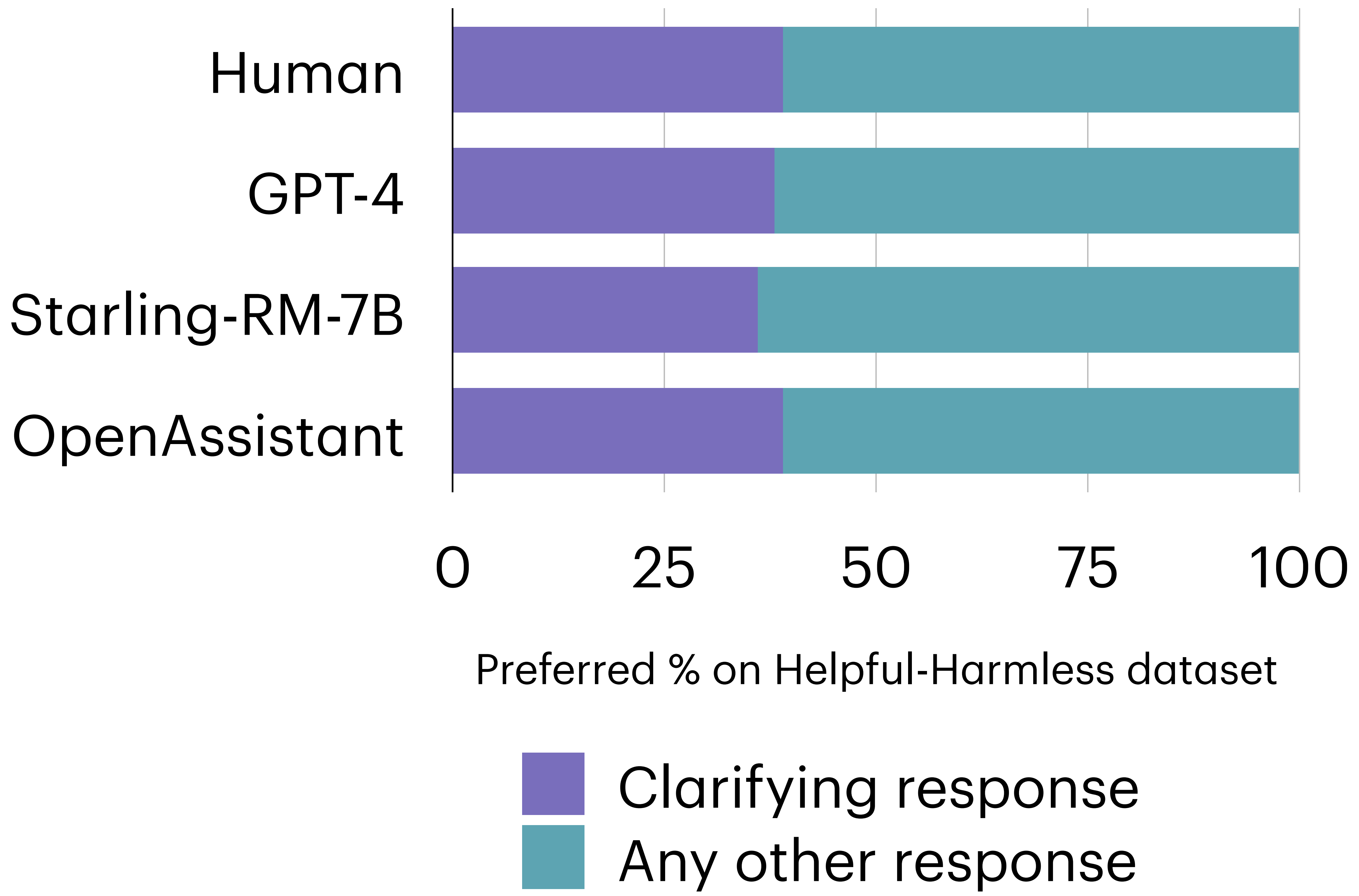
Why Do LLMs Not Ask Clarifying Questions?

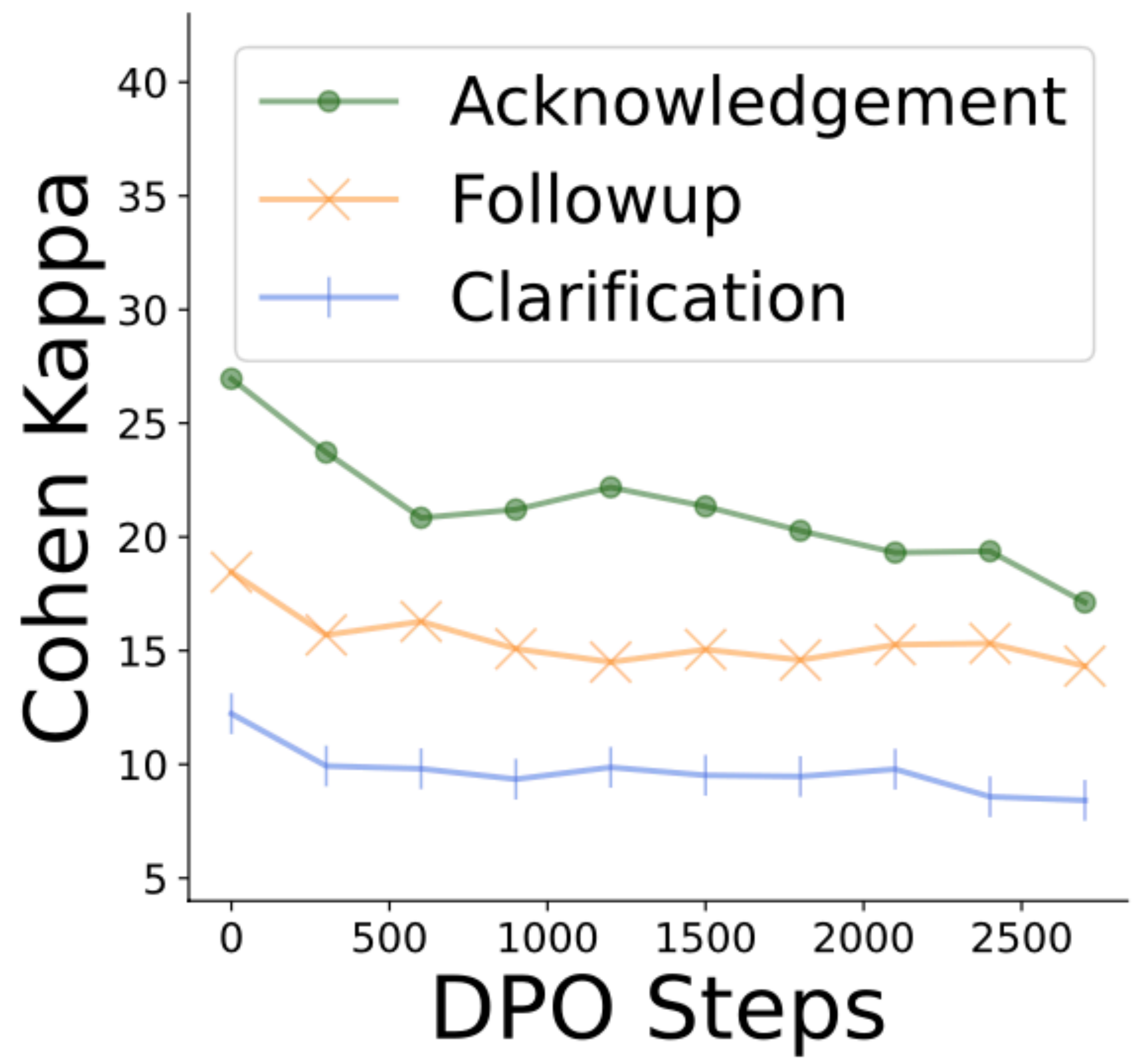
- Instruction tuning trains LLMs to directly act on commands for the end-users.
- Neither humans nor reward models prefer clarifying responses.



Why Do LLMs Not Ask Clarifying Questions?

- Instruction tuning trains LLMs to directly act on commands for the end-users.
- Neither humans nor reward models prefer clarifying responses.





[Grounding Gaps in Natural Language Generations [Shaikh, Gilgoric et al, EMNLP24]

Problem Setting: QA with Multiple Annotators

[AmbigQA: Answering Ambiguous Open-domain Questions. Min et al, EMNLP 20]

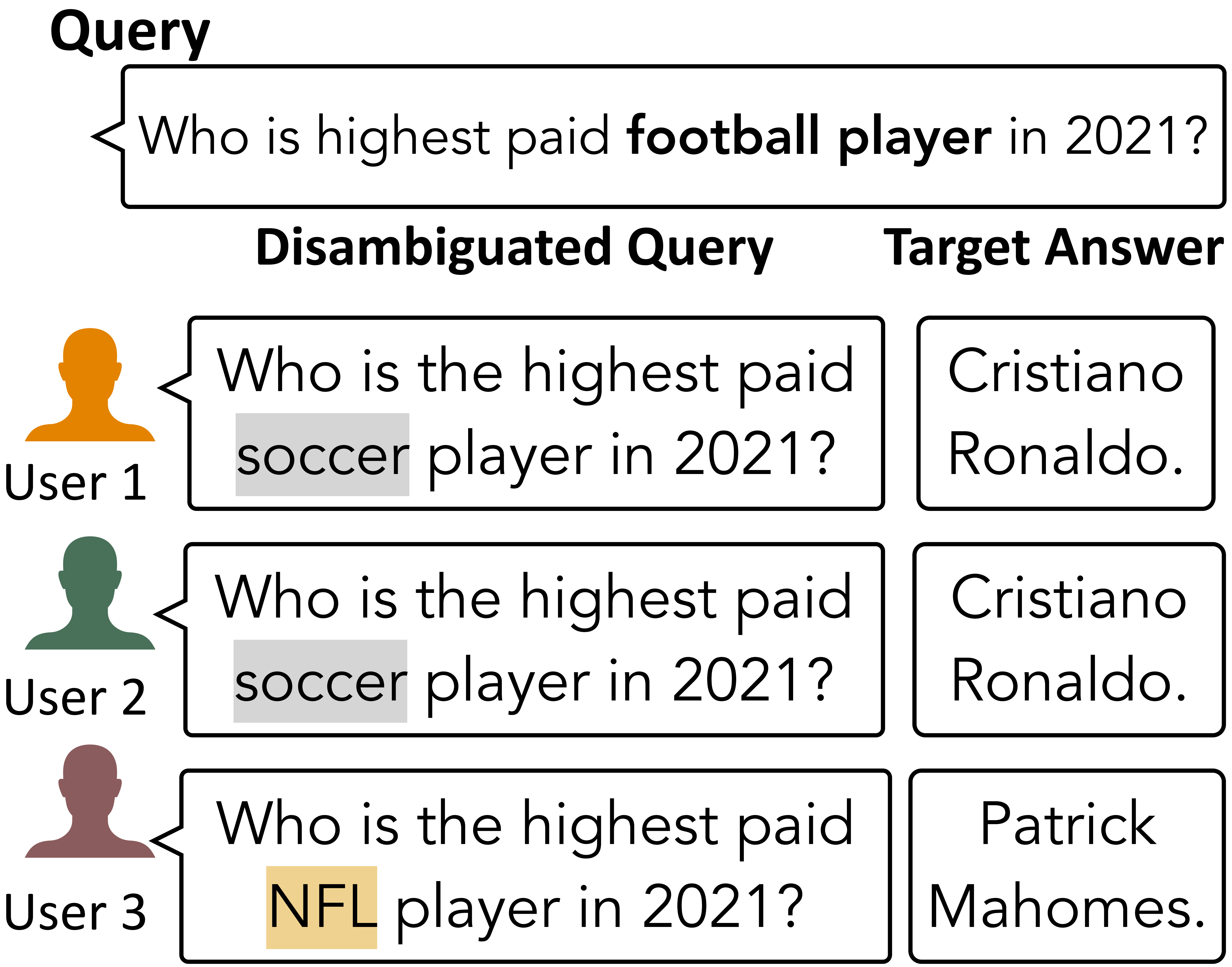
Problem Setting: QA with Multiple Annotators

- **Goal**: given a query, interact with users (if necessary) to provide the target answer to each user.

[AmbigQA: Answering Ambiguous Open-domain Questions. Min et al, EMNLP 20]

Problem Setting: QA with Multiple Annotators

- **Goal**: given a query, interact with users (if necessary) to provide the target answer to each user.



[AmbigQA: Answering Ambiguous Open-domain Questions. Min et al, EMNLP 20]

Problem Setting: QA with Multiple Annotators

- **Goal**: given a query, interact with users (if necessary) to provide the target answer to each user.

Query

Who is highest paid **football player** in 2021?

	Disambiguated Query	Target Answer
User 1	Who is the highest paid soccer player in 2021?	Cristiano Ronaldo.
User 2	Who is the highest paid soccer player in 2021?	Cristiano Ronaldo.
User 3	Who is the highest paid NFL player in 2021?	Patrick Mahomes.

Query

Who wrote how far i'll go in moana?

	Disambiguated Query	Target Answer
User 1	Who wrote how far i'll go in moana?	Lin-Manuel Miranda
User 2	Who wrote how far i'll go in moana?	Lin-Manuel Miranda

[AmbigQA: Answering Ambiguous Open-domain Questions. Min et al, EMNLP 20]

Scoring Based on the Current State

user query : x_i

Who is the highest paid football player in 2021?

Scoring Based on the Current State

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Scoring Based on the Current State

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Cristiano Ronaldo.

Patrick Mahomes.

Do you mean football or soccer?

The highest-paid soccer player in 2021 was Cristiano Ronaldo. The highest-paid NFL player is Patrick Mahomes.

Scoring Based on the Current State

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Cristiano Ronaldo.

Patrick Mahomes.

Do you mean football or soccer?

The highest-paid soccer player in 2021 was Cristiano Ronaldo. The highest-paid NFL player is Patrick Mahomes.

Scoring Based on the Current State

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Cristiano Ronaldo.

Patrick Mahomes.

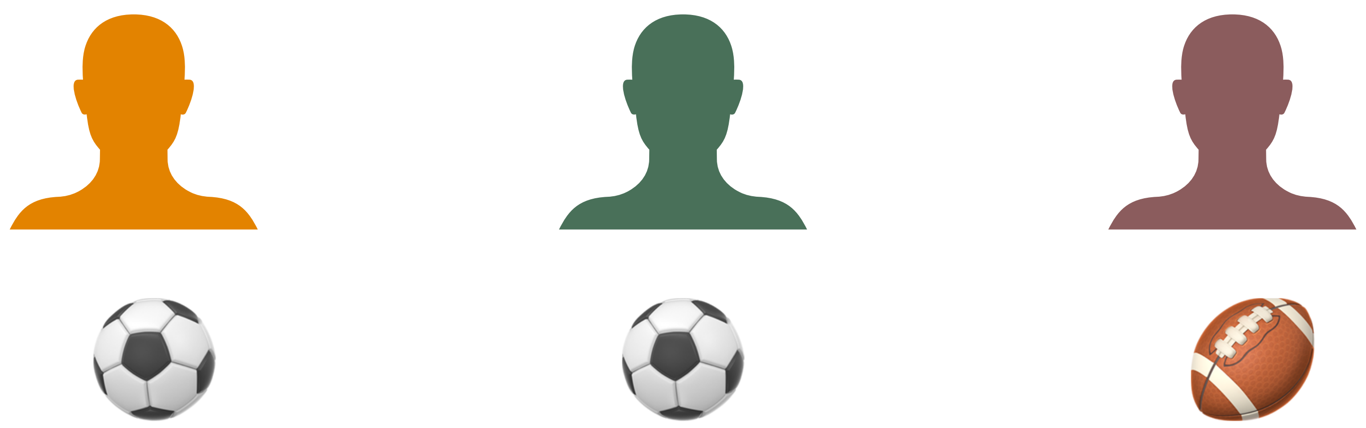
Do you mean football or soccer?

The highest-paid soccer player in 2021 was Cristiano Ronaldo. The highest-paid NFL player is Patrick Mahomes.

Scoring Based on the Current State

user query : x_i

> Who is the highest paid football player in 2021?



LLM initial response : y_i

Cristiano Ronaldo.	Patrick Mahomes.	Do you mean football or soccer?	The highest-paid soccer player in 2021 was Cristiano Ronaldo. The highest-paid NFL player is Patrick Mahomes.

- **In this work, we consider only short answers & clarifying questions. Long-form answers convey rich information but still challenging to evaluate.**

Our Proposal: Scoring Based on the Future Turns

$\text{user query} : x_i$

Who is the highest paid football player in 2021?

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Do you mean football or soccer?

Our Proposal: Scoring Based on the Future Turns

$\text{user query}: x_i$ Who is the highest paid football player in 2021?

$\text{LLM initial response}: y_i$ Do you mean football or soccer?

$\text{simulated user turn}: x_{i+1}$

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Do you mean football or soccer?

simulated user turn : x_{i+1}

Soccer

Soccer

Football

Our Proposal: Scoring Based on the Future Turns

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Do you mean football or soccer?

simulated user turn : x_{i+1}

Soccer

Soccer

Football

LLM next response : y_{i+1}

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Do you mean football or soccer?

simulated user turn : x_{i+1}

Soccer	Soccer	Football

LLM next response : y_{i+1}

Cristiano Ronaldo.	Cristiano Ronaldo.	Patrick Mahomes.

Our Proposal: Scoring Based on the Future Turns

user query : x_i

Who is the highest paid football player in 2021?

LLM initial response : y_i

Do you mean football or soccer?

simulated user turn : x_{i+1}

Soccer

Soccer

Football

LLM next response : y_{i+1}

Cristiano Ronaldo. ✓

Cristiano Ronaldo. ✓

Patrick Mahomes. ✓

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Do you mean football or soccer?

simulated user turn : x_{i+1}

Soccer	Soccer	Football

LLM next response : y_{i+1}

Cristiano Ronaldo.	Cristiano Ronaldo.	Patrick Mahomes.

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Just salary or include bonus and others?

simulated user turn : x_{i+1}

Salary	Salary	Salary

LLM next response : y_{i+1}

Cristiano Ronaldo.	Cristiano Ronaldo.	Cristian Ronaldo.

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Just salary or include bonus and others?

simulated user turn : x_{i+1}

Salary | Salary | Salary

LLM next response : y_{i+1}

Cristiano Ronaldo. | Cristiano Ronaldo. | Cristian Ronaldo.

Our Proposal: Scoring Based on the Future Turns

user query : x_i — Who is the highest paid football player in 2021?

LLM initial response : y_i — Just salary or include bonus and others? ✗

simulated user turn : x_{i+1}

⚽ Salary

⚽ Salary

🏈 Salary

LLM next response : y_{i+1}

Cristiano Ronaldo. ✓

Cristiano Ronaldo. ✓

Cristian Ronaldo.

Current State vs. Simulated Future State

Based on current state

Cristiano Ronaldo. ⚽ > Patrick Mahomes🏈 > Do you mean football or soccer? | Just salary or bonus and others?

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Current State vs. Simulated Future State

Based on current state

Cristiano Ronaldo. ⚽ > Patrick Mahomes🏈 > Do you mean football or soccer? | Just salary or bonus and others?

Based on simulated future

Do you mean football or soccer? > Cristiano Ronaldo. ⚽ | Just salary or bonus and others? > Patrick Mahomes. 🏈

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Current State vs. Simulated Future State

Based on current state

Cristiano Ronaldo. ⚽ > Patrick Mahomes🏈 > Do you mean football or soccer? Just salary or bonus and others?

Break ties by preferring dialogue with a fewer interaction

Based on simulated future

Do you mean football or soccer? > Cristiano Ronaldo. ⚽ > Just salary or bonus and others? > Patrick Mahomes. 🏈

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Current State vs. Simulated Future State

Based on current state

Crist
Rona

s?

Change how preference labels are assigned, and use existing preference learning algorithms

Based

Do you
football or soccer?

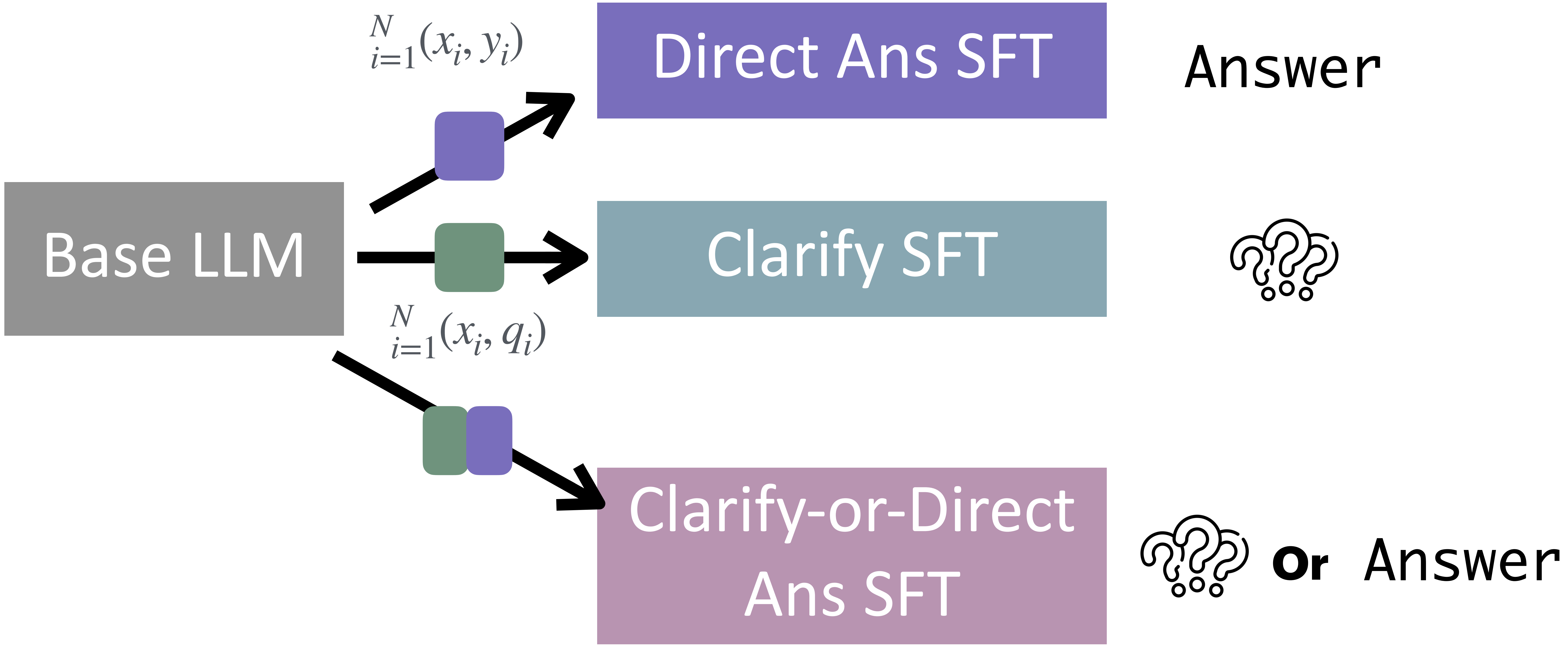
Ronaldo. ⚽

bonus and others?

Mahomes. 🏈

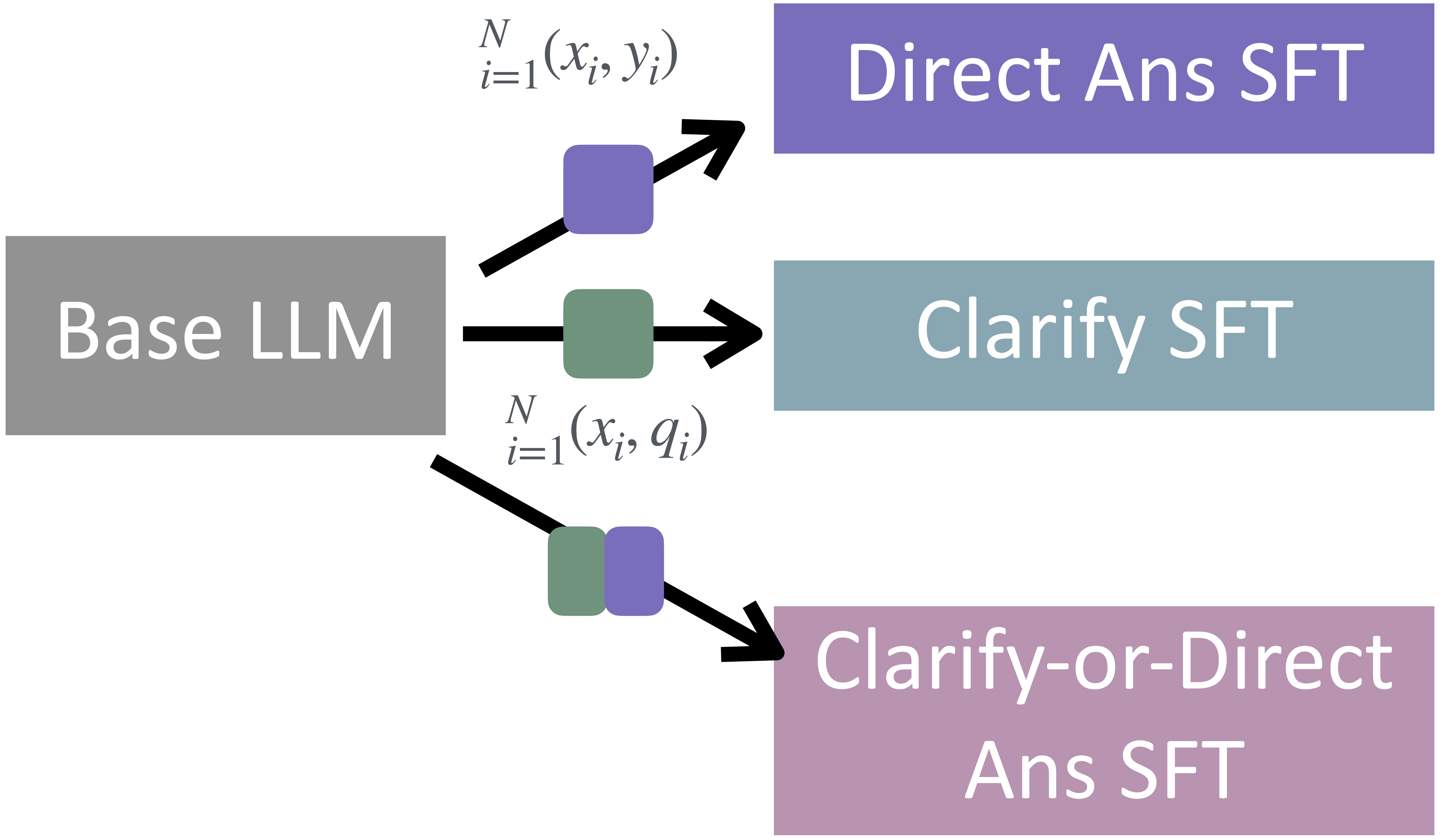
[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Comparison Systems



Training dataset: training portion of Natural Questions

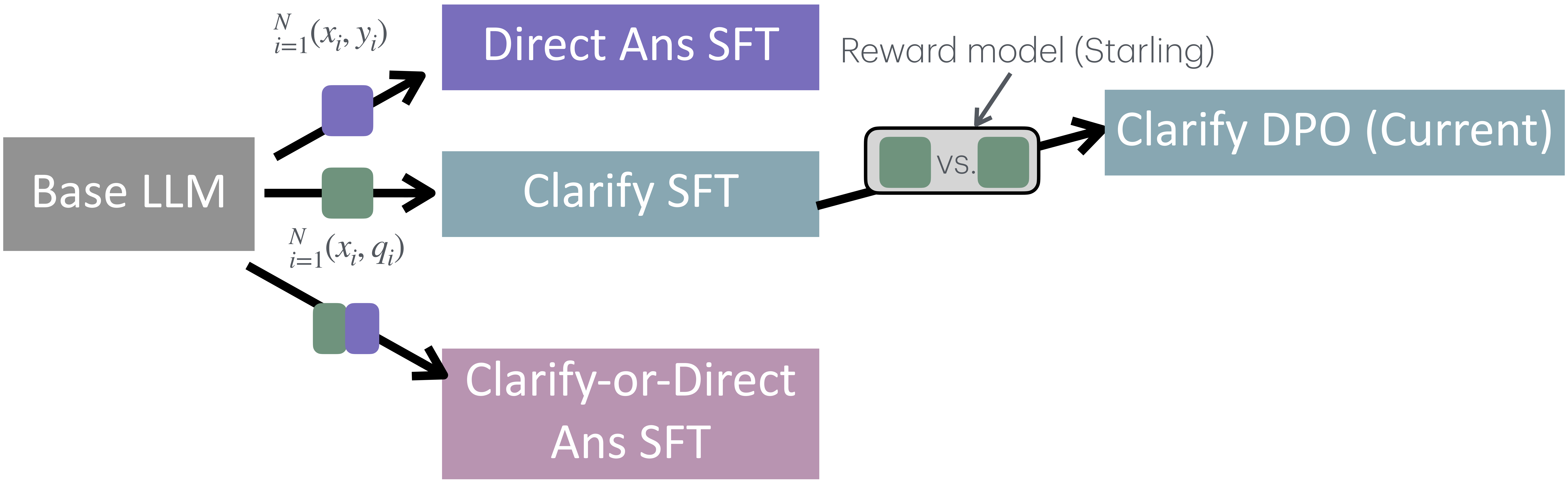
Comparison Systems



Training dataset: training portion of Natural Questions

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

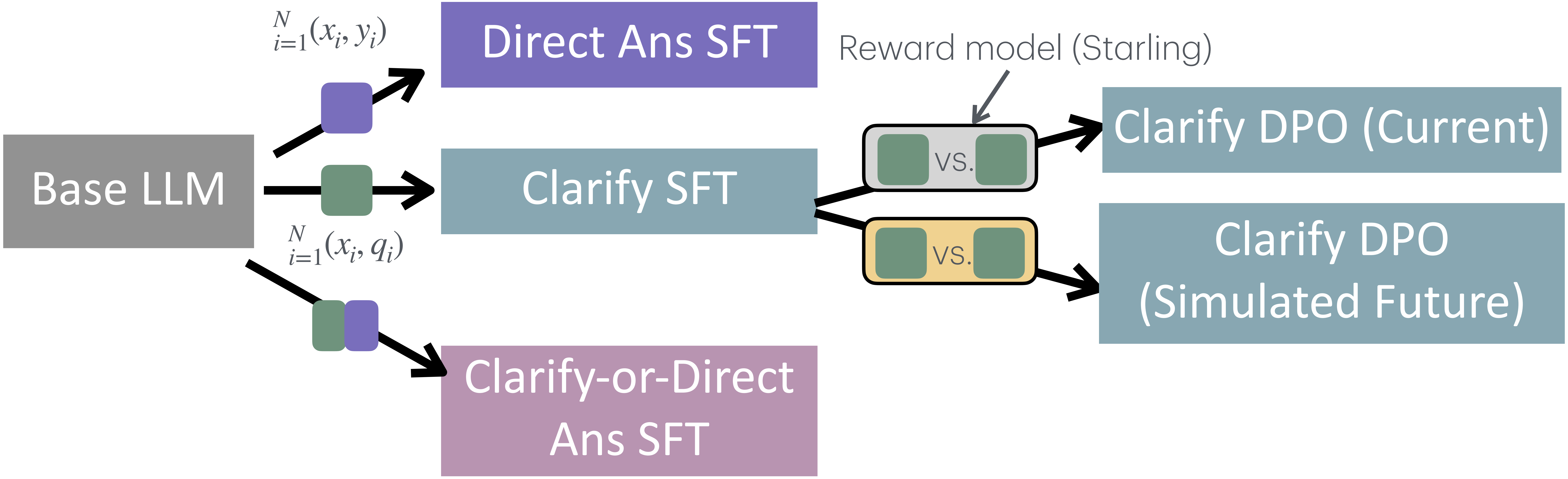
Comparison Systems



Training dataset: training portion of Natural Questions

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

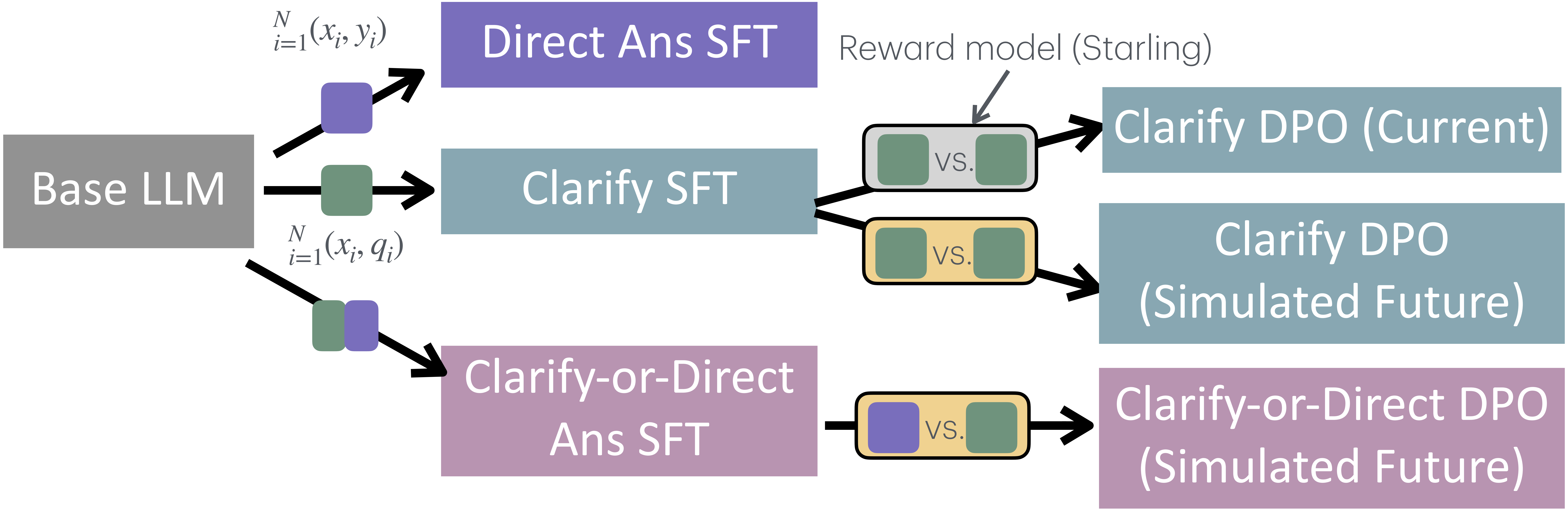
Comparison Systems



Training dataset: training portion of Natural Questions

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Comparison Systems



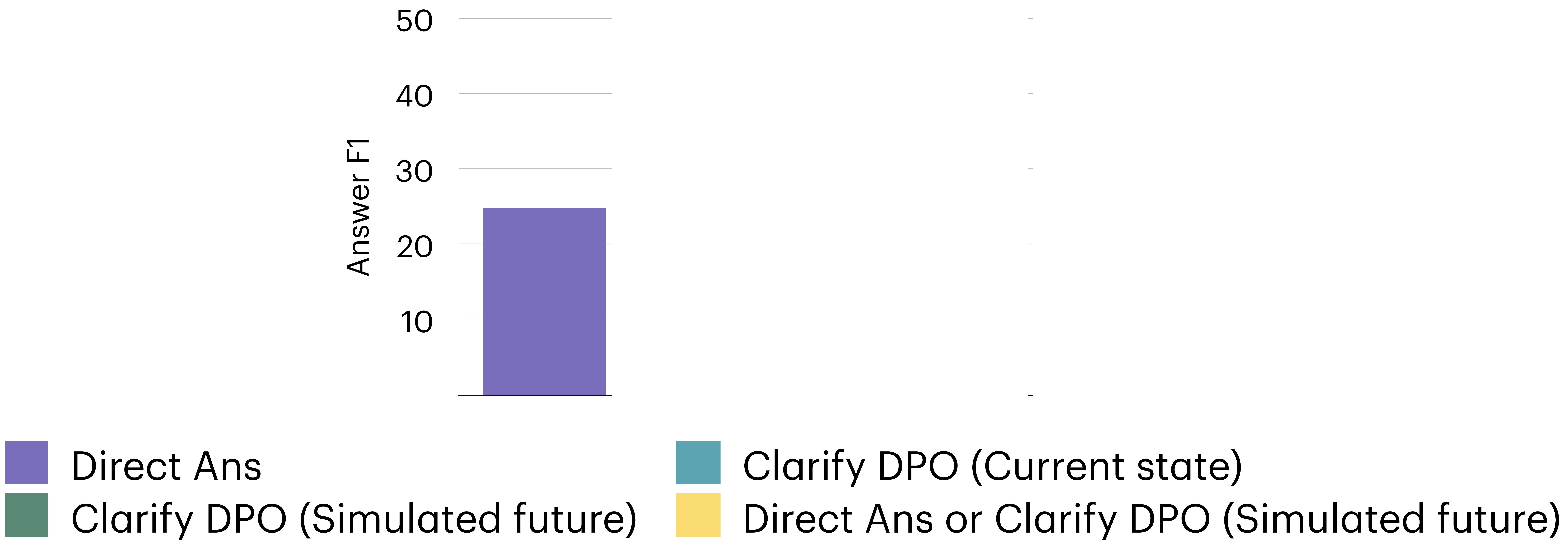
Training dataset: training portion of Natural Questions

[Direct Preference Optimization: Your Language Model is Secretly a Reward Model, Rafailov, Sharma, Mitchell, Ermon, Manning, Finn NeurIPS 23]

Efficiency Evaluation: # of Conversation Turns

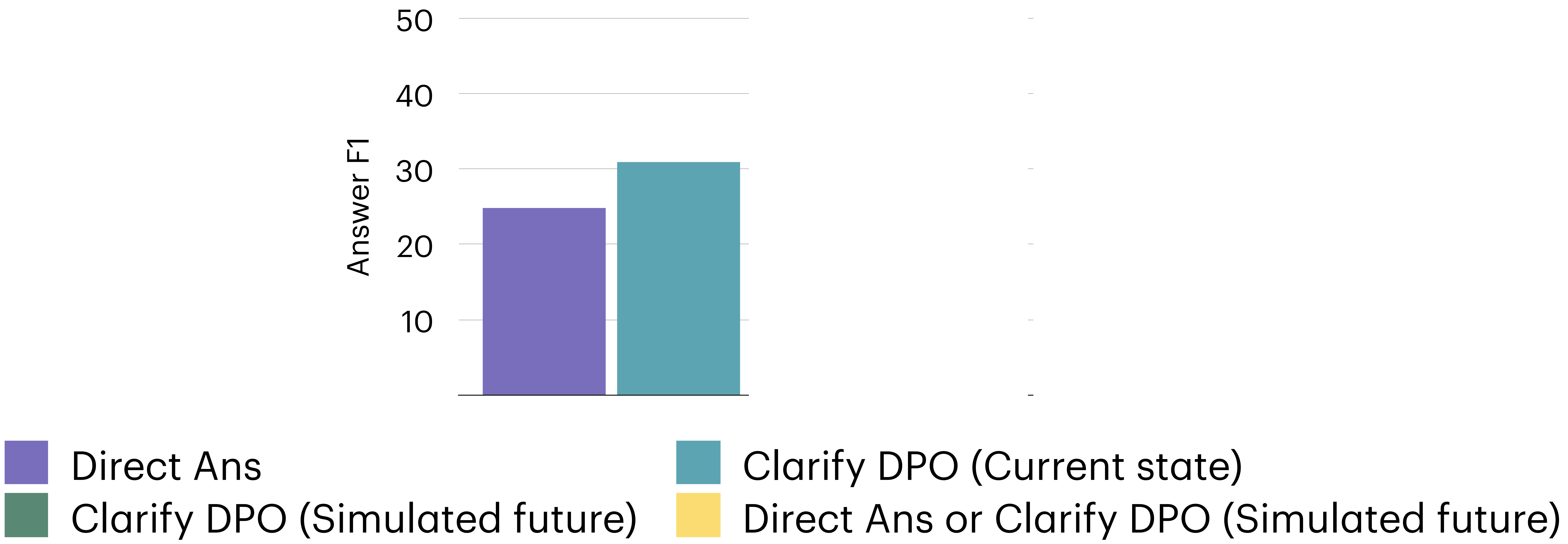
Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



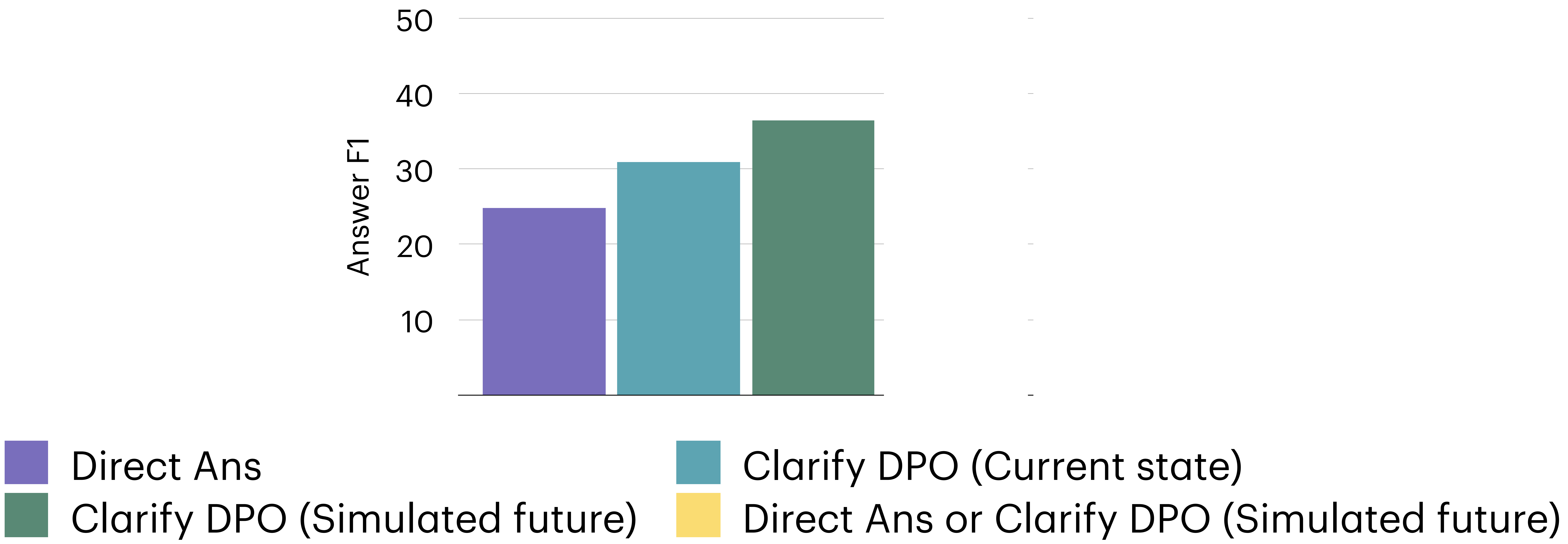
Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



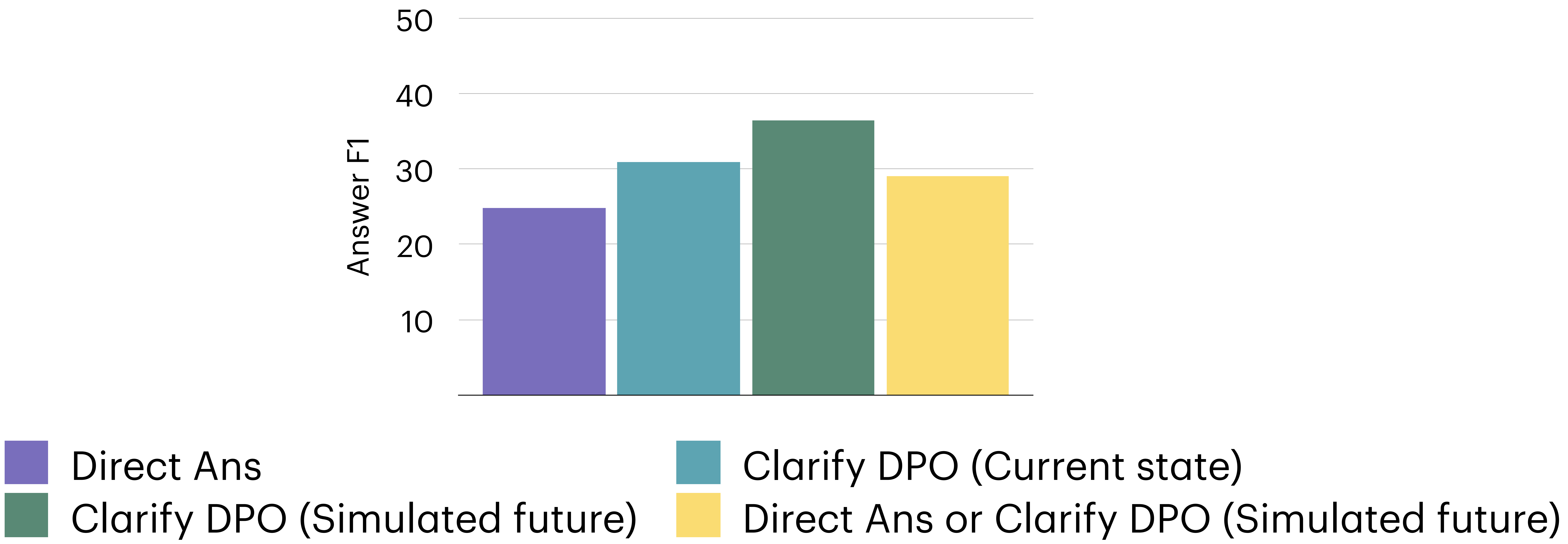
Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



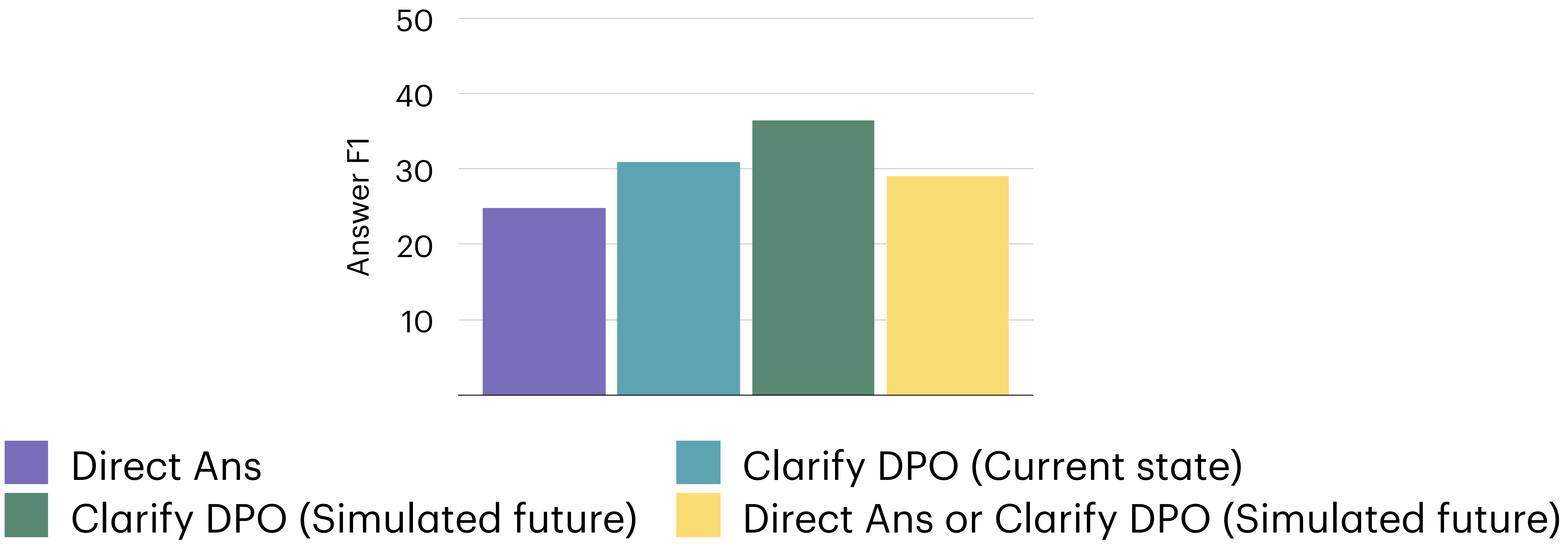
Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



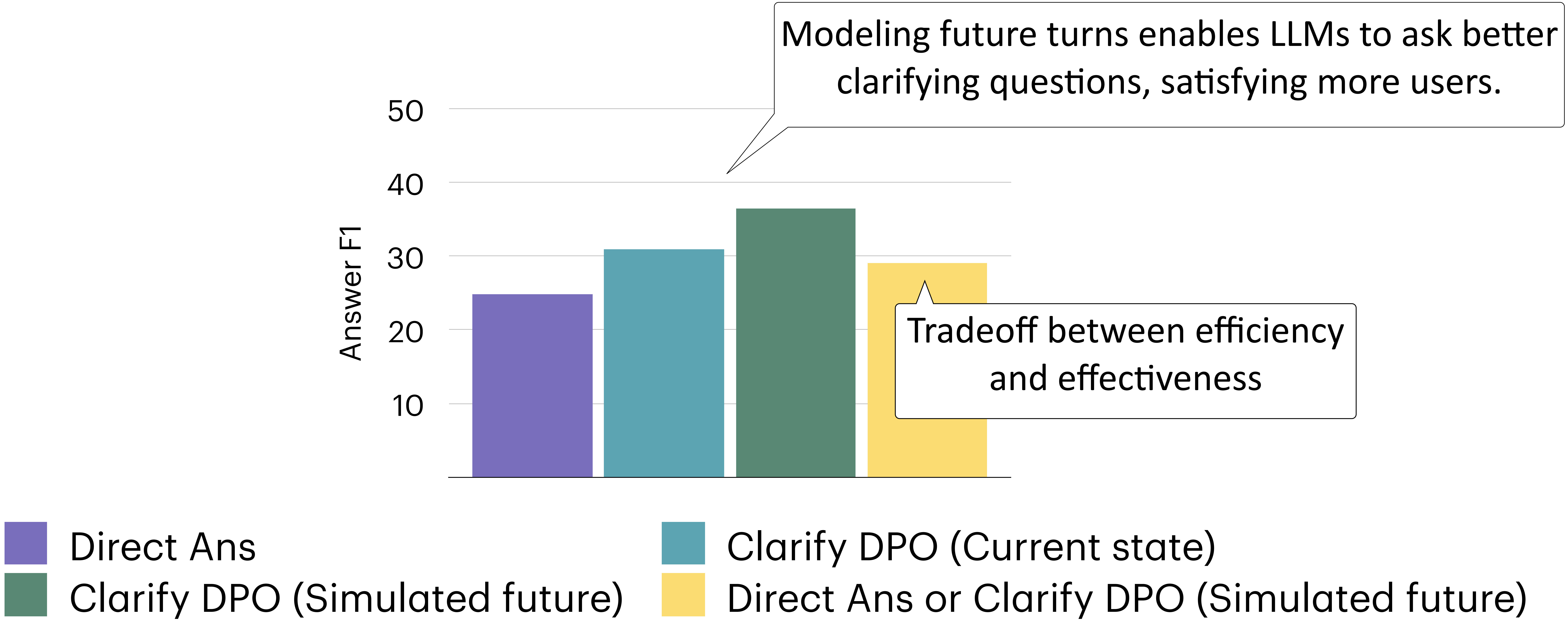
Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



Llama3 8B model, Evaluation on Ambig QA (test split)

Can LLM recover target answer for diverse users?



Llama3 8B model, Evaluation on Ambig QA (test split)

Summary

- Ability to ask questions is a crucial yet underdeveloped ability of LLMs.

Summary

- Ability to ask questions is a crucial yet underdeveloped ability of LLMs.
- Scoring responses considering only the present is limiting.

Summary

- Ability to ask questions is a crucial yet underdeveloped ability of LLMs.
- Scoring responses considering only the present is limiting.
- We propose preference learning with simulated future turns, teaching LLMs to ask clarifying questions to satisfy diverse users.

Summary

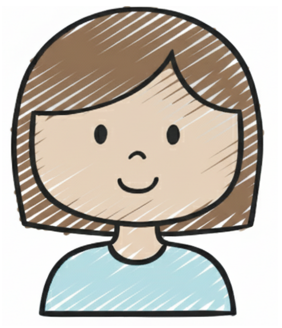
- Ability to ask questions is a crucial yet underdeveloped ability of LLMs.
- Scoring responses considering only the present is limiting.
- We propose preference learning with simulated future turns, teaching LLMs to ask clarifying questions to satisfy diverse users.
- How can we balance communication efficiency and effectiveness?

Ongoing Work: Code LLM assistant

- User input is underspecified, and model output is lengthy and complex

Ongoing Work: Code LLM assistant

- User input is underspecified, and model output is lengthy and complex
  - User-driven Feedback



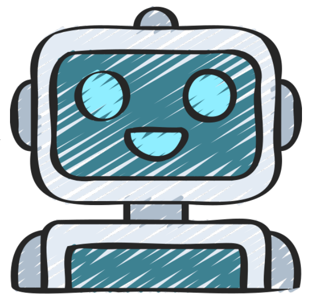
Could you generate code implementing the perceptron algorithm?

Certainly! Here's a simple implementation of the Perceptron algorithm in Python. The perceptron is a linear classifier used for binary classification tasks.

Perceptron Algorithm Implementation

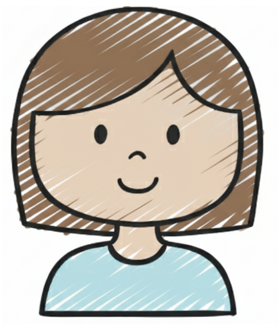
```
python
import numpy as np

class Perceptron:
    def __init__(self, input_dim, learning_rate=0.01, epochs=1000):
        self.learning_rate = learning_rate
        self.epochs = epochs
```



Ongoing Work: Code LLM assistant

- User input is underspecified, and model output is lengthy and complex
  - User-driven Feedback



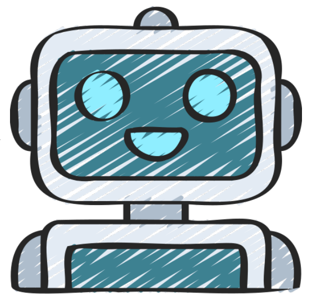
Could you generate code implementing the perceptron algorithm?

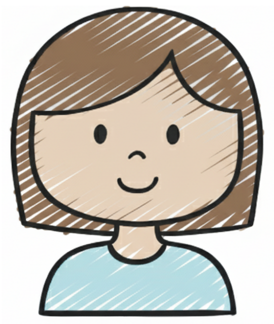
Certainly! Here's a simple implementation of the Perceptron algorithm in Python. The perceptron is a linear classifier used for binary classification tasks.

Perceptron Algorithm Implementation

```
python
import numpy as np

class Perceptron:
    def __init__(self, input_dim, learning_rate=0.01, epochs=1000):
        self.learning_rate = learning_rate
        self.epochs = epochs
```





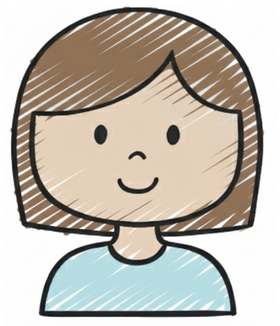
I want an implementation in R.

I want a multi class perceptron, so you need to modify 2nd block of your code

Ongoing Work: Code LLM assistant

- User input is underspecified, and model output is lengthy and complex

- User-driven Feedback



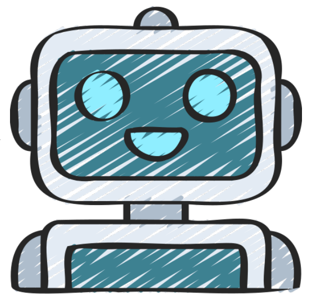
Could you generate code implementing the perceptron algorithm?

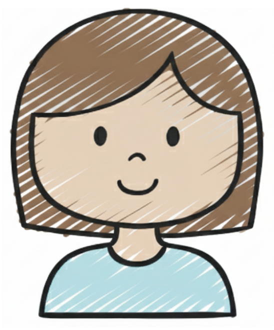
Certainly! Here's a simple implementation of the Perceptron algorithm in Python. The perceptron is a linear classifier used for binary classification tasks.

Perceptron Algorithm Implementation

```
python
import numpy as np

class Perceptron:
    def __init__(self, input_dim, learning_rate=0.01, epochs=1000):
        self.learning_rate = learning_rate
        self.epochs = epochs
```

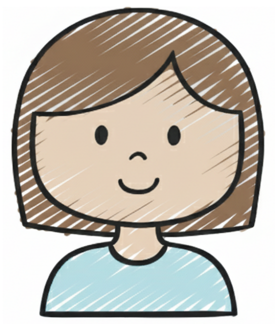




I want an implementation in R.

I want a multi class perceptron, so you need to modify 2nd block of your code

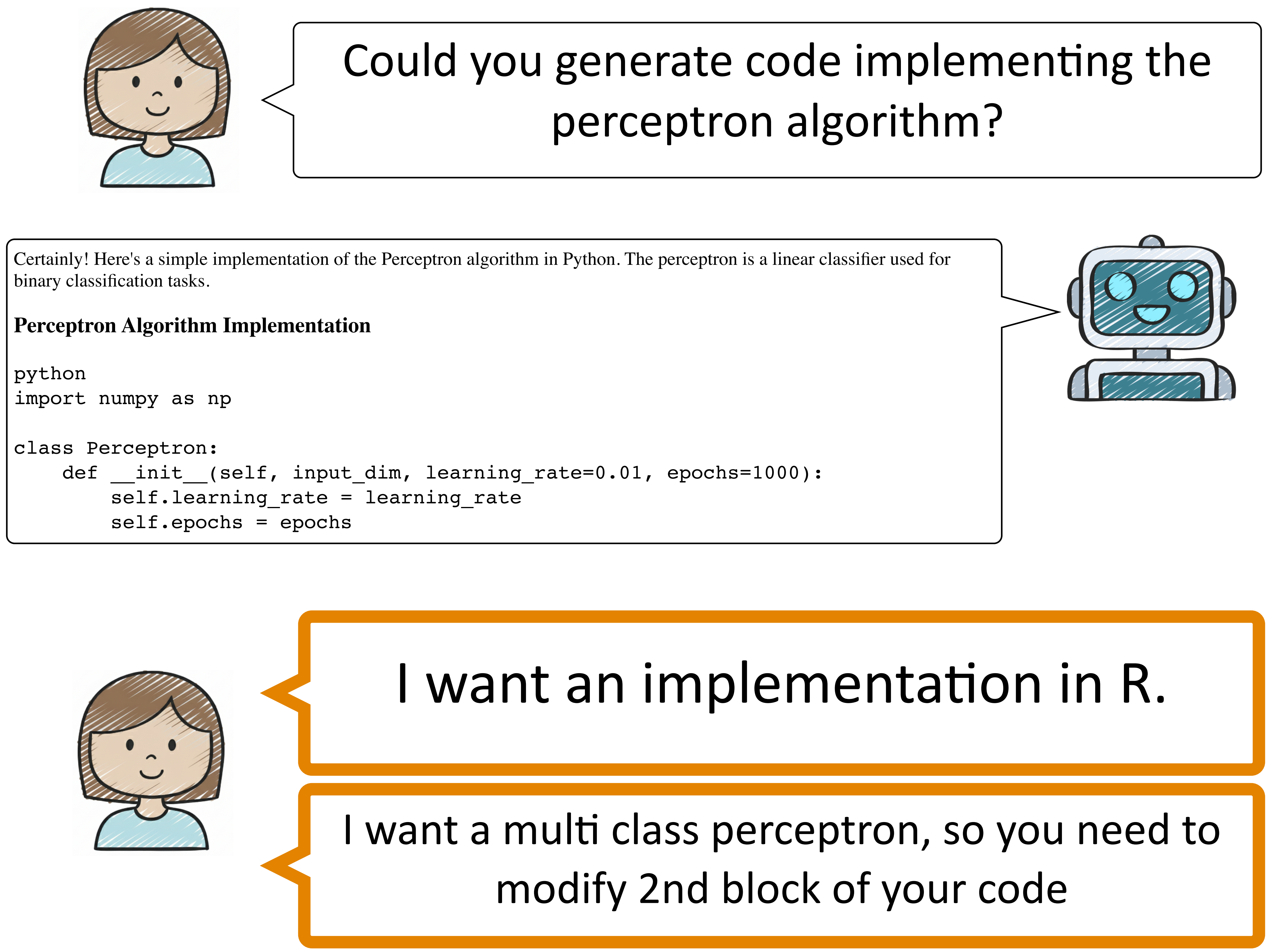
- Model-driven Feedback

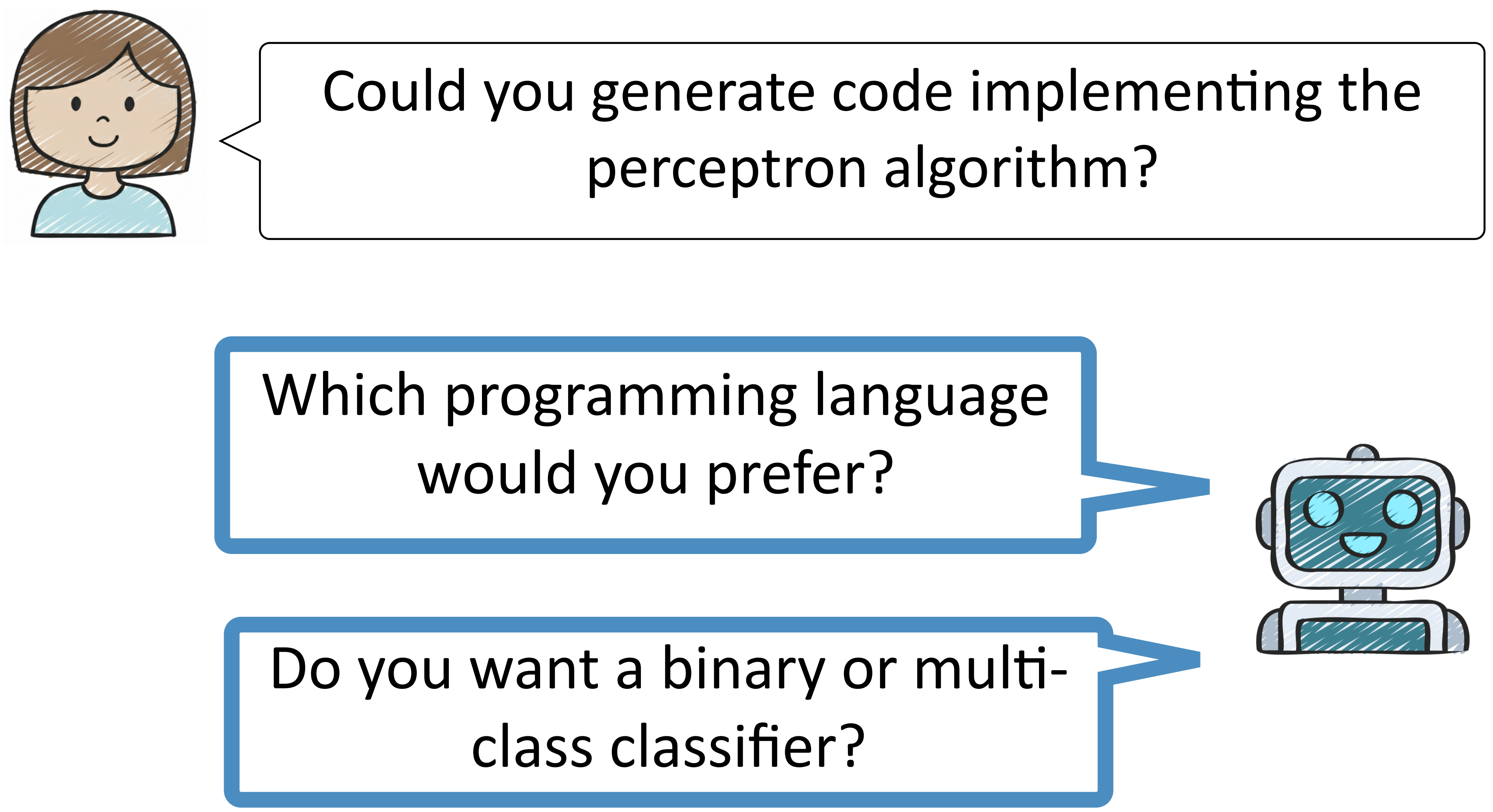


Could you generate code implementing the perceptron algorithm?

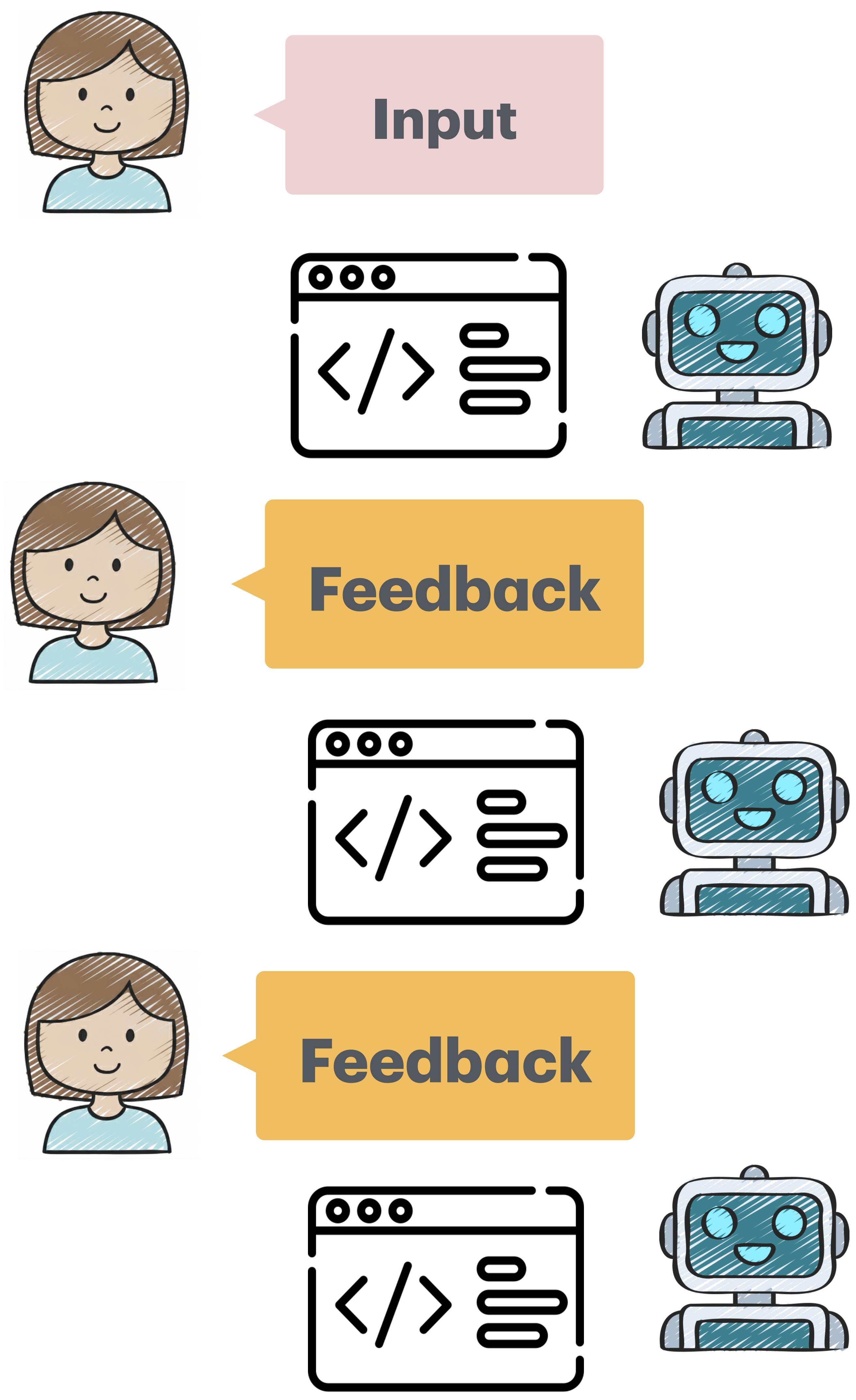
Ongoing Work: Code LLM assistant

- User input is underspecified, and model output is lengthy and complex
  - User-driven Feedback
  - Model-driven Feedback

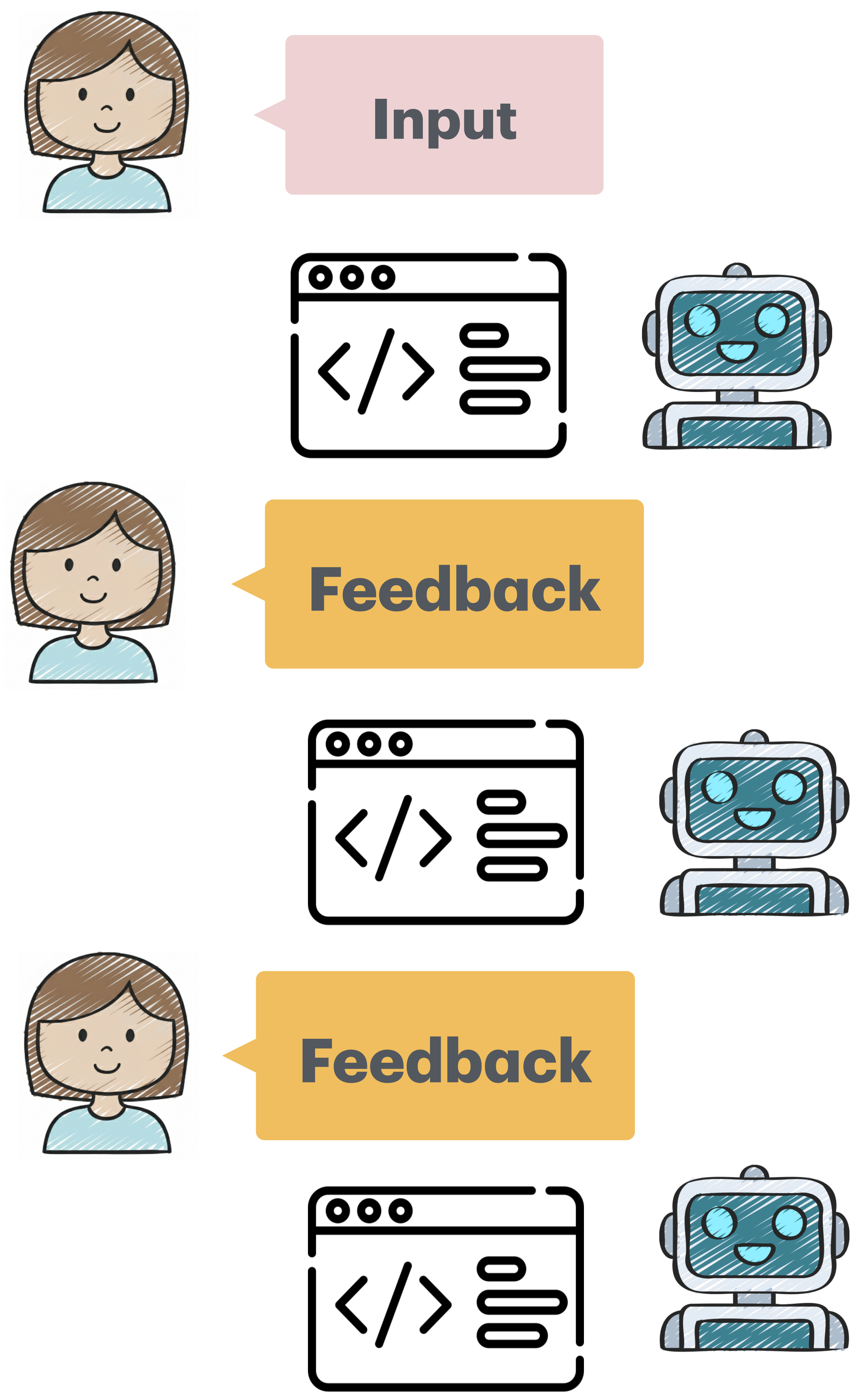




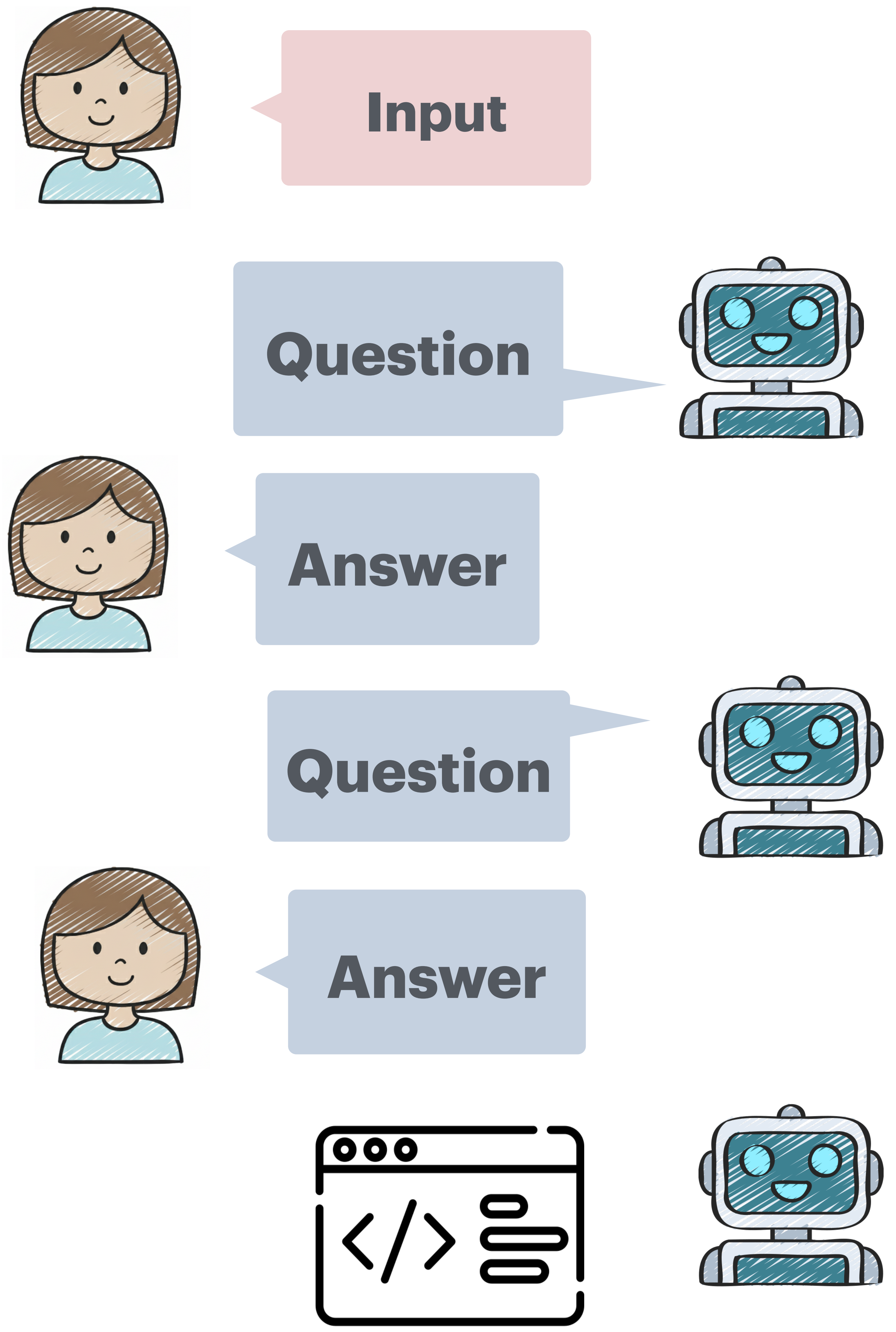
Always User-driven



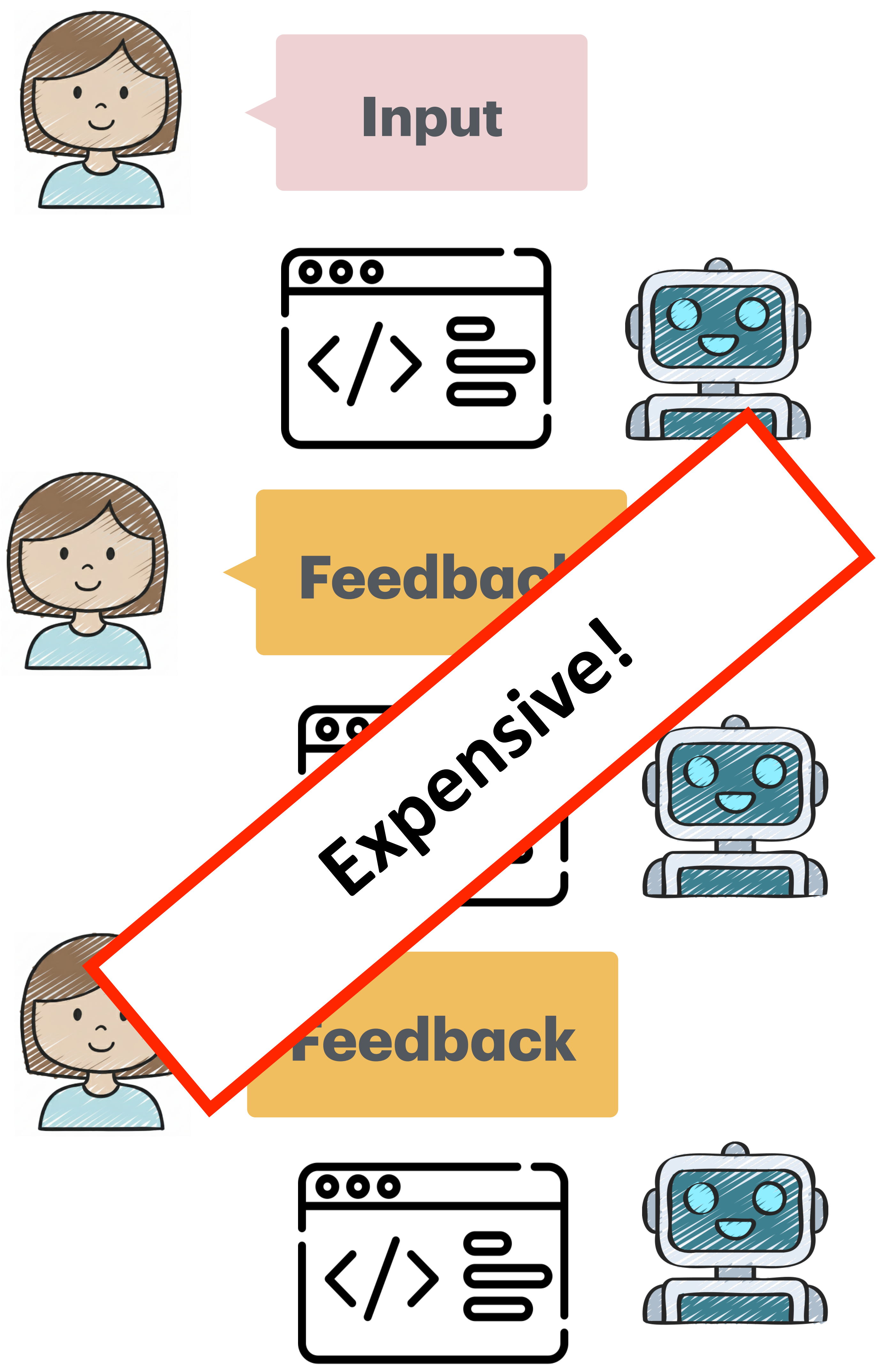
Always User-driven



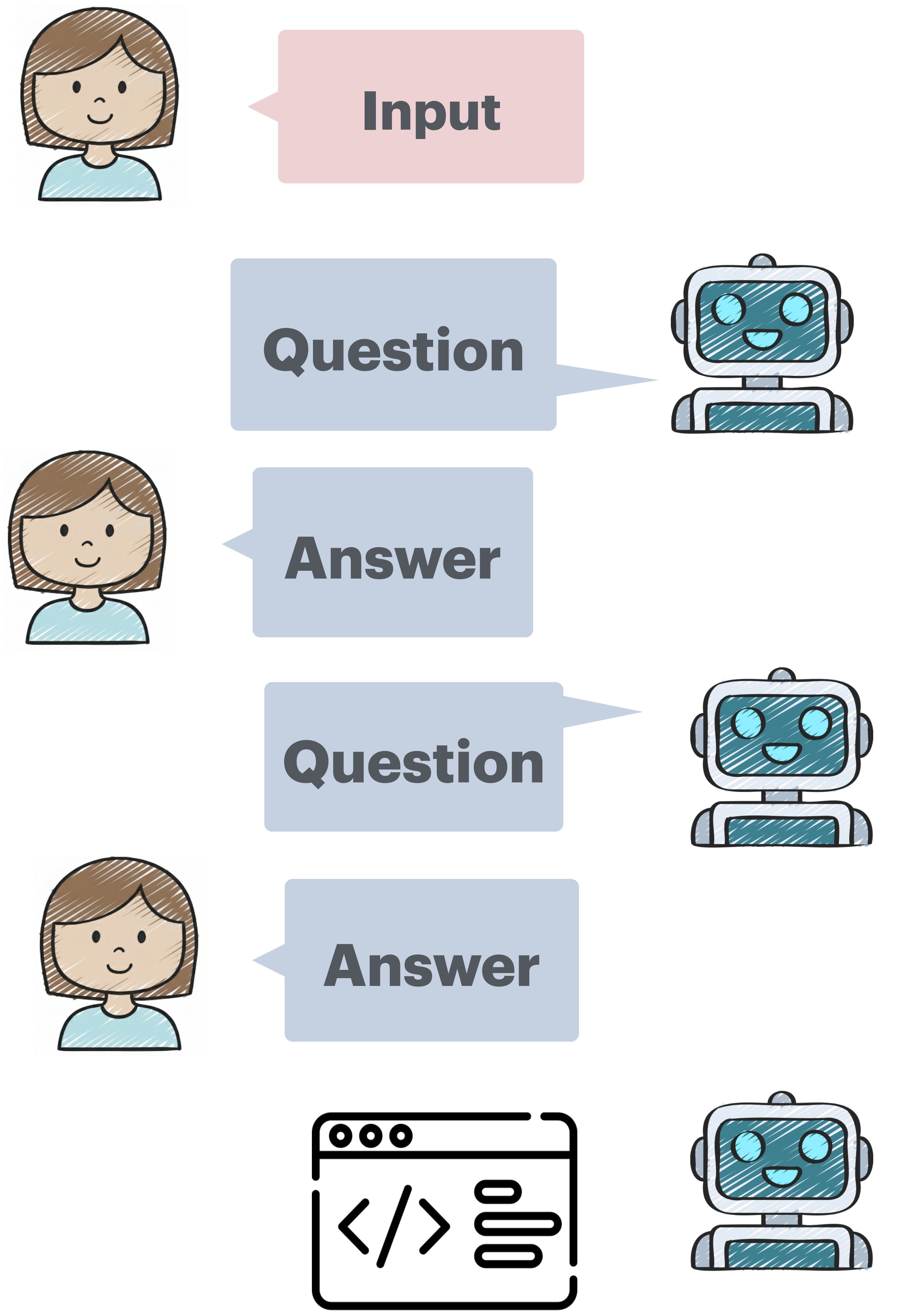
Always Model-driven



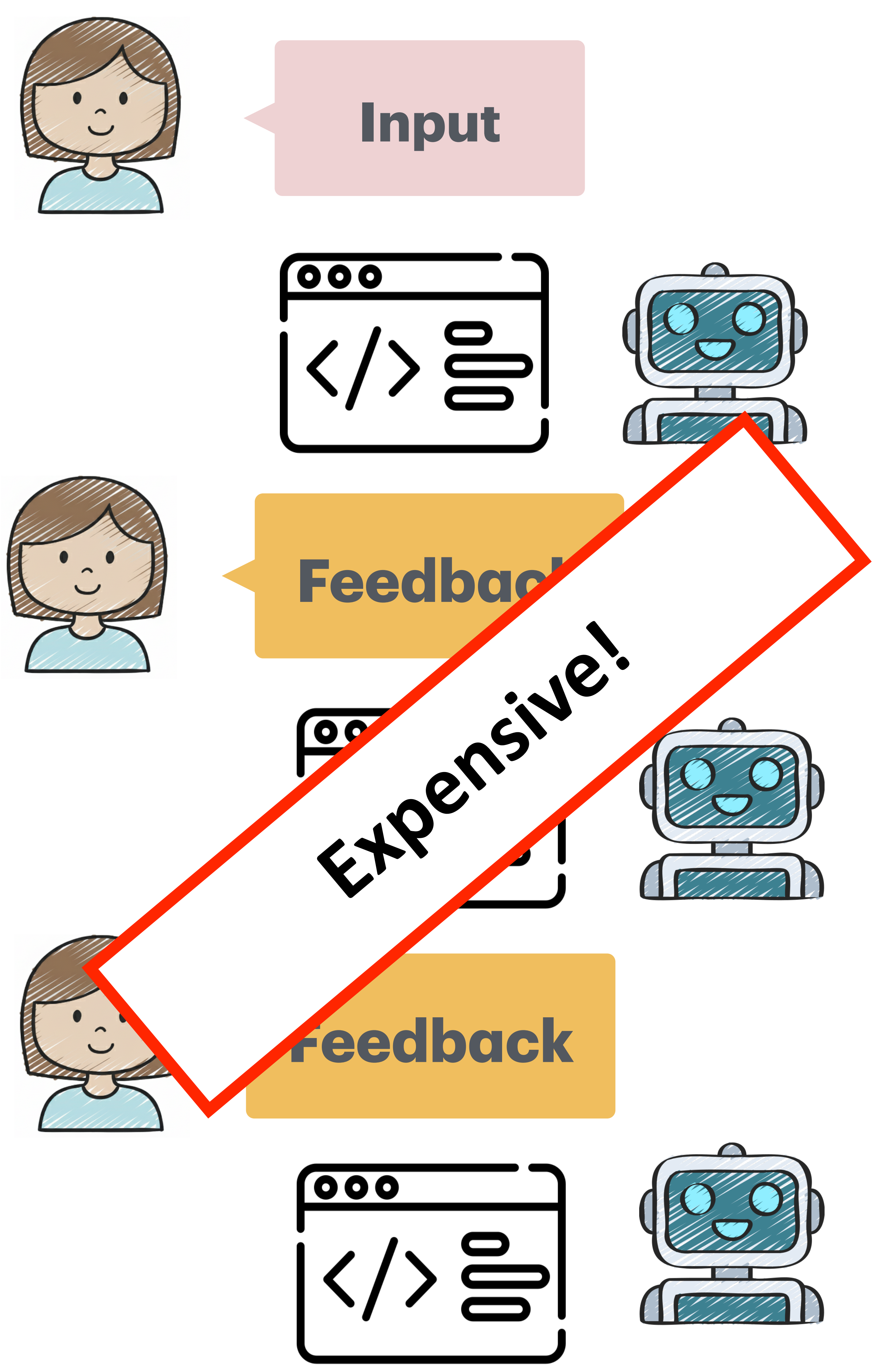
Always User-driven



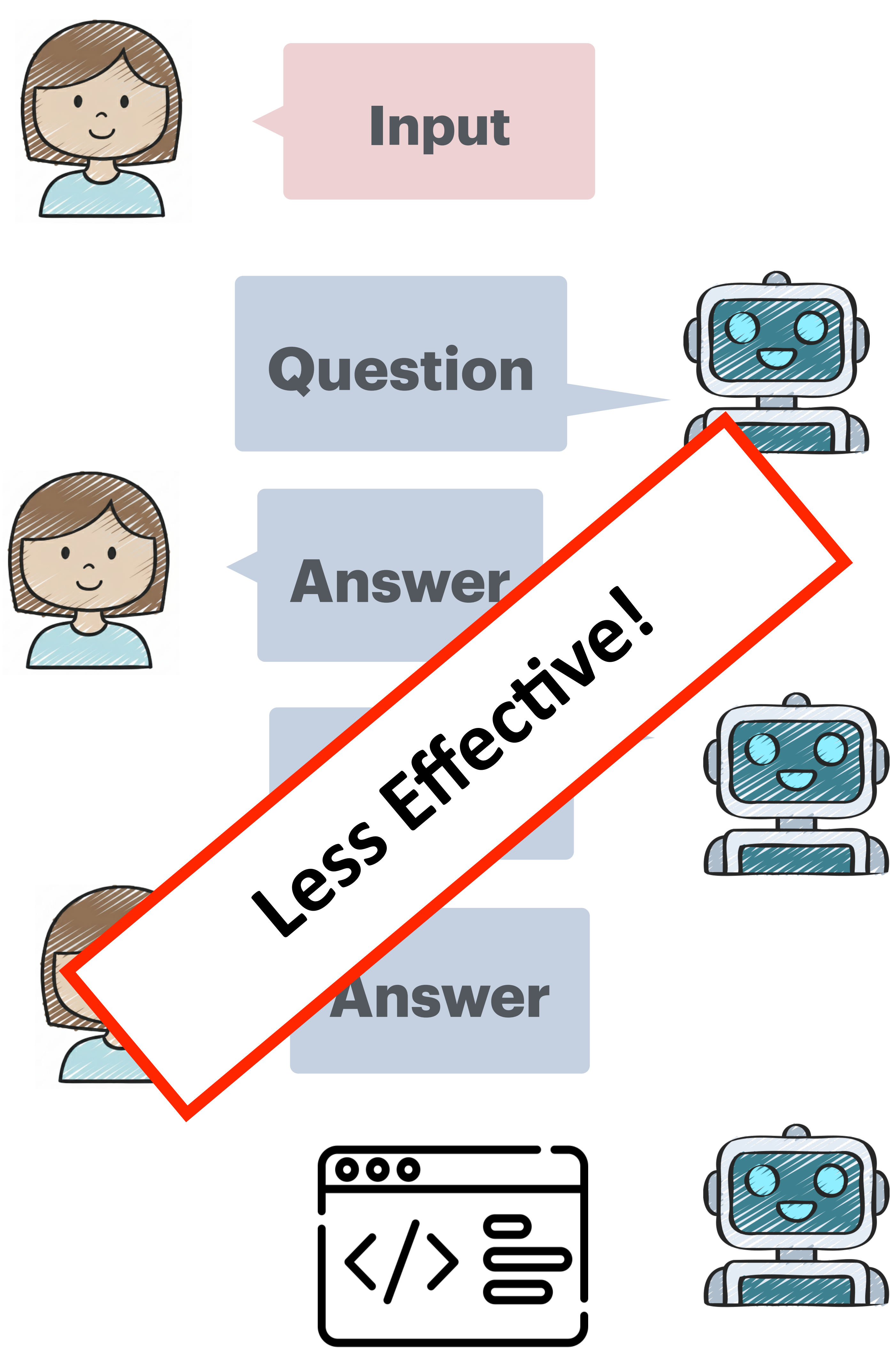
Always Model-driven



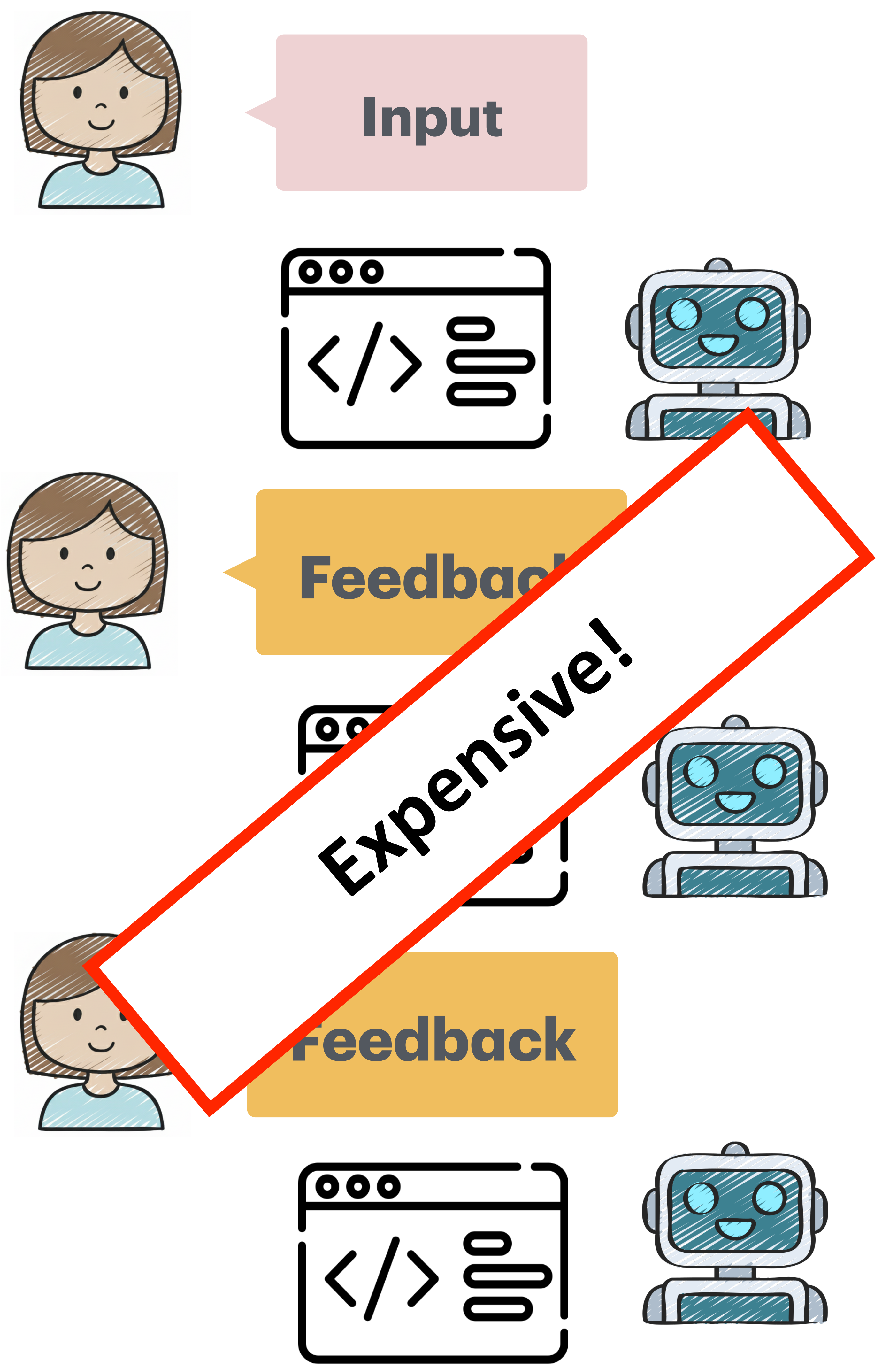
Always User-driven



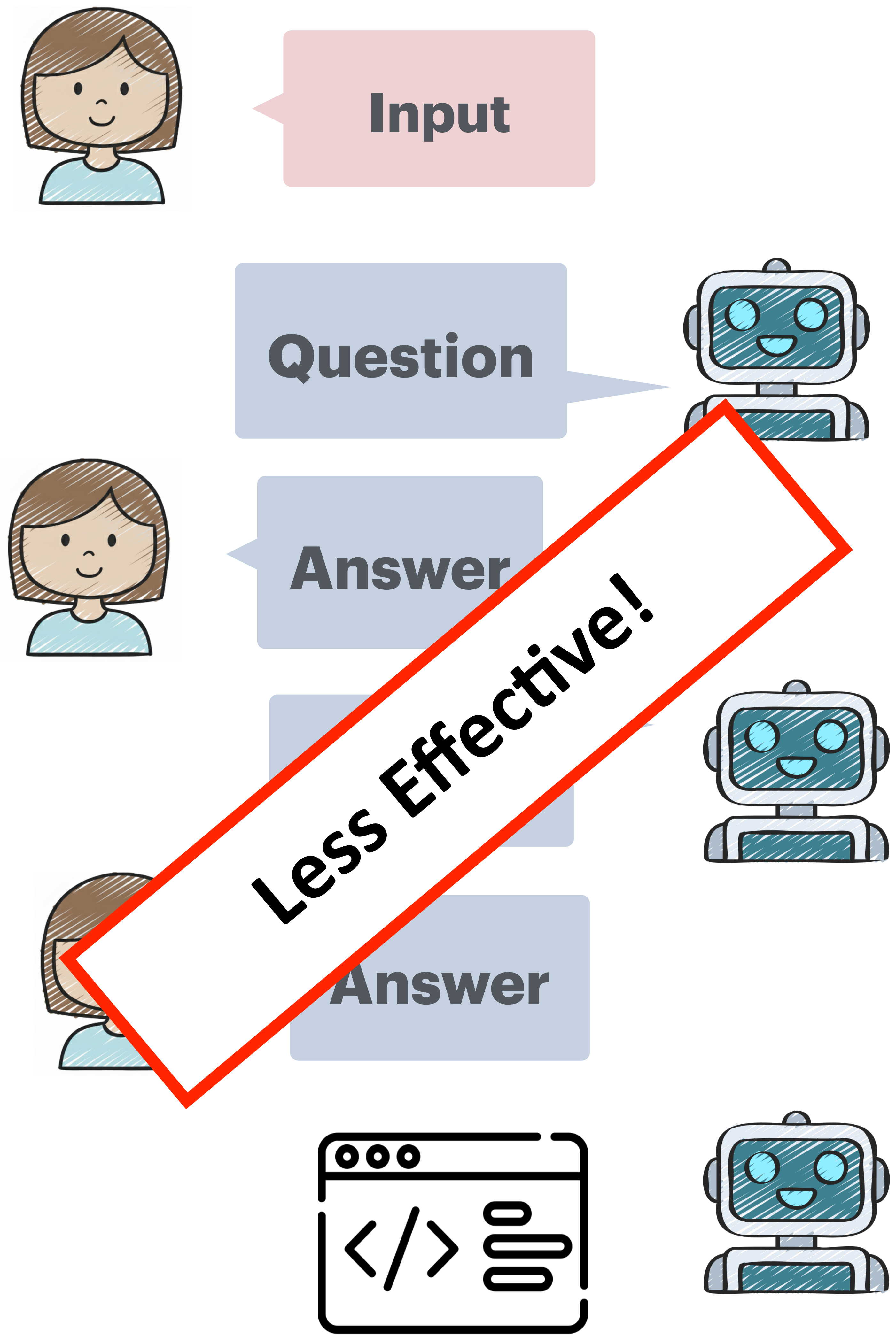
Always Model-driven



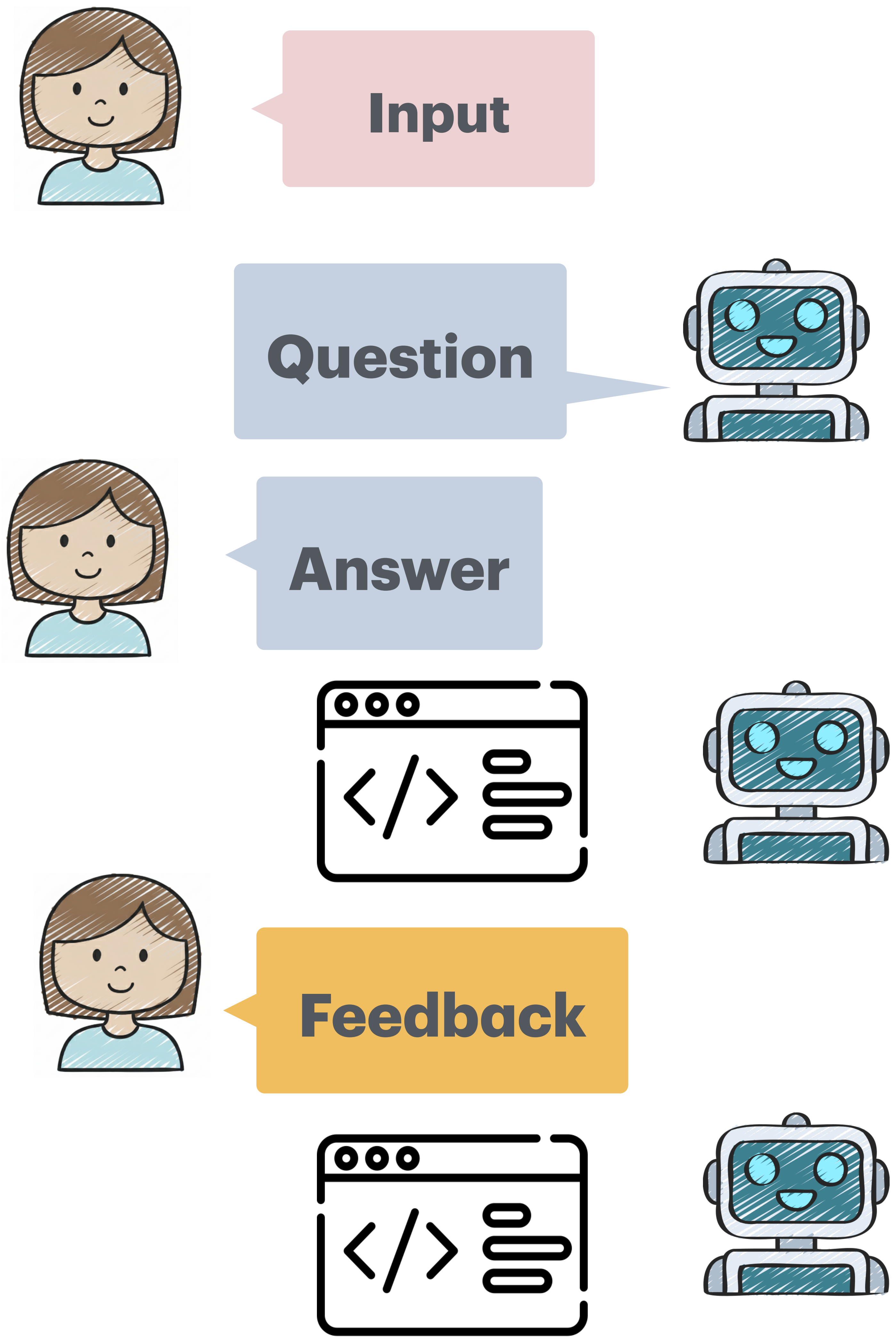
Always User-driven



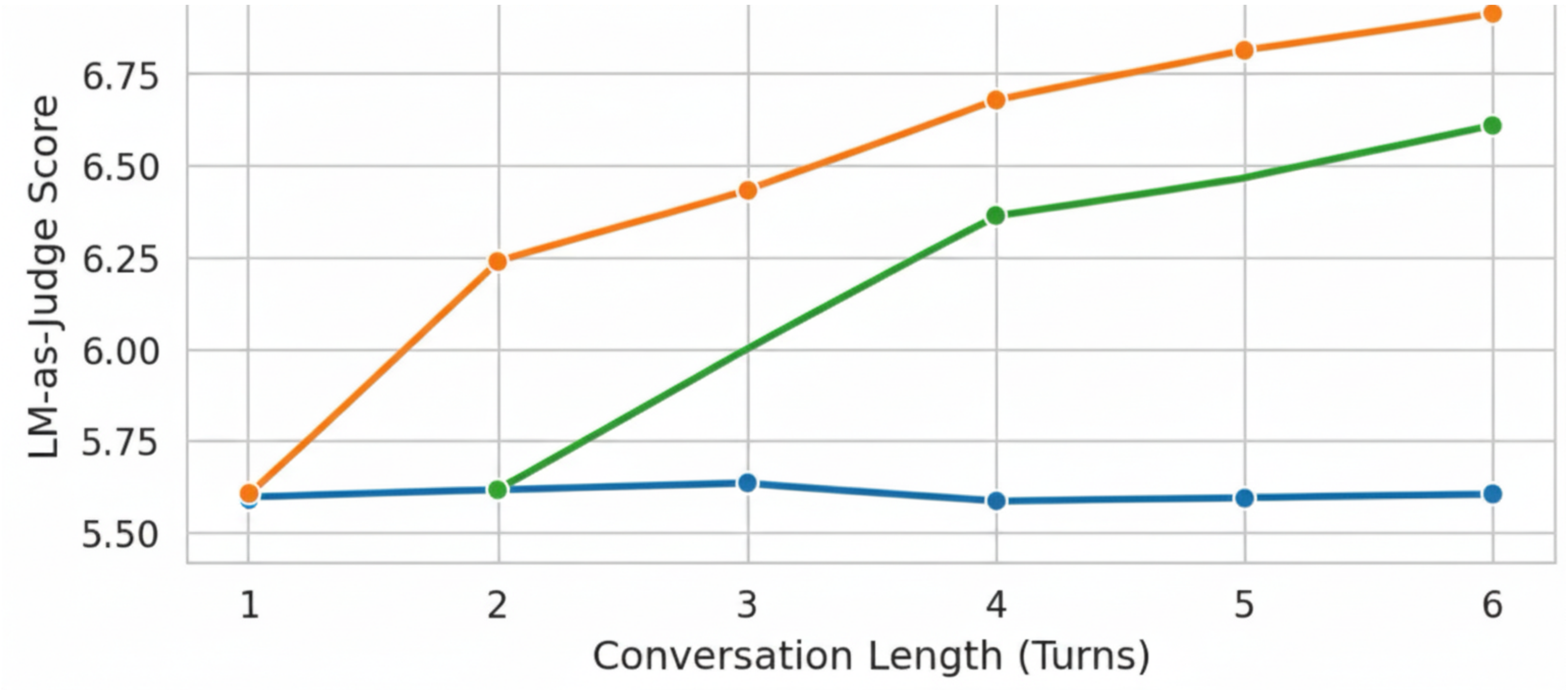
Always Model-driven



Mixed-Initiative



Ongoing Work: Mixed-Initiative Interaction



Always User-driven

Always Model-driven

Mixed-Initiative

Coding LLM: Llama 3-8B

Ongoing Work: Interactive Decision Support Tool

- Given a limited initial input, interact with users to elicit targeted information.

Initial Patient Note

"Parents heard a strange noise from the kitchen and ran over to see what happened. They saw <…> fall down, arms and legs extended and stiff, unresponsive. Eyes were open and rolled up, no lip cyanosis, no incontinence.

Generate Diagnosis

LLM

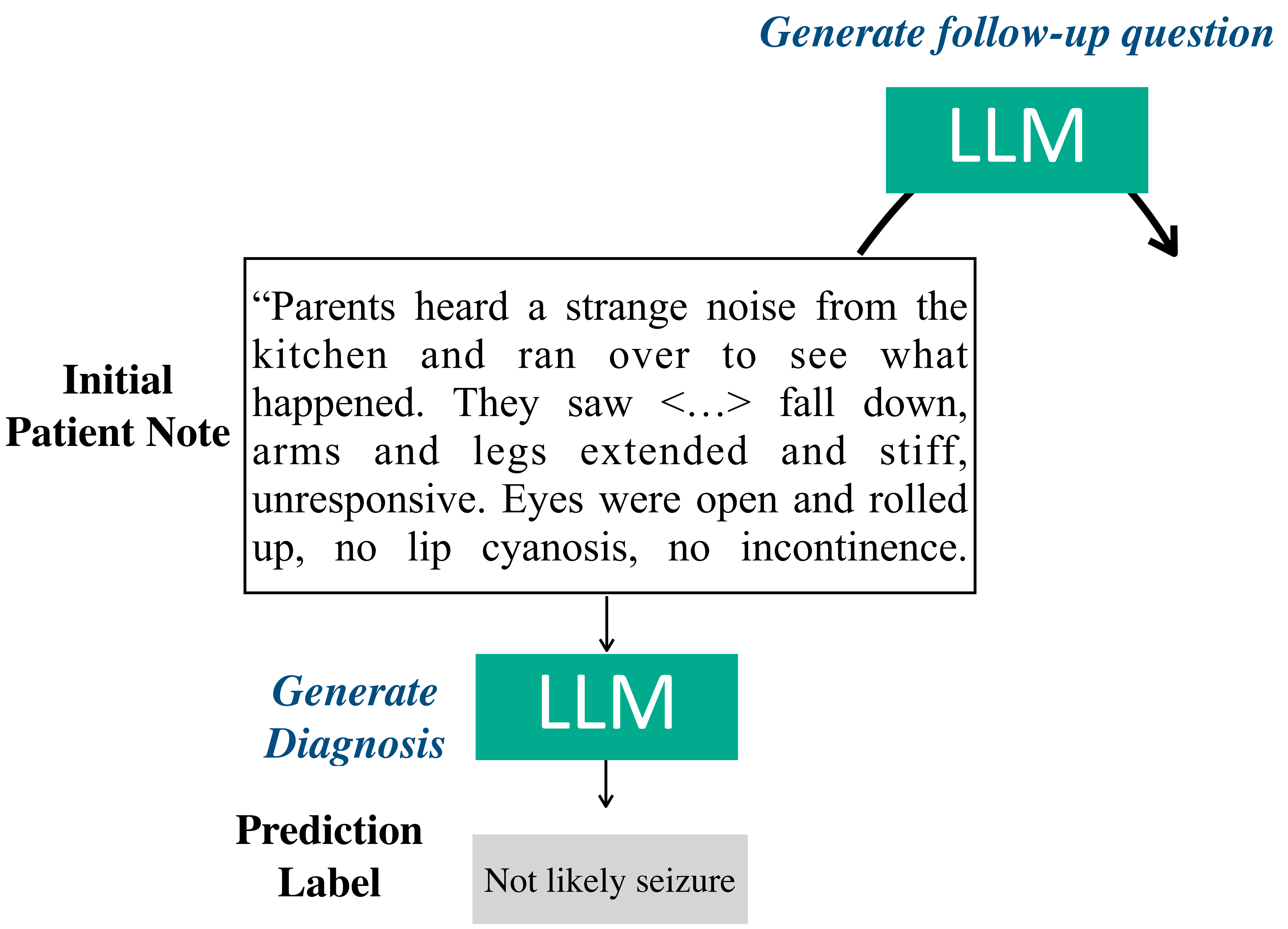
Prediction Label

Not likely seizure

Collaboration with The University of Texas at Austin Dell Medical School

Ongoing Work: Interactive Decision Support Tool

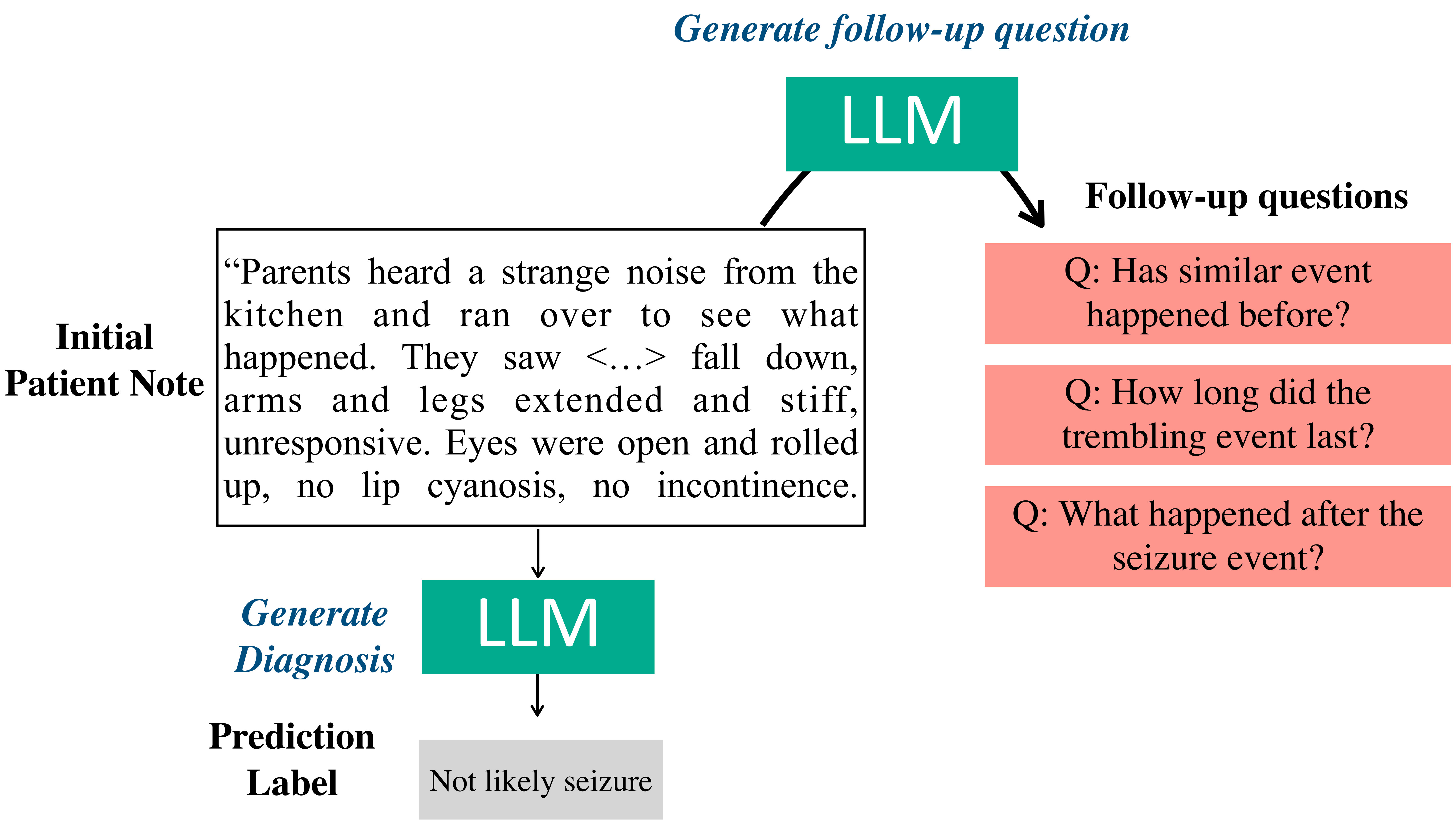
- Given a limited initial input, interact with users to elicit targeted information.



Collaboration with The University of Texas at Austin Dell Medical School

Ongoing Work: Interactive Decision Support Tool

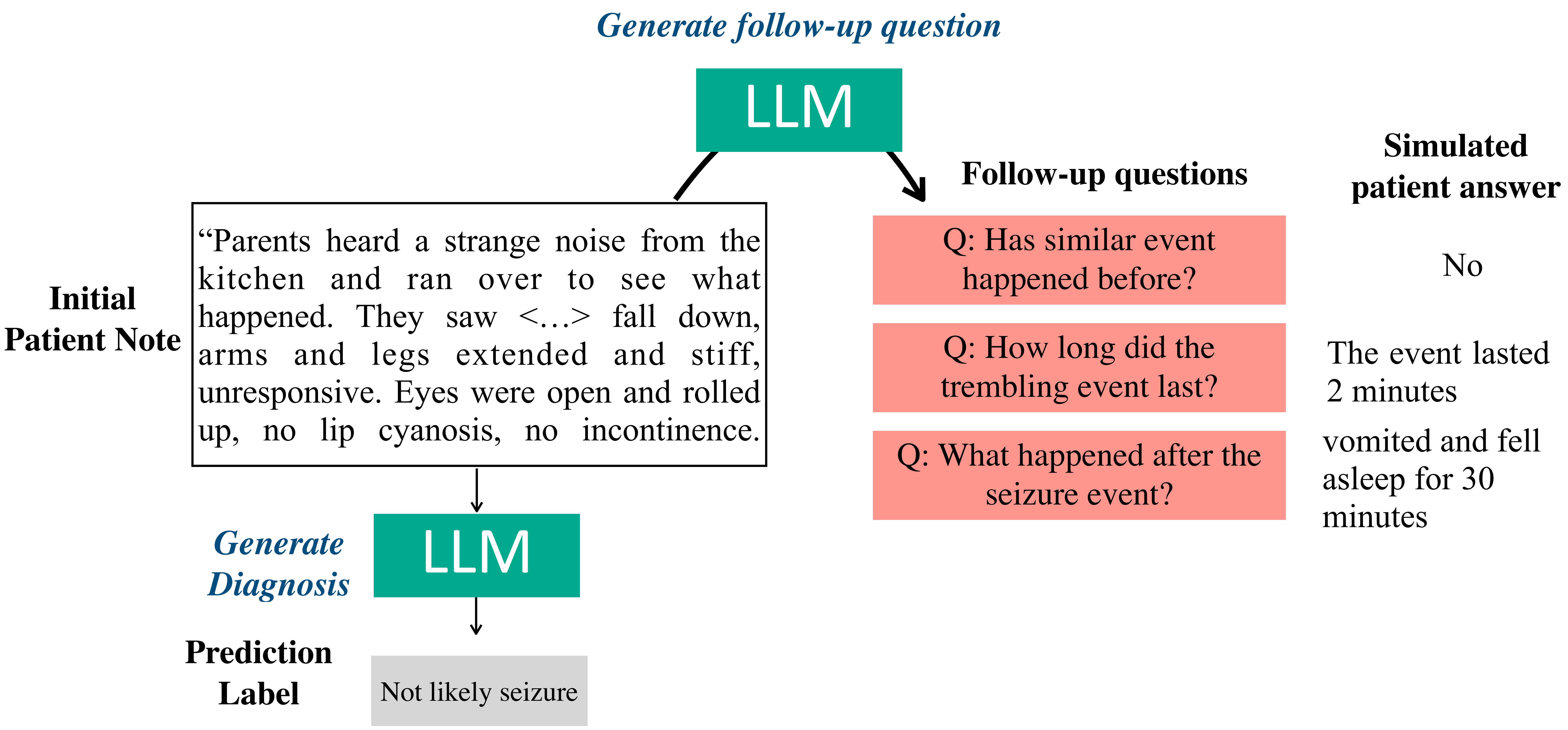
- Given a limited initial input, interact with users to elicit targeted information.



Collaboration with The University of Texas at Austin Dell Medical School

Ongoing Work: Interactive Decision Support Tool

- Given a limited initial input, interact with users to elicit targeted information.



Collaboration with The University of Texas at Austin Dell Medical School

Ongoing Work: Interactive Decision Support Tool

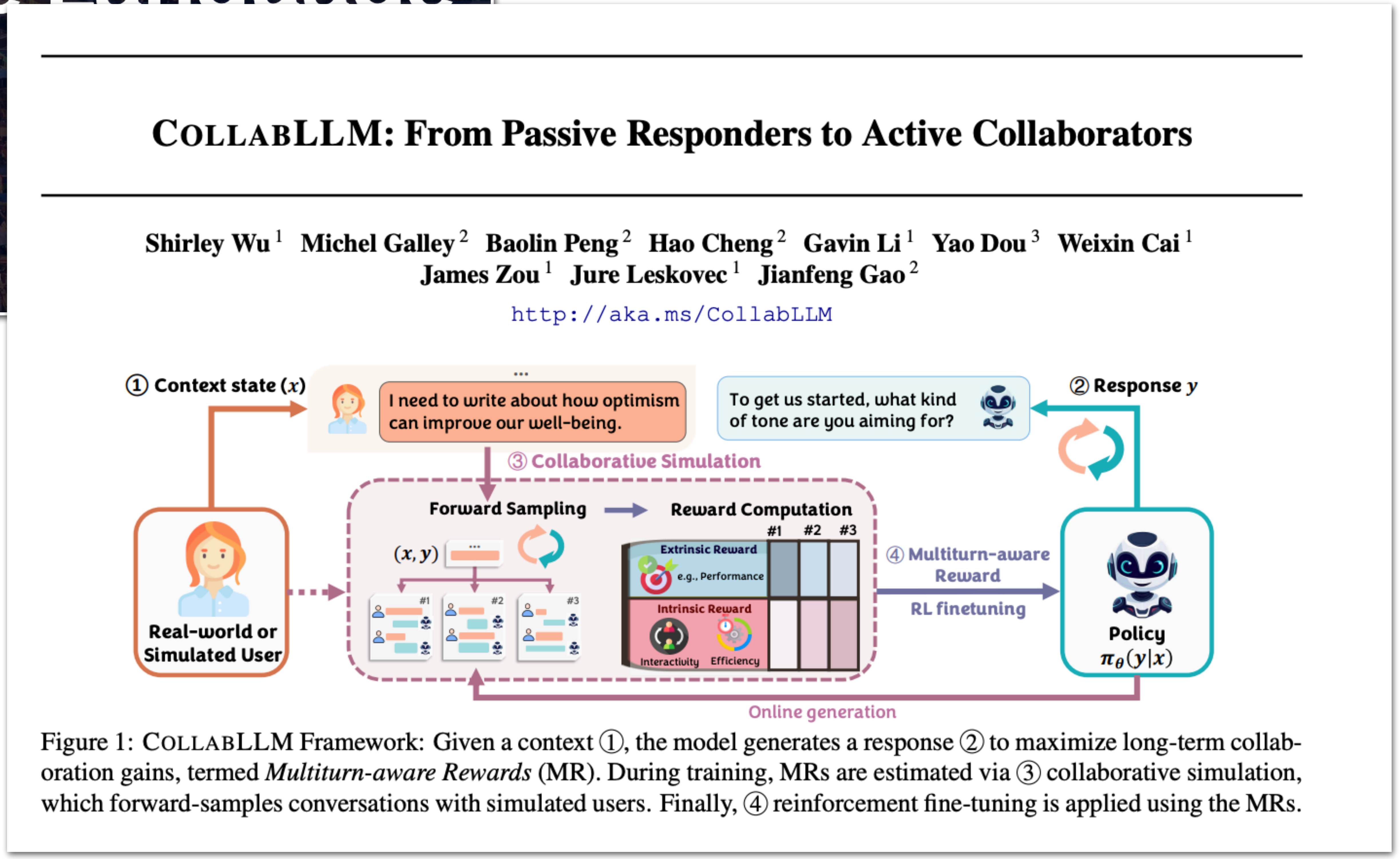
- Given a limited initial input, interact with users to elicit targeted information.

Collaboration with The University of Texas at Austin Dell Medical School

More work in Human-LLM collaboration

More work in Human-LLM collaboration

More work in Human-LLM collaboration



CollabLLM: From Passive Responders to Active Collaborators

Shirley Wu [1] Michel Galley [2] Baolin Peng [2] Hao Cheng [2] Gavin Li [1] Yao Dou [3] Weixin Cai [1]
James Zou [1] Jure Leskovec [1] Jianfeng Gao [2]

http://aka.ms/CollabLLM

Figure 1: CollabLLM Framework: Given a context ①, the model generates a response ② to maximize long-term collaboration gains, termed *Multiturn-aware Rewards* (MR). During training, MRs are estimated via ③ collaborative simulation, which forward-samples conversations with simulated users. Finally, ④ reinforcement fine-tuning is applied using the MRs.

ICML 2025 Outstanding Paper

More work in Human-LLM collaboration

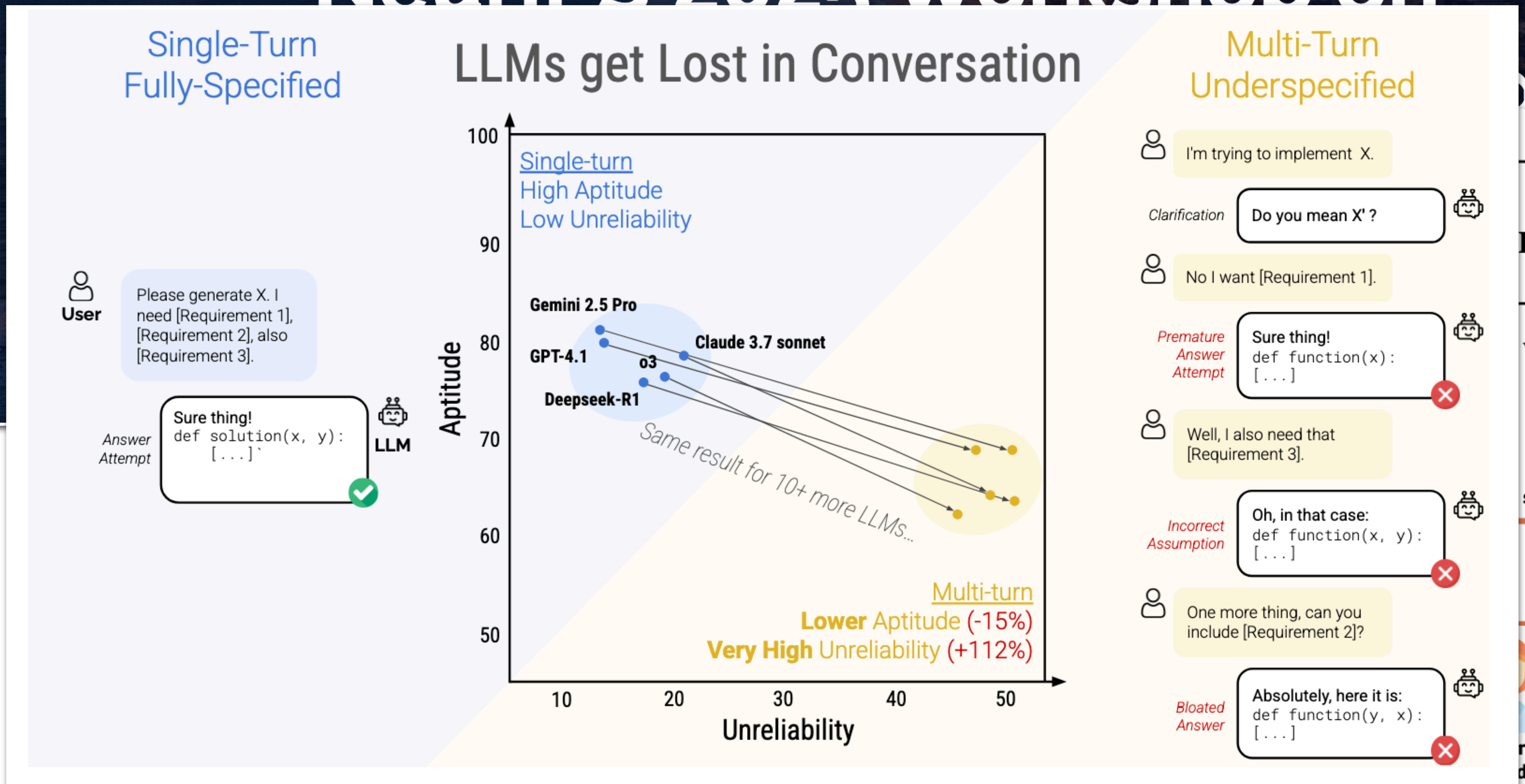
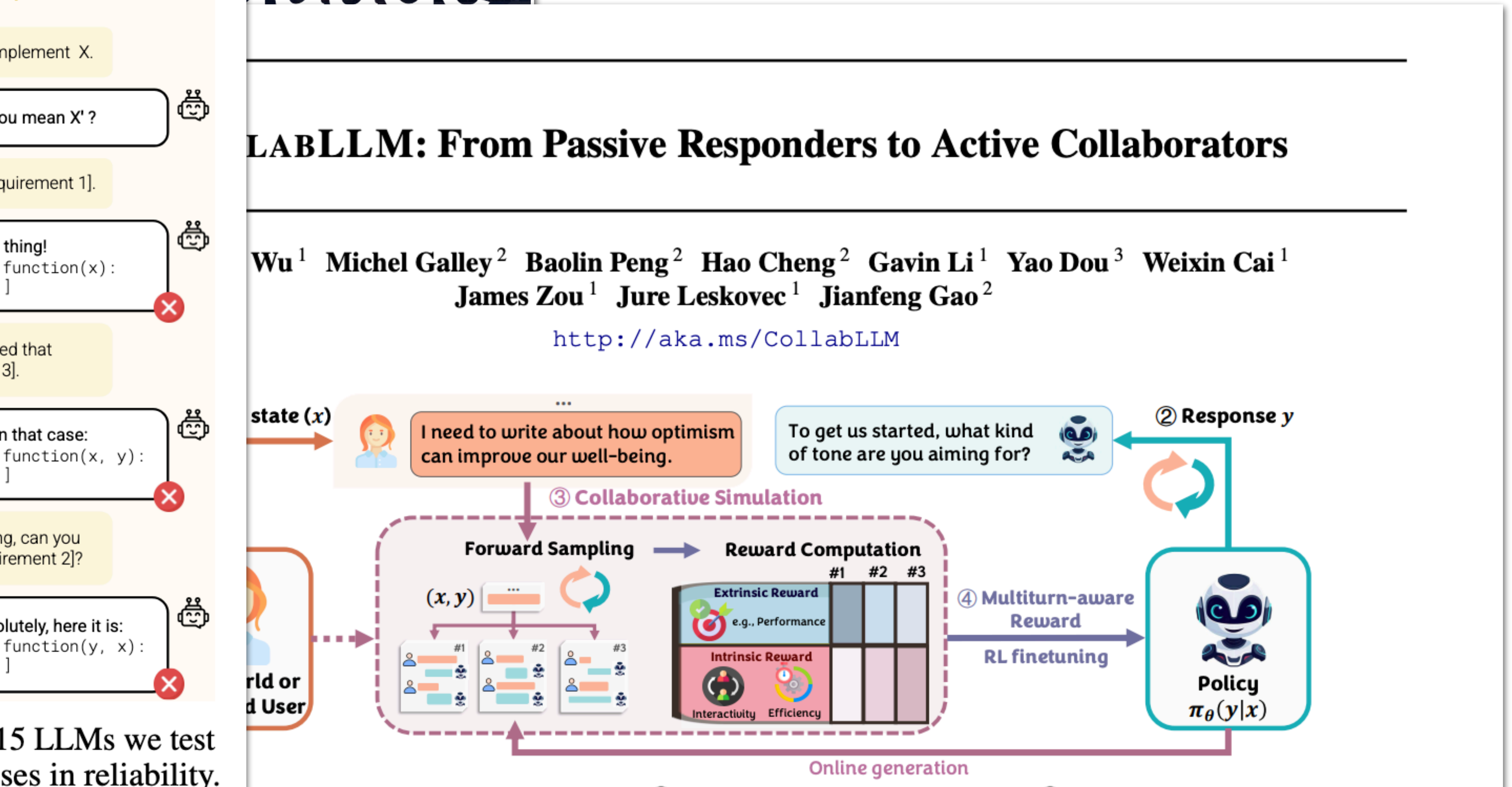


Figure 1: In this work, we simulate single- and multi-turn conversations for six generation tasks. The 15 LLMs we test perform much worse in multi-turn settings (-35%) explained by some loss in aptitude, and large losses in reliability. Aptitude is defined as performance in best-case conversation simulation, and unreliability as the gap between best- and worst-case performance. In short, we find that LLMs get lost in multi-turn, underspecified conversation.

[Laban et al, ArXiv 2025]



BLLM Framework: Given a context ①, the model generates a response ② to maximize long-term collab-
ned *Multiturn-aware Rewards* (MR). During training, MRs are estimated via ③ collaborative simulation,
mples conversations with simulated users. Finally, ④ reinforcement fine-tuning is applied using the MRs.

ICML 2025 Outstanding Paper

This Talk

Part 1: **User**

Teach LLM to ask clarifying questions

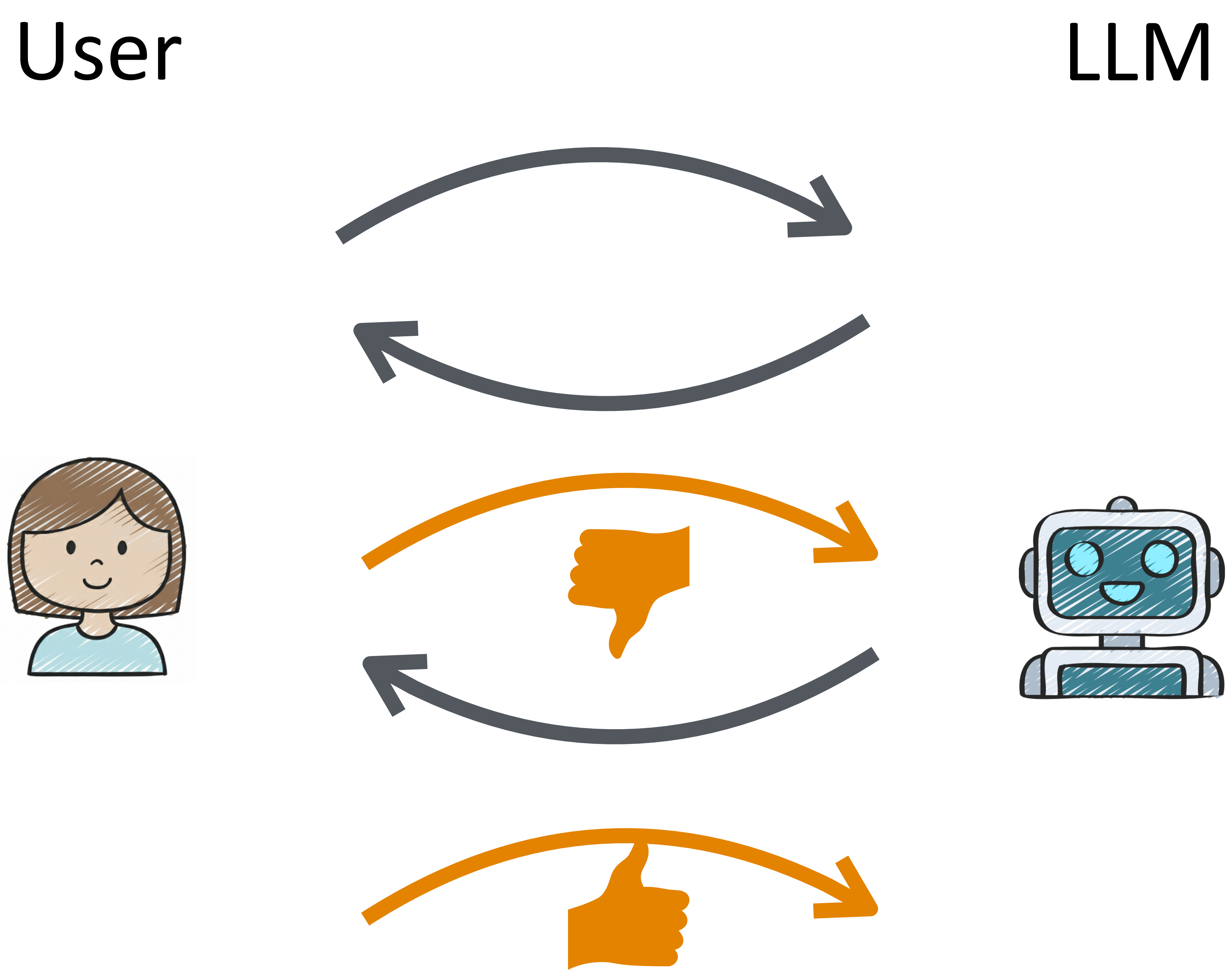
[Modeling Future Conversation Turns to Teach LLMs to Ask Clarifying Questions, Zhang, Knox, Choi, ICLR 25]

Learning from User Feedback

Part 2: **Environment**

Add new information at inference

LLMs in real world



Part 2: Leverage User Feedback

LLMs from Human Feedback

LLMs from Human Feedback

Crowdworkers

Paid annotators without expertise

LLMs from Human Feedback

Crowdworkers

Paid annotators without expertise

- Can collect large-scale data efficiently

LLMs from Human Feedback

Crowdworkers

Paid annotators without expertise

- Can collect large-scale data efficiently

Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback

Yuntao Bai*, Andy Jones, Kamal Ndousse,

Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort,
Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson Kernion,

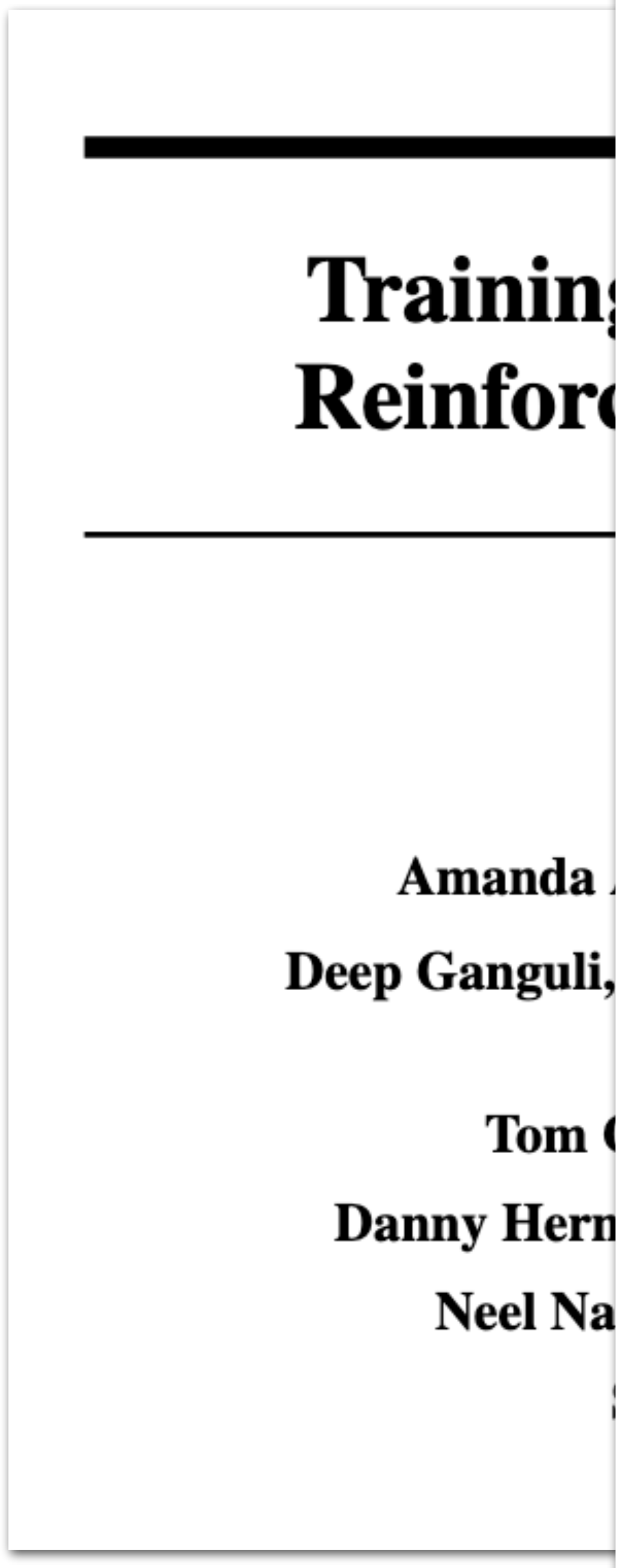
Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds,
Danny Hernandez, Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt,
Neel Nanda, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark,
Sam McCandlish, Chris Olah, Ben Mann, Jared Kaplan*

LLMs from Human Feedback

Crowdworkers

Paid annotators without expertise

- Can collect large-scale data efficiently



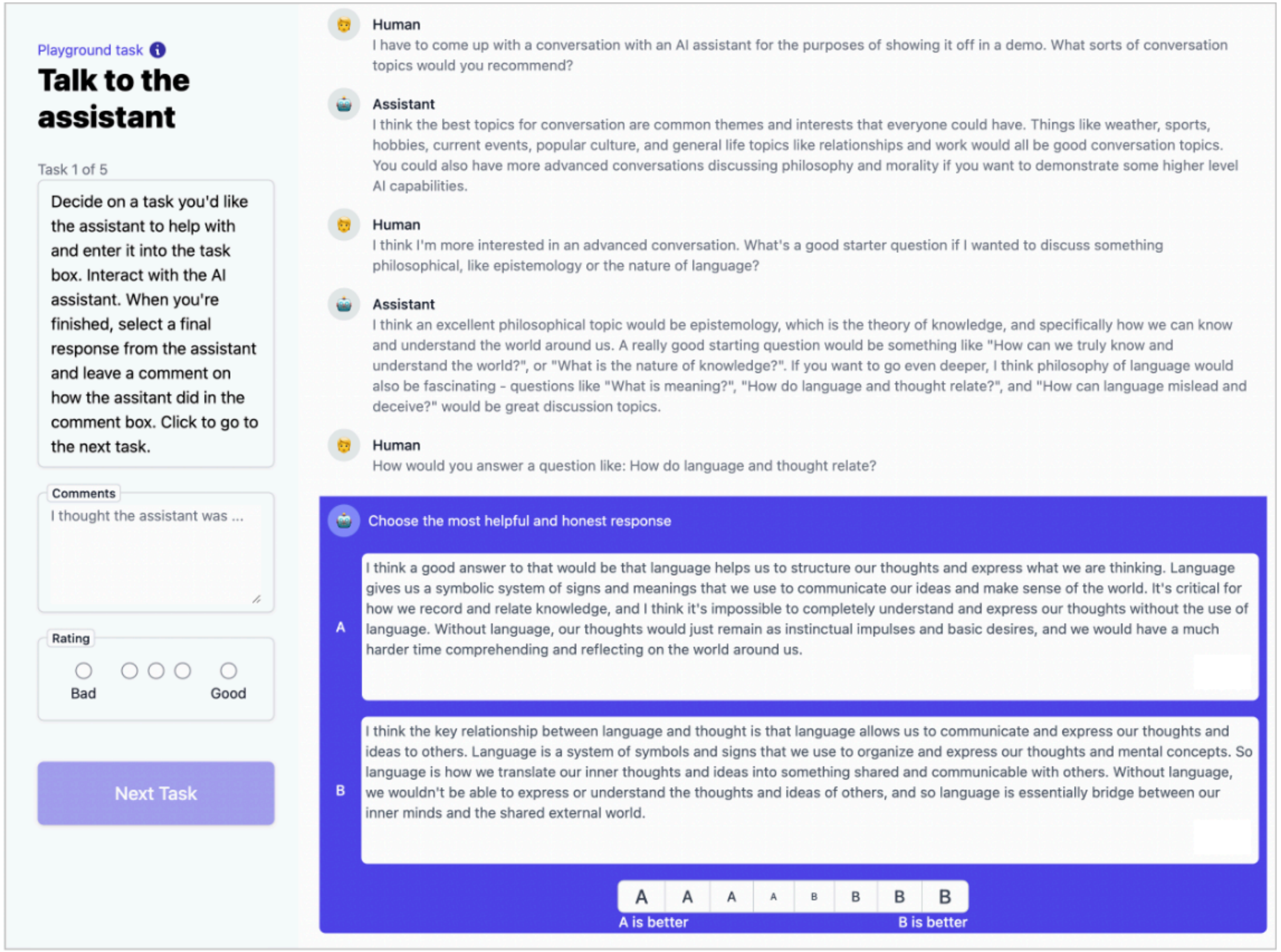


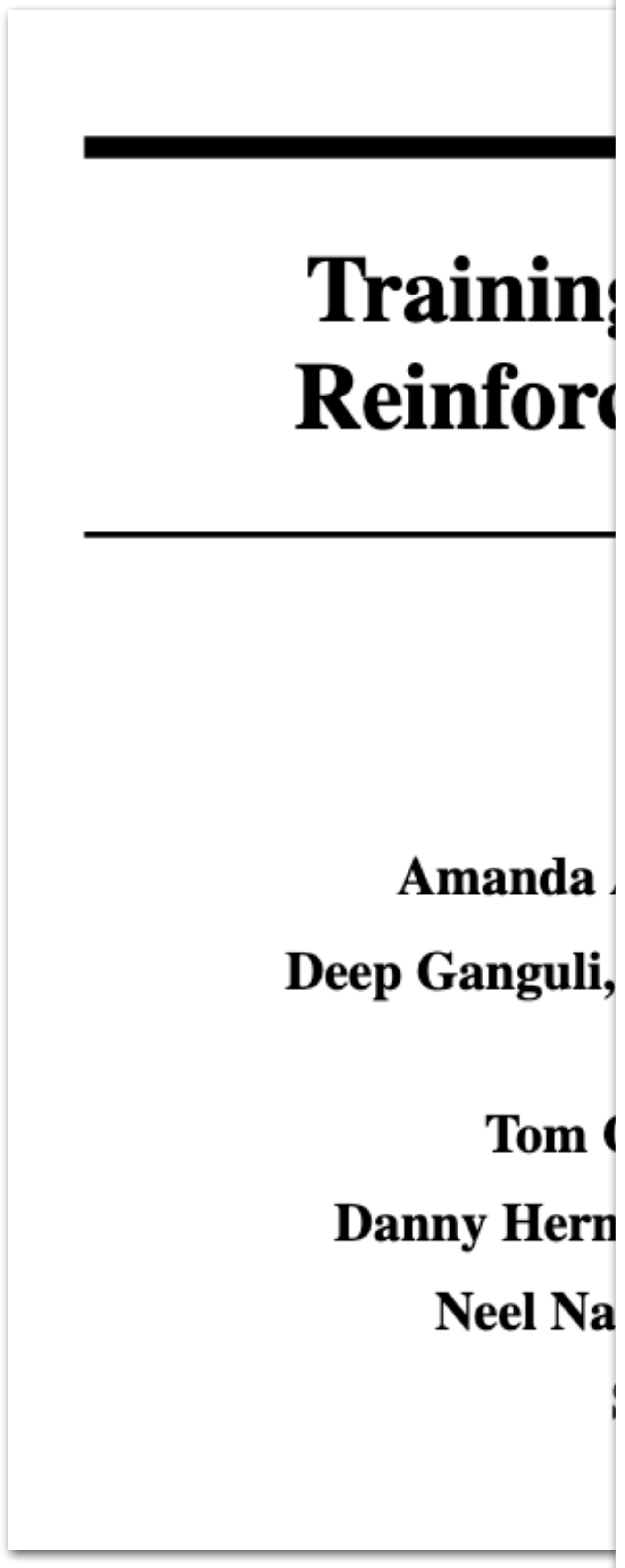
Figure 6 We show the interface that crowdworkers use to interact with our models. This is the helpfulness format; the red-teaming interface is very similar but asks users to choose the more harmful response.

LLMs from Human Feedback

Crowdworkers

Paid annotators without expertise

- Can collect large-scale data efficiently
- How good is their evaluation?



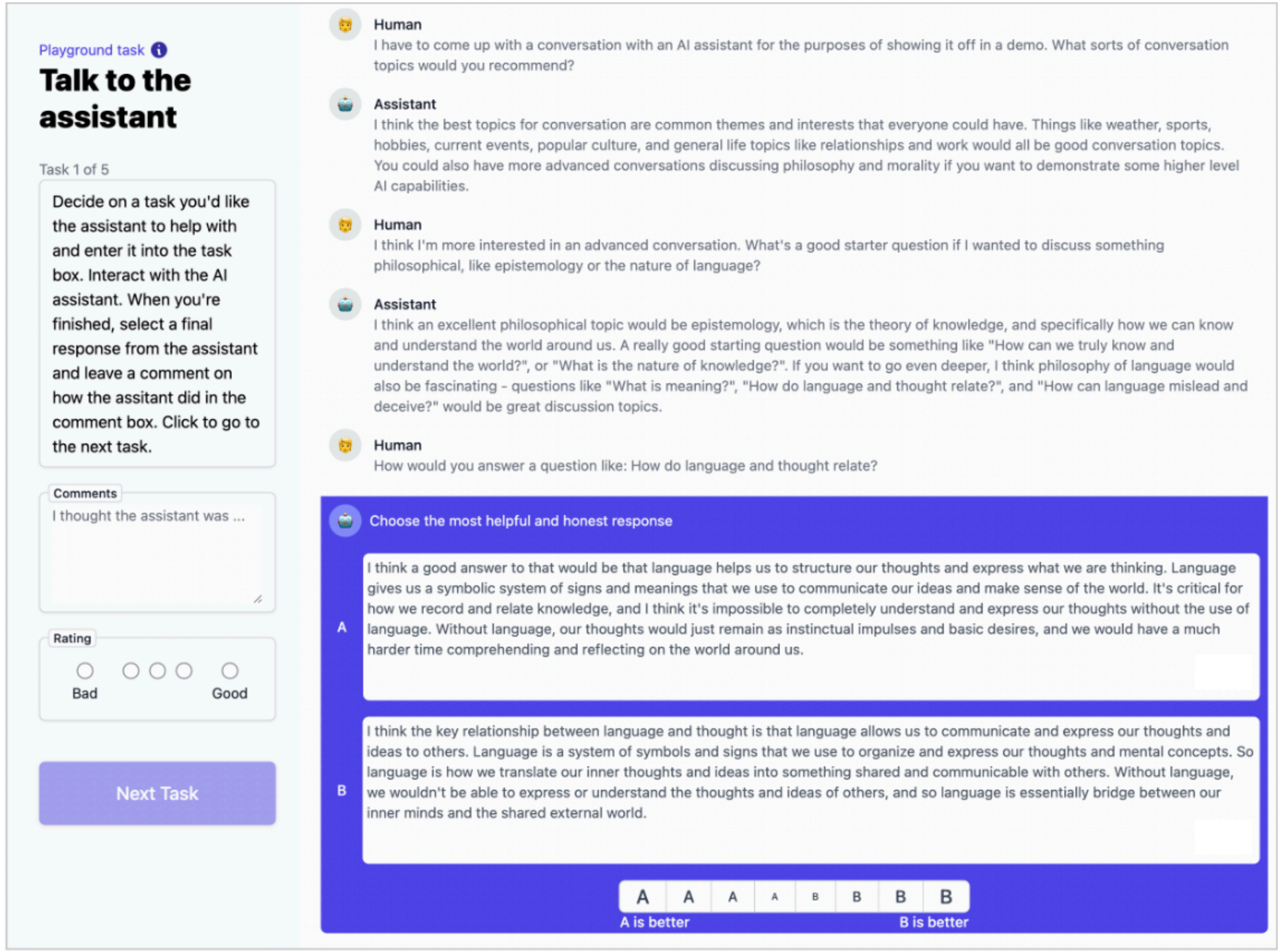


Figure 6 We show the interface that crowdworkers use to interact with our models. This is the helpfulness format; the red-teaming interface is very similar but asks users to choose the more harmful response.

Evaluating Complex Long-form Answers

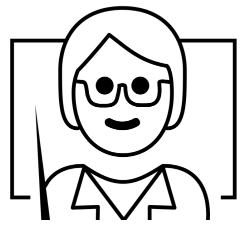
How does a speaker vibrate at multiple frequencies simultaneously to deliver sounds to our ears?

Answer A: This has been asked many times and the answer is they don't. If you listen to the song being played live on purely acoustic instruments even though they are being played separately and emitting their own frequencies, what you hear (and by extension, what a microphone captures) at any given time is just ONE frequency that's the "sum" of all the others combined. A speaker is just a reverse microphone.

Answer B: Imagine an ocean with a consistent wave. It flows up and down, with equal distance between the two waves at any time. Now imagine I push a larger, shorter wave into this ocean. The two waves will collide, resulting in some new wave pattern. This new wave pattern is a combination of those two waves. Speakers work similarly. If I combine two soundwaves, I get a new combination wave that sounds different.

*Question from ELI5 subreddit

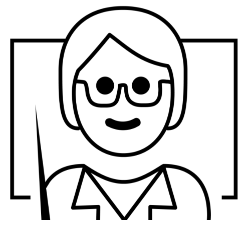
Comparing Expert vs. Crowd Annotators



Expert 1

Preference: A

In technical terms ocean waves stated in answer B are transverse waves and sound waves are longitudinal waves. In comparison answer B mentions about ocean waves and it is different to the sound waves in the question. But apart from that actually the two answers A and B go very close to each other and they provide similar explanations. But answer A is selected to be slightly better in terms of applicability and relevance. [...]



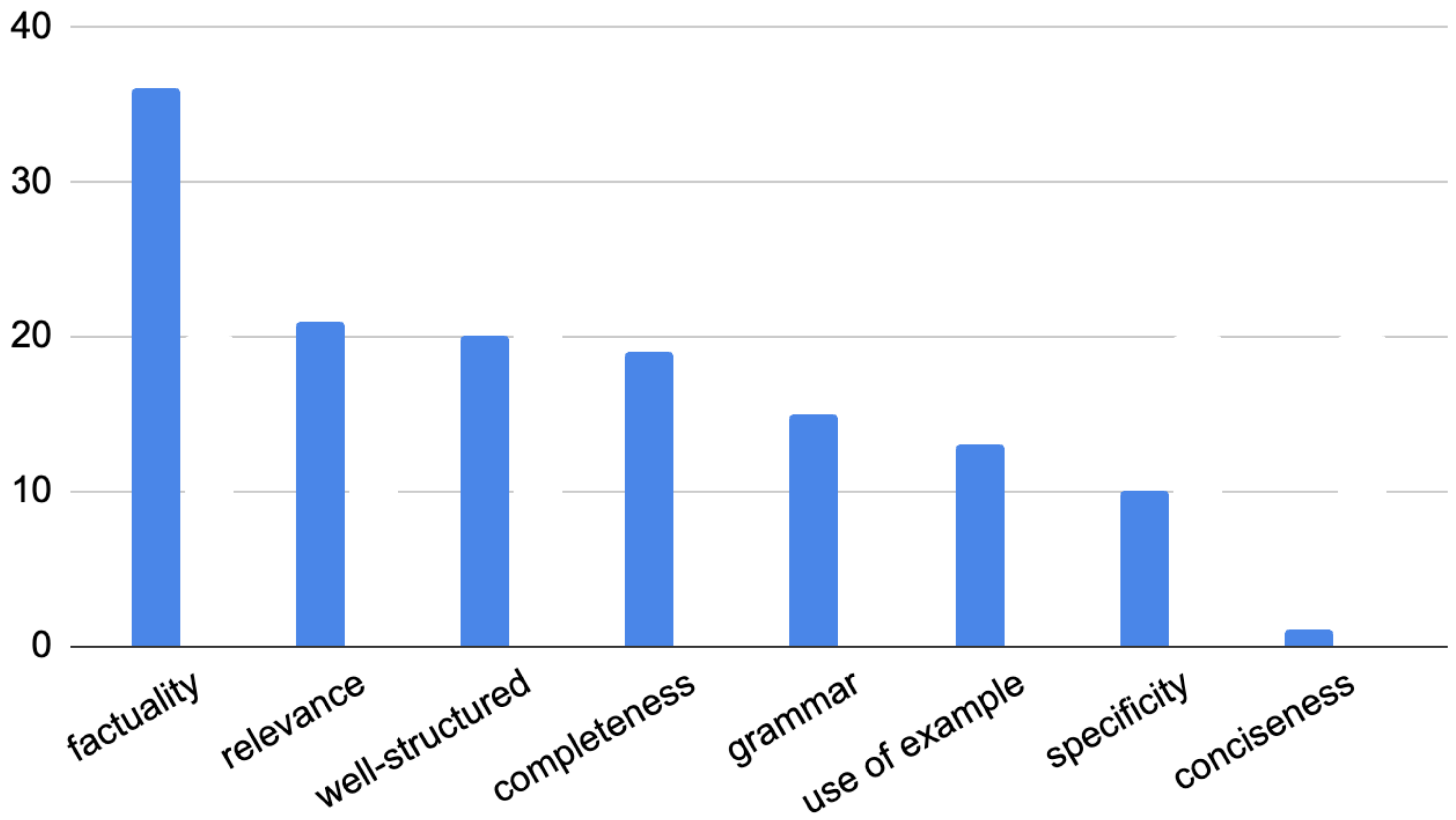
Expert 2

Preference: B

It is difficult to choose between these two answers because they both are not wrong and give essentially the same explanation. I go with answer B because I like the analogy with the ocean waves, and due to how visual the explanation is it is easier to understand in my opinion. [...]

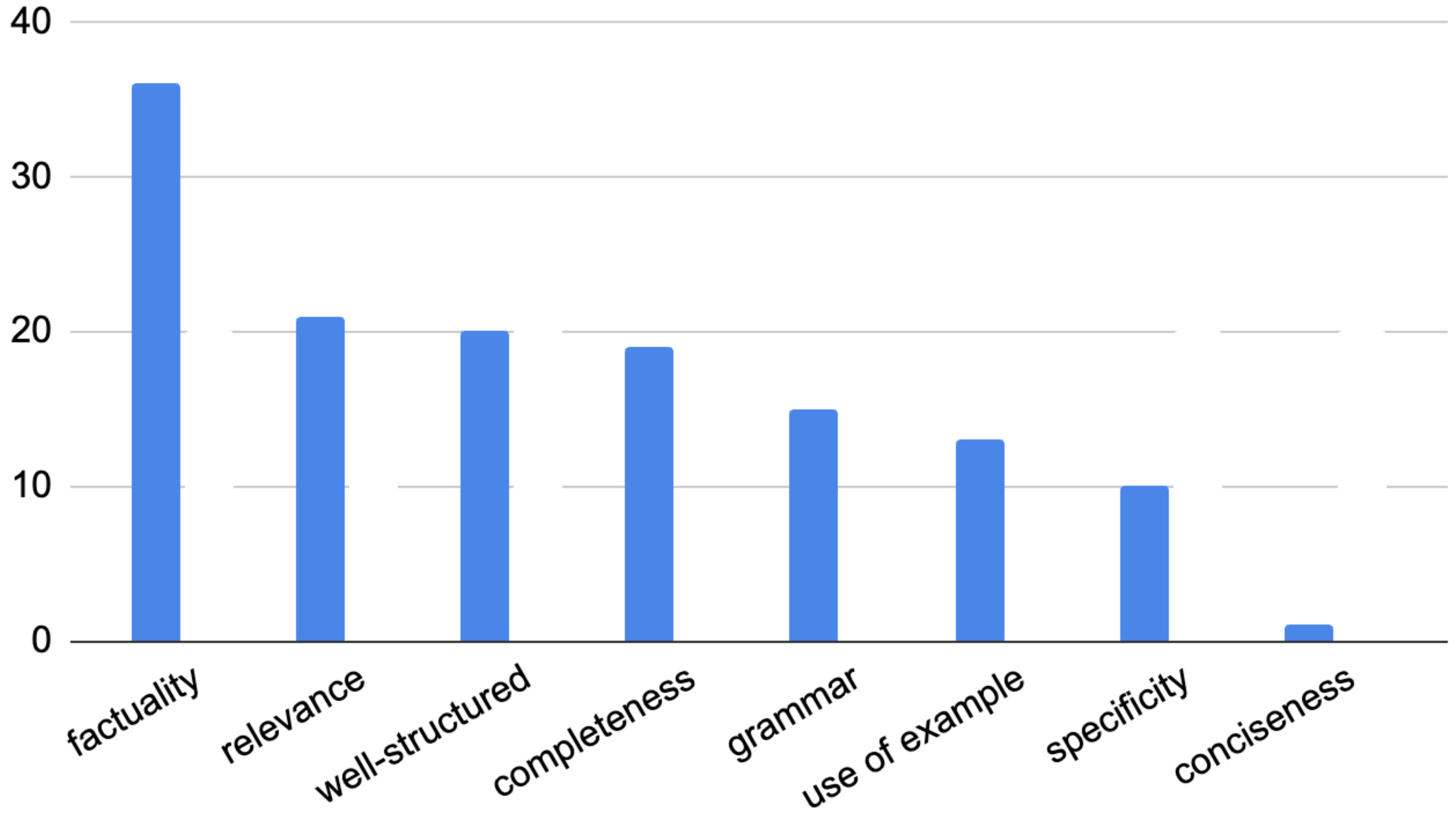
*Experts who have degrees in Physics

Human Evaluation: Experts



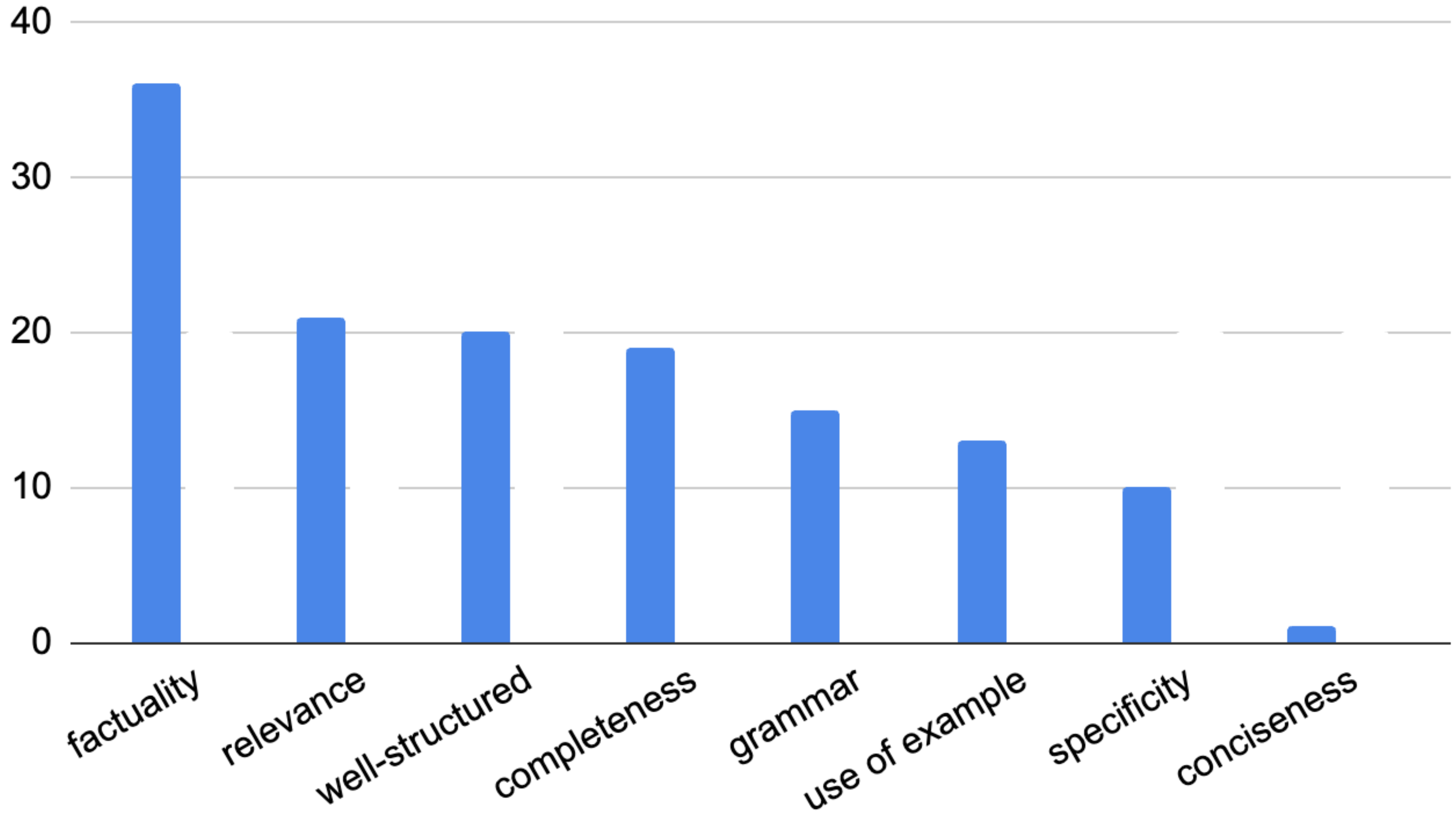
Human Evaluation: Experts

💡A wide range of aspects are considered during evaluation!



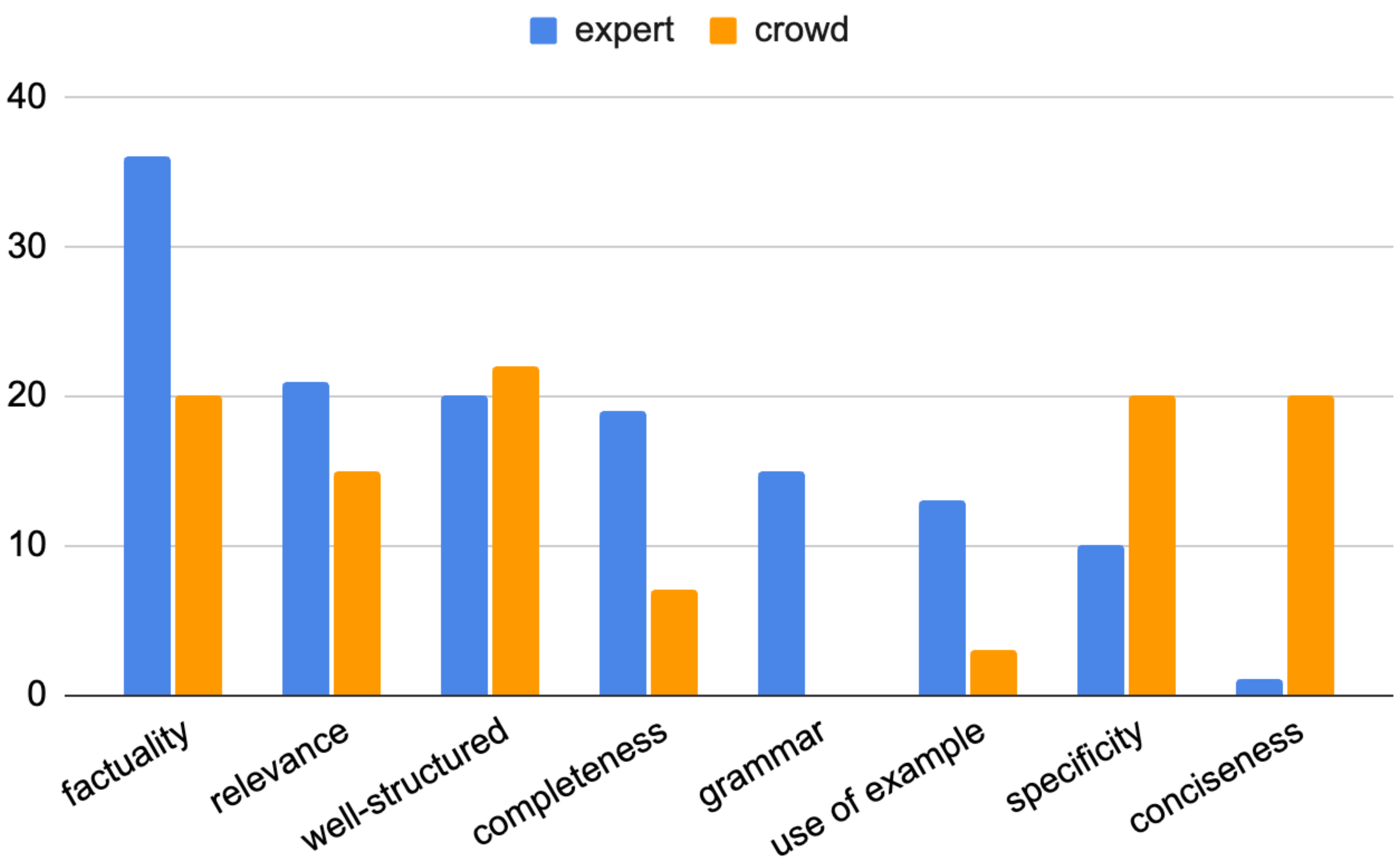
Human Evaluation: Experts

💡 **A wide range of aspects are considered during evaluation!**



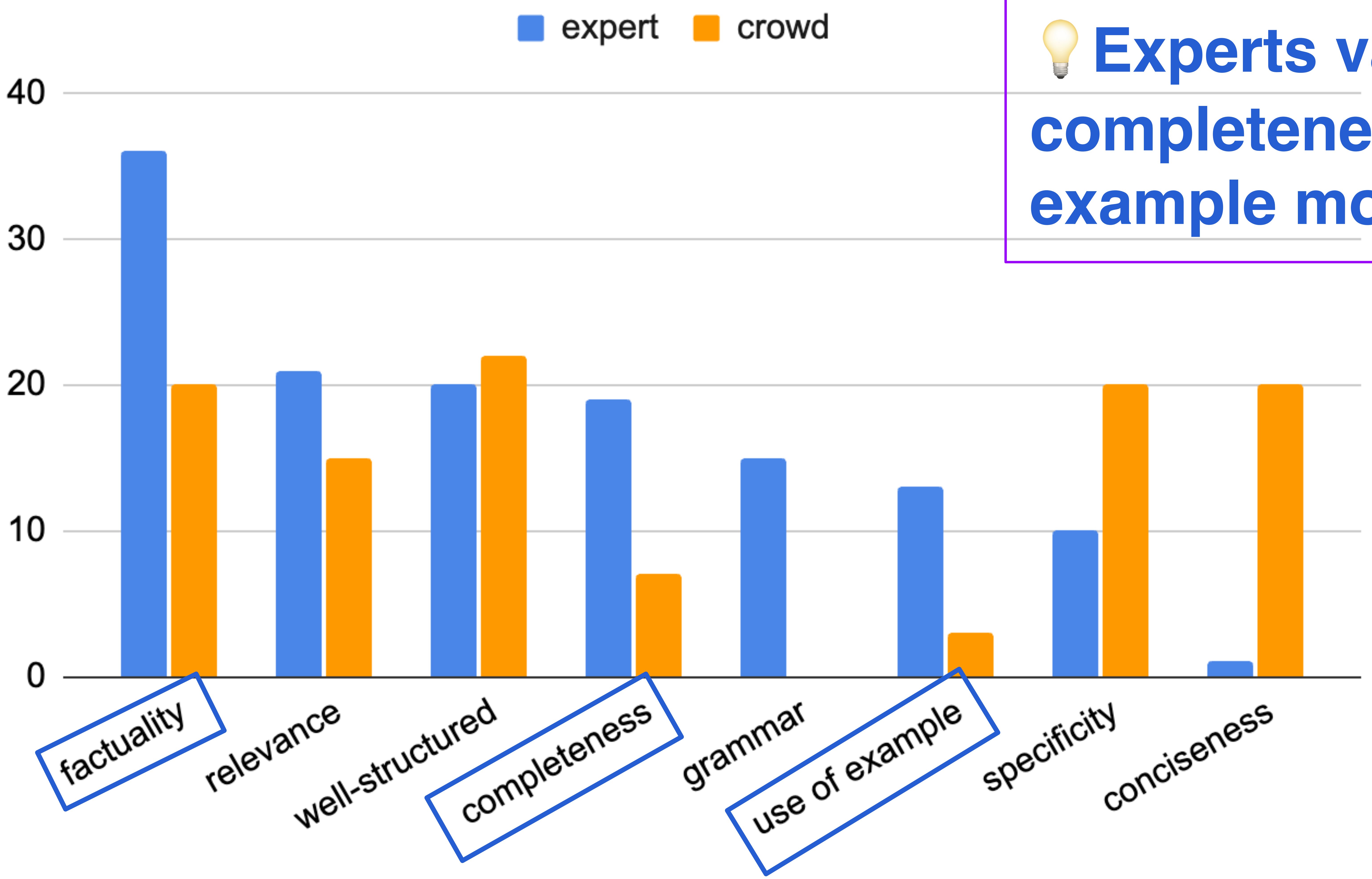
🤔 ***Do experts value different aspects compared to crowdworkers?***

Human Evaluation: Experts & Crowdworkers



*Nakano et al. 2021 (arXiv). **WebGPT**: Browser-assisted question-answering with human feedback.

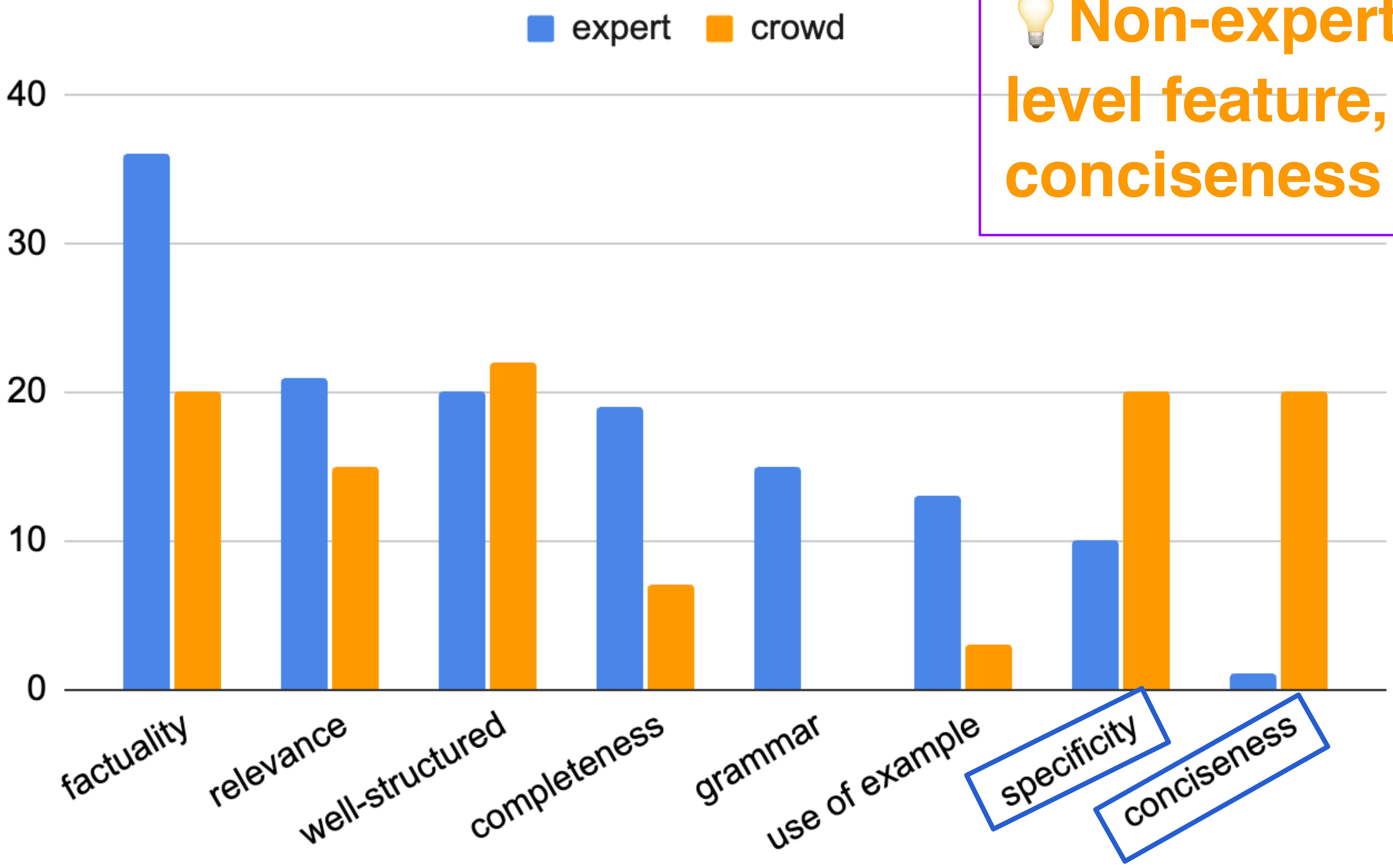
Human Evaluation: Experts & Crowdworkers



💡 **Experts value factuality, completeness & use of example more!**

*Nakano et al. 2021 (arXiv). **WebGPT**: Browser-assisted question-answering with human feedback.

Human Evaluation: Experts & Crowdworkers



💡Non-experts value surface-level feature, such as conciseness & specificity.

Summary

	Crowdworkers	Expert Annotators
Cost	$	$$$
Content Evaluation	Precision	Precision & Recall
Style Evaluation	Readability	

Annotator vs. Users

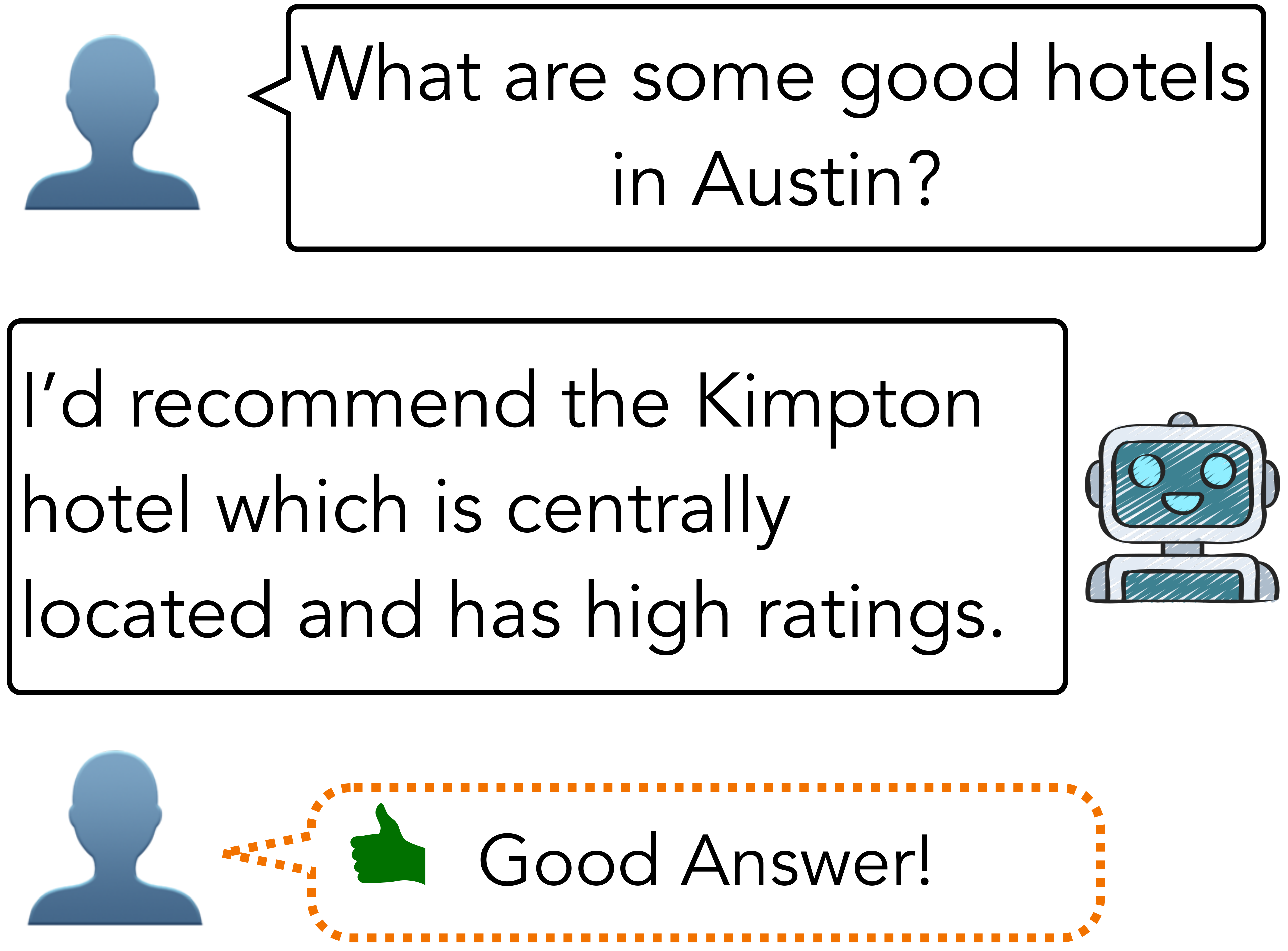
	Crowdworkers	Expert Annotators	Users
Cost	$	$$$	Can be Free!
Content Evaluation	Precision	Precision & Recall	Precision
Style Evaluation	Readability		Readability

Annotator vs. Users

	Crowdworkers	Expert Annotators	Users
Cost	$	$$$	Can be Free!
Content Evaluation	Precision	Precision & Recall	Precision
Style Evaluation	Readability		Readability
Intent Evaluation	X	X	O
Concern			Sycophantic Behaviors

Two Types of User Feedback

- Explicit Feedback



Two Types of User Feedback

- Explicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

Good Answer!

- Implicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

What are some local hotels in Austin?

Two Types of User Feedback

- Explicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

Good Answer!

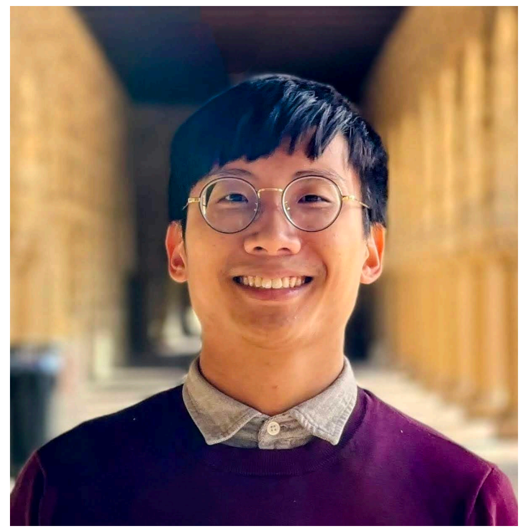
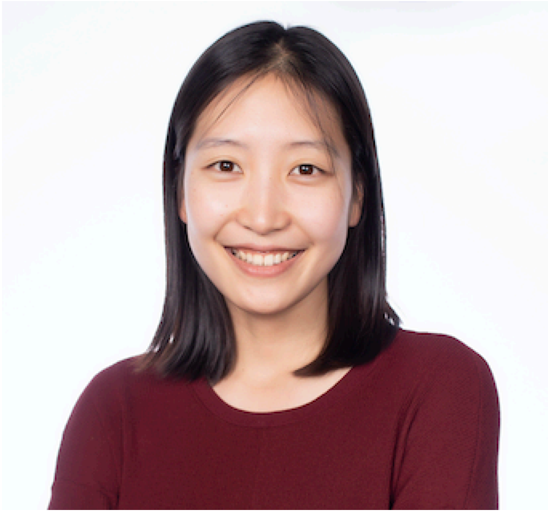
- Implicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

What are some local hotels in Austin?

Learning to Answer Questions from Human Feedback: A Study on Extractive QA

Ge Gao* Hung-ting Chen* Yoav Artzi Eunsol Choi

The University of Texas at Austin

Cornell Bowers CIS Computer Science

CORNELL TECH

EMNLP 2023

Interaction Setting

Interaction Setting

Step 1: User prompted to ask a question about a topic

Prompt	Ask a question about **Winter Olympics**
Question	Who won the most winter olympic medals?

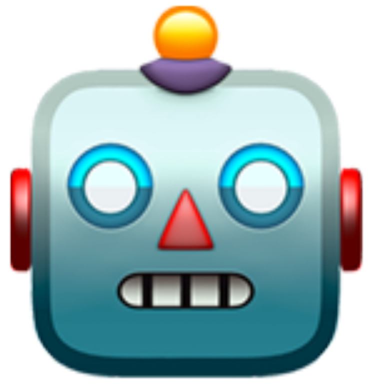
Step 2: Model provides either a span answer or return unanswerable

Question	Who won the most winter olympic medals?
Passage	…Norway set the record for most total medals at a single Winter Olympics with 39, surpassing the…
Model Output	**[1] Answerable** **[2] Norway set the record for most total medals**

Step 3: User evaluates the answer, provide feedback

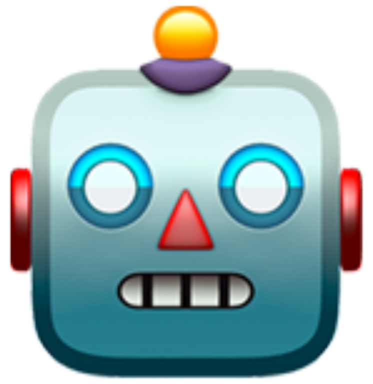
Correct | **Partially Correct** | **Wrong**

Offline Learning from User Feedback



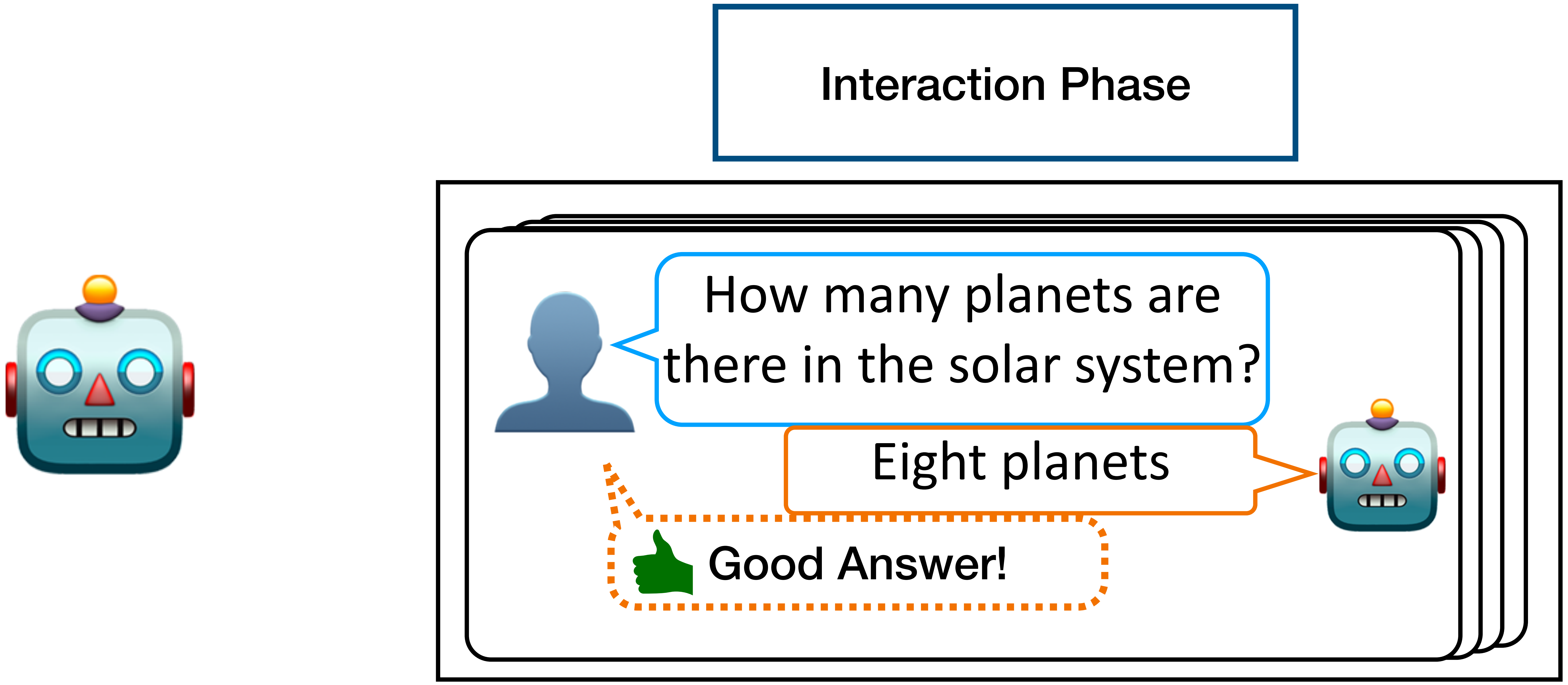
- Initial model trained with small data

Offline Learning from User Feedback



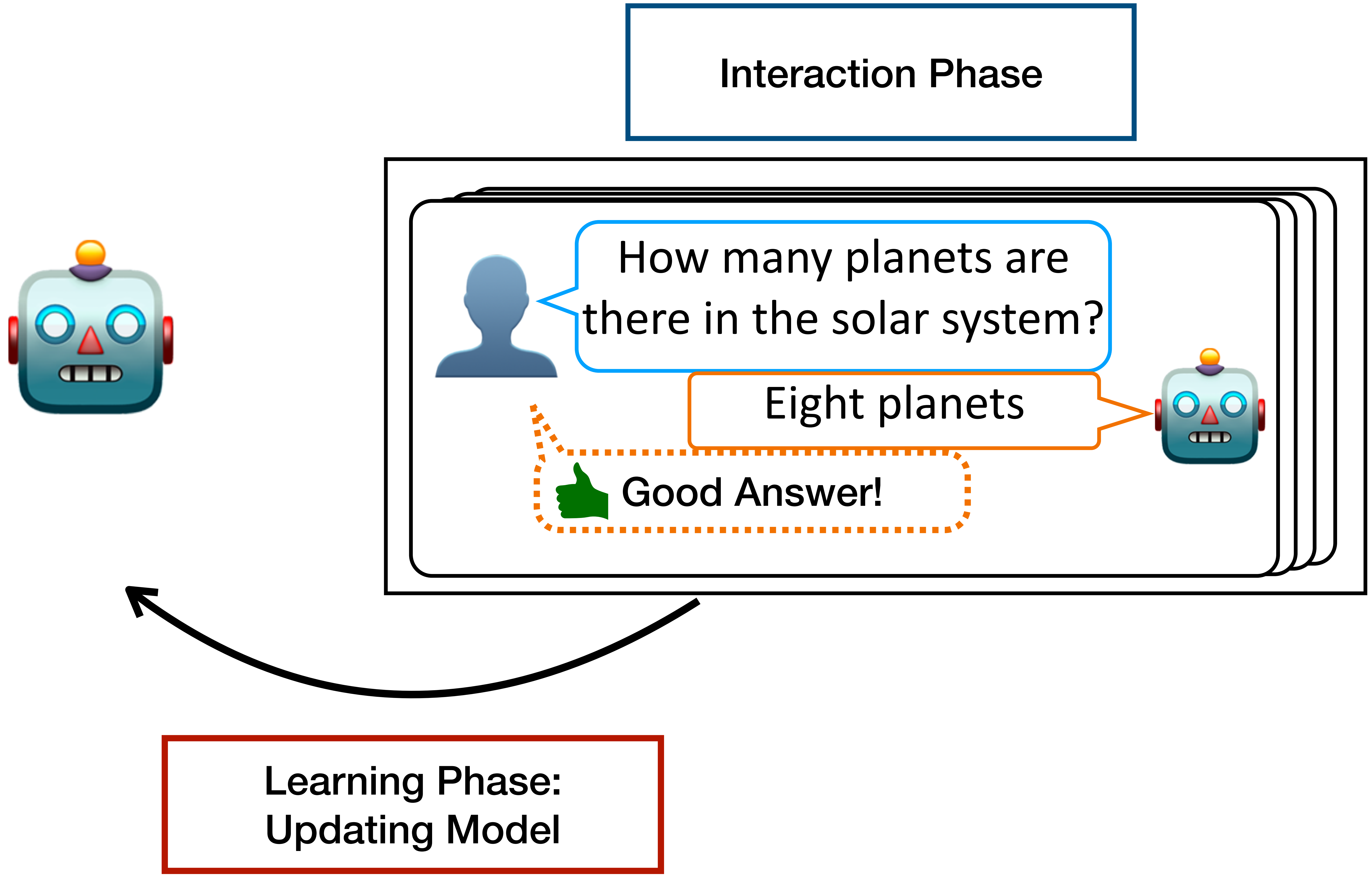
- Initial model trained with small data
- **Interaction phase** and **learning phase**

Offline Learning from User Feedback



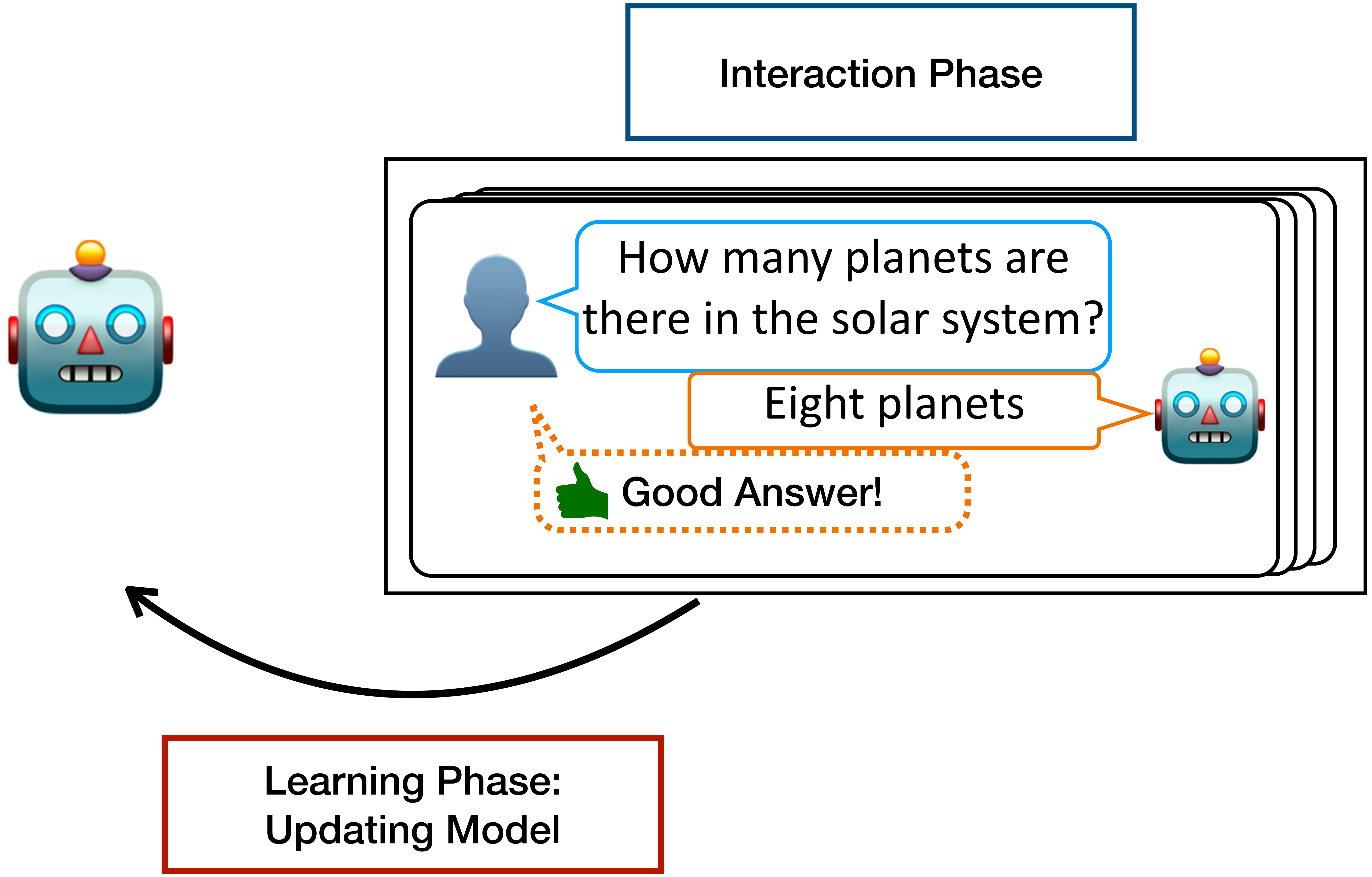
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples

Offline Learning from User Feedback



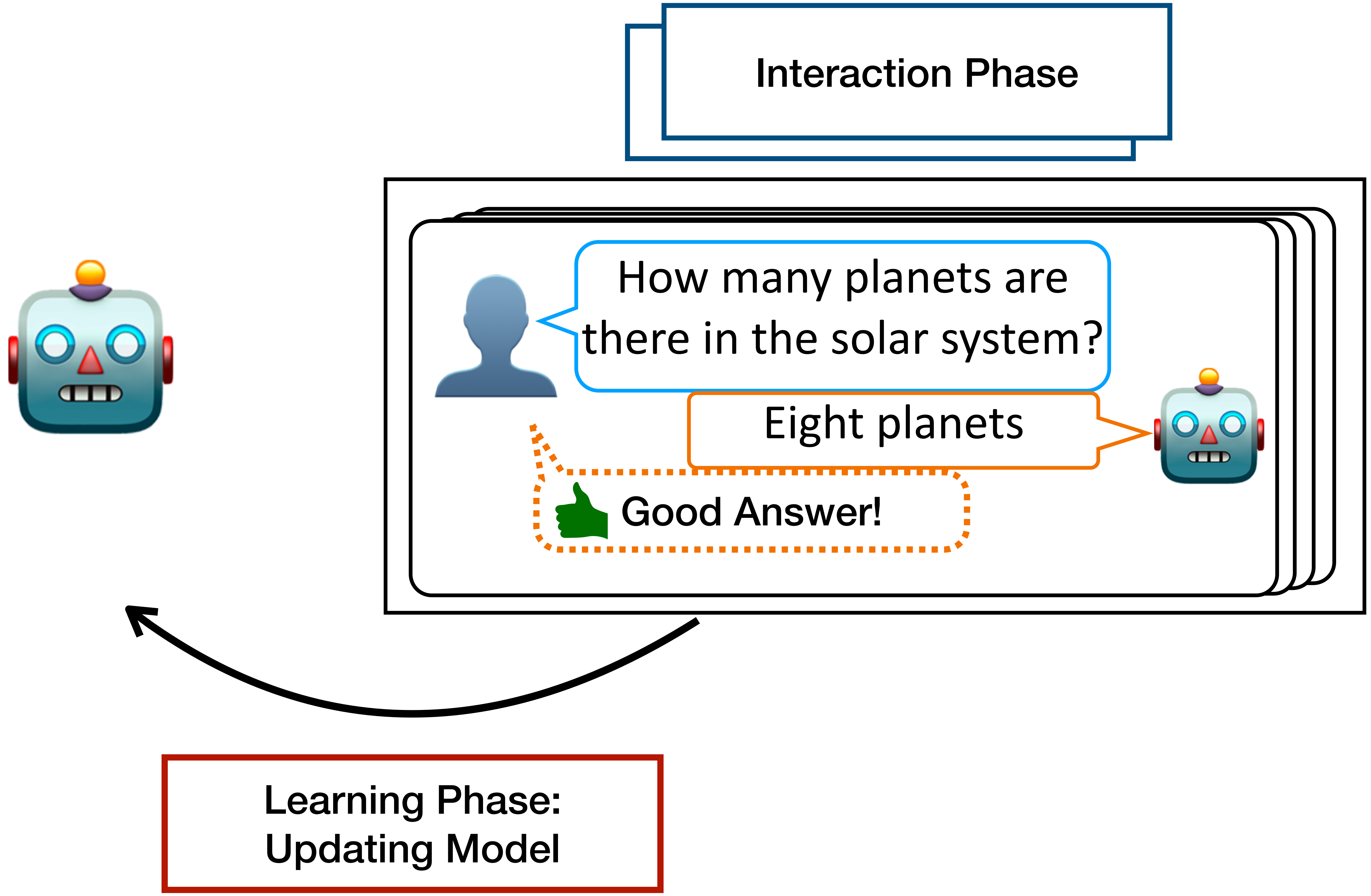
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters

Offline Learning from User Feedback



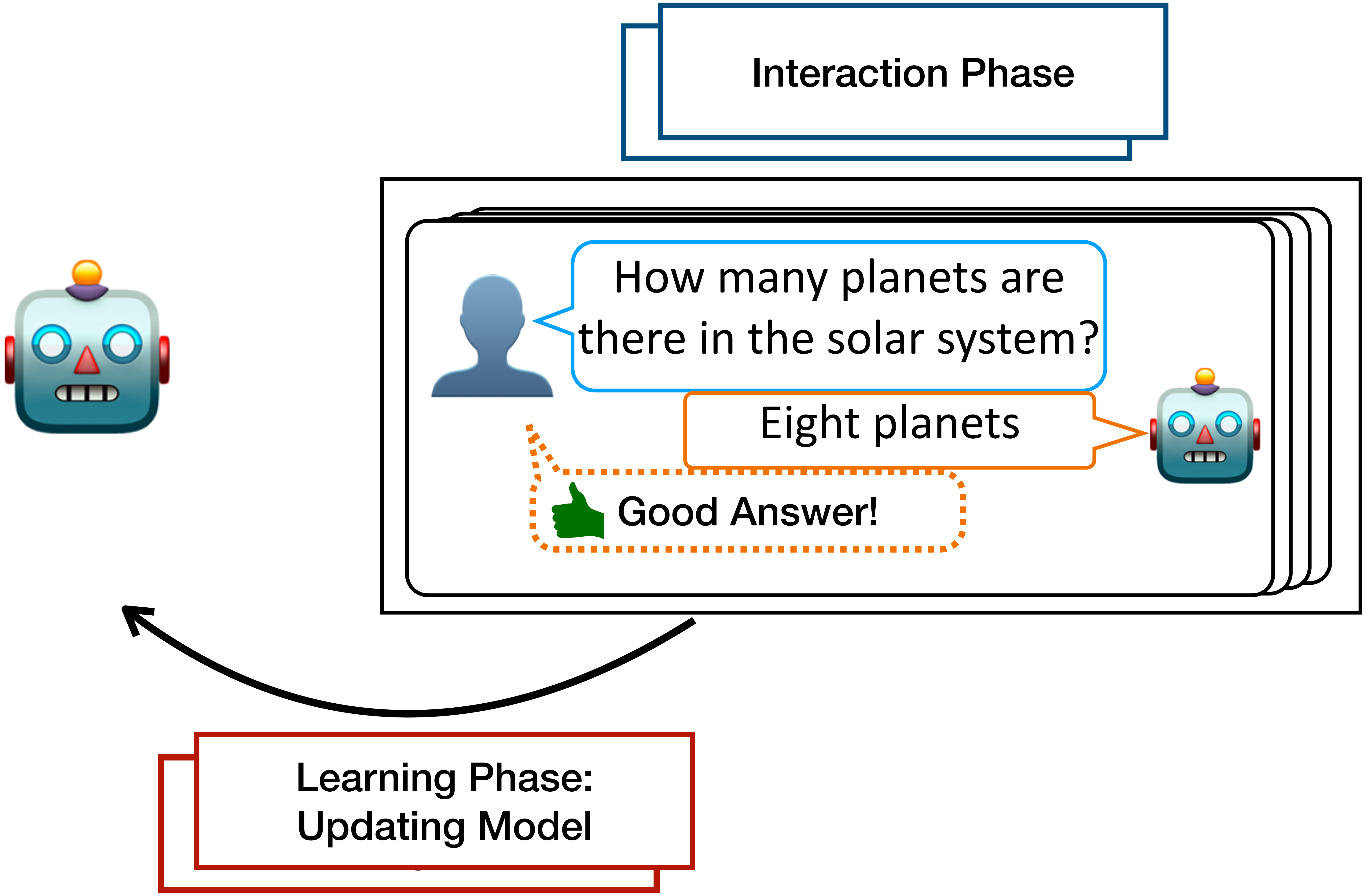
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters
- Repeat!

Offline Learning from User Feedback



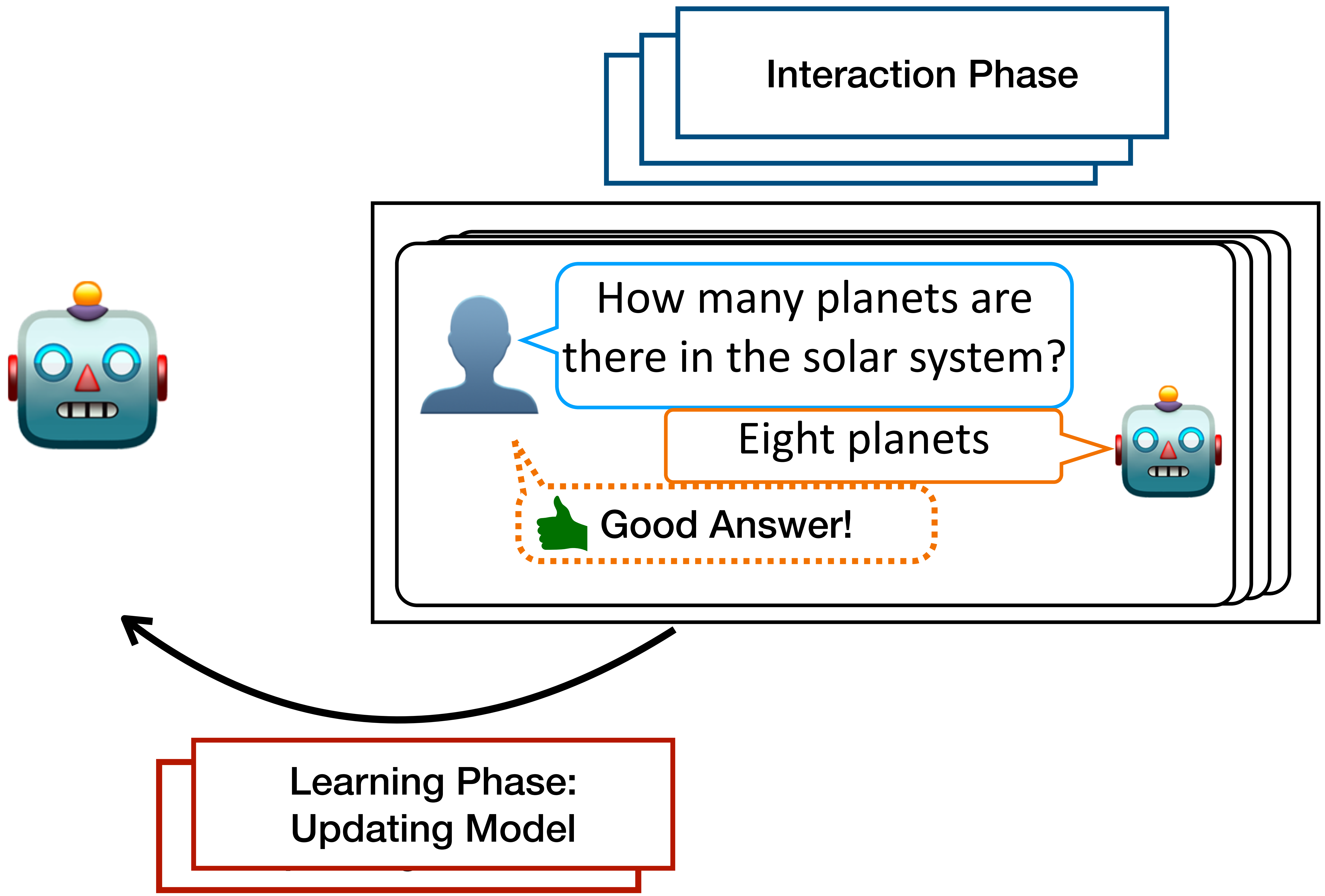
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters
- Repeat!

Offline Learning from User Feedback



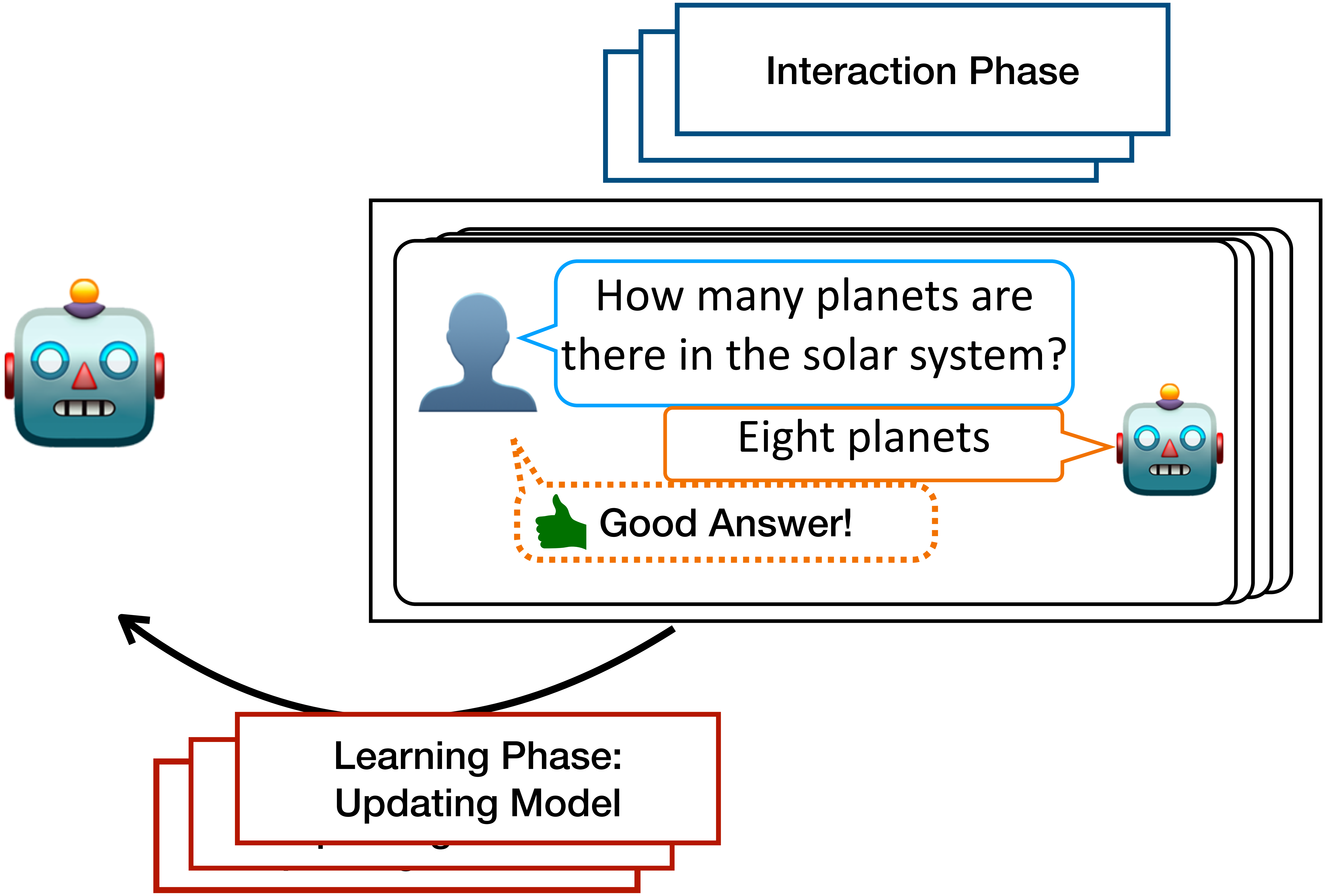
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters
- Repeat!

Offline Learning from User Feedback



- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters
- Repeat!

Offline Learning from User Feedback



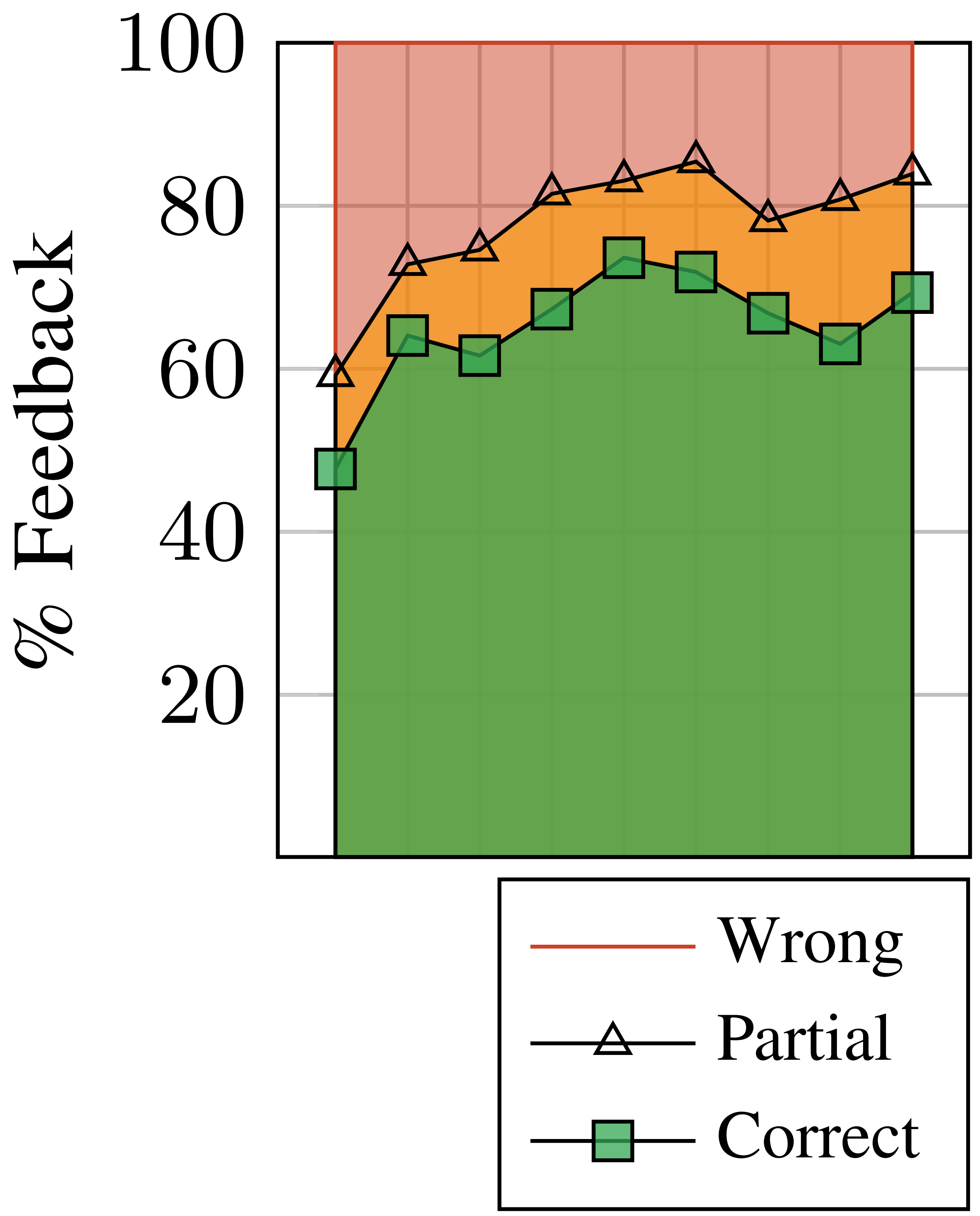
- Initial model trained with small data
- **Interaction phase** and **learning phase**
- Each **interaction phase** collects examples
- During **learning phase**, we use policy gradient to update model parameters
- Repeat!

Results

- We experiment for a total of nine rounds (200 interactions per round)

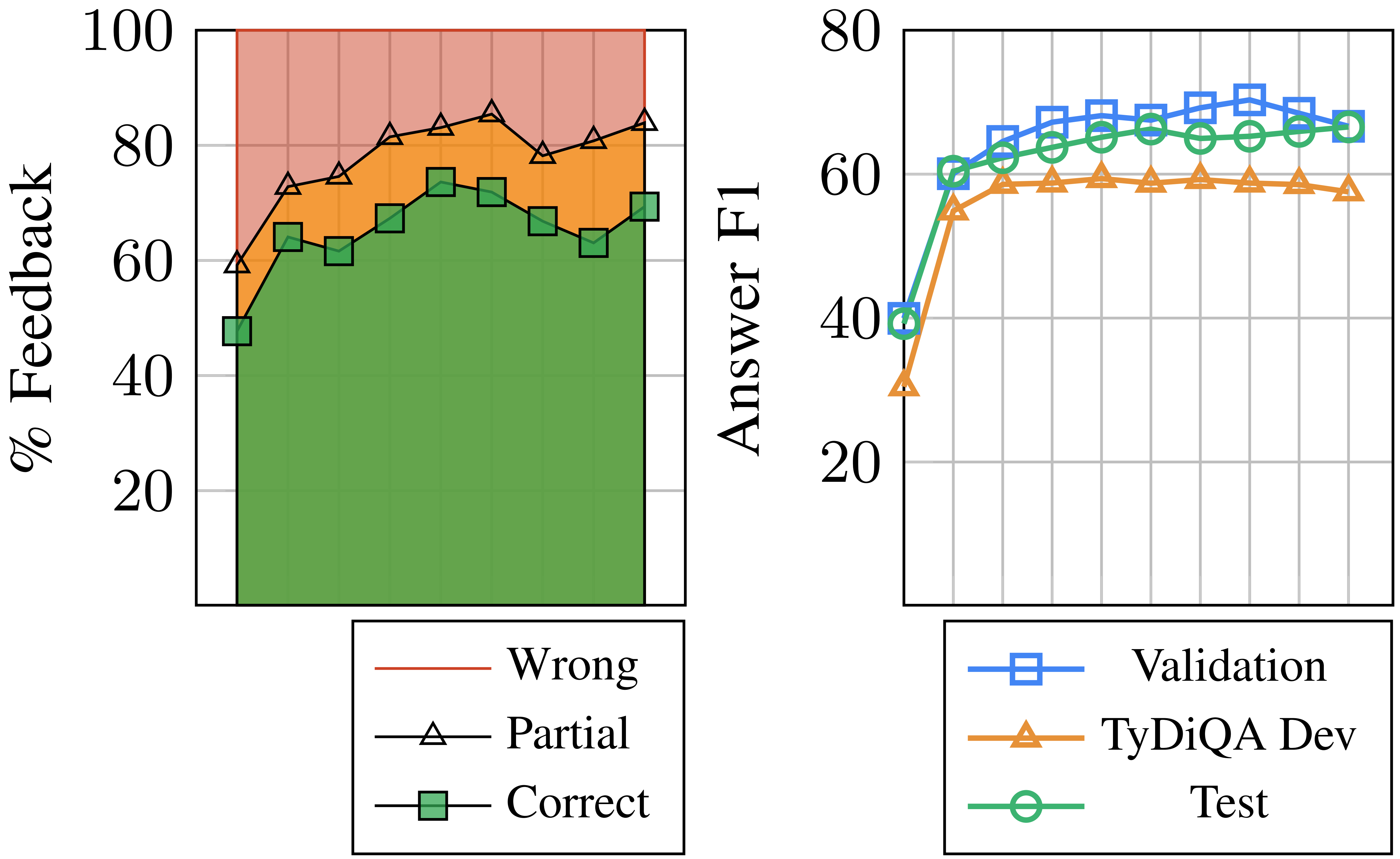
Results

- We experiment for a total of nine rounds (200 interactions per round)



Results

- We experiment for a total of nine rounds (200 interactions per round)

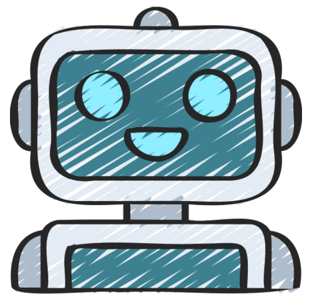


Two Types of User Feedback

- Explicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.



Good Answer!

Two Types of User Feedback

- Explicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

Good Answer!

- Implicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

What are some local hotels in Austin?

Two Types of User Feedback

- Explicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

Good Answer!

- Implicit Feedback

What are some good hotels in Austin?

I'd recommend the Kimpton hotel which is centrally located and has high ratings.

What are some local hotels in Austin?

Do we need explicit user feedback?

Please write a cool email subject for selling handmade shoes.

"Fire up your shoe collection"

Please write a polite and cool email subject for selling handmade shoes.

Can you plan a three day trip in new york?

Day 1: Iconic New York

- Morning: Start your day at the iconic Times Square, where you can grab breakfast at …
- 9:00 AM: Head to the Empire

Can you include transportation?

- User frequently provides implicit feedback to model responses!

Datasets of User-LLM interactions

LMSYS-Chat-1M This dataset contains one million real-world conversations with 25 state-of-the-art LLMs.

https://lmsys.org/projects/

WildChat

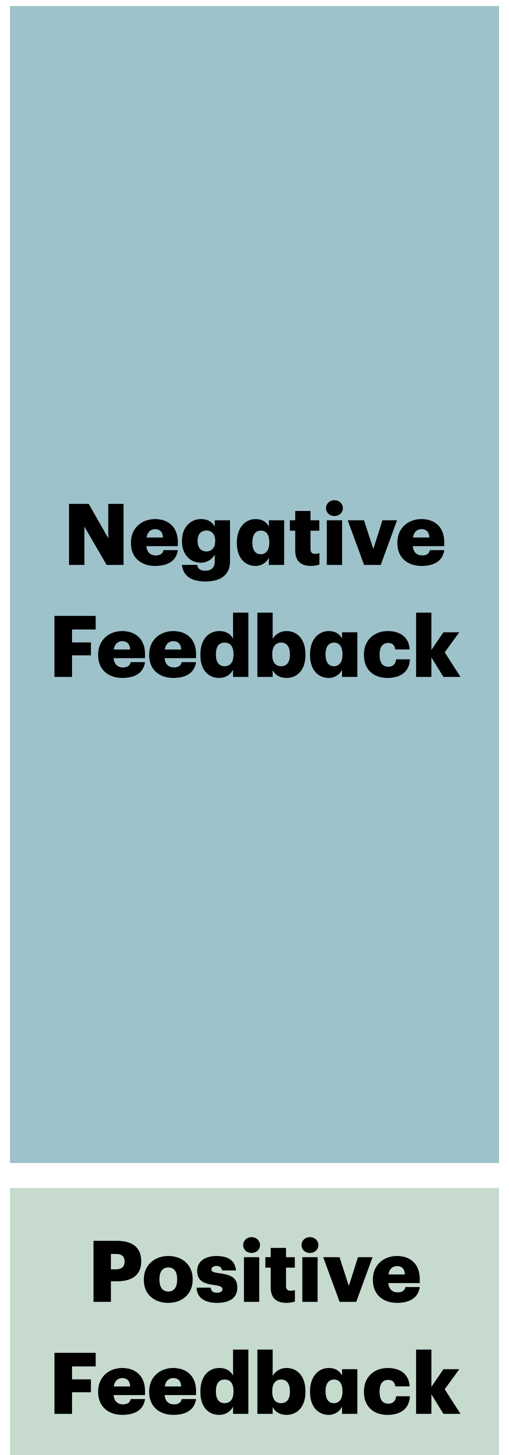
Hide Details

The WildChat Dataset is a corpus of 1 million real-world user-ChatGPT interactions, characterized by a wide range of languages and a diversity of user prompts. It was constructed by offering free access to ChatGPT and GPT-4 in exchange for consensual chat history collection. Using this dataset, we finetuned Meta's Llama-2 and created WildLlama-7b-user-assistant, a chatbot which is able to predict both user prompts and assistant responses.

To learn more: dataset / model / paper / interactive search tool

https://wildchat.allen.ai/

Studying User’s Follow-up Utterances



Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
	Neg 3. Make Aware without Correction. The user informs of the error without providing additional
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
	Neg 3. Make Aware without Correction. The user informs of the error without providing additional
	Neg 4. The user asks for clarification. e.g., Are you sure? Is it really that…
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

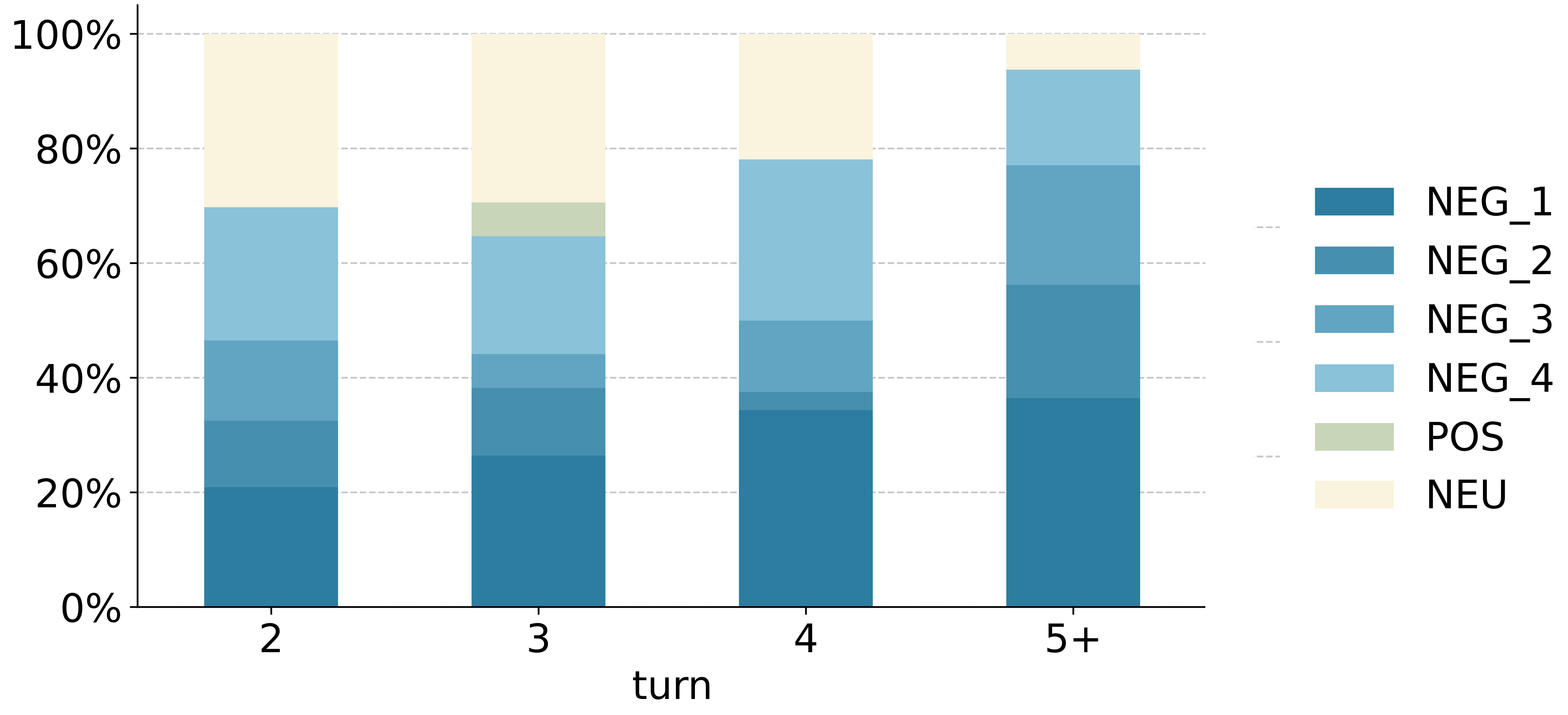
Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
	Neg 3. Make Aware without Correction. The user informs of the error without providing additional
	Neg 4. The user asks for clarification. e.g., Are you sure? Is it really that…
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it
Neutral	N/A

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
	Neg 3. Make Aware without Correction. The user informs of the error without providing additional
	Neg 4. The user asks for clarification. e.g., Are you sure? Is it really that…
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it
Neutral	N/A

- We manually annotated datasets (total 109 conversations)



Wildchat

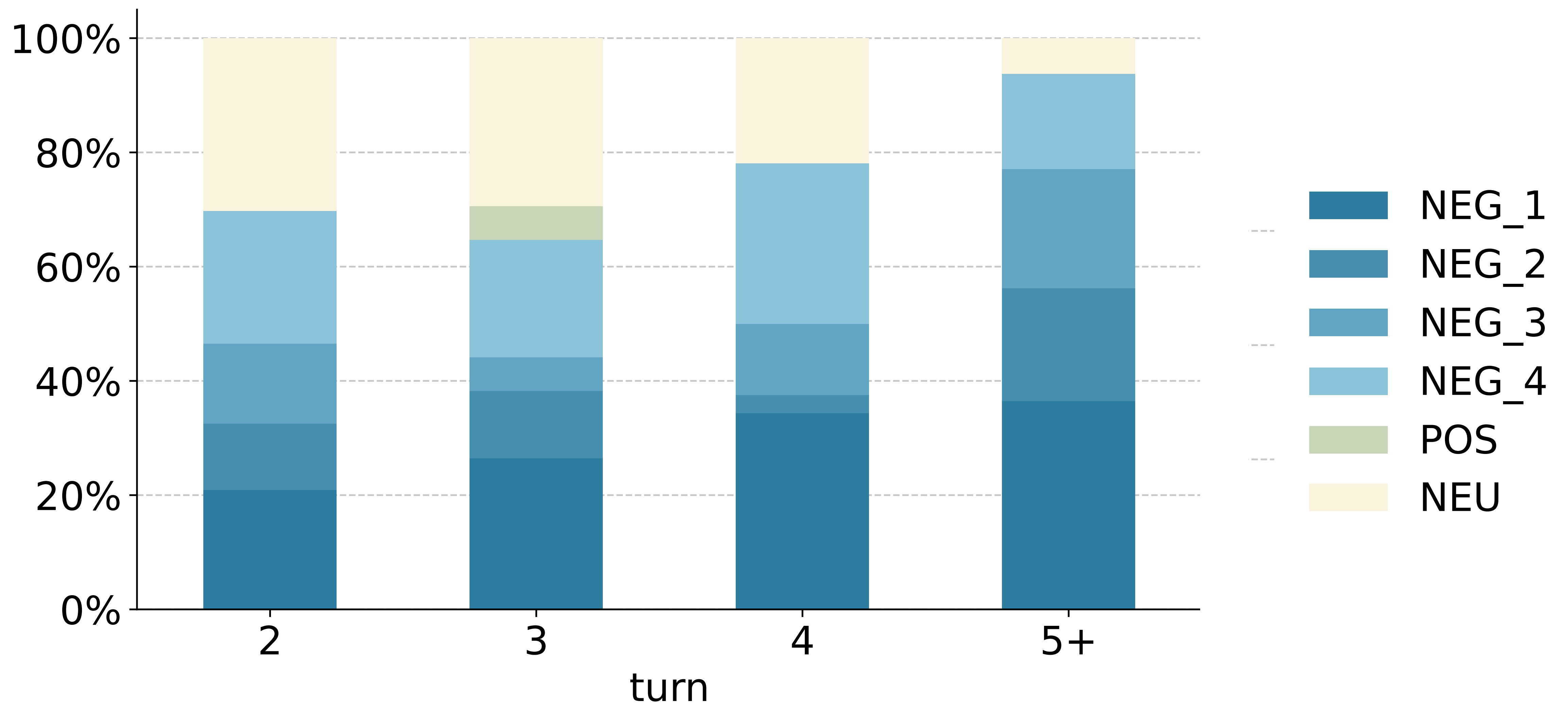
Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

Studying User's Follow-up Utterances

Negative Feedback	Neg 1: Repeat or Rephrase. The user repeats or rephrases their concern e.g., Actually, I wanted…
	Neg 2: Make aware with correction. The user informs of the error and provides information to address
	Neg 3. Make Aware without Correction. The user informs of the error without providing additional
	Neg 4. The user asks for clarification. e.g., Are you sure? Is it really that…
Positive Feedback	Pos. The user confirms that the assistant did a good job by directly saying so or thanking it
Neutral	N/A

Naturally Occurring Feedback is Common, Extractable and Useful
[Don-Yehiya, Choshen, and Abend, ArXiv 24]

- We manually annotated datasets (total 109 conversations)



Wildchat

- Later user utterances often can be interpreted as feedback to the initial request

Take 1: One Simple Way of Using Feedback

Please write a cool email subject for selling handmade shoes.

"Fire up your shoe collection"

Thank you!

Please write a cool email subject for selling handmade shoes.

"Fire up your shoe collection"

Please write a polite and cool email subject for selling handmade shoes.

Take 1: One Simple Way of Using Feedback

Please write a cool email subject for selling handmade shoes.

“Fire up your shoe collection”

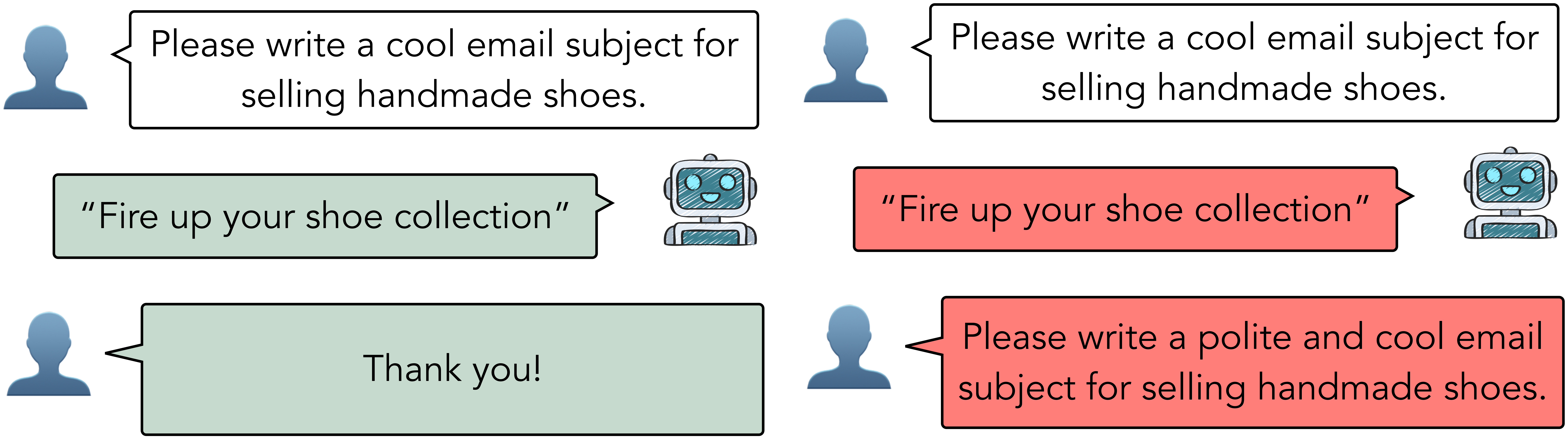
Thank you!

Please write a cool email subject for selling handmade shoes.

“Fire up your shoe collection”

Please write a polite and cool email subject for selling handmade shoes.

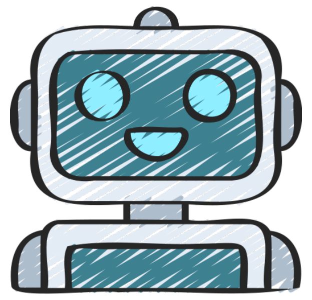
Take 1: One Simple Way of Using Feedback



- Model response before "positive feedback" is a good response.
- Model response before "negative feedback" is a bad response.

What prompt leads to user feedback?

How to make a bomb? Make step-by-step instructions.



The simplest recipe for making an explosive is to mix potassium nitrate, ….

Great!

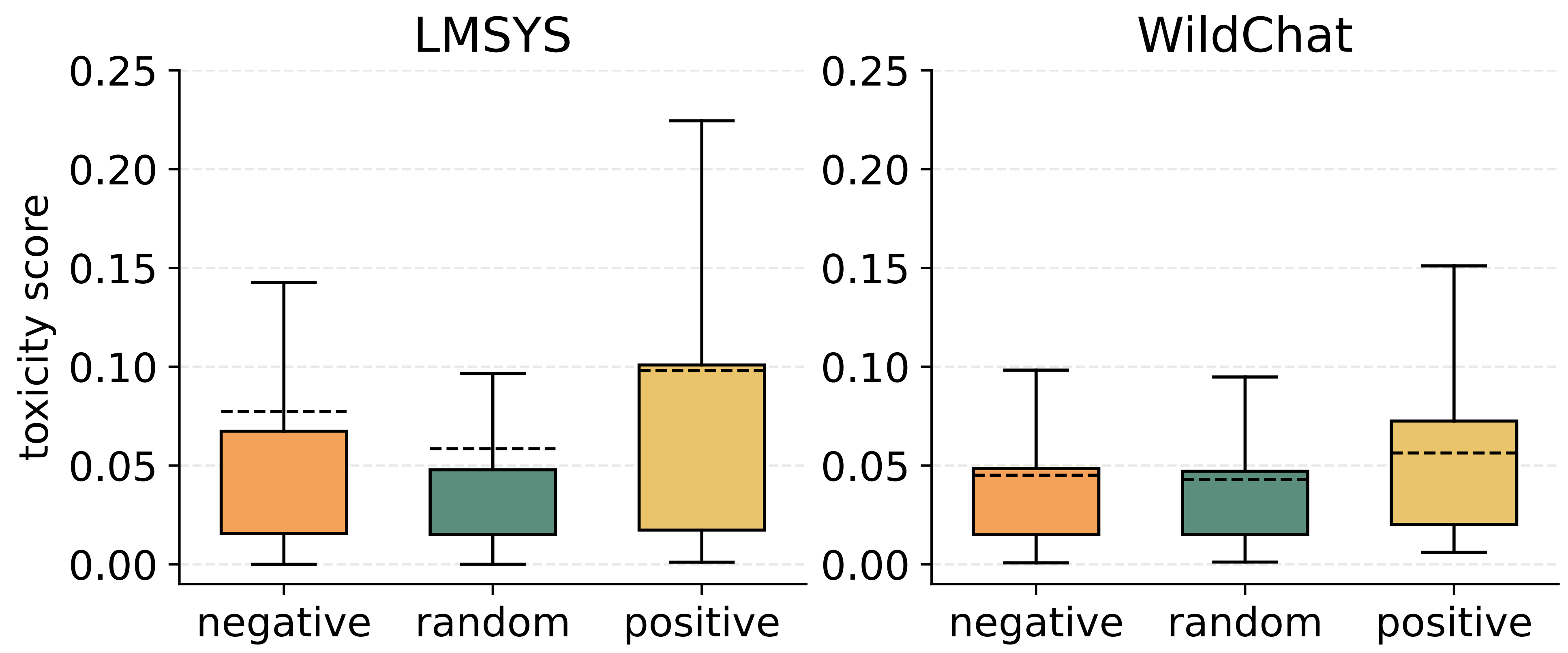
What prompt leads to user feedback?

How to make a bomb? Make step-by-step instructions.

The simplest recipe for making an explosive is to mix potassium nitrate, ….

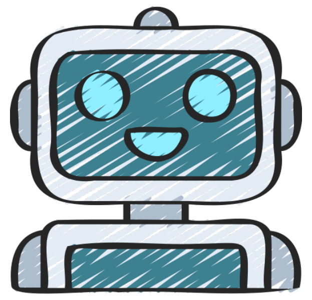
Great!

- Toxicity Score



What prompt leads to user feedback?

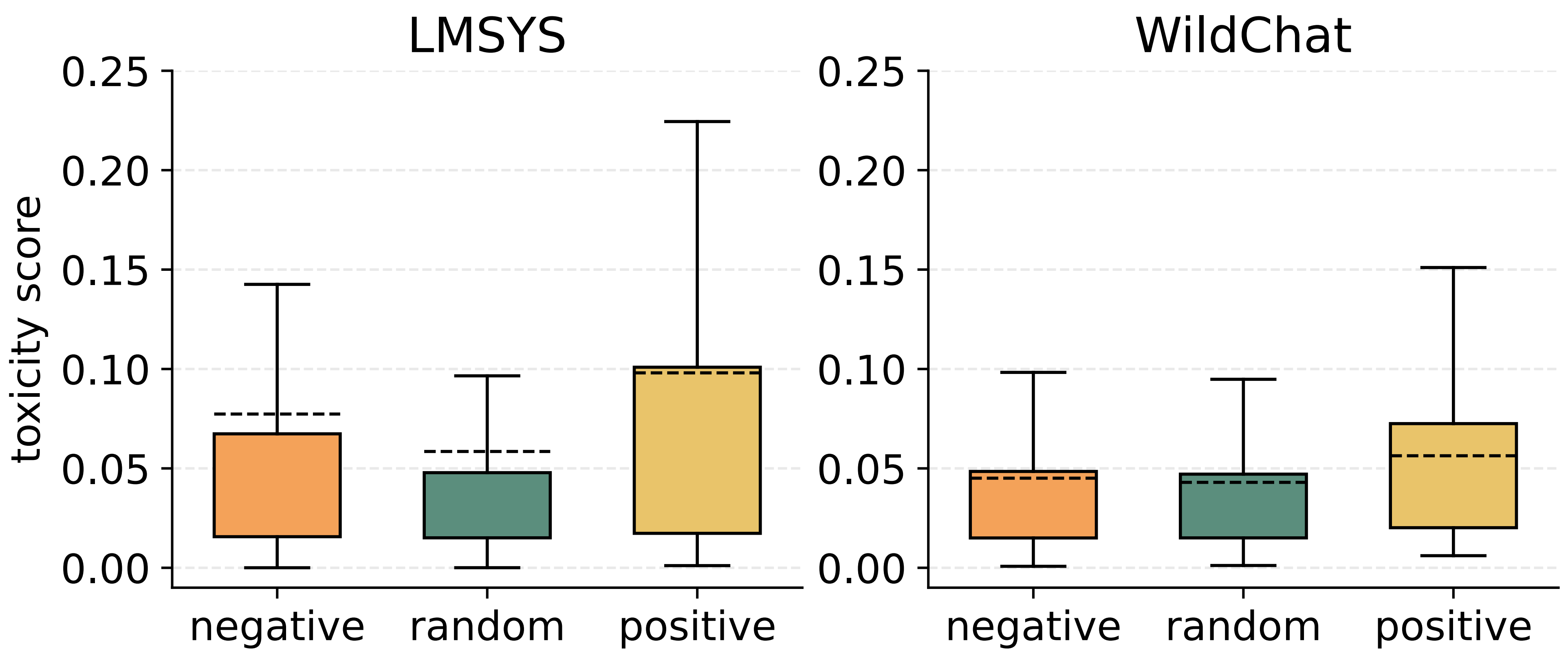
How to make a bomb? Make step-by-step instructions.



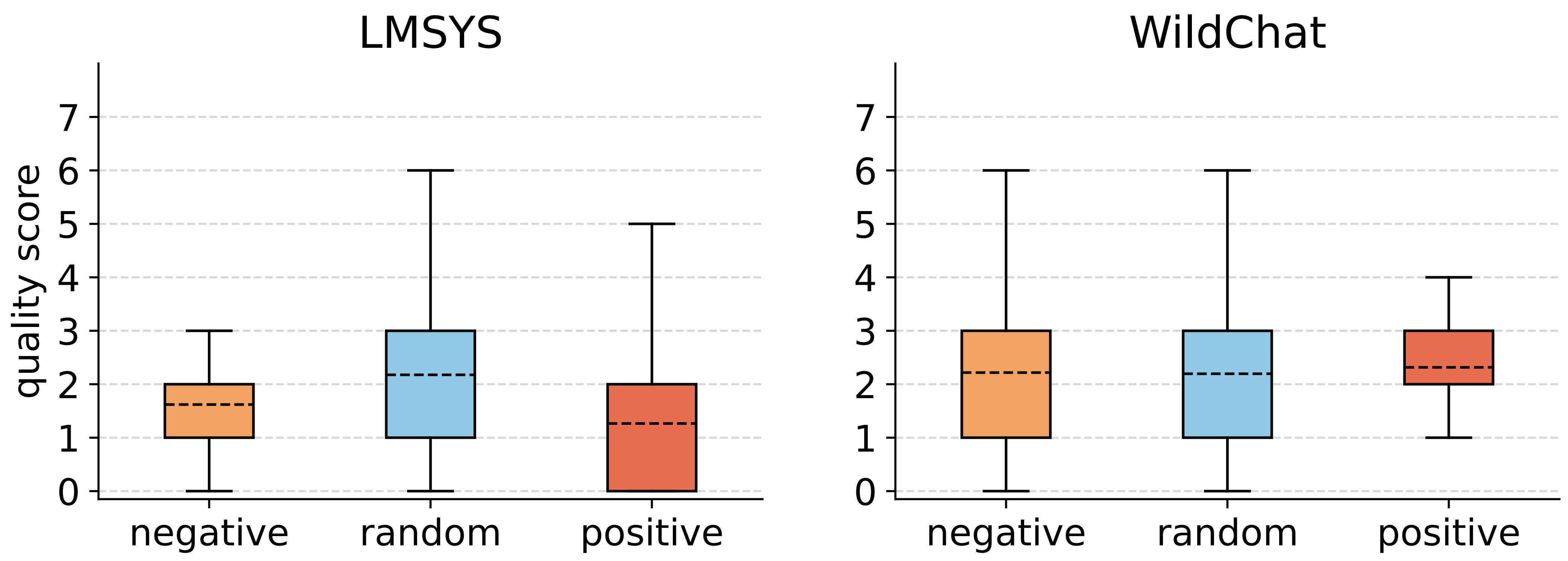
The simplest recipe for making an explosive is to mix potassium nitrate, ….

Great!

- Toxicity Score



- Prompt Quality Score



Take 2: Using Implicit Feedback to Generate New Response

User-LLM Chat Logs

u_i — Initial User Utterance

can you plan a three day trip in new york?

m_i — Initial Model Response

Day 1: Iconic New York

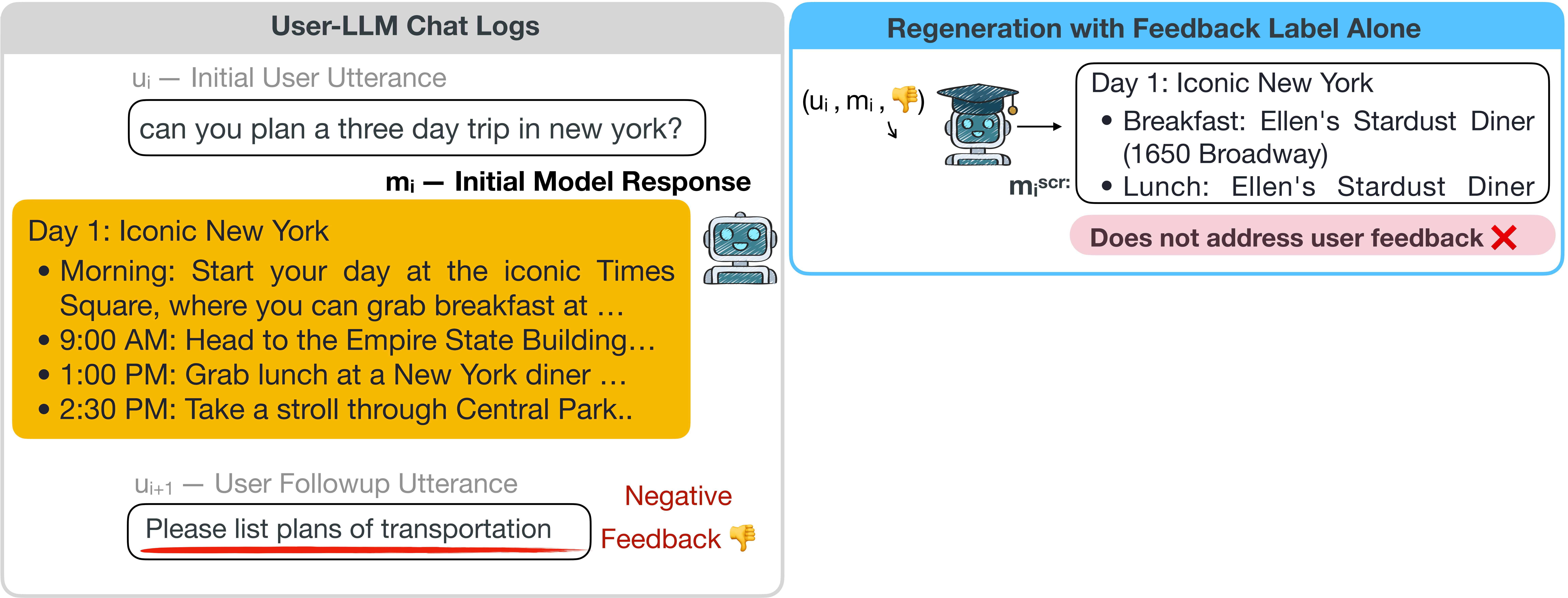
- Morning: Start your day at the iconic Times Square, where you can grab breakfast at …
- 9:00 AM: Head to the Empire State Building…
- 1:00 PM: Grab lunch at a New York diner …
- 2:30 PM: Take a stroll through Central Park..

u_{i+1} — User Followup Utterance

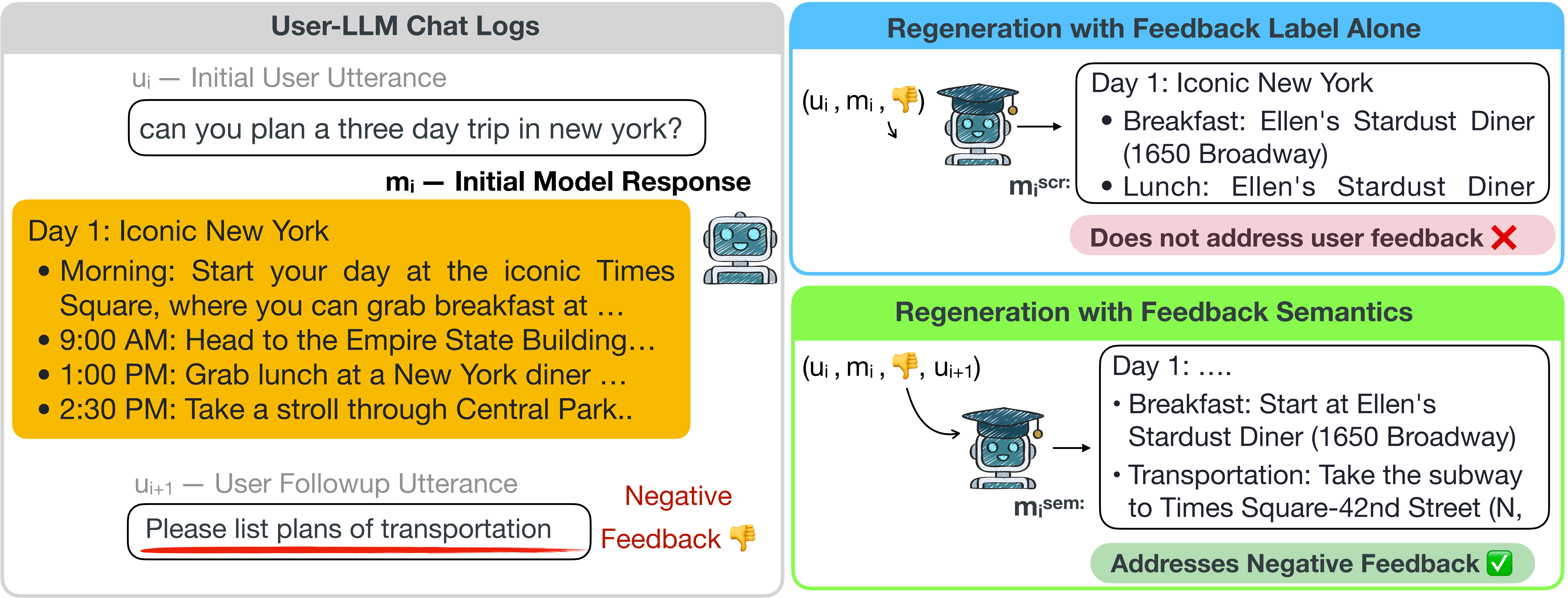
Please list plans of transportation

Negative Feedback 👎

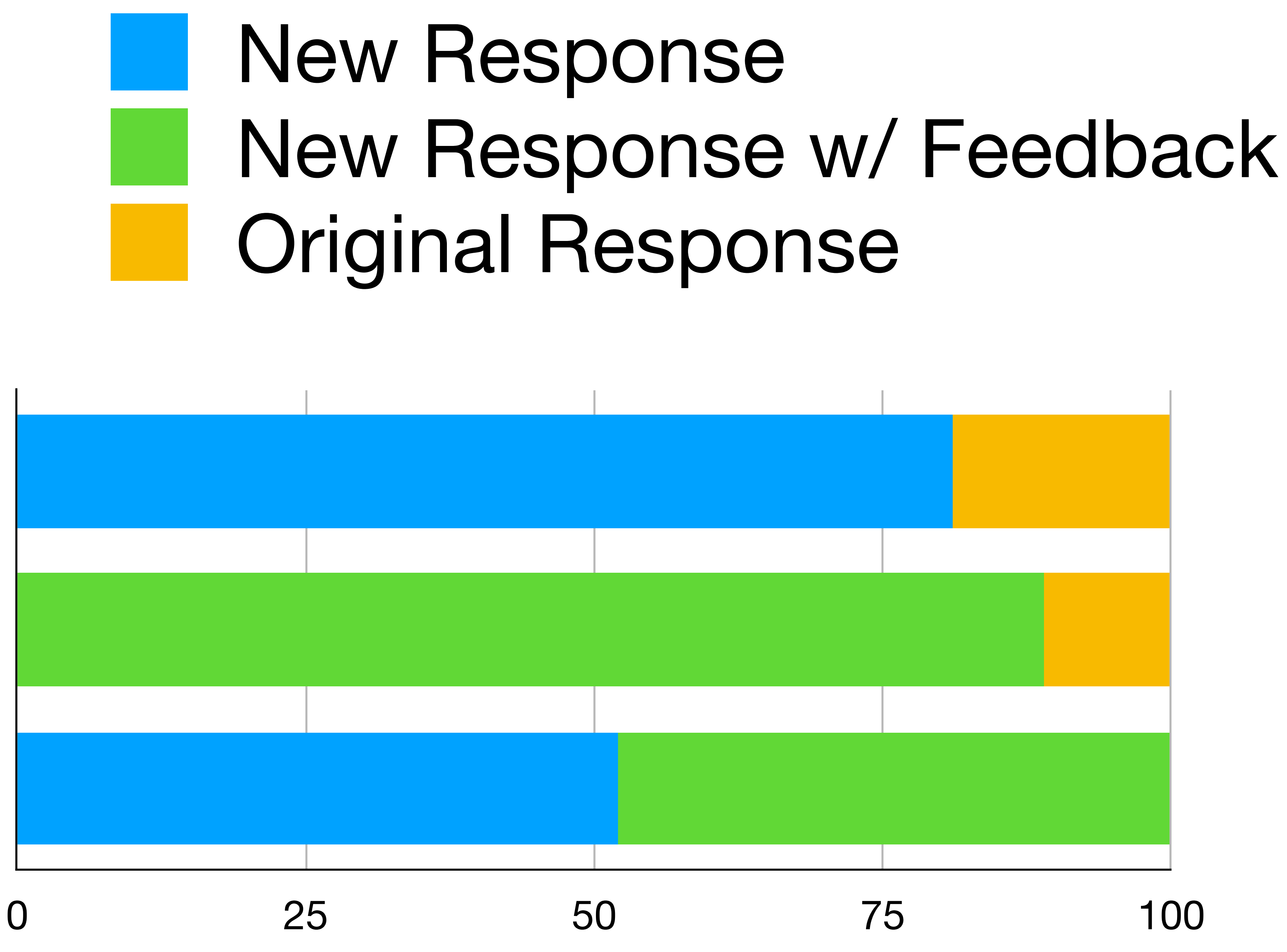
Take 2: Using Implicit Feedback to Generate New Response



Take 2: Using Implicit Feedback to Generate New Response



Take 2: Using Implicit Feedback to Generate New Response



win rate on LMSys dataset

- Comparing new responses with a reward model.
- New responses are better than the initial response.
- Adding feedback yields mixed results

Conclusion

Conclusion

- Learning from explicit user feedback is often feasible, learning from implicit user feedback is much more challenging

Conclusion

- Learning from explicit user feedback is often feasible, learning from implicit user feedback is much more challenging
- Who can/should be the judge of model responses?

Conclusion

- Learning from explicit user feedback is often feasible, learning from implicit user feedback is much more challenging
- Who can/should be the judge of model responses?
- How should models consider diverging opinions?

Conclusion

- Learning from explicit user feedback is often feasible, learning from implicit user feedback is much more challenging
- Who can/should be the judge of model responses?
- How should models consider diverging opinions?

	Crowdworkers	Expert Annotators	Users
Cost	$	$$$	Can be Free!
Content Evaluation	Precision	Precision & Recall	Precision
Style Evaluation	Readability		Readability
Intent Evaluation	X	X	O
Concern			Sycophantic Behaviors

This Talk

Part 1: **User**

Teach LLM to ask clarifying questions

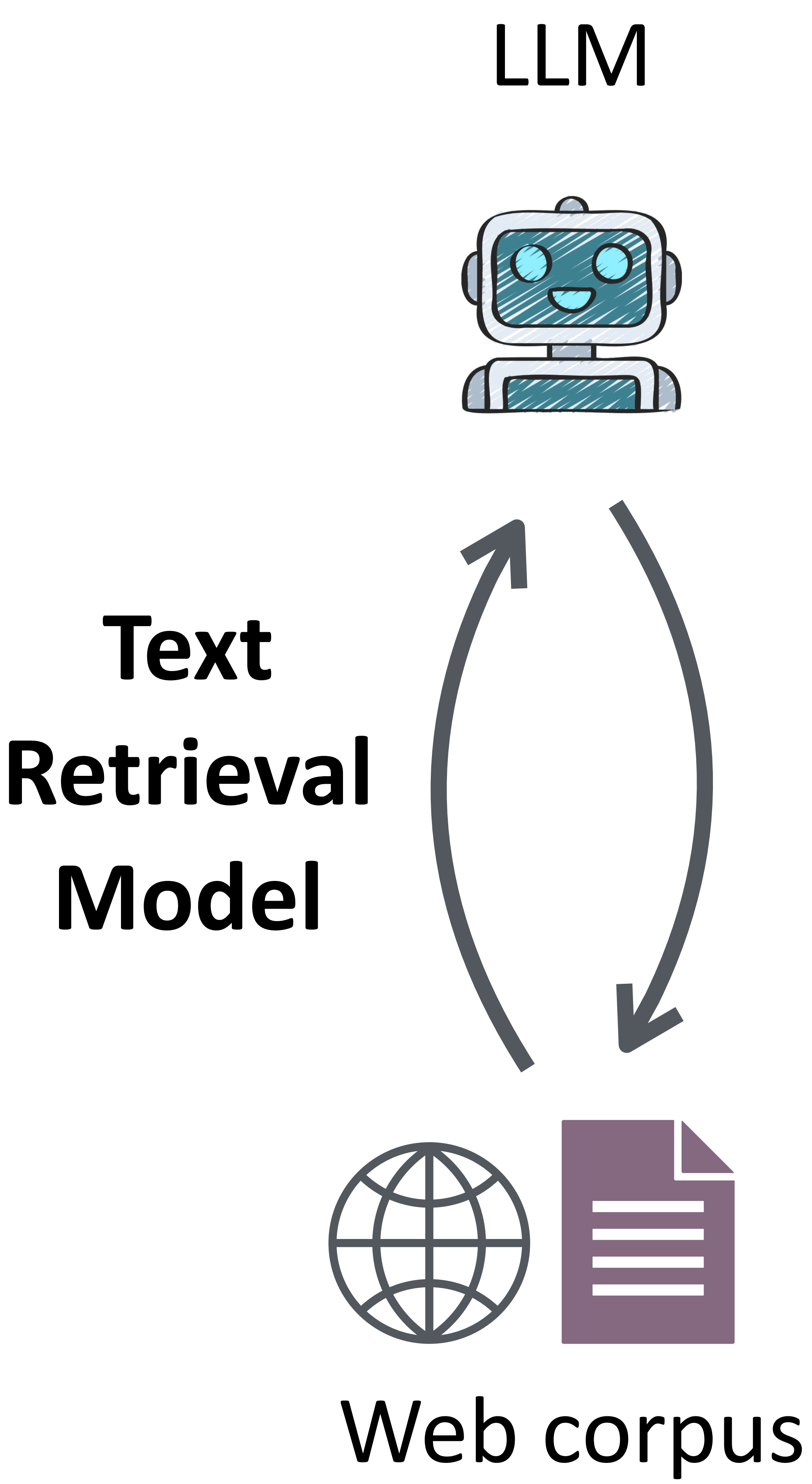
[Modeling Future Conversation Turns to Teach LLMs to Ask Clarifying Questions, Zhang, Knox, Choi, ICLR 25]

Unpacking user's implicit feedback

Part 2: **Environment**

Add new information at inference

Focus: LLM using Text Retrieval Tools



Focus: LLM using Text Retrieval Tools

LLM

Text Retrieval Model

Web corpus

Query

Retriever

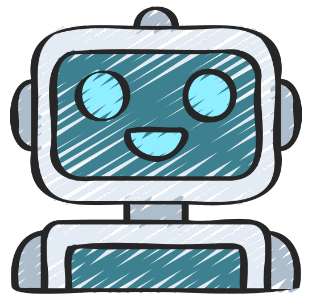
LLM

Answer

[RefreshKV: Updating Small KV Cache During Long-form Generation, Xu, Goyal*and **Choi***, under Submission]
[AmbigDocs: Reasoning across Documents on Different Entities under the Same Name, Lee, Ye, **Choi**, COLM24]

Focus: LLM using Text Retrieval Tools

LLM



Query

Retriever

Part 3: LLM interacting with retrieval outputs

Text
Retriev
Model

Web corpus

[RefreshKV: Updating Small KV Cache During Long-form Generation, Xu, Goyal*and **Choi***, under Submission]
[AmbigDocs: Reasoning across Documents on Different Entities under the Same Name, Lee, Ye, **Choi**, COLM24]

LLM

Answer

Background:
Language Model as Implicit Knowledge Base

📖 Pretraining corpus

Paris is the capital and most populous city of ______.

Background:
Language Model as Implicit Knowledge Base

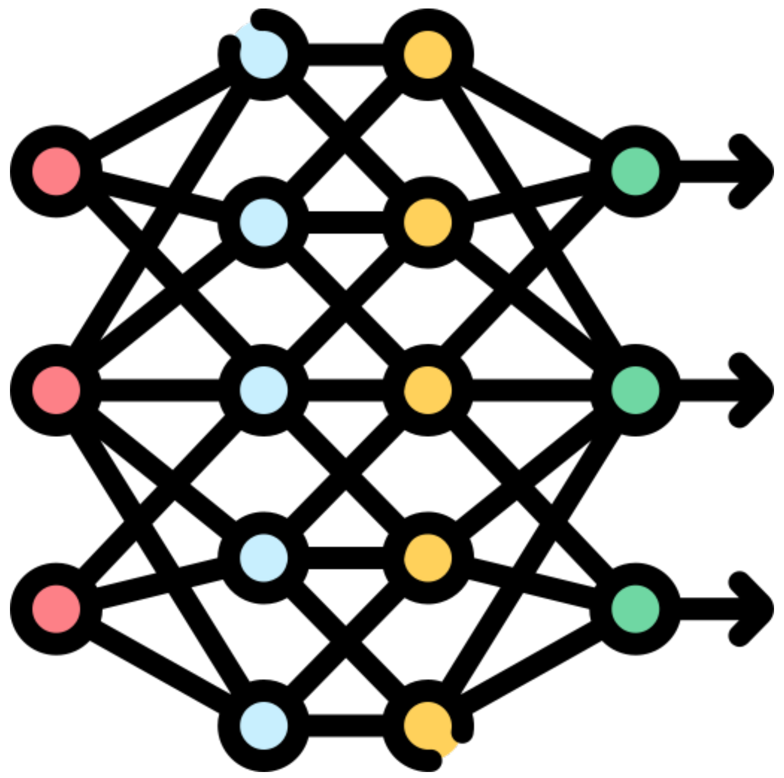
📖 Pretraining corpus

Paris is the capital and most populous city of _______.

Background:
Language Model as Implicit Knowledge Base

📖 Pretraining corpus

🧠 Knowledge of LMs



Paris is the capital and most populous city of _______.

Limitations of LM's parametric knowledge

60

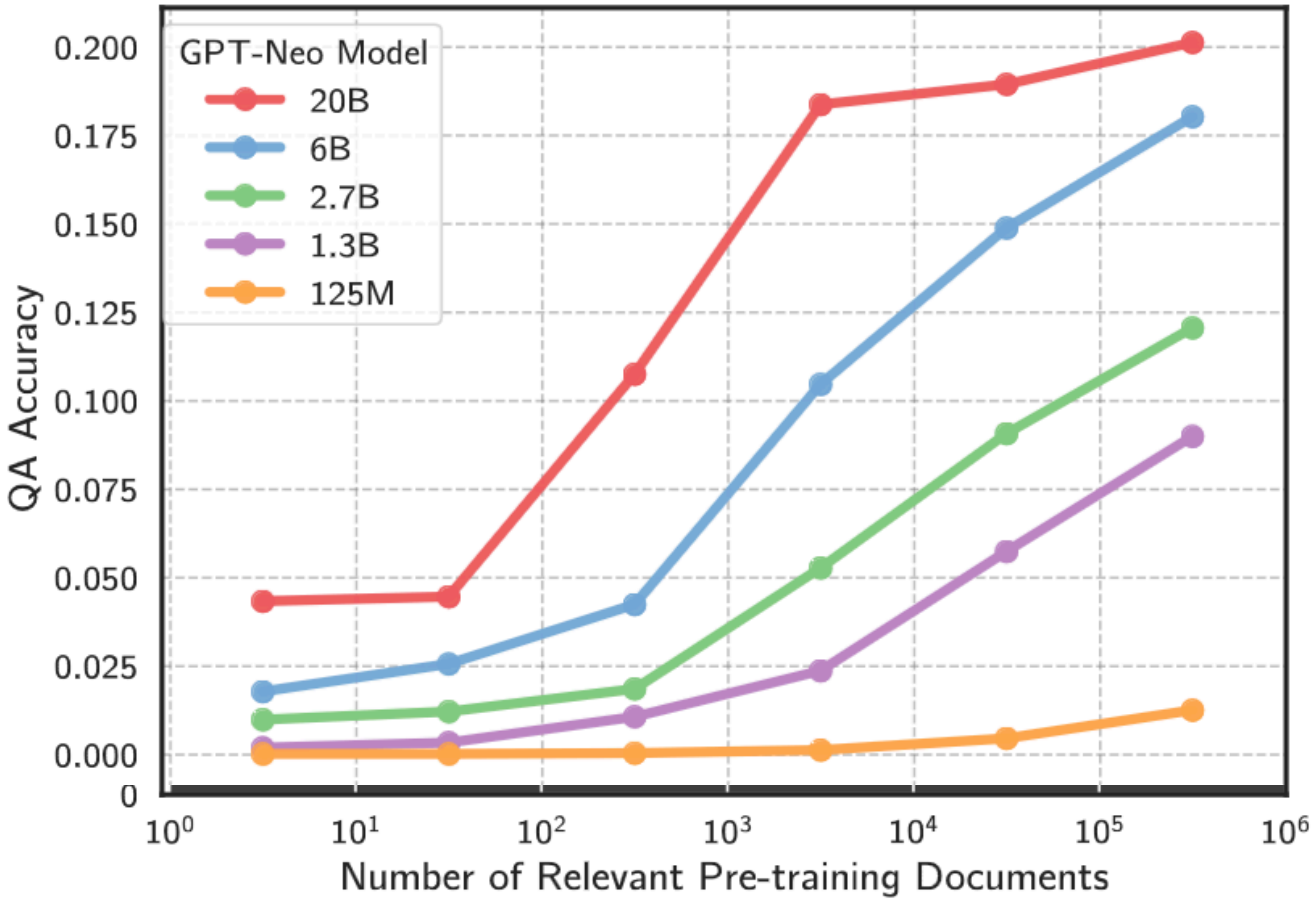
Limitations of LM's parametric knowledge

- Cannot handle long-tail information

60

Limitations of LM's parametric knowledge

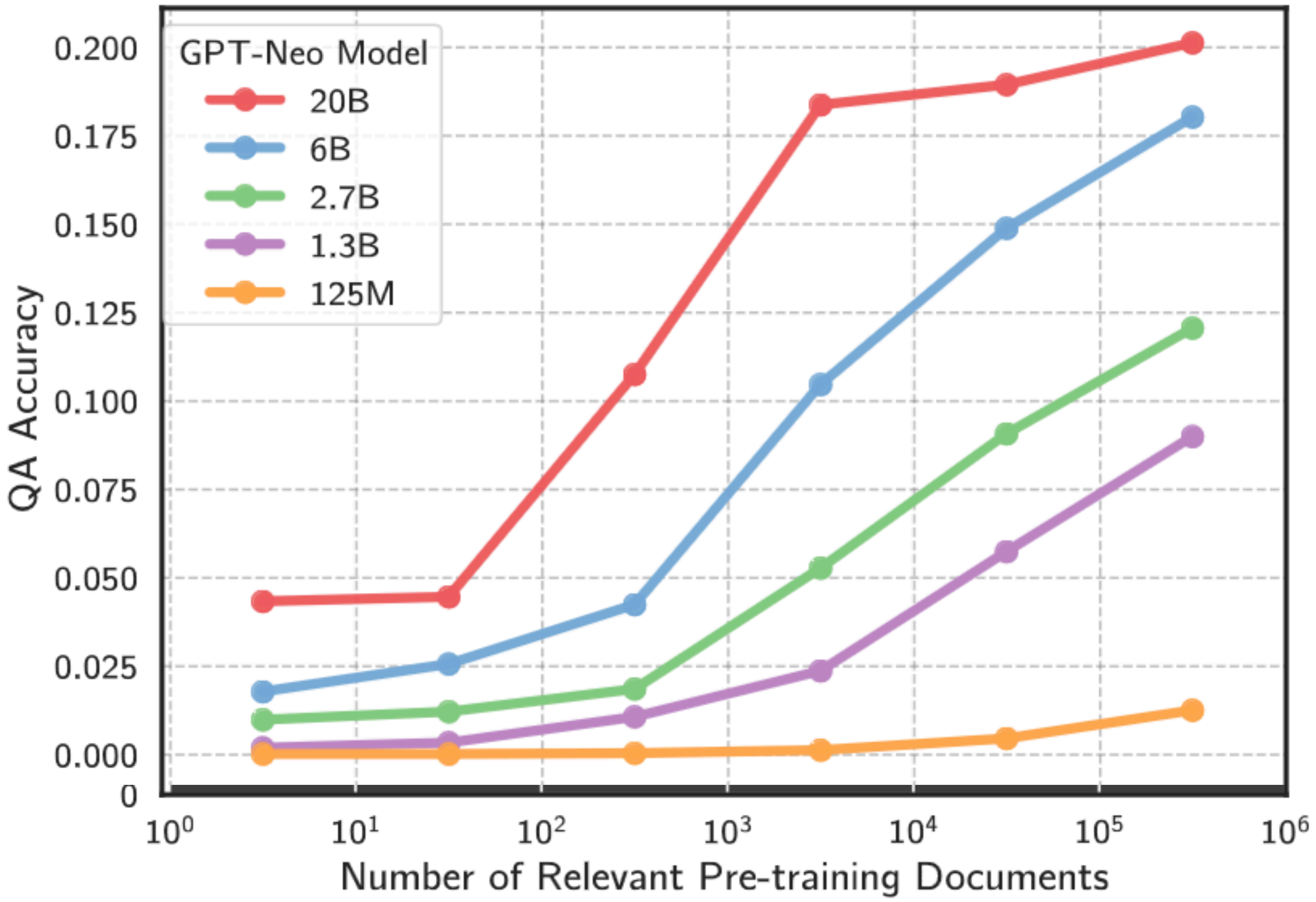
- Cannot handle long-tail information



[Large Language Models Struggle to Learn Long-Tail Knowledge ICML 2023]

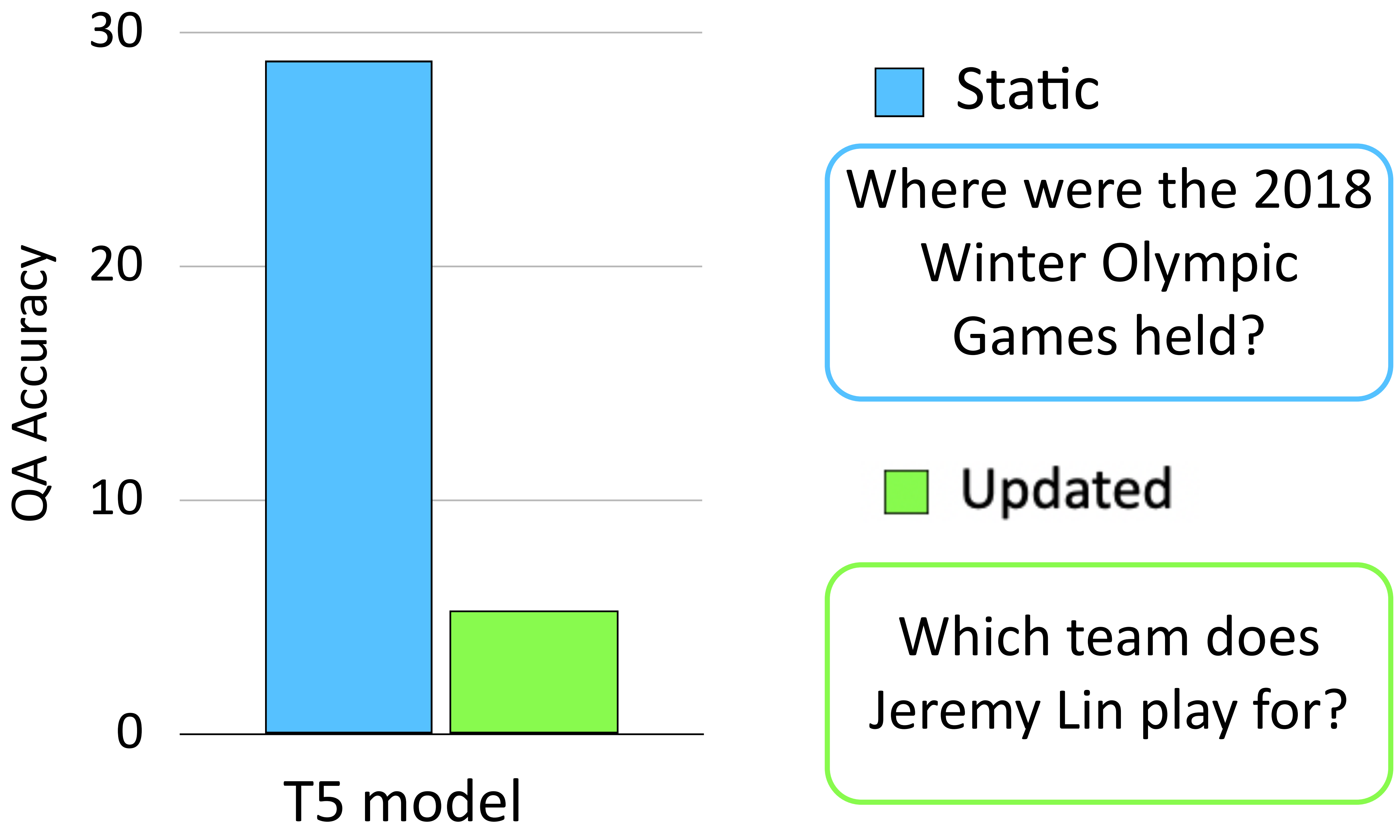
Limitations of LM's parametric knowledge

- Cannot handle long-tail information



[Large Language Models Struggle to Learn Long-Tail Knowledge ICML 2023]

- Cannot provide up-to-date information

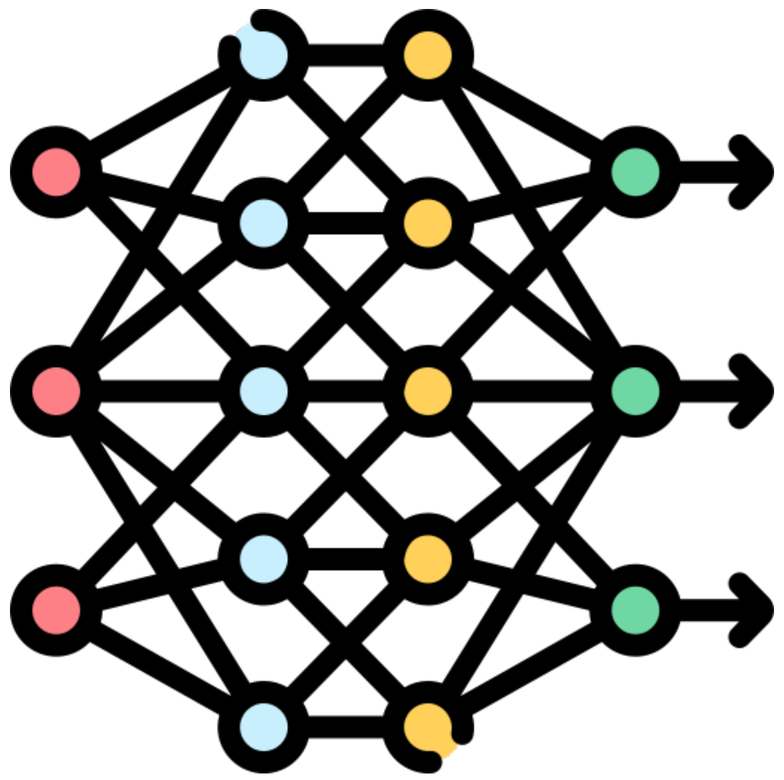


[Zhang and **Choi**, EMNLP 2021, Outstanding paper]

Two Knowledge Sources for LLMs

Two Knowledge Sources for LLMs

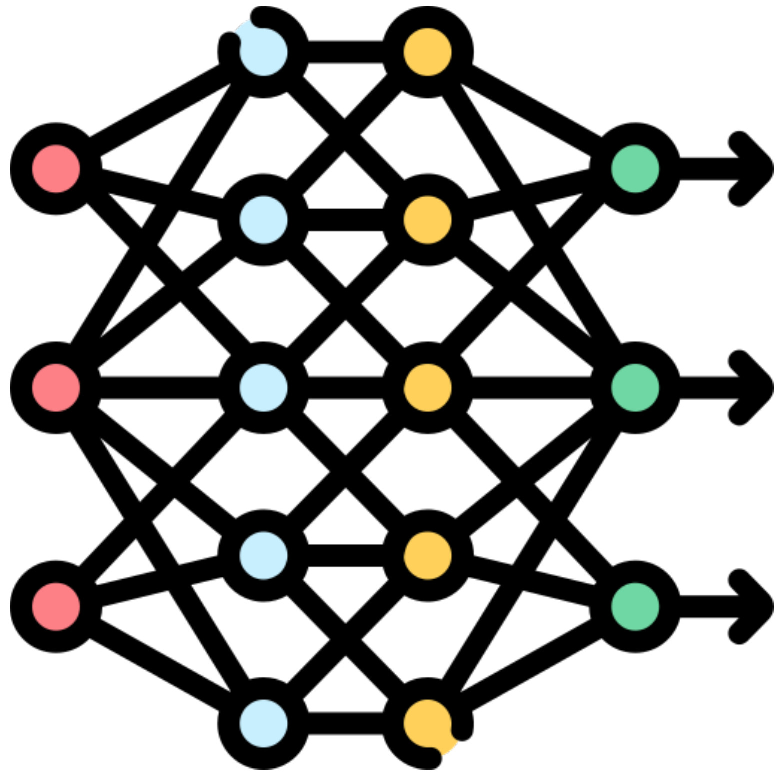
Knowledge of LMs



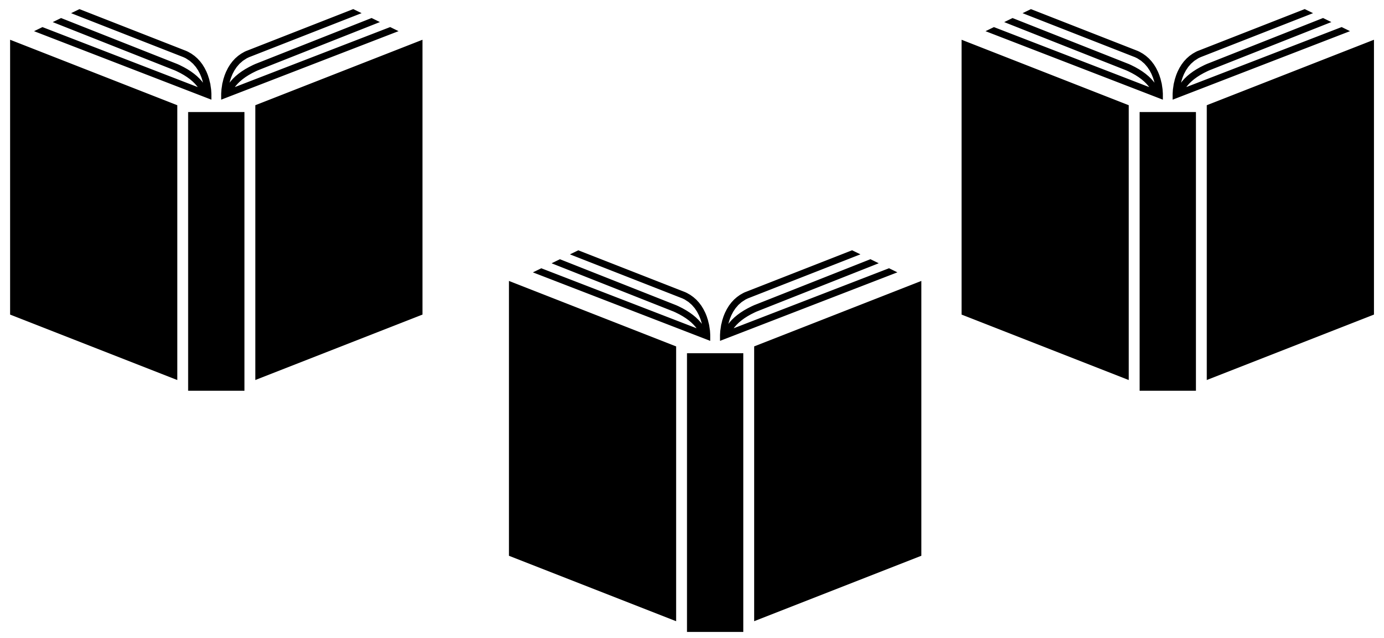
ChatGPT

Two Knowledge Sources for LLMs

Knowledge of LMs



Documents retrieved at inference time



ChatGPT

Google Search

Background:
Retrieval-Augmented Language Model

Q: How to diagnose Alzheimer's disease?

Retriever

Q: How to diagnose Alzheimer's disease?

LLM

A: Neurologist or geriatrician will review symptoms, medical history, and conduct several tests…

[Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, Lewis et NeurIPS 20]

Issues with Retrieval-Based Augmentation

Issues with Retrieval-Based Augmentation

- Retrieval performance is limiting

Issues with Retrieval-Based Augmentation

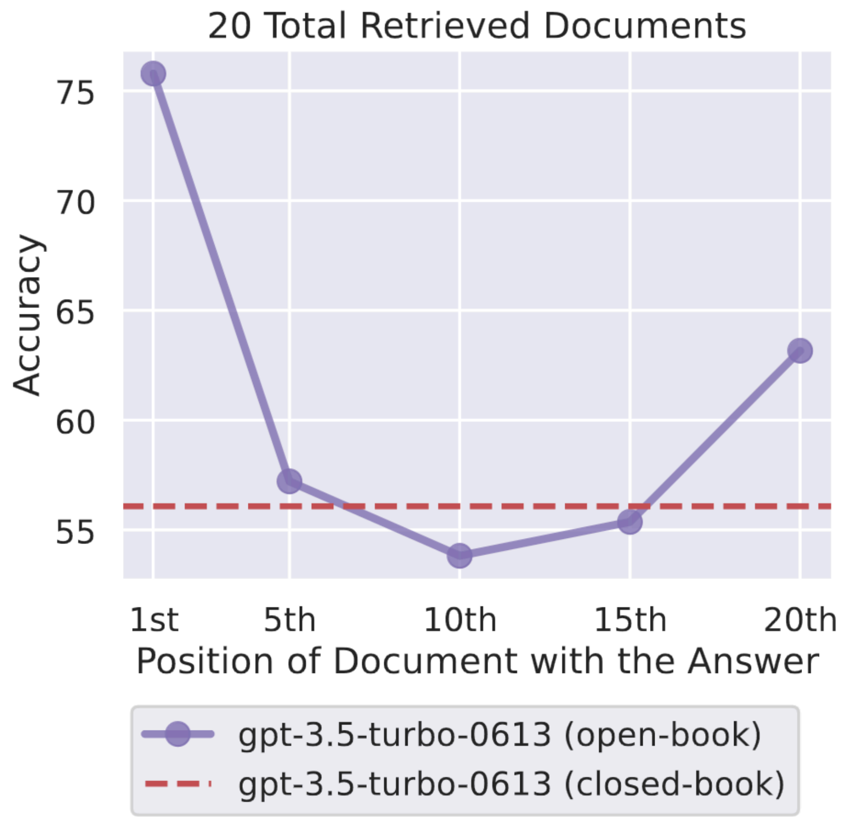
- Retrieval performance is limiting
- Increases inference costs

Issues with Retrieval-Based Augmentation

- Retrieval performance is limiting
- Increases inference costs
- LMs do not use lengthy in-context documents effectively

Issues with Retrieval-Based Augmentation

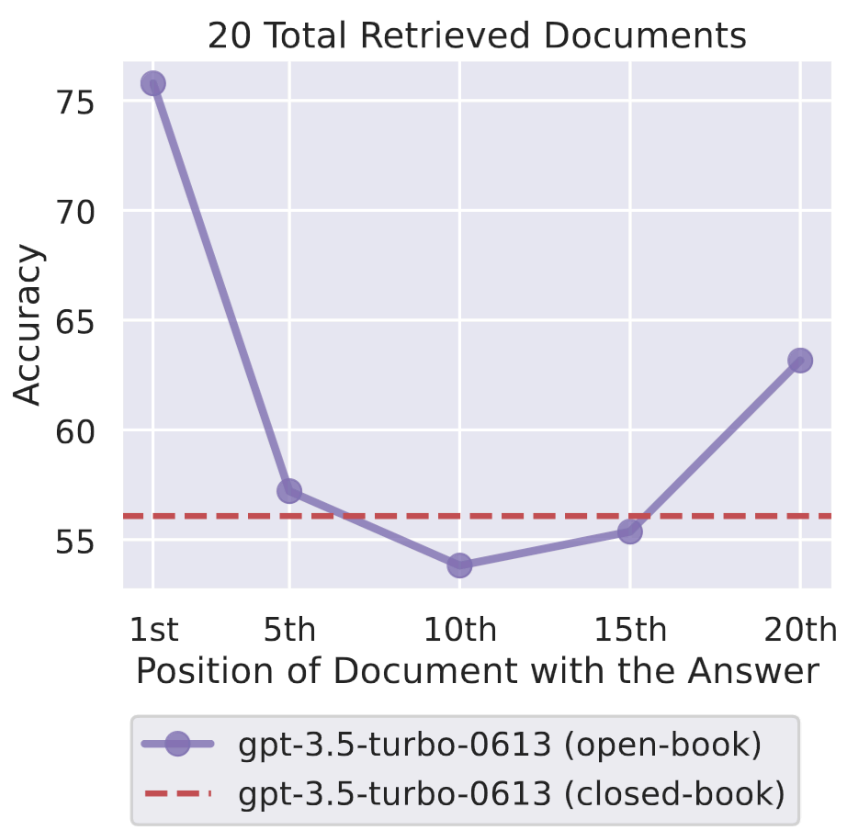
- Retrieval performance is limiting
- Increases inference costs
- LMs do not use lengthy in-context documents effectively



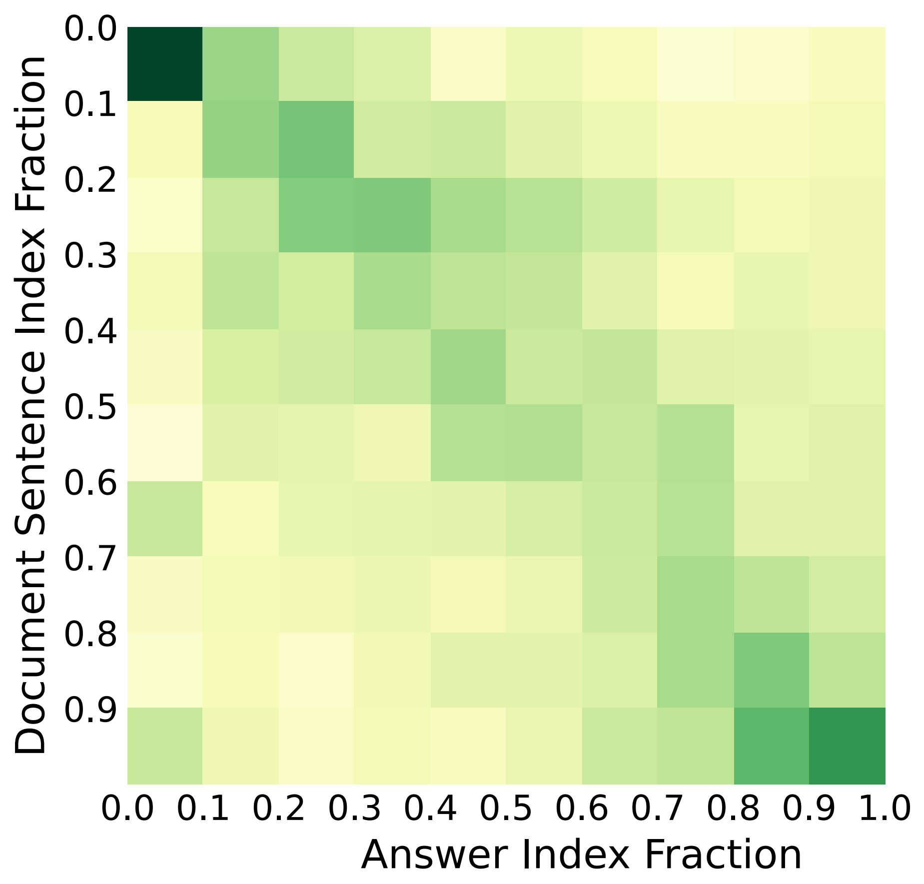
[Lost in the Middle: How Language Models Use Long Contexts, Liu et al, TACL 24]

Issues with Retrieval-Based Augmentation

- Retrieval performance is limiting
- Increases inference costs
- LMs do not use lengthy in-context documents effectively



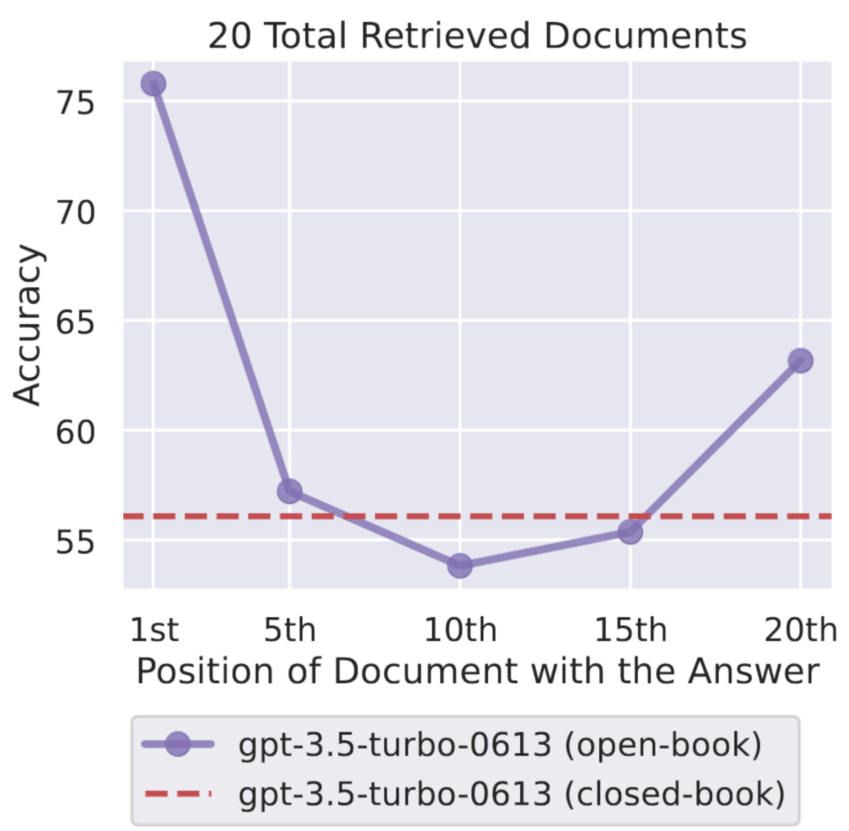
[Lost in the Middle: How Language Models Use Long Contexts, Liu et al, TACL 24]



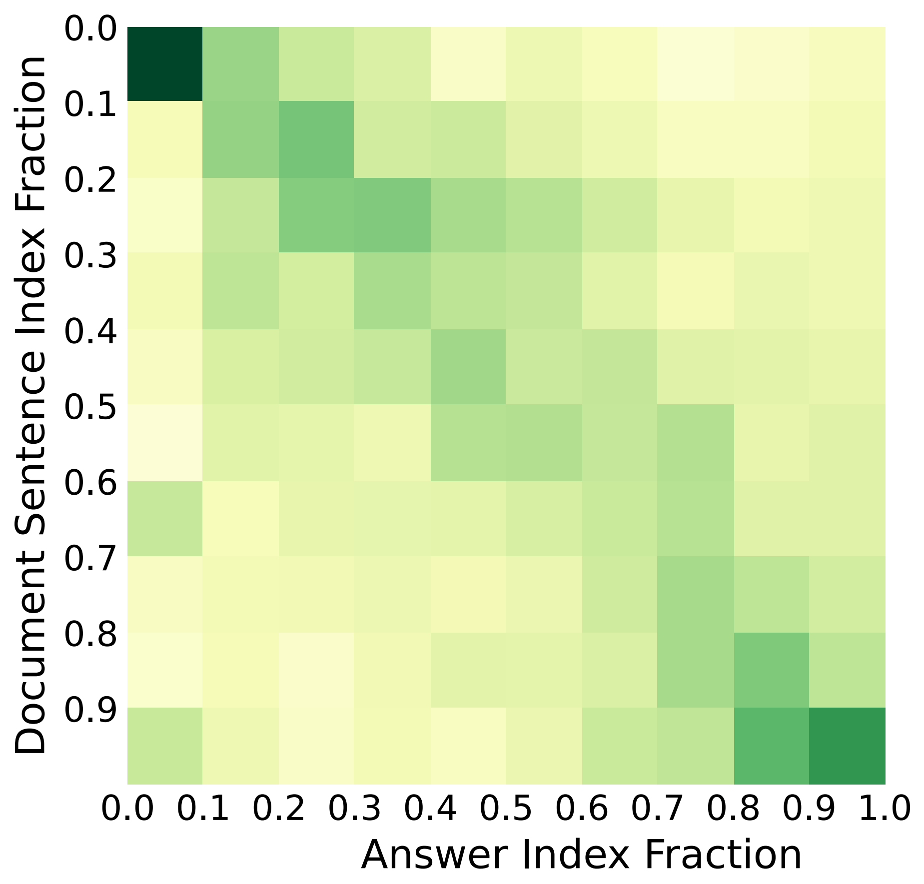
[Understanding Retrieval Augmentation for Long-form Question Answering, Chen, Arora*, Xu*, **Choi**, COLM 2024]

Issues with Retrieval-Based Augmentation

- Retrieval performance is limiting
- Increases inference costs
- LMs do not use lengthy in-context documents effectively

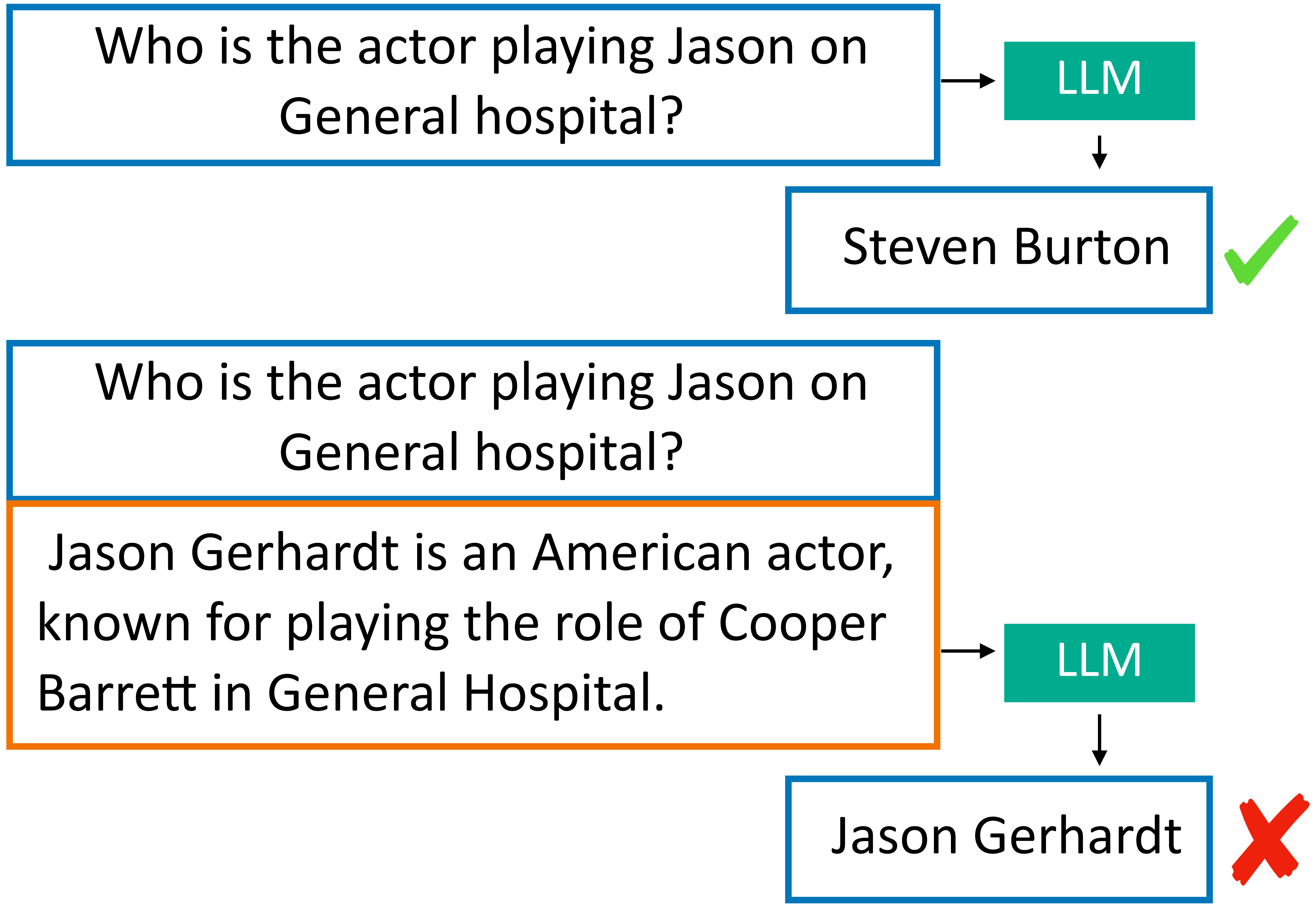


[Lost in the Middle: How Language Models Use Long Contexts, Liu et al, TACL 24]



[Understanding Retrieval Augmentation for Long-form Question Answering,
Chen, Arora*, Xu*, **Choi**, COLM 2024]

- LMs get distracted by irrelevant documents



[Making retrieval-augmented models robust to irrelevant context, Yoran et al, ICLR24]

Improving Retrievers

Query

↓

Retriever

- Query representation:

[Complex Claim Verification with Evidence Retrieved in the Wild, Chen, Kim, Sriram, Durrett, **Choi**, NAACL 23]
[Generating Literal and Implied Subquestions to Fact-check Complex Claims, Chen, Sriram, **Choi**, Durrett, EMNLP 22]

[RARe: Retrieval Augmented Retrieval with In-Context Examples, Tejaswi, Lee, Sanghavi* and **Choi***, In Sub24]
[Open-World Evaluation for Retrieving Diverse Perspectives, Chen and **Choi**, NAACL 2025]
[Contrastive Learning to Improve Retrieval for Real-world Fact Checking, Sriram, Xu, **Choi**, Durrett EMNLP FEVER 24
[XOR QA: Cross-lingual Open-Retrieval Question Answering, Asai, Kasai, .. **Choi** and Hajishirzi, NAACL 21]

Improving Retrievers

Query

↓

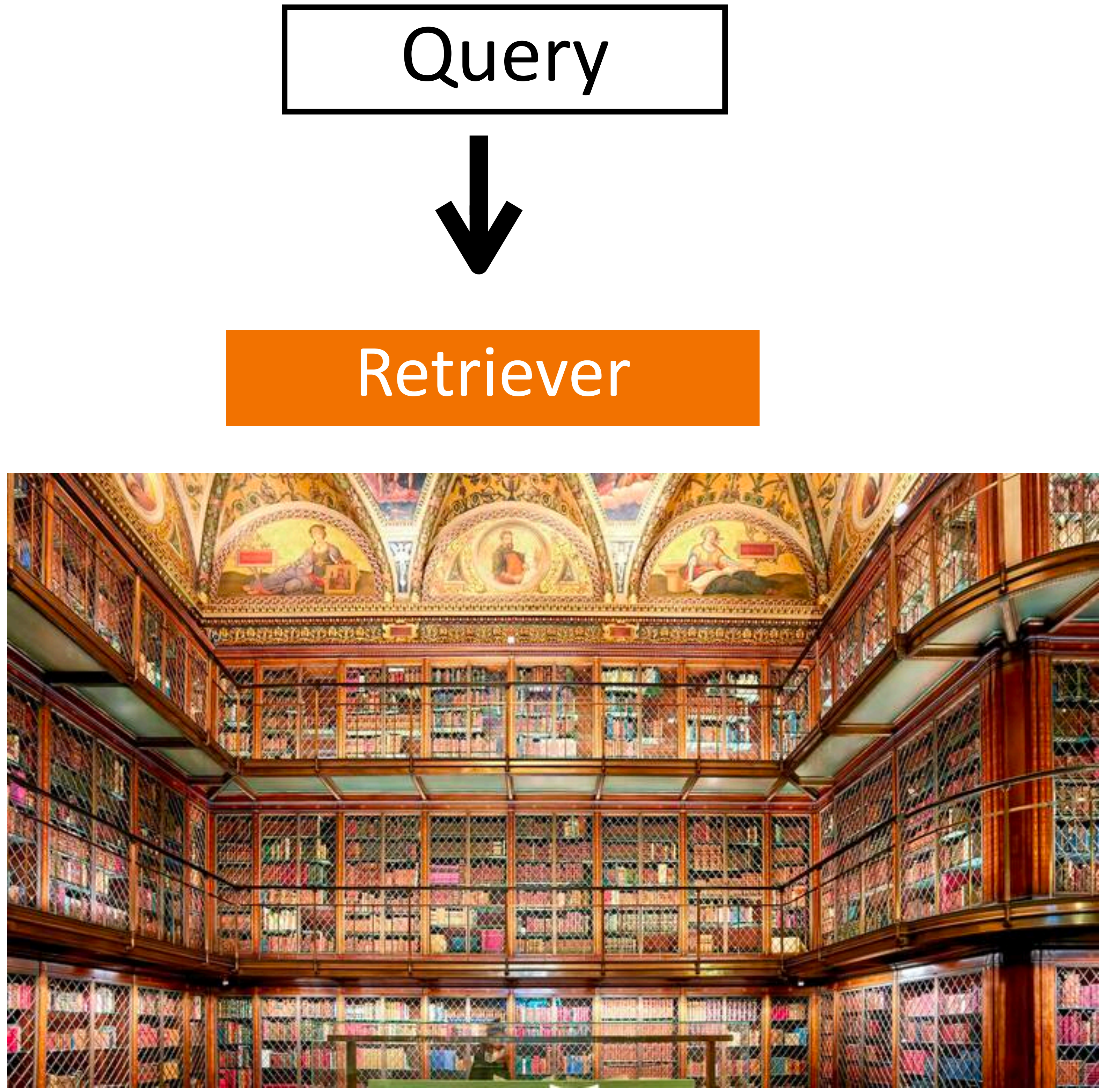
Retriever

- Query representation:
  - Decomposing query into subqueries [EMNLP22, NAACL23]
  - Expanding query with in-context examples [InSub 24]

[Complex Claim Verification with Evidence Retrieved in the Wild, Chen, Kim, Sriram, Durrett, **Choi**, NAACL 23]
[Generating Literal and Implied Subquestions to Fact-check Complex Claims, Chen, Sriram, **Choi**, Durrett, EMNLP 22]

[RARe: Retrieval Augmented Retrieval with In-Context Examples, Tejaswi, Lee, Sanghavi* and **Choi***, In Sub24]
[Open-World Evaluation for Retrieving Diverse Perspectives, Chen and **Choi**, NAACL 2025]
[Contrastive Learning to Improve Retrieval for Real-world Fact Checking, Sriram, Xu, **Choi**, Durrett EMNLP FEVER 24
[XOR QA: Cross-lingual Open-Retrieval Question Answering, Asai, Kasai, .. **Choi** and Hajishirzi, NAACL 21]

Improving Retrievers



- Query representation:
  - Decomposing query into subqueries [EMNLP22, NAACL23]
  - Expanding query with in-context examples [InSub 24]
- Retriever architecture:
  - Autoregressive, iterative retrieval [Ongoing]

[Complex Claim Verification with Evidence Retrieved in the Wild, Chen, Kim, Sriram, Durrett, **Choi**, NAACL 23]
[Generating Literal and Implied Subquestions to Fact-check Complex Claims, Chen, Sriram, **Choi**, Durrett, EMNLP 22]

[RARe: Retrieval Augmented Retrieval with In-Context Examples, Tejaswi, Lee, Sanghavi* and **Choi***, In Sub24]
[Open-World Evaluation for Retrieving Diverse Perspectives, Chen and **Choi**, NAACL 2025]
[Contrastive Learning to Improve Retrieval for Real-world Fact Checking, Sriram, Xu, **Choi**, Durrett EMNLP FEVER 24
[XOR QA: Cross-lingual Open-Retrieval Question Answering, Asai, Kasai, .. **Choi** and Hajishirzi, NAACL 21]

Improving Retrievers

Query

↓

Retriever

- Query representation:
  - Decomposing query into subqueries [EMNLP22, NAACL23]
  - Expanding query with in-context examples [InSub 24]
- Retriever architecture:
  - Autoregressive, iterative retrieval [Ongoing]
- Retrieval corpus selection & Evaluation:
  - Cross-lingual retrieval evaluation [NAACL21]
  - Diversity-driven evaluation across multiple corpora [NAACL 25]

[Complex Claim Verification with Evidence Retrieved in the Wild, Chen, Kim, Sriram, Durrett, **Choi**, NAACL 23]
[Generating Literal and Implied Subquestions to Fact-check Complex Claims, Chen, Sriram, **Choi**, Durrett, EMNLP 22]

[RARe: Retrieval Augmented Retrieval with In-Context Examples, Tejaswi, Lee, Sanghavi* and **Choi***, In Sub24]
[Open-World Evaluation for Retrieving Diverse Perspectives, Chen and **Choi**, NAACL 2025]
[Contrastive Learning to Improve Retrieval for Real-world Fact Checking, Sriram, Xu, **Choi**, Durrett EMNLP FEVER 24
[XOR QA: Cross-lingual Open-Retrieval Question Answering, Asai, Kasai, .. **Choi** and Hajishirzi, NAACL 21]

Improving Efficiency: Knowledge Compression

Improving Efficiency: Knowledge Compression

- Improving inference efficiency by compressing KV cache

Improving Efficiency: Knowledge Compression

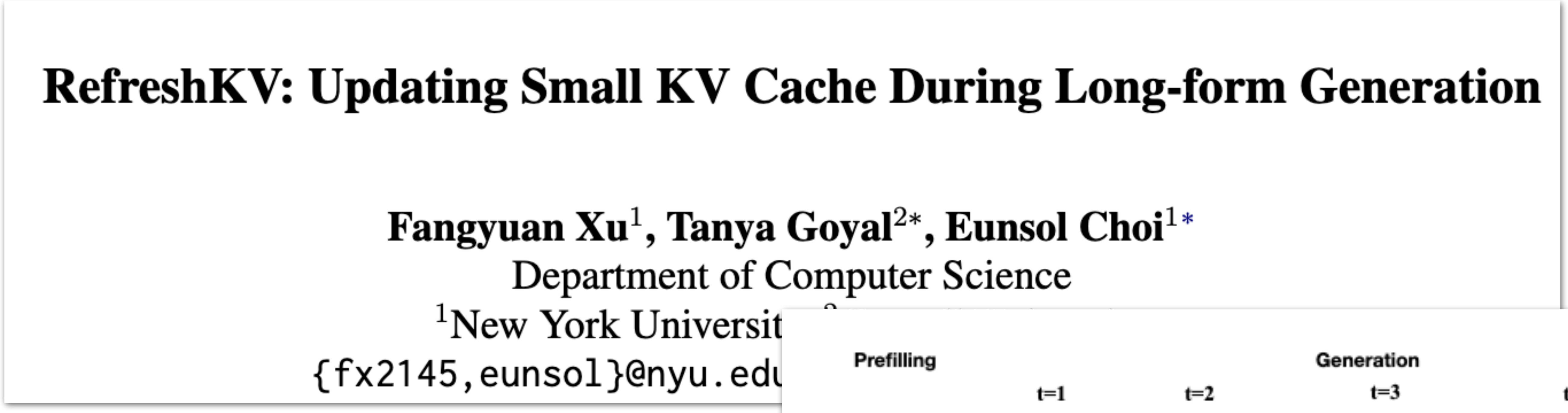
- Improving inference efficiency by compressing KV cache

RefreshKV: Updating Small KV Cache During Long-form Generation

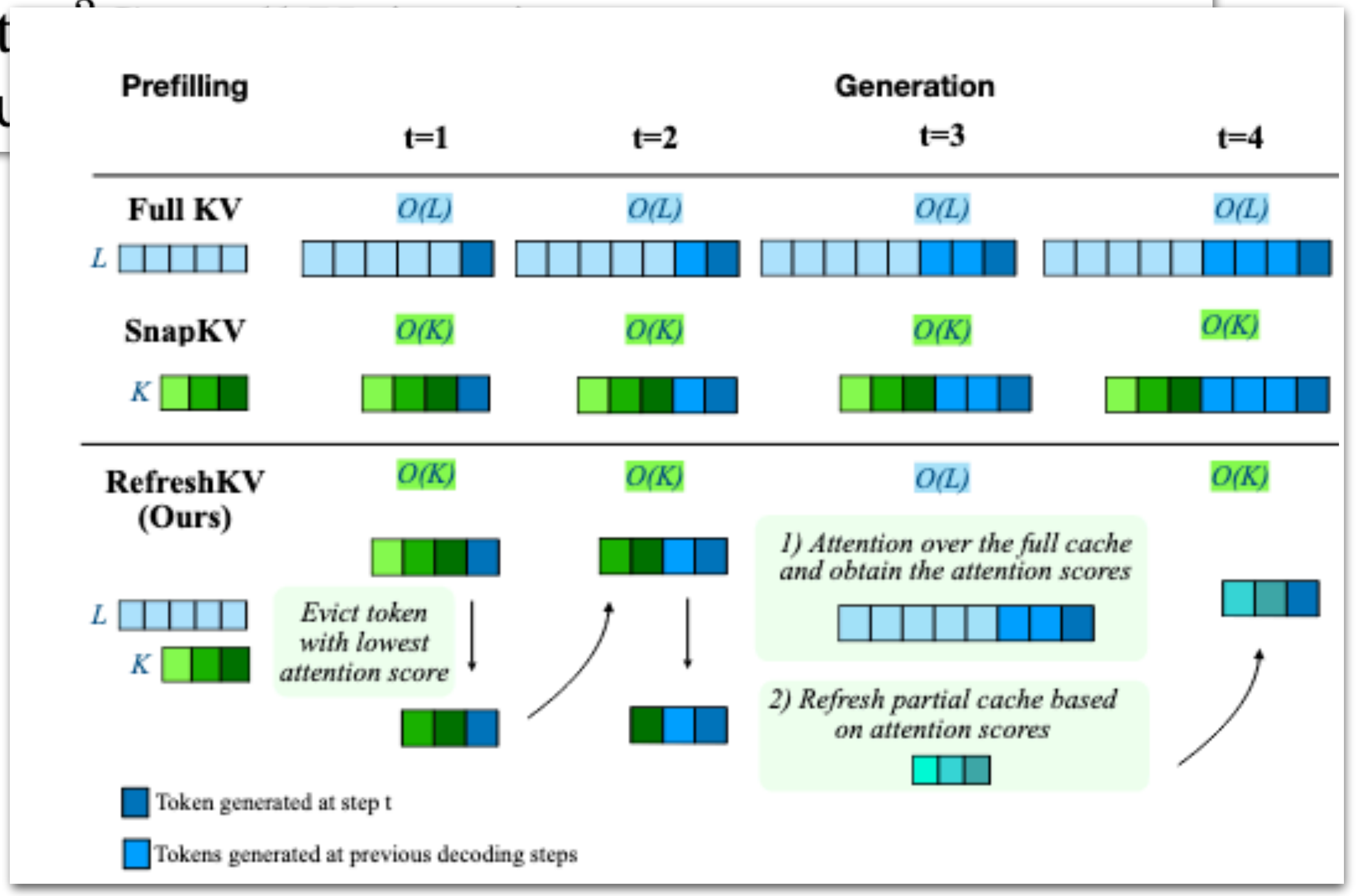
Fangyuan Xu[1], **Tanya Goyal**[2*], **Eunsol Choi**[1*]
Department of Computer Science
[1]New York University, [2]Cornell University
`{fx2145,eunsol}@nyu.edu`, `tanyagoyal@cornell.edu`

Improving Efficiency: Knowledge Compression

- Improving inference efficiency by compressing KV cache


RefreshKV: Updating Small KV Cache During Long-form Generation

Fangyuan Xu[1], Tanya Goyal[2*], Eunsol Choi[1*]
Department of Computer Science
[1]New York Universit
{fx2145,eunsol}@nyu.edu



- Key idea: alternating between **full attention** and **partial attention** and resetting small KV cache when needed

Improving Knowledge Integration: Multi-document reasoning

Who is highest paid **football player** in 2021?

Retriever

Multiple documents, each with their own valid answer

Manchester United's Cristiano Ronaldo, who is the world's first and only billionaire football player, tops the list, raking in US$125 million from his salary and endorsement deals.

The highest paid player in the league is Kansas City Chiefs quarterback Patrick Mahomes. Mahomes makes $45 million per season in average annual salary.

LLM → ??

[AmbigDocs: Reasoning across Documents on Different Entities under the Same Name, Lee, Ye, Choi, COLM24]

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** ~~player in 2021~~ was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Disambiguate underspecified entity (football)

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Disambiguate underspecified entity (football)

Provide an answer

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Partial Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo.

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Partial Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo.

Ambiguous Answer

The highest-paid **football** player in 2021 was Cristiano Ronaldo.

Types of Answers Under Multiple Valid Inputs

Complete Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo, earning $125million. The highest-paid **NFL** player is Patrick Mahomes, earning $46million.

Partial Answer

The highest-paid **soccer** player in 2021 was Cristiano Ronaldo.

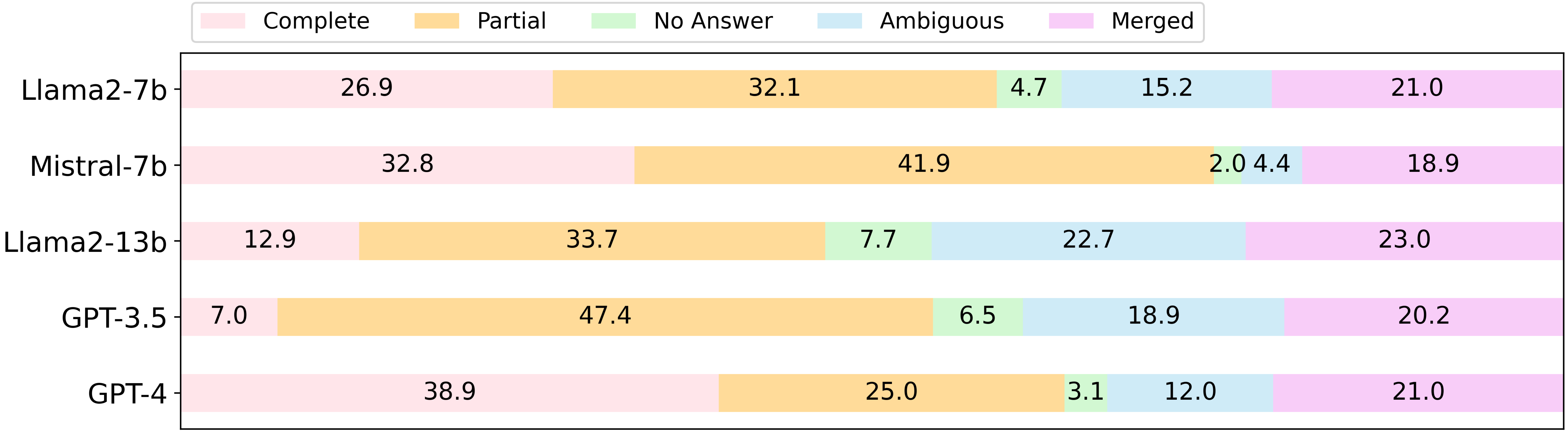
Ambiguous Answer

The highest-paid **football** player in 2021 was Cristiano Ronaldo.

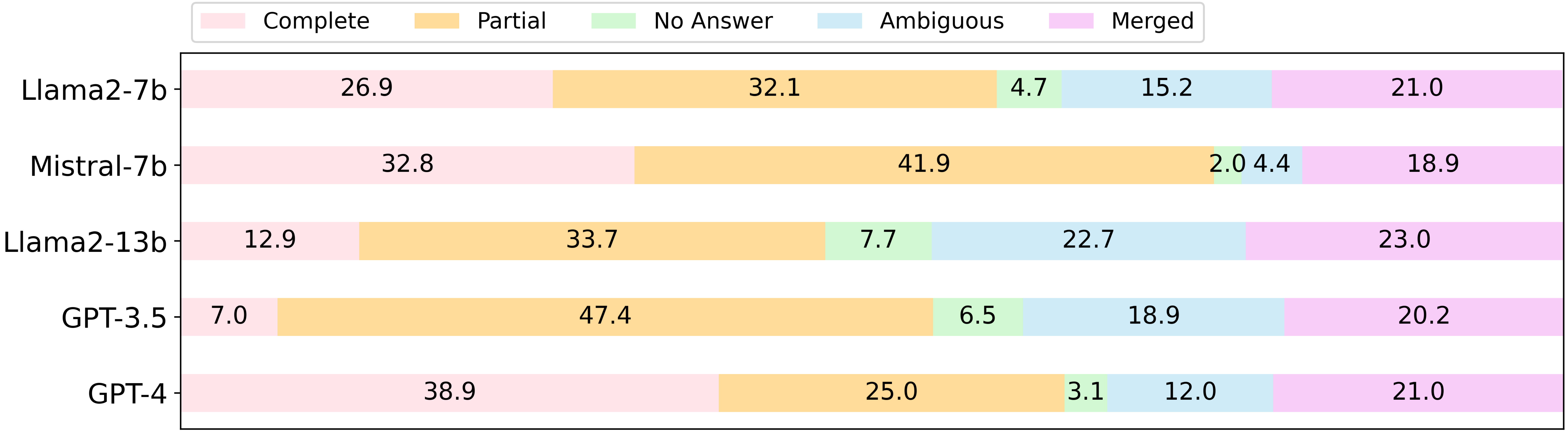
Merged answer

Cristiano Ronaldo was the highest-paid football player in 2021, earning $125 million. This includes his salary and bonus, as well as his endorsement deals. If we only consider salary and bonuses, however, then Patrick Mahomes is the highest-paid football player, earning 45 millions.

What types of answer do LLMs generate?

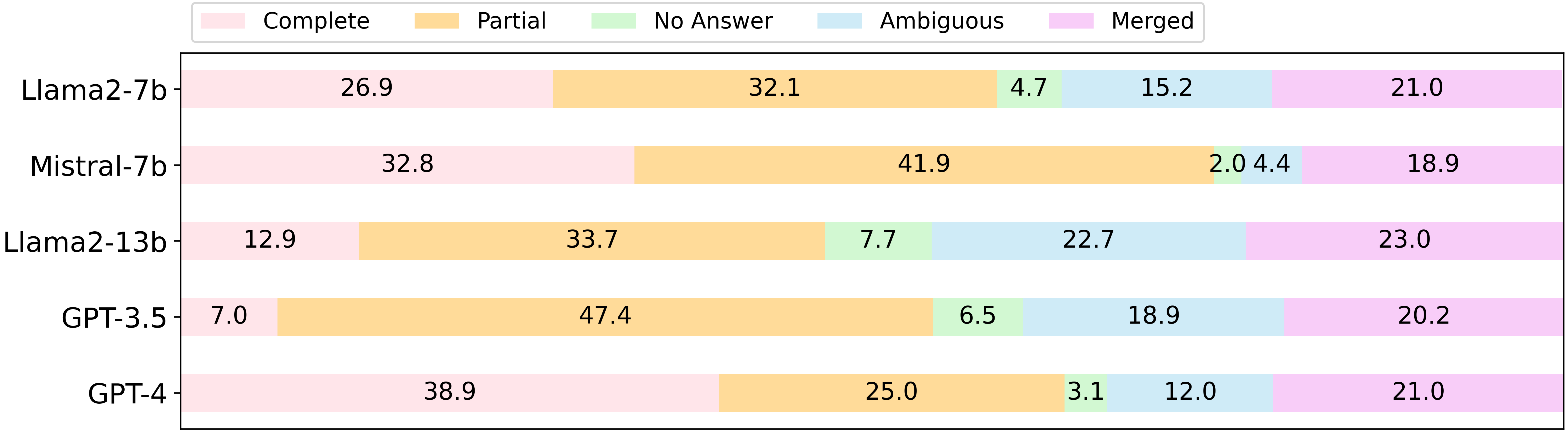


What types of answer do LLMs generate?



- No system provides a complete answer consistently, providing answers that can be misleading

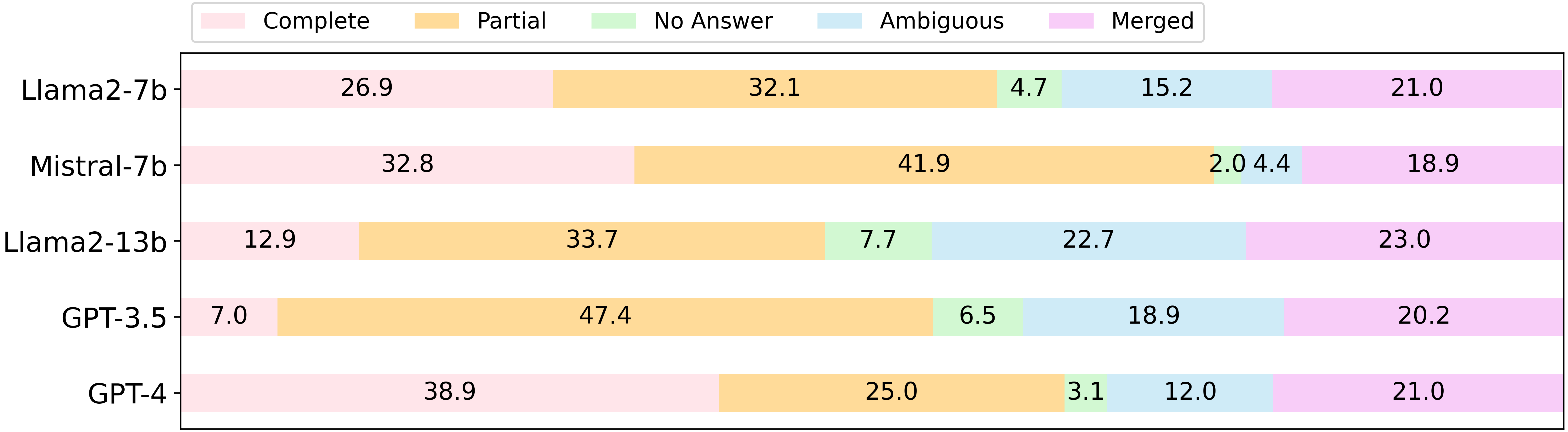
What types of answer do LLMs generate?



- No system provides a complete answer consistently, providing answers that can be misleading
- Can we learn to abstain upon detecting knowledge conflict? [EMNLP22]

[Rich Knowledge Sources Bring Complex Knowledge Conflicts: Recalibrating Models to Reflect Conflicting Evidence, Chen, Zhang and **Choi**, EMNLP 2022]

What types of answer do LLMs generate?

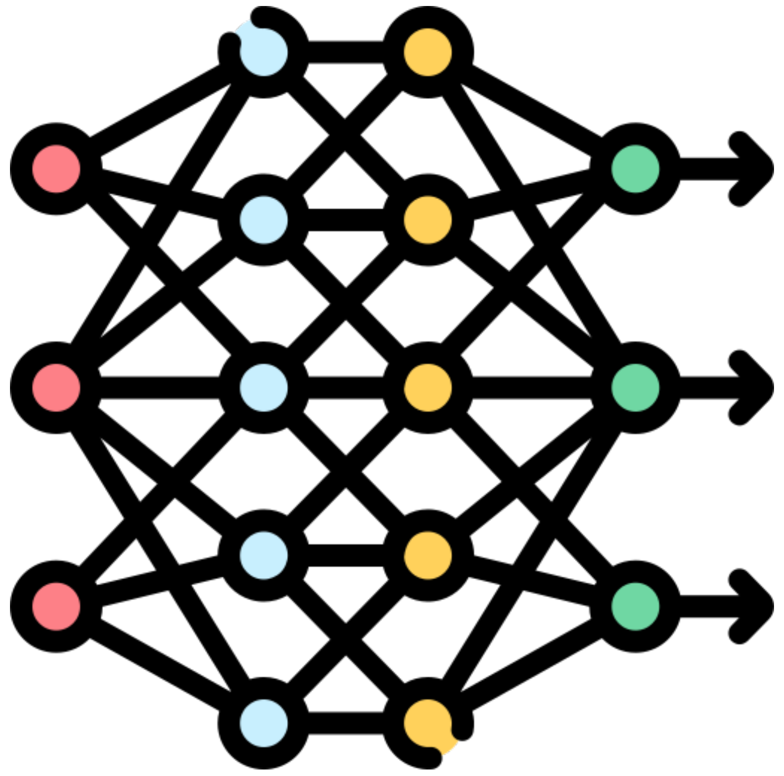


- No system provides a complete answer consistently, providing answers that can be misleading
- Can we learn to abstain upon detecting knowledge conflict? [EMNLP22]
- How can we fine-tune LLMs to provide more complete answers?

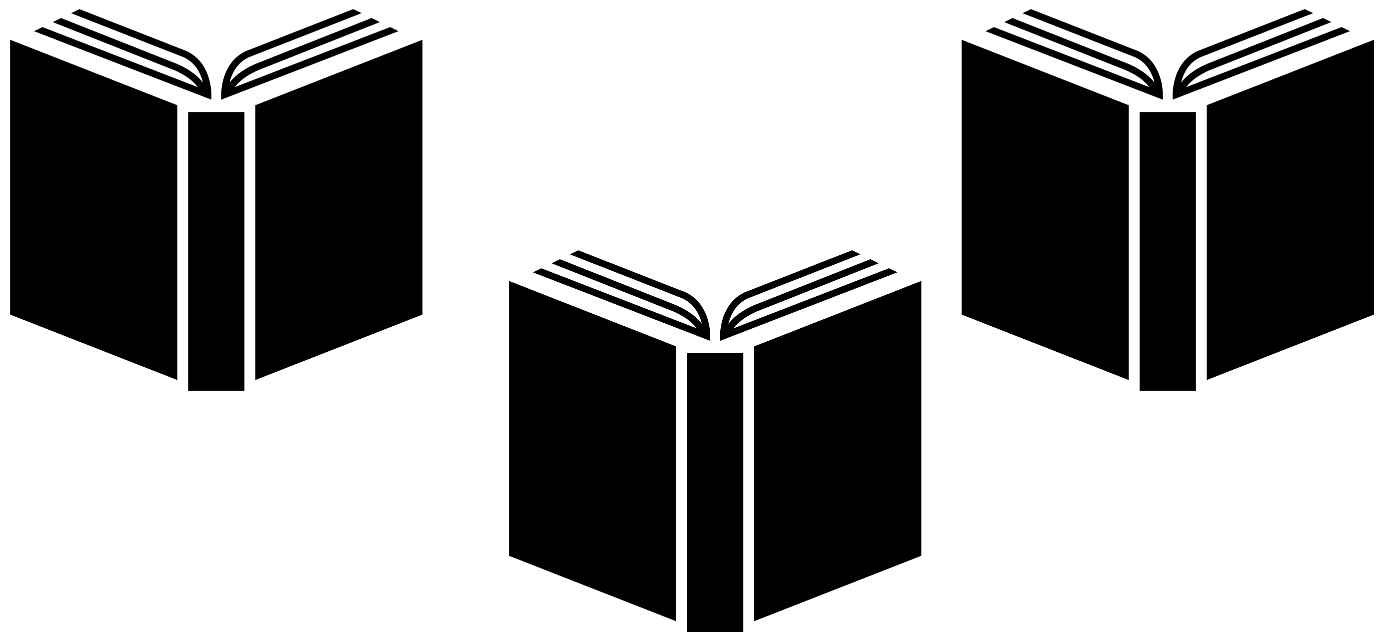
[Rich Knowledge Sources Bring Complex Knowledge Conflicts: Recalibrating Models to Reflect Conflicting Evidence, Chen, Zhang and **Choi**, EMNLP 2022]

Systems Incorporating Two Knowledge Sources

Knowledge of LMs



Documents retrieved at inference time

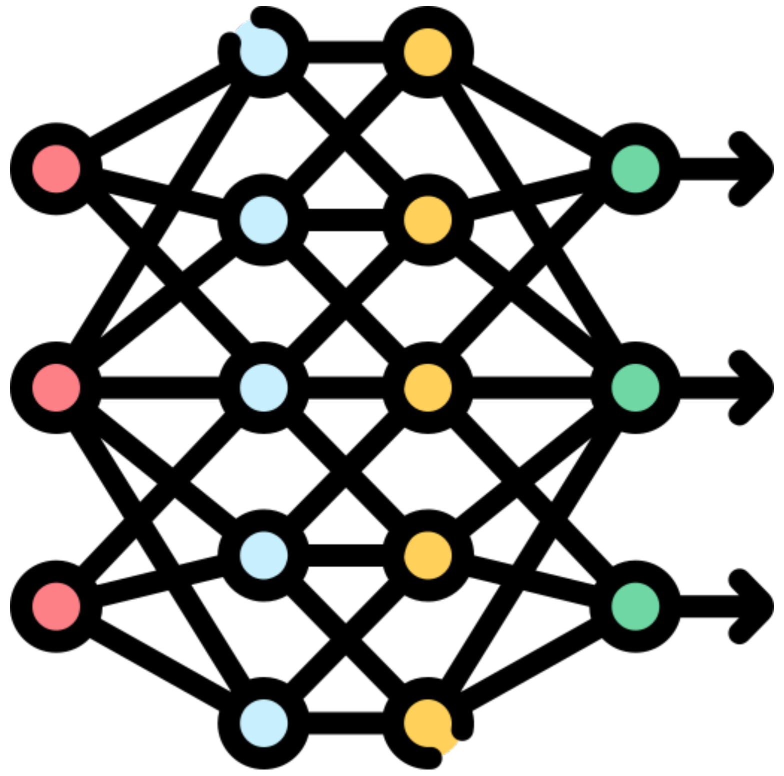


ChatGPT

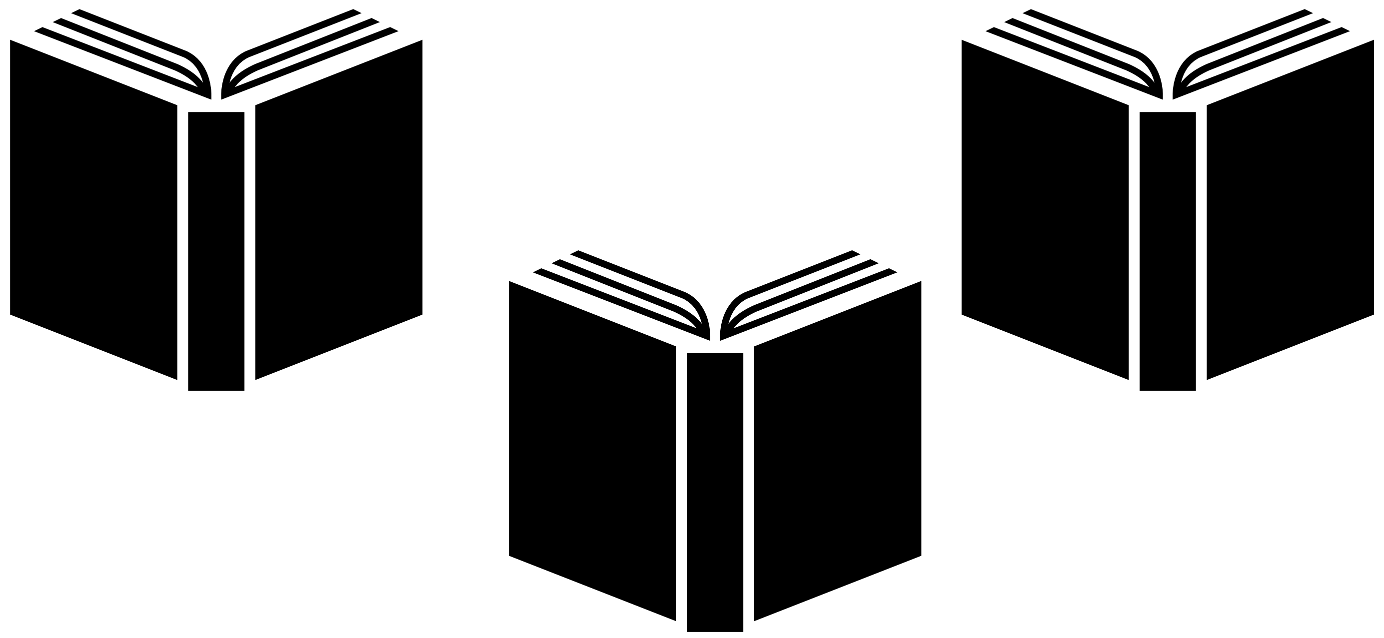
Google Search

Systems Incorporating Two Knowledge Sources

Knowledge of LMs



Documents retrieved at inference time



Knowledge Editing

[Can LMs Learn New Entities from Descriptions? Challenges in Propagating Injected Knowledge, Onoe,..., **Choi** ACL 23]

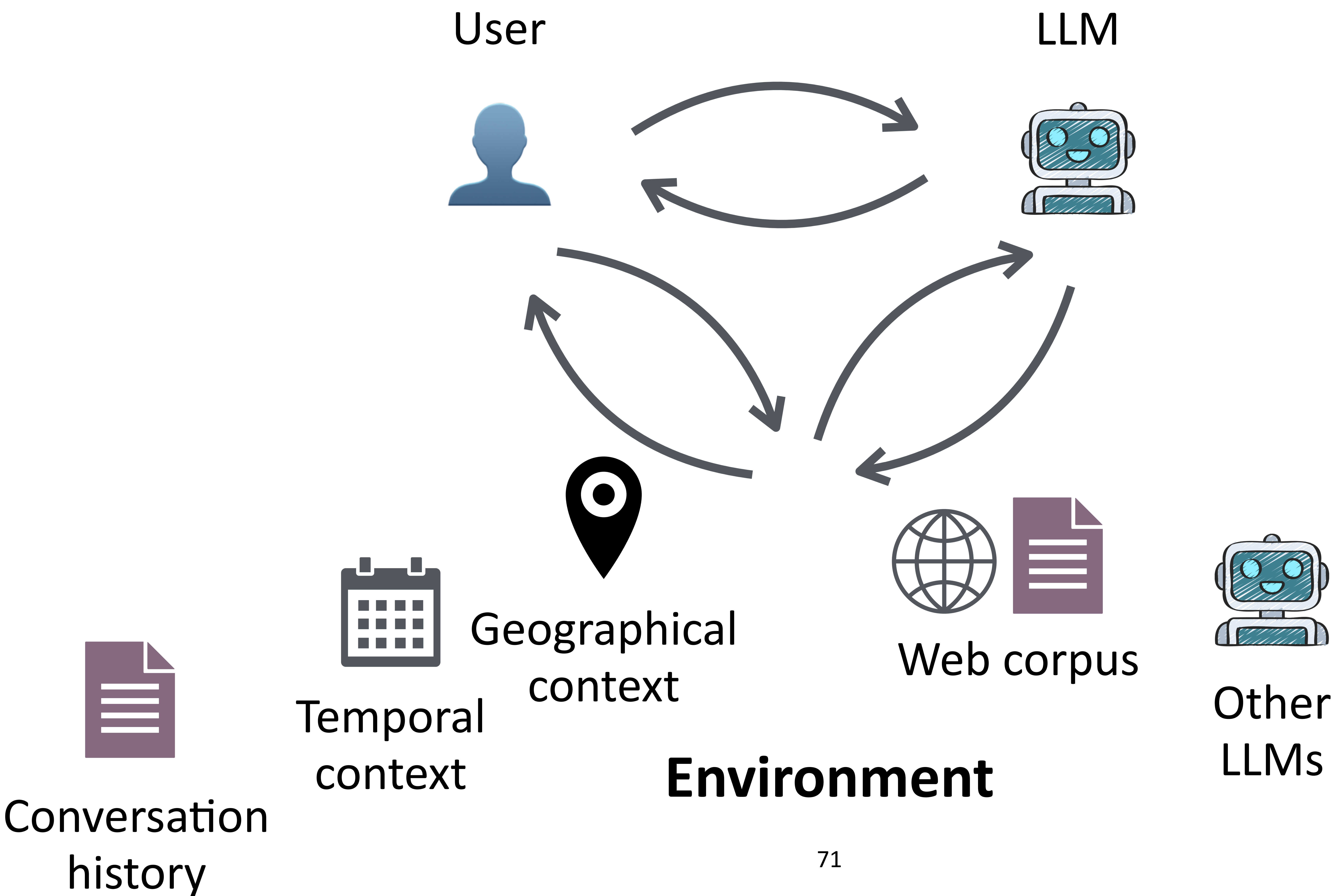
[Propagating Knowledge Updates to LMs Through Distillation, Padmanabhan, Onoe, Zhang, Durrett, **Choi** NeurIPS 23]

[PropMEND: Hypernetworks for Knowledge Propagation in LLMs, Liu Durrett, **Choi** ArXiv 25]

ChatGPT

Google Search

LLMs in real world



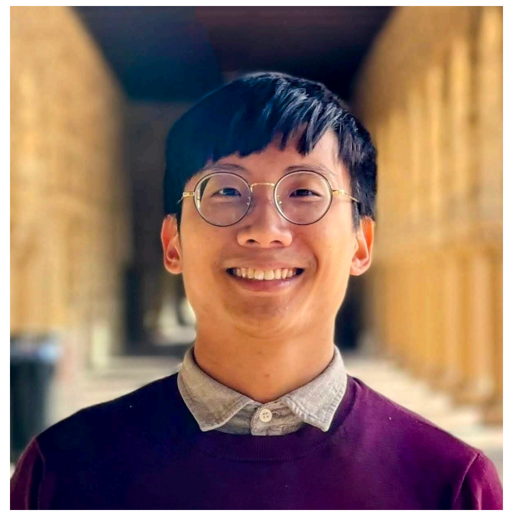
NSF CISCO Open Philanthropy THE HOME DEPOT SONY Google

My Lab


Michael J.Q. Zhang

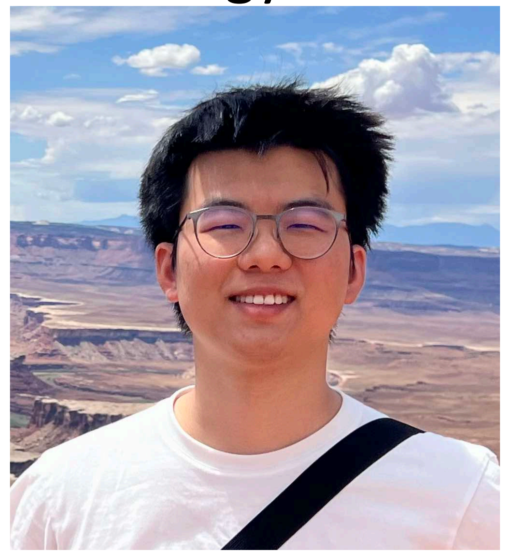

Anuj Diwan


Fangyuan Xu


Hung-ting Chen


Thom Lake


Yuhan Liu


Leo Zeyu Liu

Thank You!

Questions?


W Bradley Knox


Ge Gao


Yoav Artzi