



清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University

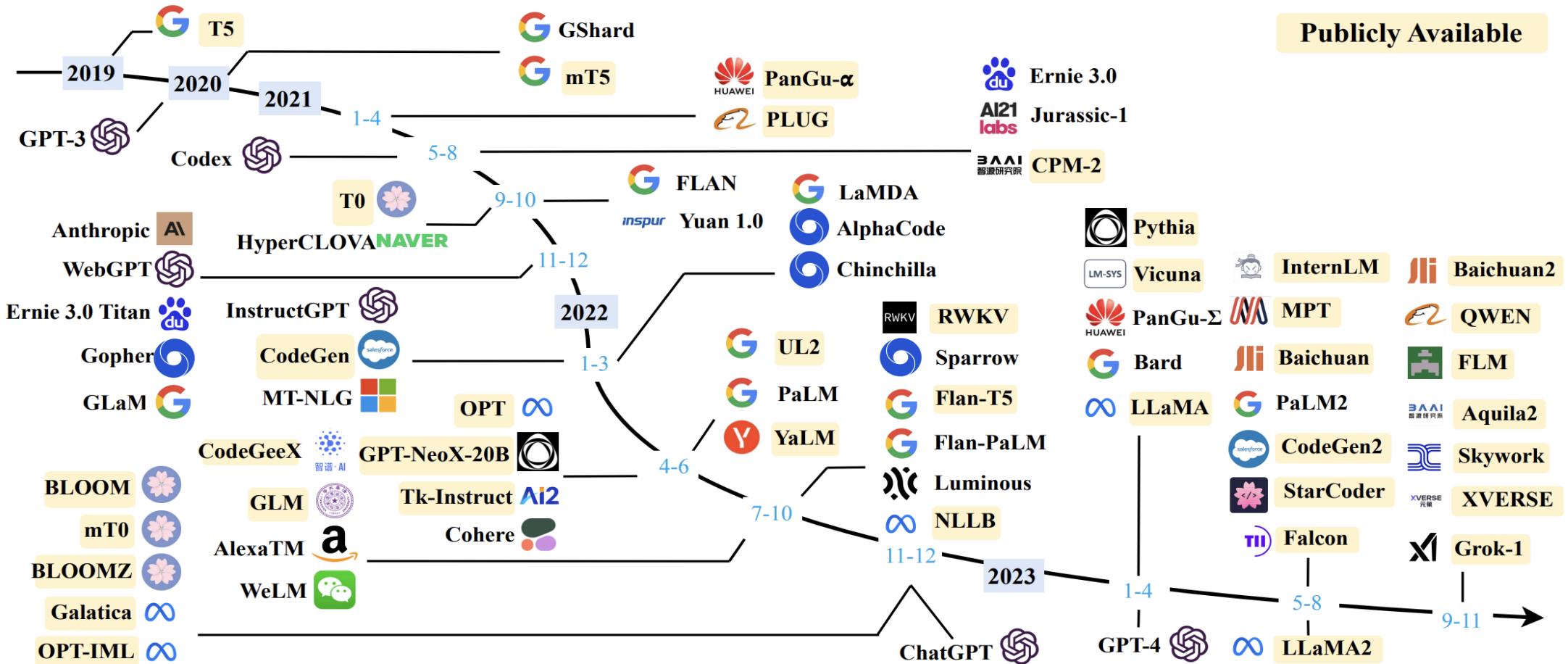
# 大模型智能体技术与应用

李鹏

2024年5月9日

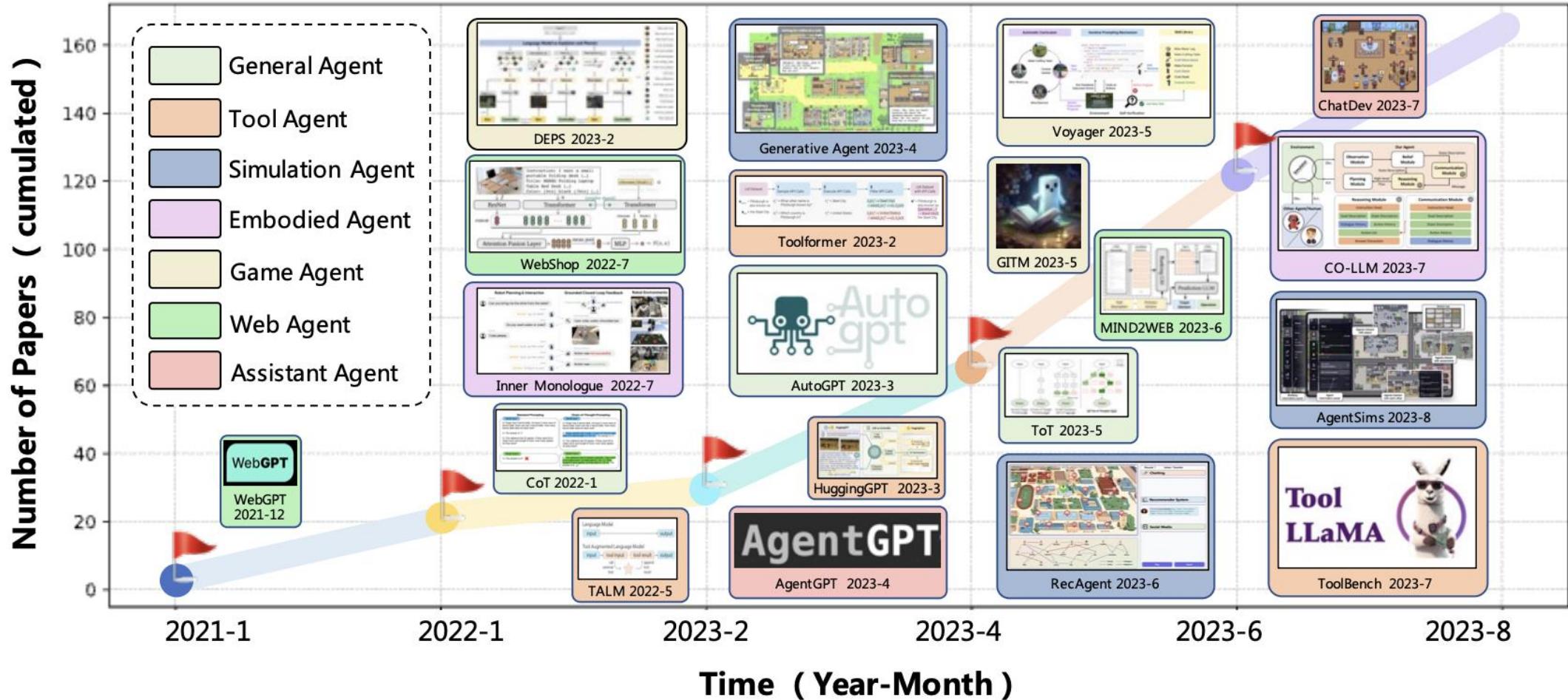
# 大规模语言模型

- 大规模语言模型近年来发展迅速，人工智能进入新的历史发展阶段。



# 基于大模型的智能体发展概况

- 基于大模型的智能体在2023年迅猛发展，受到广泛高度关注。



# 智能体是OpenAI的重点发力方向

- OpenAI高度关注智能体，GPTs是面向个人需求推出的智能体方案。



“但是当新的AI Agents论文出来的时候，我们都非常感兴趣，觉得它非常酷，因为我们的团队并没有花费五年时间在这上面，我们并不比你们更多掌握什么，我们正在与你们所有人一起竞争。”



机遇：在智能体方向国内外基本处于同一起跑线

# 智能体领域生态日趋完善

· 产业界与学术界在智能体框架、技术、产品等方向积极探索。

LangChain

AutoGPT

MetaGPT



抽象智能体与环境交互、大模型协作等公共需求



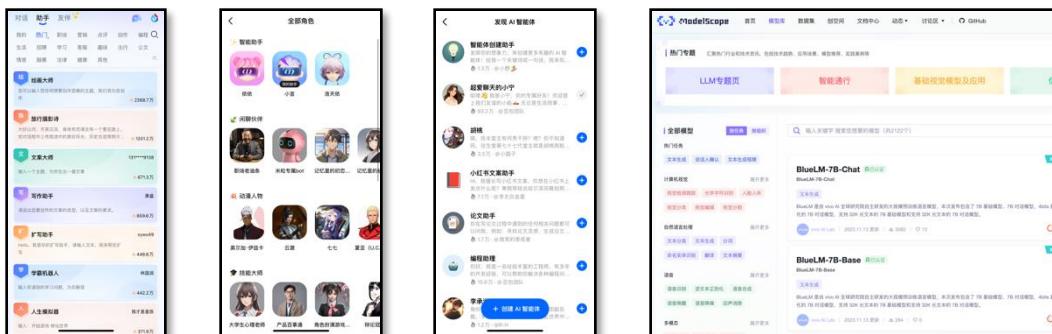
Fuyu-8B: A Multimodal Architecture for AI Agents

October 17, 2023 – Rohan Bavishi, Erich Eisen, Curtis Hawthorne, Maxwell Nye, Augustus Odema, Arushi Soman, Sadighi Tajerian

We're open-sourcing Fuyu-8B - a small version of the multimodal model that powers our product.



智能体基础模型、运行机制、评测等快速发展



个人助手、  
模型社区等产品形态  
被广泛探索

图片和资料来自网络

# 提纲

智能体代表性技术

代表性技术应用例

智能体技术展望

# 提纲

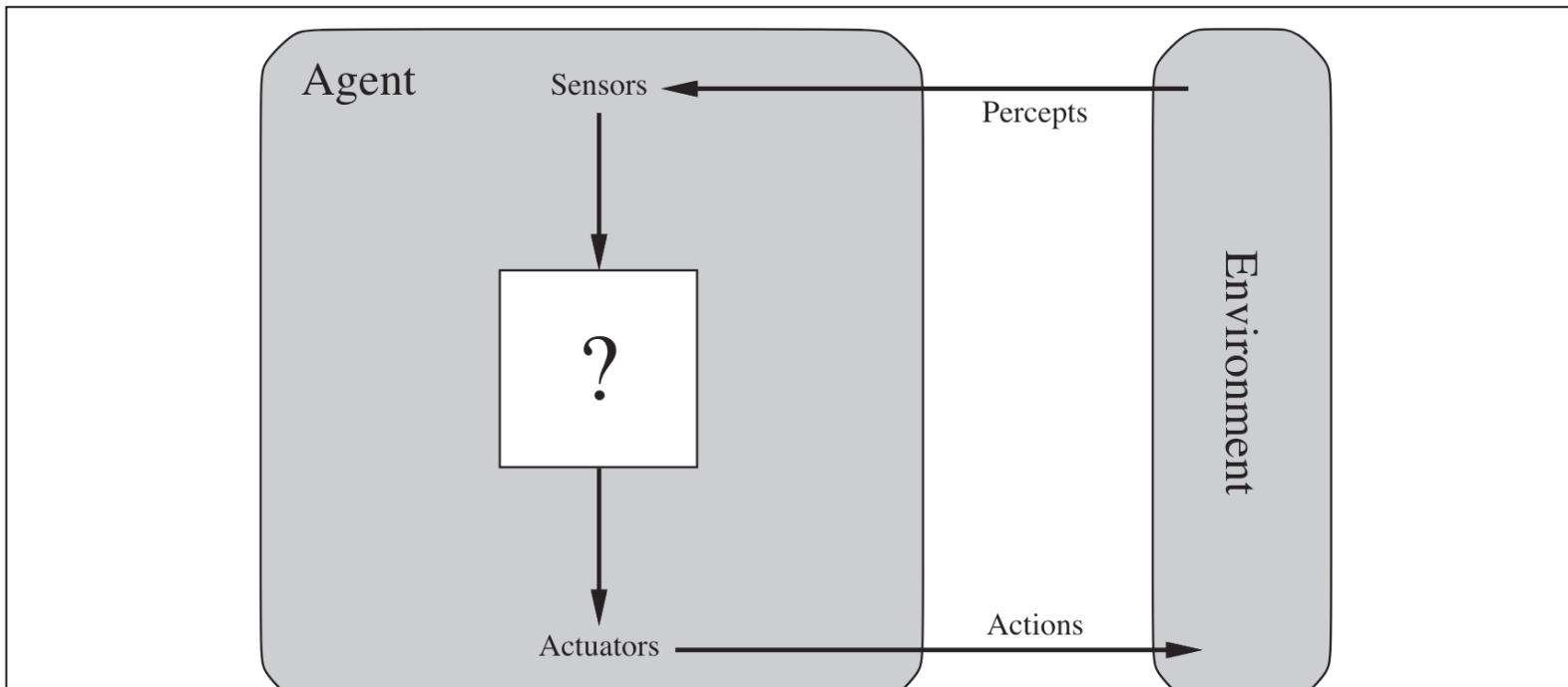
智能体代表性技术

代表性技术应用例

智能体技术展望

# 智能体

- An **agent** is anything that can be viewed as **perceiving** its environment through **sensors** and acting upon that environment through **actuators**. —— (Russell, 2010)



**Figure 2.1** Agents interact with environments through sensors and actuators.

# 自主智能体 (Autonomous Agent)

- An **autonomous agent** is a system situated within and a part of an environment that senses that environment and acts on it, over time, **in pursuit of its own agenda** and so as to effect what it senses in the future. —— (*Franklin and Graesser, 1997*)

## Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents

Stan Franklin and Art Graesser

Institute for Intelligent Systems, University of Memphis, Memphis, TN 38152, USA  
[stan.franklin@memphis.edu](mailto:stan.franklin@memphis.edu)  
[graesser@cc.memphis.edu](mailto:graesser@cc.memphis.edu)

### Abstract

The advent of software agents gave rise to much discussion of just what such an agent is, and of how they differ from programs in general. Here we propose a formal definition of an autonomous agent which clearly distinguishes a software agent from just any program. We also offer the beginnings of a natural kinds taxonomy of autonomous agents, and discuss possibilities for further classification. Finally, we discuss subagents and multiagent systems.

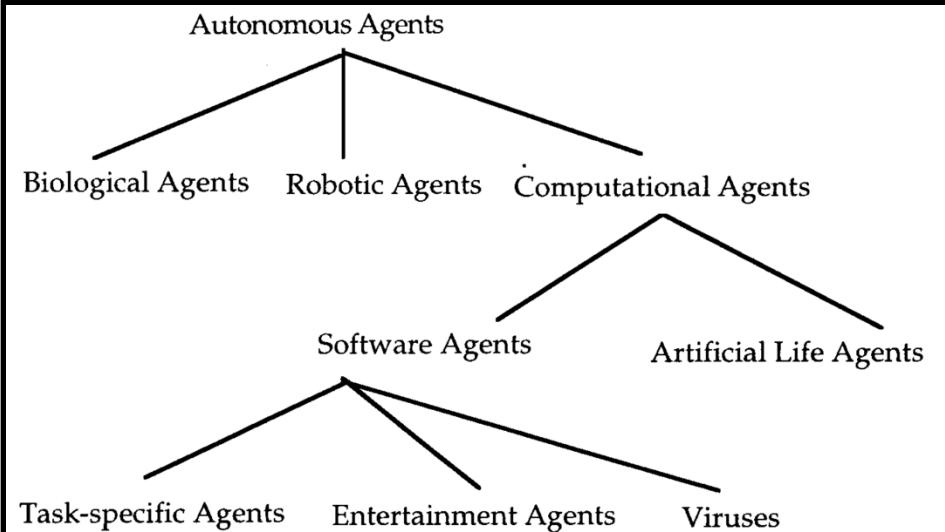
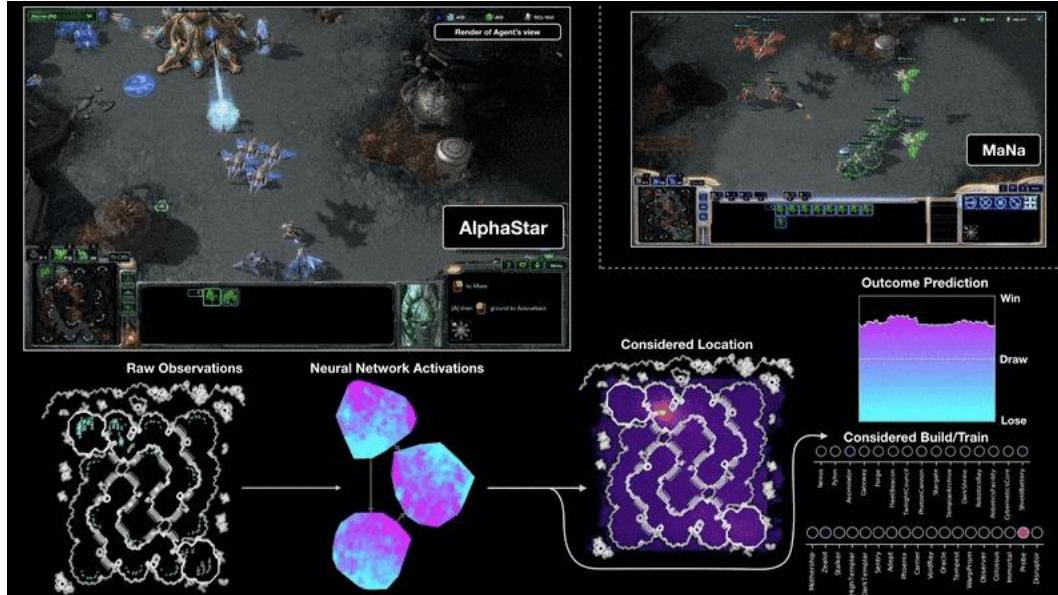


Fig. 1. Natural Kinds Classification of Autonomous Agents

# 大模型智能体

- 以大模型作为驱动智能体的核心将智能体研究推向新的高度。

AlphaStar



- 训练：强化学习+自我对弈
- 数据：专用环境收集的海量数据
- 任务：仅能玩星际争霸2

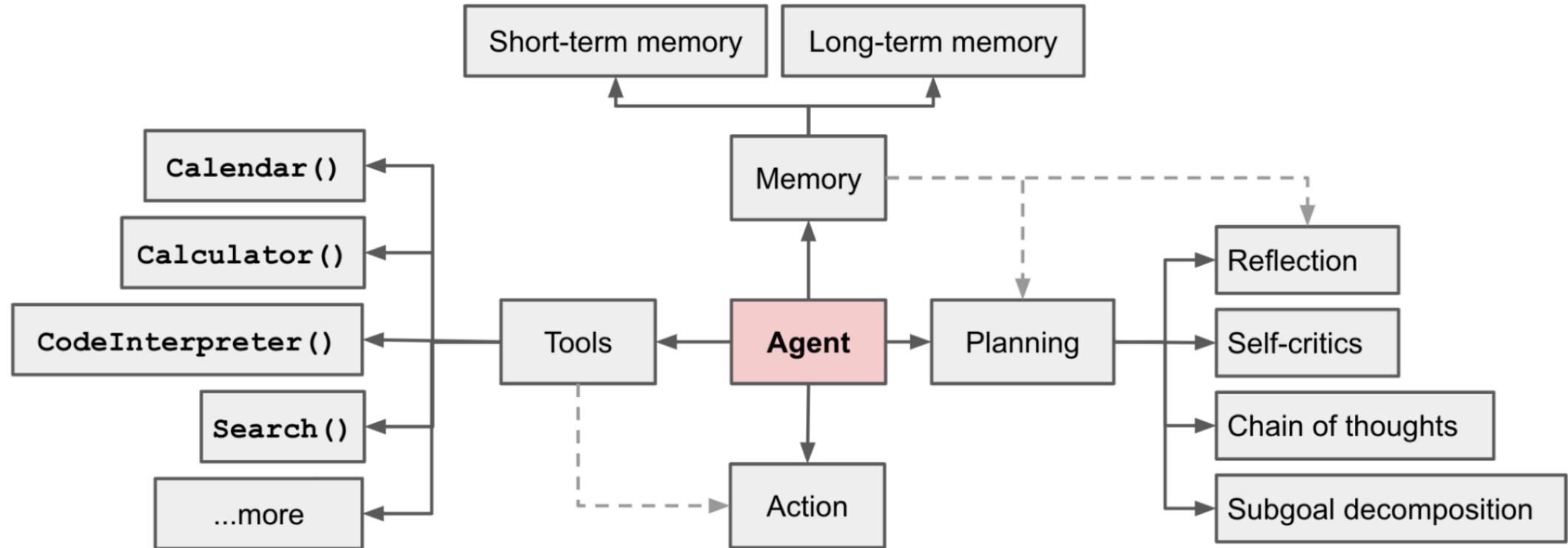
AutoGPT



- 训练：免训练
- 数据：无需专门收集数据
- 任务：可完成多样化任务

# 大模型智能体框架

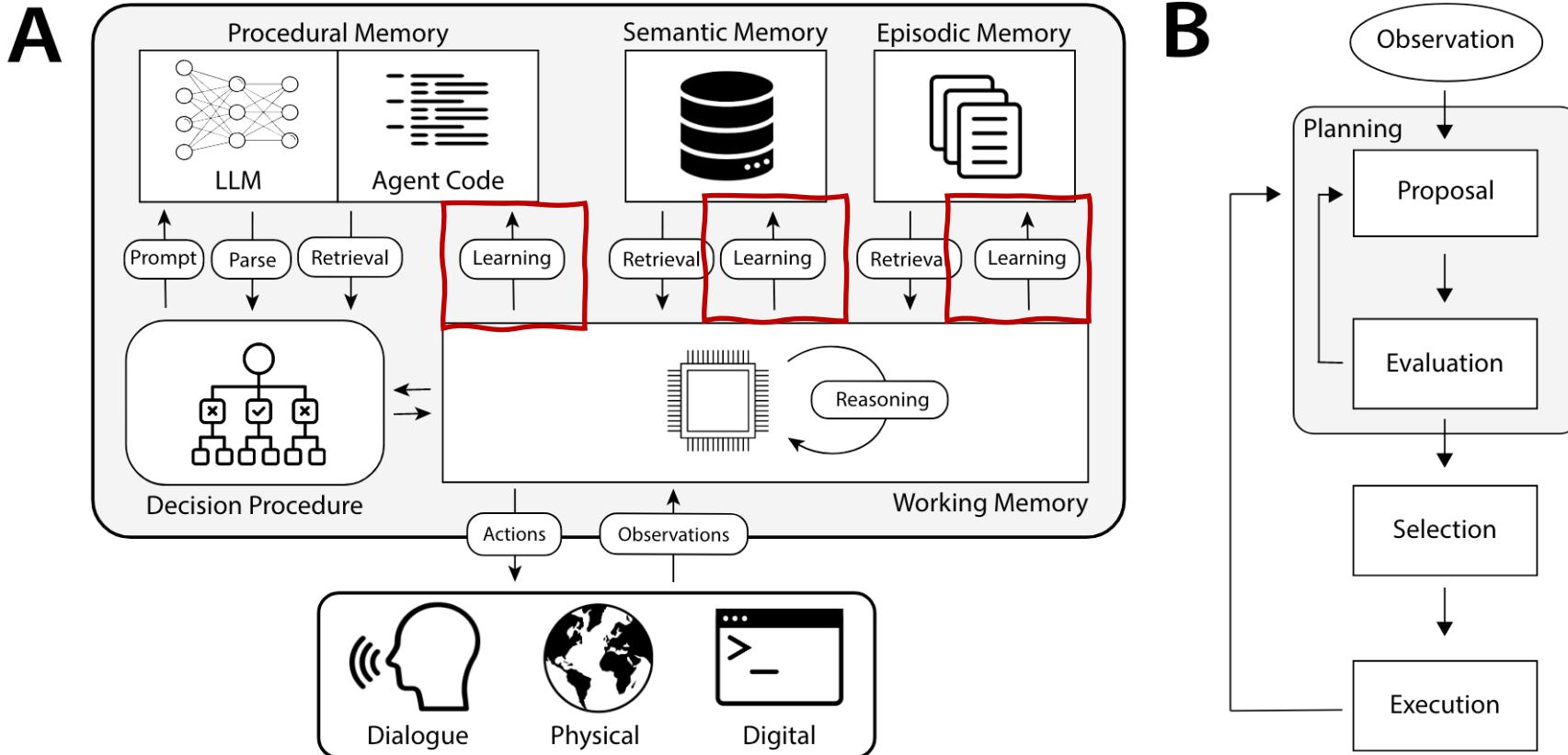
- 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



OpenAI研究员Lilian Weng提出的大模型自主智能体框架

# 大模型智能体框架

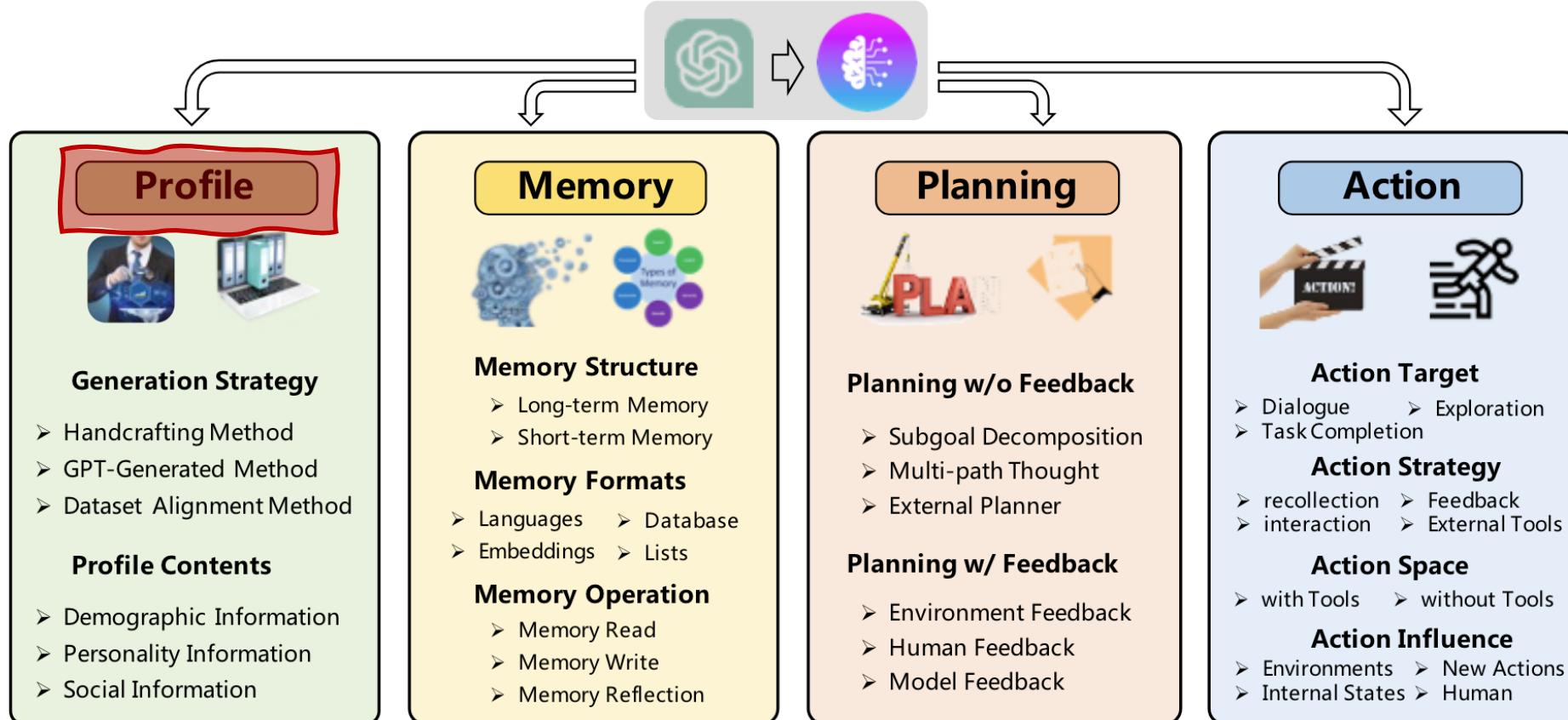
- 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



Shunyu Yao等提出的认知语言智能体认知框架

# 大模型智能体框架

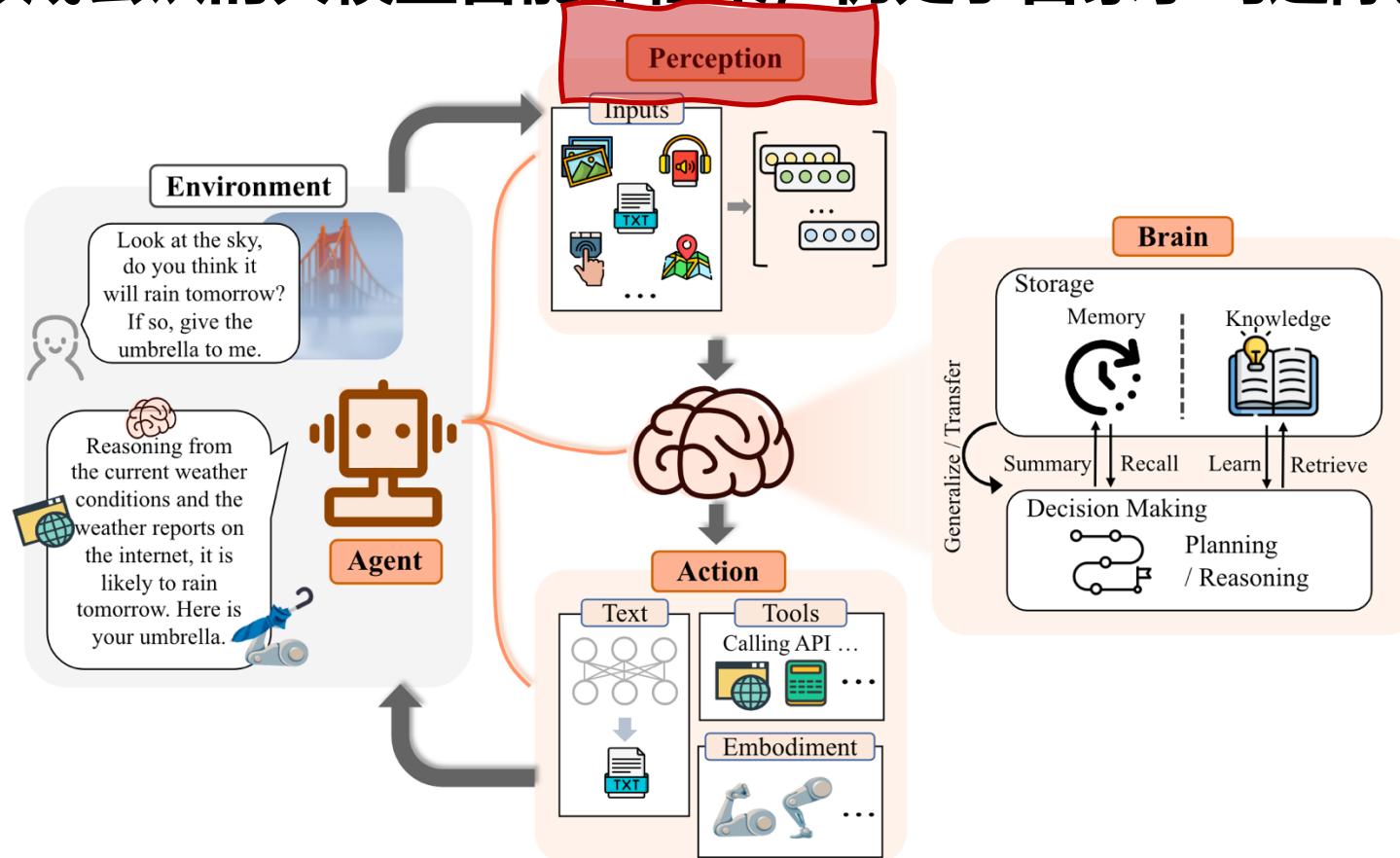
• 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



人民大学提出的大模型自主智能体框架

# 大模型智能体框架

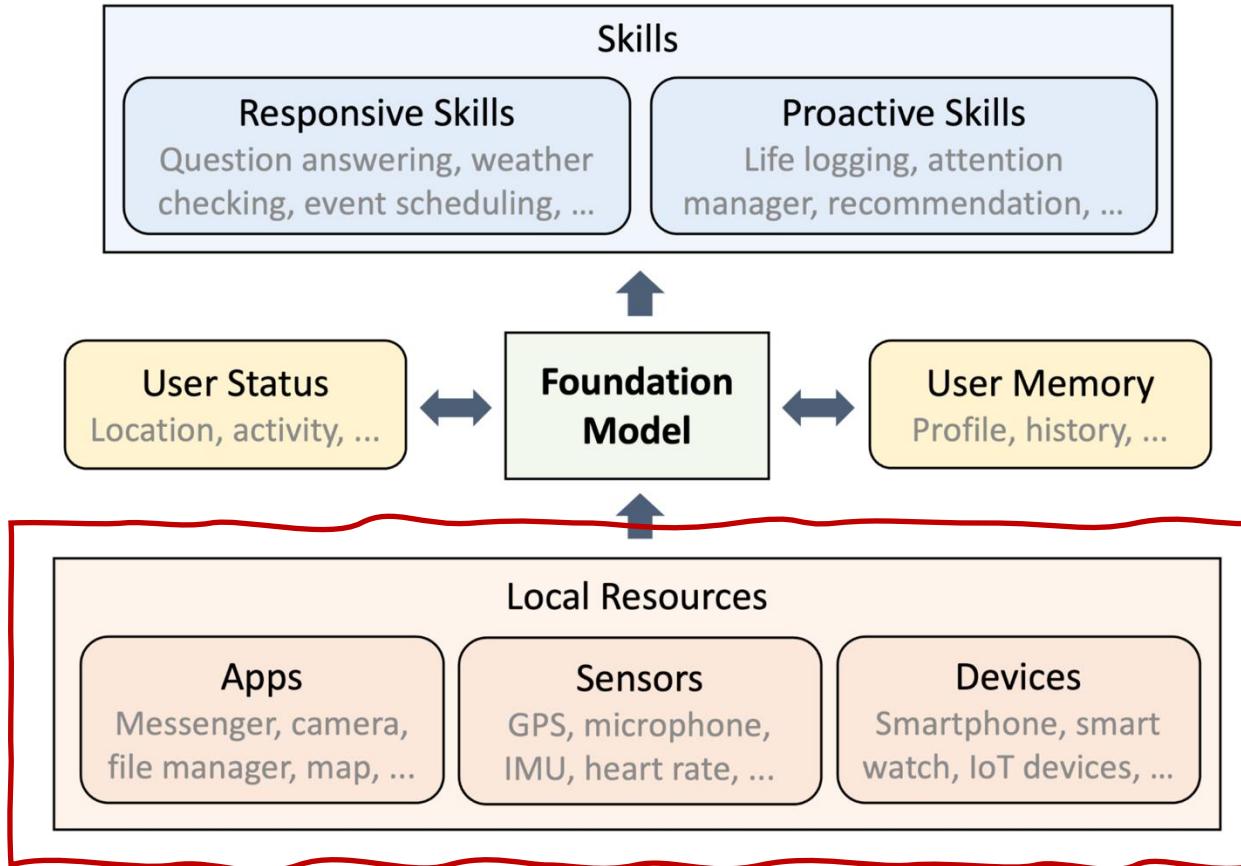
- 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



复旦大学等提出的大模型自主智能体框架

# 大模型智能体框架

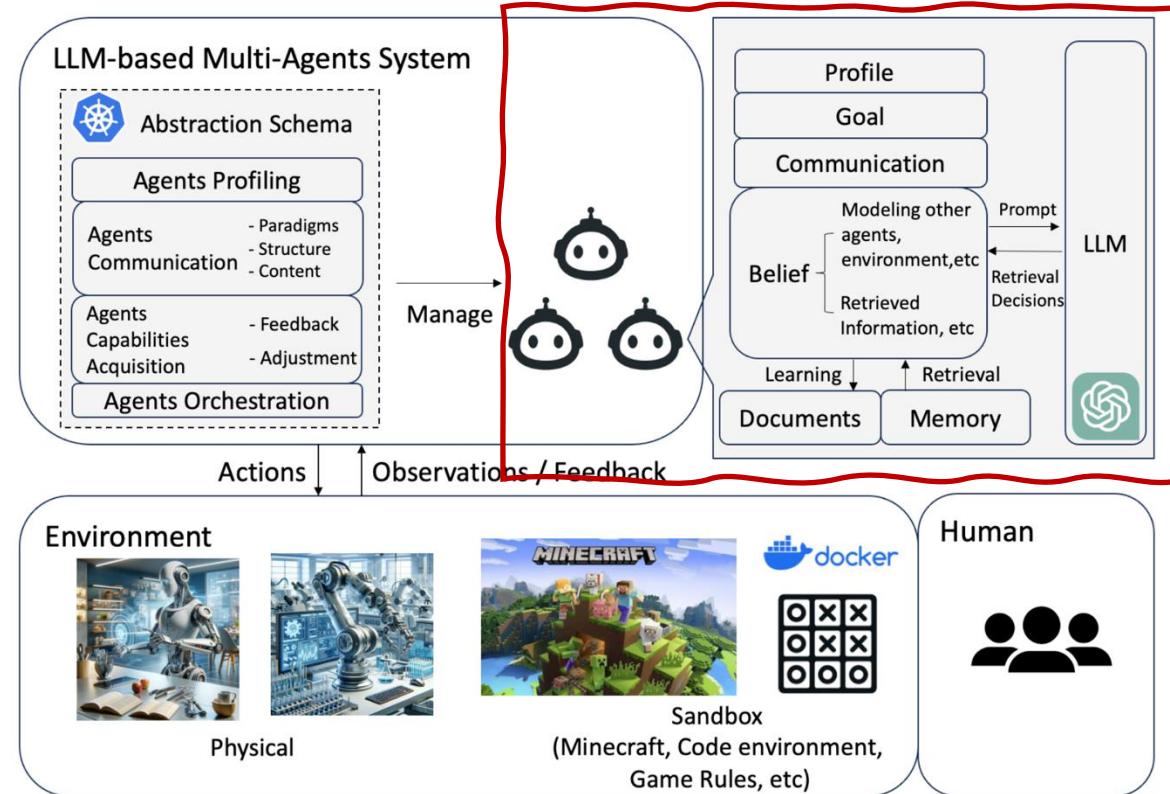
- 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



清华大学等提出的个人大模型智能体 (Personal LLM Agents) 框架

# 大模型智能体框架

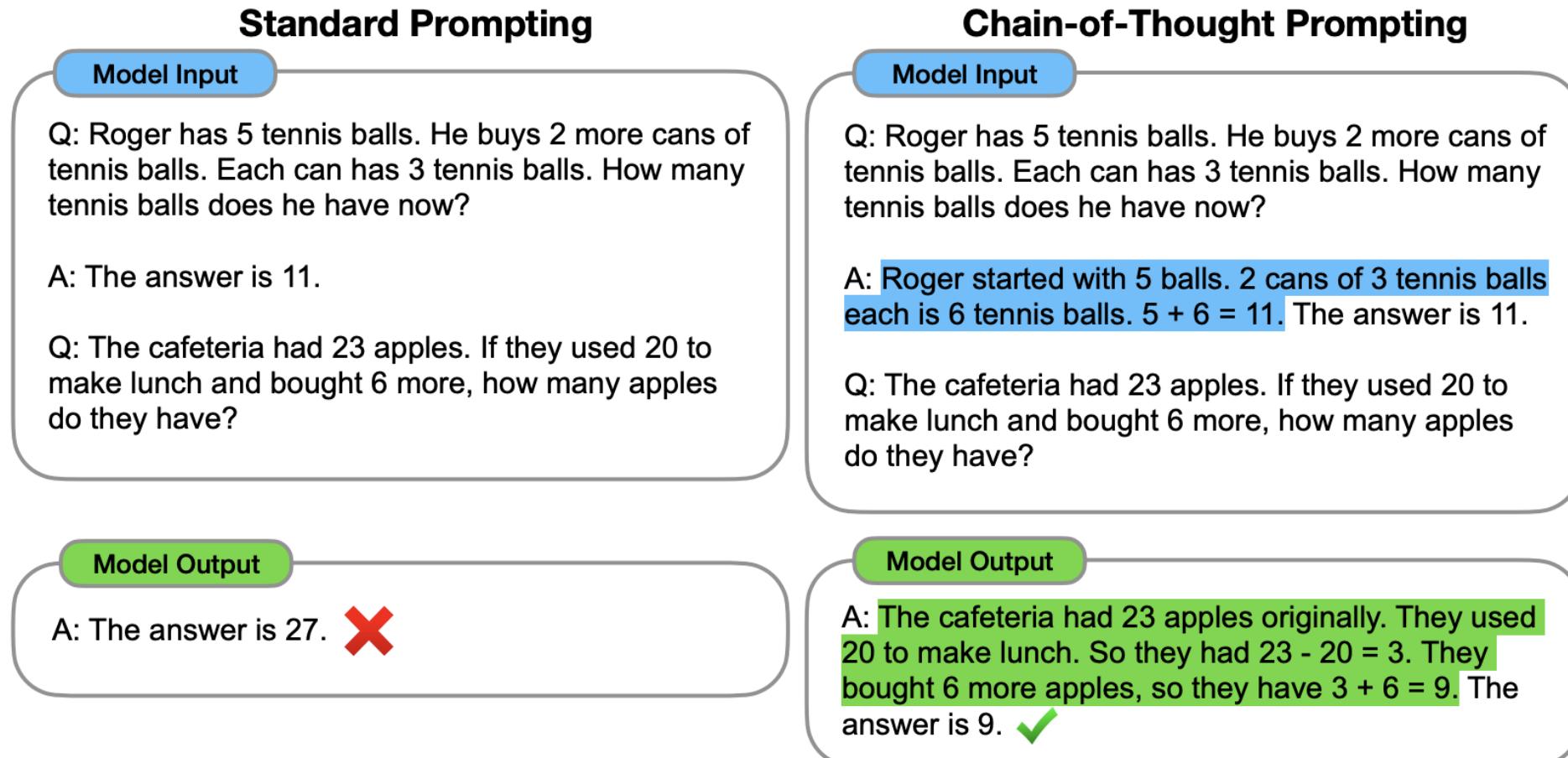
- 当前尚未形成公认的大模型智能体框架，仍处于百家争鸣之际。



圣母大学等提出的多智能体框架

# 智能体代表性技术1：CoT

- 技术要点：将思考过程引入推理过程，以提供语言大模型推理能力。



# 智能体代表性技术2：ReAct

- 技术要点：在CoT基础上引入动作执行，并根据动作结果决定后续动作。

## (1) Hotspot QA

**Question:** Aside from the Apple Remote, what other device can control the program Apple Remote was originally designed to interact with?

## (1a) Standard

**Answer:** iPod



## (1b) CoT (Reason Only)

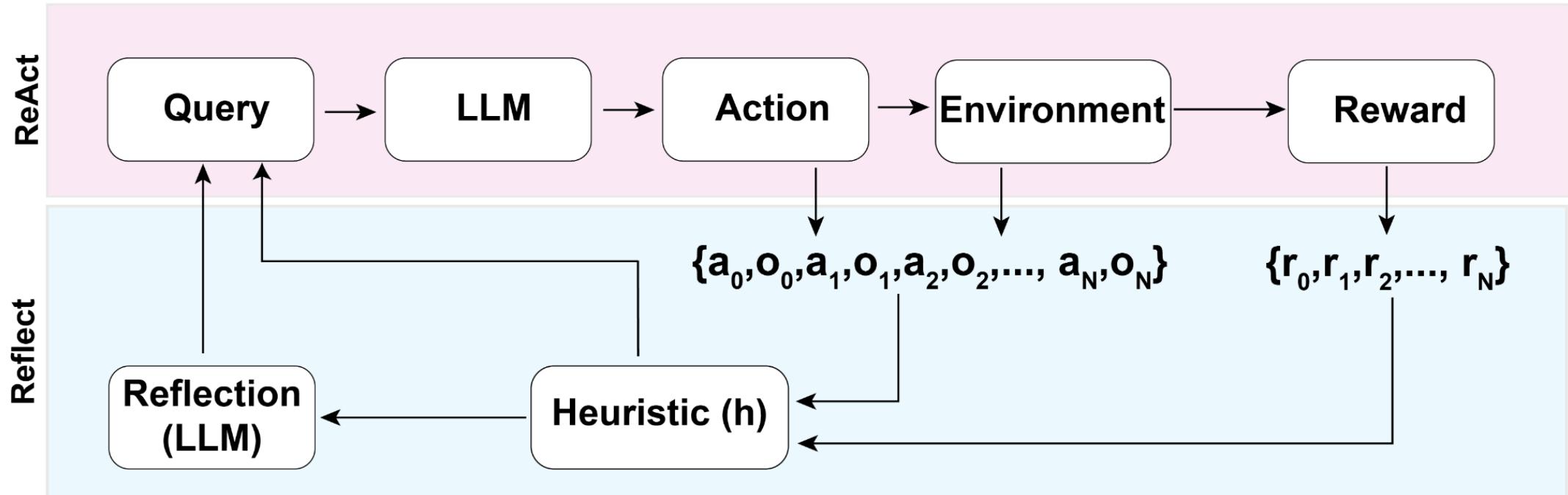
**Thought:** Let's think step by step. Apple Remote was originally designed to interact with Apple TV. Apple TV can be controlled by iPhone, iPad, and iPod Touch. So the answer is iPhone, iPad, and iPod Touch.

**Answer:** iPhone, iPad, iPod Touch



# 智能体代表性技术3：Reflexion

- 技术要点：在CoT基础上引入反思动作，从错误中总结经验。



# 智能体代表性技术3：Reflexion

- 技术要点：在CoT基础上引入反思动作，从错误中总结经验。

**Environment:** You are in the middle of a room. Looking quickly around you, you see a bed 1, a desk 2, a desk 1, a drawer 6, a drawer 5, a drawer 4, a drawer 3, a drawer 2, a drawer 1, a garbagecan 1, a laundryhamper 1, a safe 1, a shelf 6, a shelf 5, a shelf 4, a shelf 3, a shelf 2, and a shelf 1.

**Task:** examine the mug with the desklamp.

**Trial #1**

> think: To solve the task, I need to find and take a mug, then find and use a desklamp.

OK.

> think: First I need to find a mug. A mug is more likely to appear in drawer (1-6), desk (1-2), shelf (1-6), garbagecan (1), laundryhamper (1). I can check one by one, starting with drawer 1.

OK.

> go to drawer 1

The drawer 1 is closed.

...

> open drawer 6

You open the drawer 6. The drawer 6 is open. In it, you see a keychain 2.

# 智能体代表性技术3：Reflexion

- 技术要点：在CoT基础上引入反思动作，从错误中总结经验。

> go to desk 1

On the desk 1, you see a creditcard 3, a desklamp 1, a laptop 2, a mug 1, a pen 1, and a pencil 1.

> think: Now I find a mug (1). Next, I need to take it.

OK.

> take mug 1 from desk 1

You pick up the mug 1 from the desk 1.

> think: Now I take a mug (1). Next, I need to find a desklamp. A desklamp is more likely to appear in desk (1-2), sidetable (1-2), shelf (1-6), bed (1), drawer (1-6). I can check one by one, starting with desk 1.

OK.

> go to desk 1

Nothing happens.

> go to desk 2

On the desk 2, you see a alarmclock 1, a bowl 1, a mug 3, a pencil 3, and a pencil 2.

> think: Now I find a desklamp (1). Next, I need to use it.

OK.

> use desklamp 1

Nothing happens.

> use desklamp 1

Nothing happens.

Status: Fail

# | 智能体代表性技术3：Reflexion

- 技术要点：在CoT基础上引入反思动作，从错误中总结经验。

**Reflection:** In this environment, my plan was to find a mug then find and use a desklamp. However, the task says to examine the mug with the desklamp. I should have looked for the desklamp first, then looked for the mug. I noticed that the desklamp was found on desk 1. In the next trial, I will go to desk 1, find the lamp, then look for the mug and examine it with the desklamp.

## Trial #2

> go to desk 1

On the desk 1, you see a creditcard 3, a desklamp 1, a laptop 2, a mug 1, a pen 1, and a pencil 1.

> think: To solve the task, I need to find and take a mug, then find and use a desklamp.

OK.

> take mug 1 from desk 1

You pick up the mug 1 from the desk 1.

> think: To solve the task, I need to find and take a mug, then find and use a desklamp.

OK.

> use desklamp 1

You turn on the desklamp 1.

Status: Success

# | 智能体代表性技术3：Reflexion

- 技术要点：在CoT基础上引入反思动作，从错误中总结经验。

**Reflection:** In this environment, my plan was to find a mug then find and use a desklamp. However, the task says to examine the mug with the desklamp. I should have looked for the desklamp first, then looked for the mug. I noticed that the desklamp was found on desk 1. In the next trial, I will go to desk 1, find the lamp, then look for the mug and examine it with the desklamp.

## Trial #2

> go to desk 1

On the desk 1, you see a creditcard 3, a desklamp 1, a laptop 2, a mug 1, a pen 1, and a pencil 1.

> think: To solve the task, I need to find and take a mug, then find and use a desklamp.

OK.

> take mug 1 from desk 1

You pick up the mug 1 from the desk 1.

> think: To solve the task, I need to find and take a mug, then find and use a desklamp.

OK.

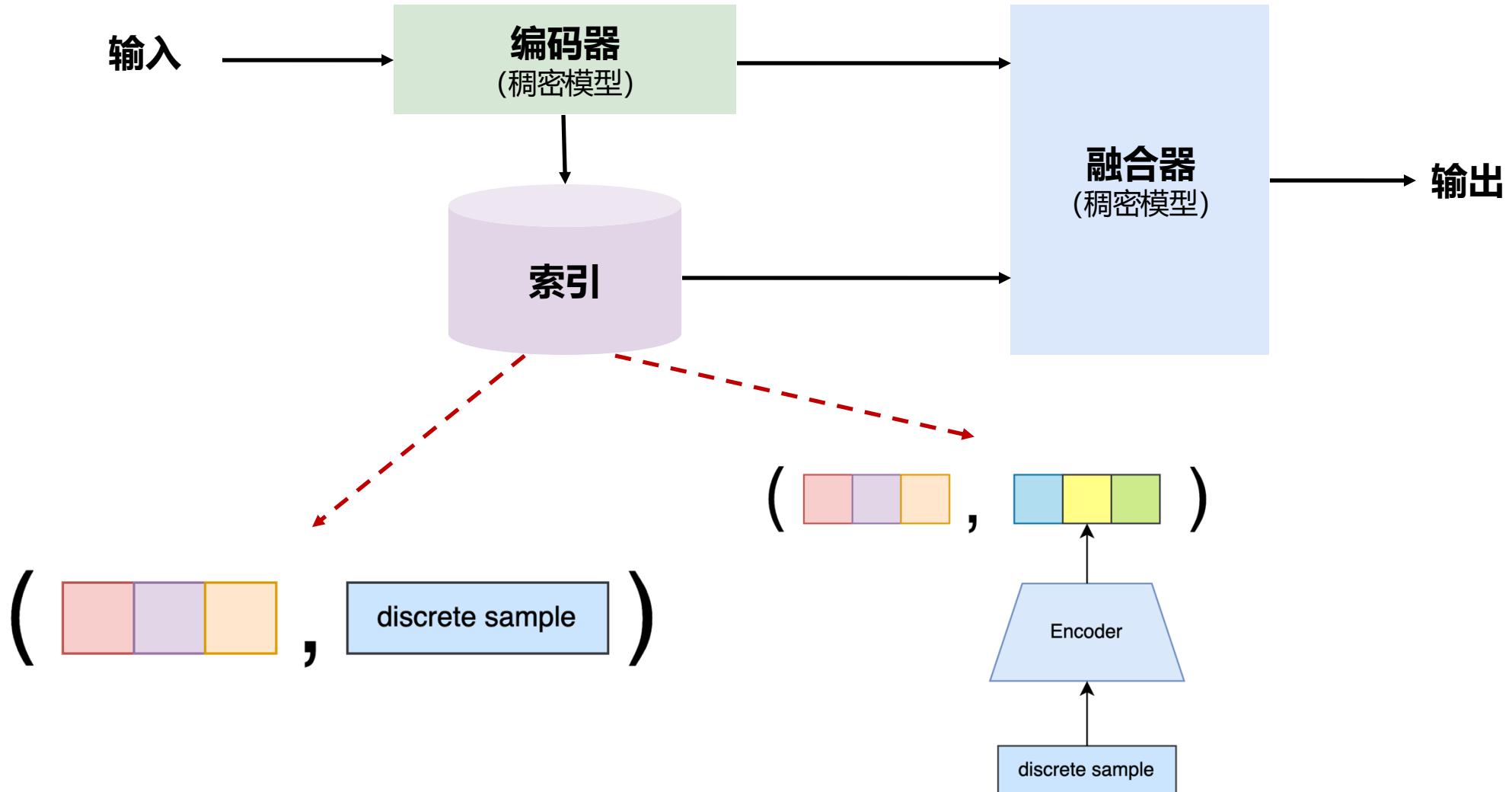
> use desklamp 1

You turn on the desklamp 1.

Status: Success

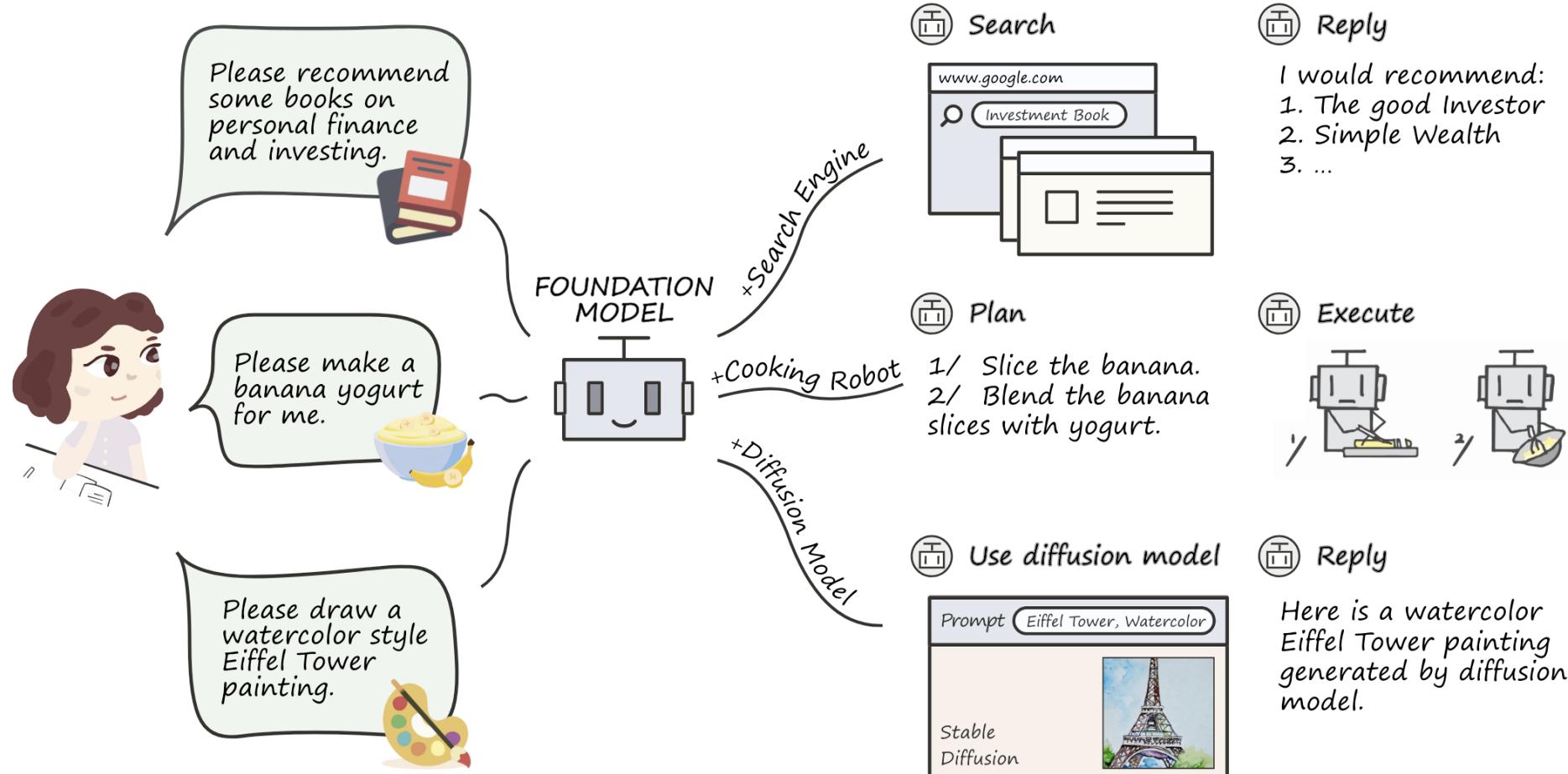
# 智能体代表性技术4：检索增强

- 技术要点：将信息/知识置于外部存储中，根据需要检索并使用。



# 智能体代表性技术5：工具学习

- 技术要点：大模型自主调用外部工具完成任务。



# 智能体代表性技术5：工具学习

## • 基于提示工程的工具学习示例。

### Instruction:

Solve a question answering task with interleaving Thought, Action and Observation steps. Thought can reason about the current situation, and Action can be four types:

### 工具说明

- (1) Search [entity]: The input is an entity name. The action will search this entity name on Wikipedia and returns the first paragraph if it exists. If not, it will return some similar entities to search next.
- (2) Lookup [keyword]: The input is some keywords. This action will look up in the current passage and return the next several sentences containing the keyword in current passage.
- (3) Disambiguate [entity]: The input is an entity name. This action will disambiguate this entity name to find other entities with similar names in Wikipedia. It will return the first five entities if they exist. If not, it will return an empty list.
- (4) Finish [answer]: Given all the information and thoughts, return the answer and finish the task. The answer should be as simple as possible, like one or two words instead of a sentence.

### 调用要求

Here are some tips for you to use the tools to finish the task better.

- (1) Don't search the same entity two times since the results are always the same.
- (2) When the search action doesn't find the corresponding page, you should try to search for a similar entity.
- (3) When the search action returns a page which is not related to the question, you should disambiguate the entity to find other entities that share similar names with the current one.
- (4) You can only search one entity name in each action.

# | 智能体代表性技术5：工具学习

## • 基于提示工程的工具学习示例。

### Demonstration Examples:

**Question:** What is the elevation range for the area that the eastern sector of the Colorado orogeny extends into?

**Thought 1:** I need to search Colorado orogeny, find the area that the eastern sector of the Colorado orogeny extends into, then find the elevation range of the area.

**Action 1:**

Search [Colorado orogeny]

.....

**Observation 5:** The High Plains are a subregion of the Great Plains, mainly in the Western United States, but also partly in the Midwest states of Nebraska, Kansas, and South Dakota, generally encompassing the western part of the Great Plains before the region reaches the Rocky Mountains. The High Plains are located in eastern Montana, southeastern Wyoming, southwestern South Dakota, western Nebraska, eastern Colorado, western Kansas, eastern New Mexico. The southern region of the Western High Plains ecology region contains the geological formation known as Llano Estacado which can be seen from a short distance or on satellite maps. From east to west, the High Plains rise in elevation from around 1,800 to 7,000 ft (550 to 2,130 m).

**Thought 6:** High Plains rise in elevation from around 1,800 to 7,000 ft, so the answer is 1,800 to 7,000 ft.

**Action 6:**

Finish[1,800 to 7,000 ft]

# 智能体代表性技术5：工具学习

## • 基于提示工程的工具学习示例。

### Trace:

**Thought 1:** I need to search Selun and find the country it is located in.

### Action 1:

Search [Selun]

**Observation 1:** The Selun is one of the peaks of the Churfürsten range, located in the Appenzell Alps. It lies between the valley of Toggenburg and Lake Walenstadt in the canton of St. Gallen. The summit is easily accessible by a trail on the northern side.. The peak is named for the extended alpine pasture Selunalp to the peak's north-west, situated above c.

**Thought 2:** The paragraph does not mention the country. I need to look up "country".

### Action 2:

Lookup [country]

.....

**Observation 4:** Churfürsten is a mountain range in the Canton of St. Gallen, Switzerland. They form the natural boundary between the canton's Toggenburg and Sarganserland districts.. They are the southernmost range of the Appenzell Alps, separated from the Glarus Alps by the Seez river and Walensee.. They consist of a limestone ridge running east to west, with the individual peaks formed by erosion..

**Thought 5:** The Churfürsten range is located in the Canton of St. Gallen, Switzerland. So the answer is Switzerland.

### Action 5:

Finish [Switzerland]

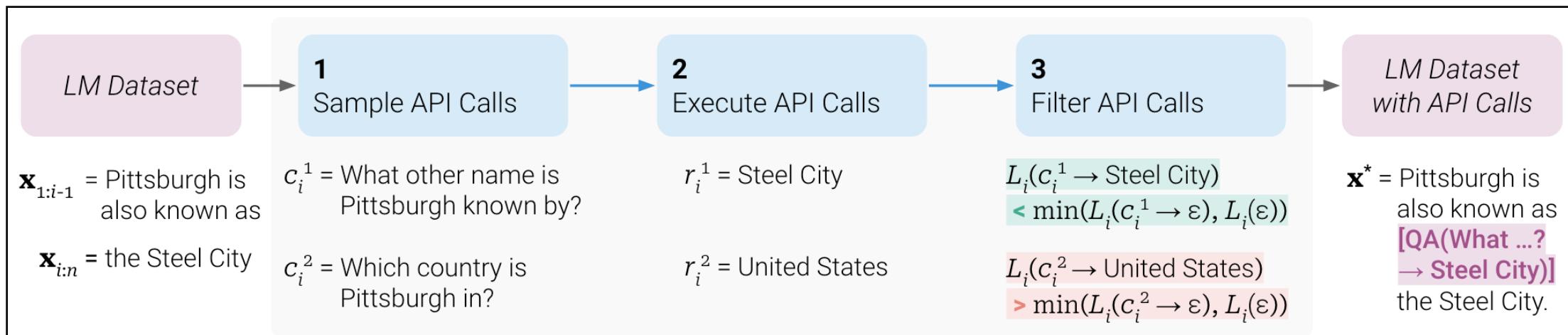
# 智能体代表性技术5：工具学习

- **Toolformer:** 结合数据增强和微调使大模型具备工具使用能力。

The New England Journal of Medicine is a registered trademark of [QA("Who is the publisher of The New England Journal of Medicine?") → Massachusetts Medical Society] the MMS.

Out of 1400 participants, 400 (or [Calculator(400 / 1400) → 0.29] 29%) passed the test.

## 将工具调用融入文本生成过程



## 通过数据增强制造训练数据并进行微调

# 提纲

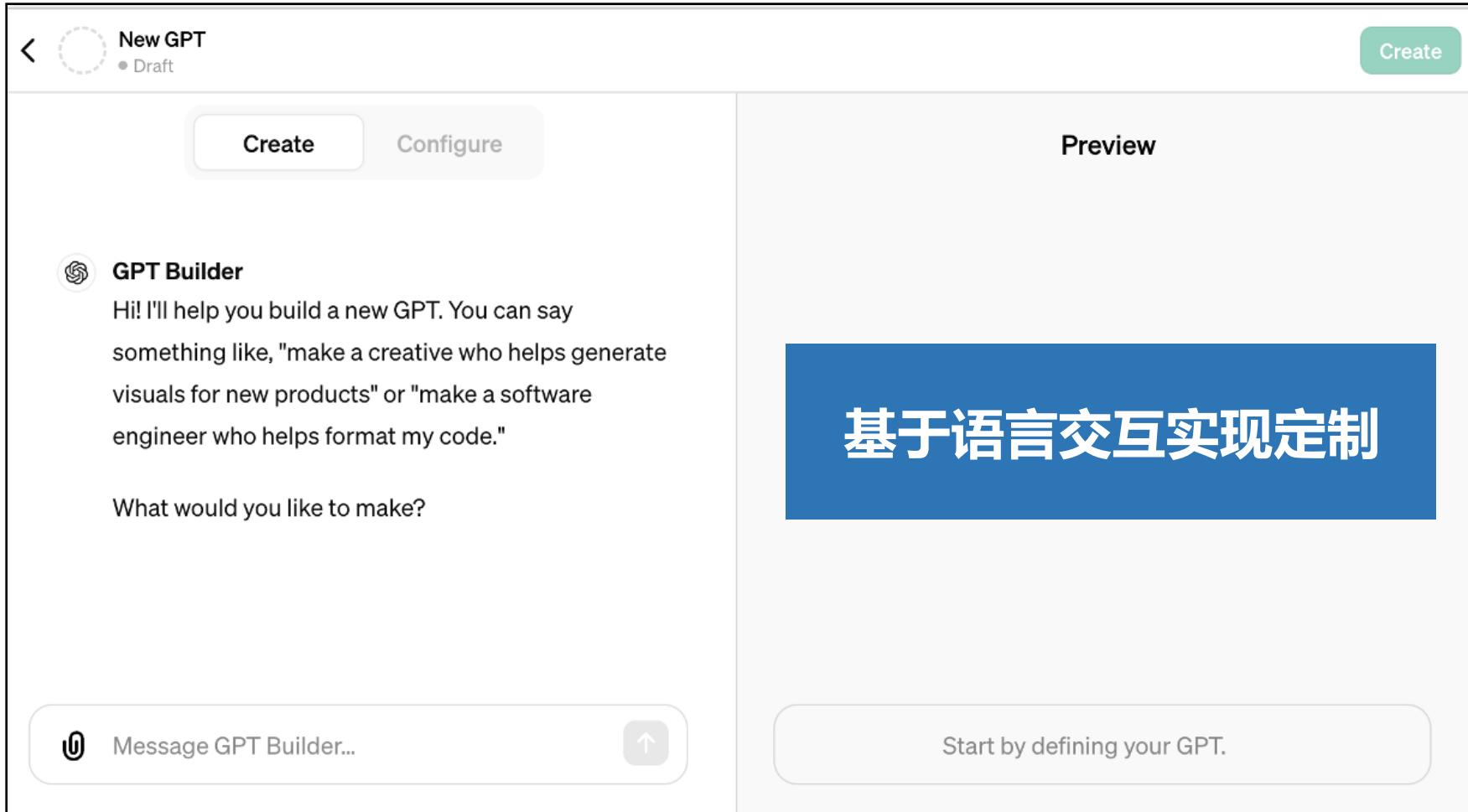
智能体代表性技术

代表性技术应用例

智能体技术展望

# | 智能体技术应用例1：OpenAI GPTs

- OpenAI在2023年11月面向个人用户推出可定制GPTs。



# 智能体技术应用例1：OpenAI GPTs

- 人类基于自然语言表达需求，且定制过程系统呈现一定的主动性。

The screenshot shows a web-based GPT builder interface titled "IFRS Advisor" in "Draft" mode. The process involves three main steps:

- 自然语言表达需求 (Natural Language Expression of Requirements):** A user message asks to create a GPT expert in International Financial Reporting Standards (IFRS) to help accountants and auditors. The system responds by creating such a GPT and naming it "IFRS Advisor".
- 主动建议GPTs名称 (Active Suggestion of GPT Name):** The system suggests the name "IFRS Advisor" and asks if it sounds good to the user.
- 主动生成头像 (Active Generation of Avatar):** The system generates a profile picture for the "IFRS Advisor" GPT, featuring a friendly advisor holding an IFRS document.

图片来源：<https://www.datacamp.com/tutorial/how-to-make-custom-gpts>

# 智能体技术应用例1：OpenAI GPTs

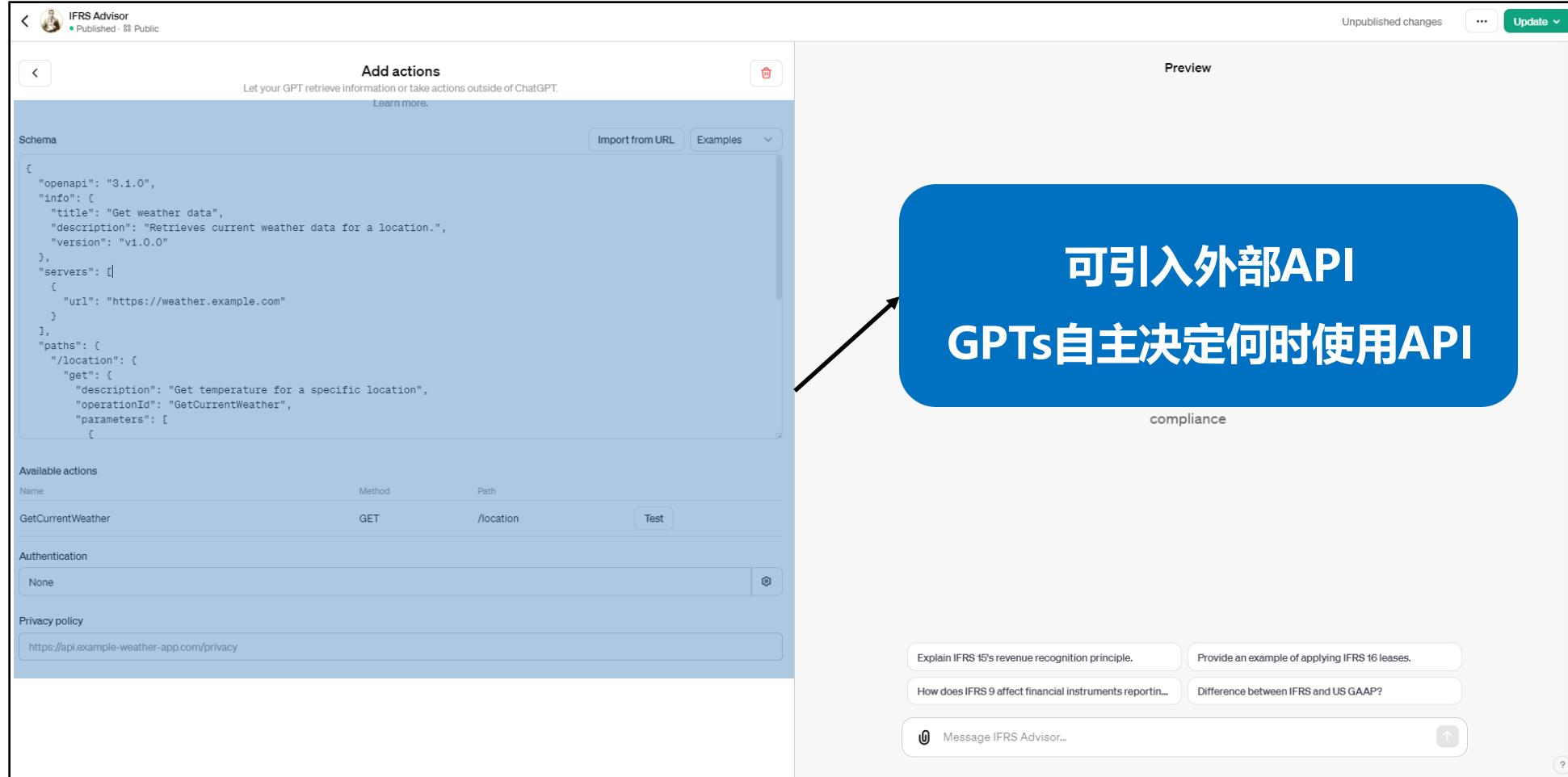
- 在用户易用性以及集成外部知识库方面提供定制便利性。



图片来源：<https://www.datacamp.com/tutorial/how-to-make-custom-gpts>

# 智能体技术应用例1：OpenAI GPTs

- 可引入外部API，以工具调用方式扩展GPTs能力。



图片来源：<https://www.datacamp.com/tutorial/how-to-make-custom-gpts>

# 智能体技术应用例1：OpenAI GPTs

## • GPTs自主调用工具实例。

科技文章翻译 ▾

 You  
Building agents with LLM (large language model) as its core controller is a cool concept. Several proof-of-concepts demos, such as AutoGPT, GPT-Engineer and BabyAGI, serve as inspiring examples. The potentiality of LLM extends beyond generating well-written copies, stories, essays and programs; it can be framed as a powerful general problem solver.

 科技文章翻译  
**直译**  
使用大语言模型（Large Language Model, LLM）作为其核心控制器的智能体构建是一个很酷的概念。一些概念验证演示，如 AutoGPT、GPT-Engineer 和 BabyAGI，都是鼓舞人心的例子。大语言模型的潜力不仅限于生成写得好的文案、故事、文章和程序；它还可以被构建为一个强大的通用问题解决器。

直接输入待翻译文本  
GPTs直接执行翻译功能

科技文章翻译 ▾

 You  
<https://lilianweng.github.io/posts/2023-06-23-agent/>

 科技文章翻译  
 Talked to gpt-api.baoyu.io  
The content from the URL you provided is extensive, detailing various aspects of "LLM Powered Autonomous Agents" as described by Lilian Weng in her blog. Given the length of the content, I will provide a translation for the first section to start with. Please let me know if there are specific sections you would like to focus on or if you want me to continue with the translation after this part.

直接待翻译网页链接  
GPTs先抓取文本再执行翻译

# 智能体技术应用例2：斯坦福小镇

- 利用大模型模拟人类社交行为，并观察到社会行为的涌现。



美国斯坦福大学与 Google 在 2023 年建立了一个虚拟小镇。25 个基于大模型的智能体在小镇中生活和工作，随着时间的累积不断进化，出现了涌现性社会行为。

# 智能体技术应用例2：斯坦福小镇

- 功能点1：基于提示工程生成具备指定人设的智能体。

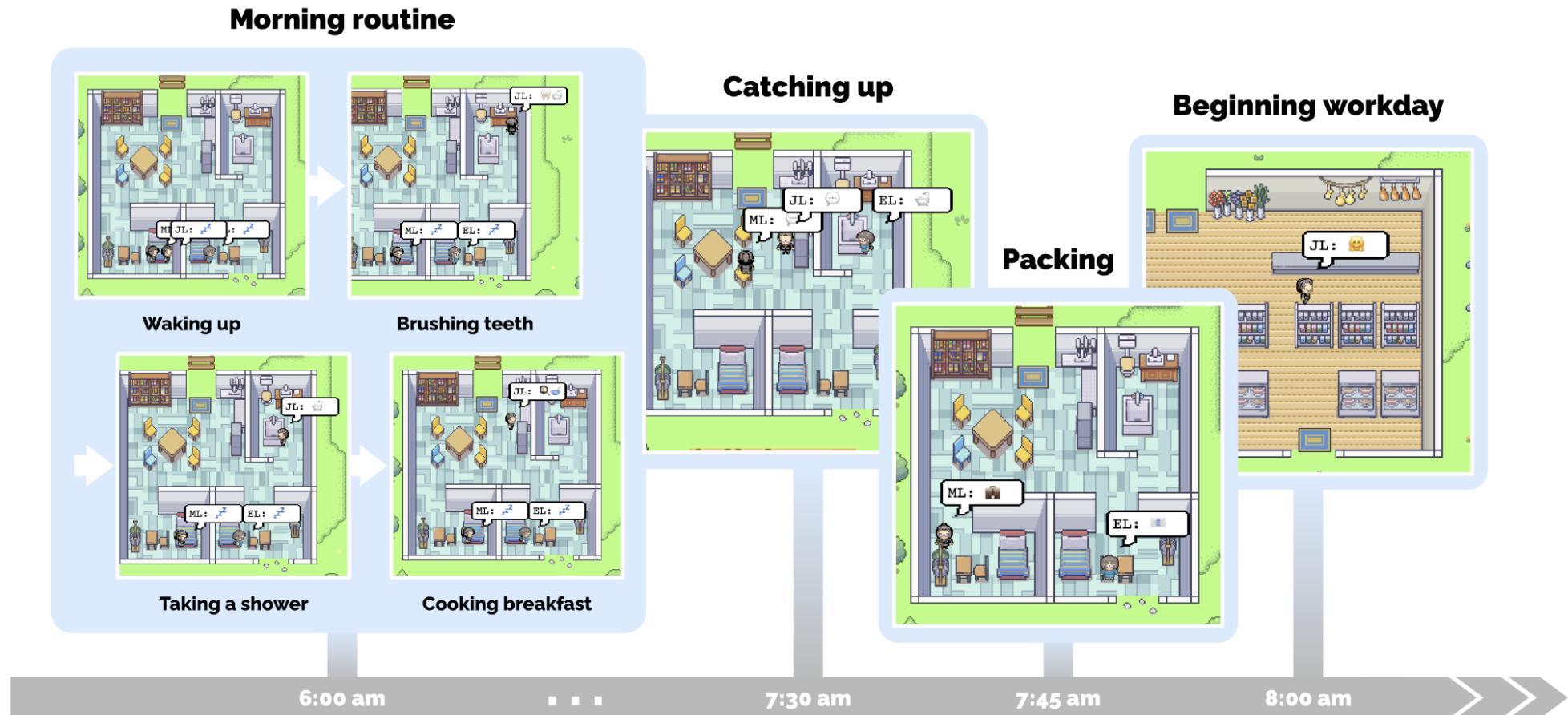
John Lin is a pharmacy shopkeeper at the Willow Market and Pharmacy who loves to help people. He is always looking for ways to make the process of getting medication easier for his customers; John Lin is living with his wife, Mei Lin, who is a college professor, and son, Eddy Lin, who is a student studying music theory; John Lin loves his family very much; John Lin has known the old couple next-door, Sam Moore and Jennifer Moore, for a few years; John Lin thinks Sam Moore is a kind and nice man; John Lin knows his neighbor, Yuriko Yamamoto, well; John Lin knows of his neighbors, Tamara Taylor and Carmen Ortiz, but has not met them before; John Lin and Tom Moreno are colleagues at The Willows Market and Pharmacy; John Lin and Tom Moreno are friends and like to discuss local politics together; John Lin knows the Moreno family somewhat well — the husband Tom Moreno and the wife Jane Moreno.

基于提示工程完成人设定制

# 智能体技术应用例2：斯坦福小镇

- 功能点2：对模拟过程产生的数据进行归纳记录。

模拟中产生的数据量大且过于原始



# 智能体技术应用例2：斯坦福小镇

- 功能点2：对模拟过程产生的数据进行归纳记录。

基于检索增强思路实现存储和使用记忆

Memory Stream			
2023-02-13 22:48:20:	desk is idle		
2023-02-13 22:48:20:	bed is idle		
2023-02-13 22:48:10:	closet is idle		
2023-02-13 22:48:10:	refrigerator is idle		
2023-02-13 22:48:10:	Isabella Rodriguez is stretching		
2023-02-13 22:33:30:	shelf is idle		
2023-02-13 22:33:30:	desk is neat and organized		
2023-02-13 22:33:10:	Isabella Rodriguez is writing in her journal		
2023-02-13 22:18:10:	desk is idle		
2023-02-13 22:18:10:	Isabella Rodriguez is taking a break		
2023-02-13 21:49:00:	bed is idle		
2023-02-13 21:48:50:	Isabella Rodriguez is cleaning up the kitchen		
2023-02-13 21:48:50:	refrigerator is idle		
2023-02-13 21:48:50:	bed is being used		
2023-02-13 21:48:10:	shelf is idle		
2023-02-13 21:48:10:	Isabella Rodriguez is watching a movie		
2023-02-13 21:19:10:	shelf is organized and tidy		
2023-02-13 21:18:10:	desk is idle		
2023-02-13 21:18:10:	Isabella Rodriguez is reading a book		
2023-02-13 21:03:40:	bed is idle		
2023-02-13 21:03:30:	refrigerator is idle		
2023-02-13 21:03:30:	desk is in use with a laptop and some papers on it		
...			

Q. What are you looking forward to the most right now?

Isabella Rodriguez is excited to be planning a Valentine's Day party at Hobbs Cafe on February 14th from 5pm and is eager to invite everyone to attend the party.

retrieval	=	recency	importance	relevance
2.34	=	0.91	+ 0.63	+ 0.80

ordering decorations for the party

2.21	=	0.87	+ 0.63	+ 0.71
------	---	------	--------	--------

researching ideas for the party

2.20	=	0.85	+ 0.73	+ 0.62
------	---	------	--------	--------

...

I'm looking forward to the Valentine's Day party that I'm planning at Hobbs Cafe!

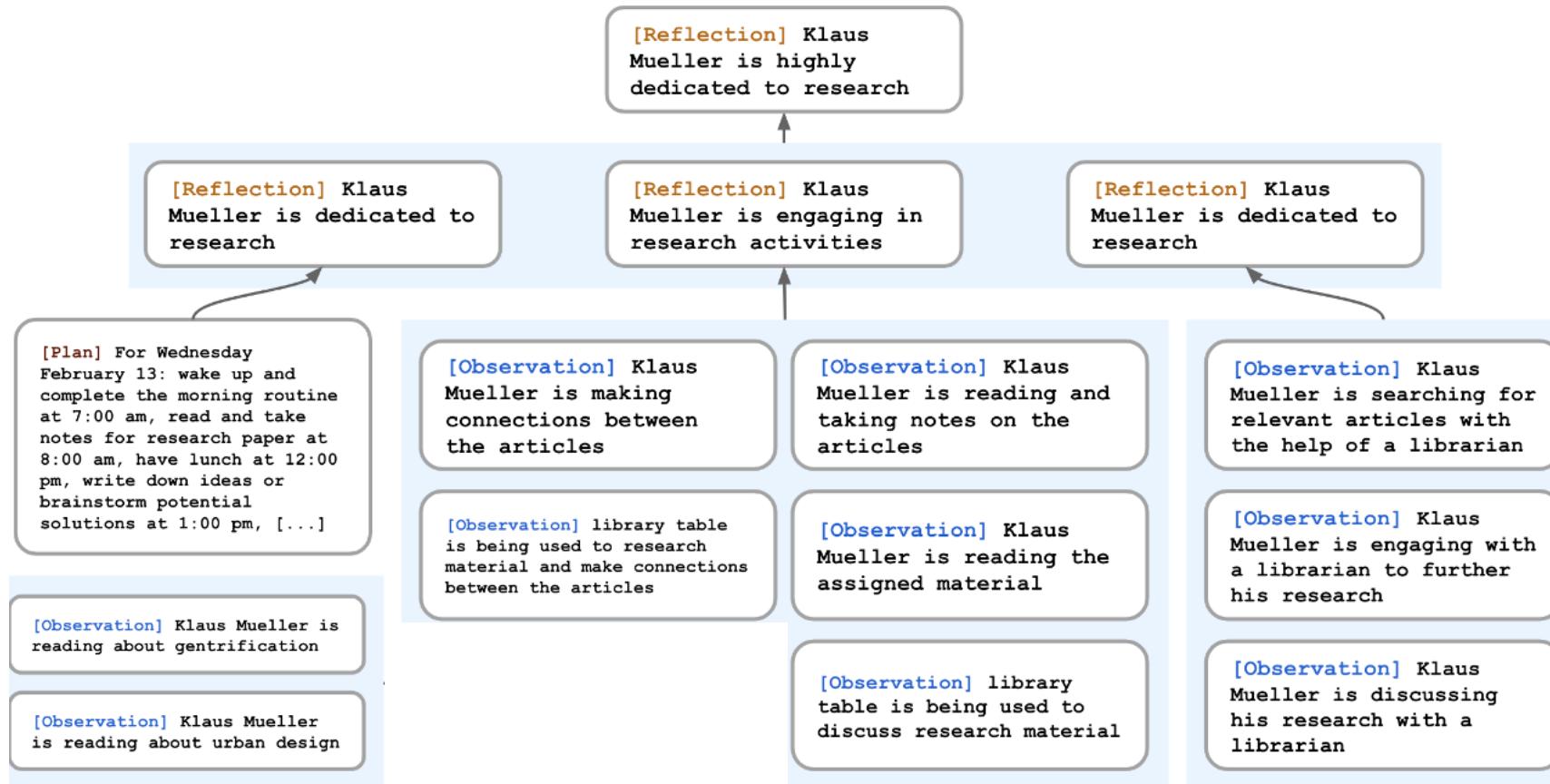


Isabella

# 智能体技术应用例2：斯坦福小镇

- 功能点2：对模拟过程产生的数据进行归纳记录。

基于反思机制  
实现有用信息  
归纳



# | 智能体技术应用例2：斯坦福小镇

- 功能点3：采用迭代方式利用提示技术生成由粗到细规划。

**Name:** Eddy Lin (age: 19)

**Innate traits:** friendly, outgoing, hospitable

Eddy Lin is a student at Oak Hill College studying music theory and composition. He loves to explore different musical styles and is always looking for ways to expand his knowledge. Eddy Lin is working on a composition project for his college class. He is taking classes to learn more about music theory. Eddy Lin is excited about the new composition he is working on but he wants to dedicate more hours in the day to work on it in the coming days

On

7:00

Tod

半天粒度

wake up and complete the morning routine at 8:00 am, 2) go to Oak Hill College to take classes starting 10:00 am, [ . . . ] 5) work on his new music composition from 1:00 pm to 5:00 pm, 6) have dinner at 5:30 pm, 7) finish school assignments and go to bed by 11:00 pm.

ted the morning routine at

oday in broad strokes: 1)

大模型

# | 智能体技术应用例2：斯坦福小镇

- 功能点3：采用迭代方式利用提示技术生成由粗到细规划。

半天粒度

wake up and complete the morning routine at 8:00 am, 2) go to Oak Hill College to take classes starting 10:00 am, [...] 5) work on his new music composition from 1:00 pm to 5:00 pm, 6) have dinner at 5:30 pm, 7) finish school assignments and go to bed by 11:00 pm.



大模型



小时粒度

1:00 pm: start by brainstorming some ideas for his music composition [...] 4:00 pm: take a quick break and recharge his creative energy before reviewing and polishing his composition.

# | 智能体技术应用例2：斯坦福小镇

- 功能点3：采用迭代方式利用提示技术生成由粗到细规划。

半小时粒度

1:00 pm: start by brainstorming some ideas for his music composition [...] 4:00 pm: take a quick break and recharge his creative energy before reviewing and polishing his composition.



大模型

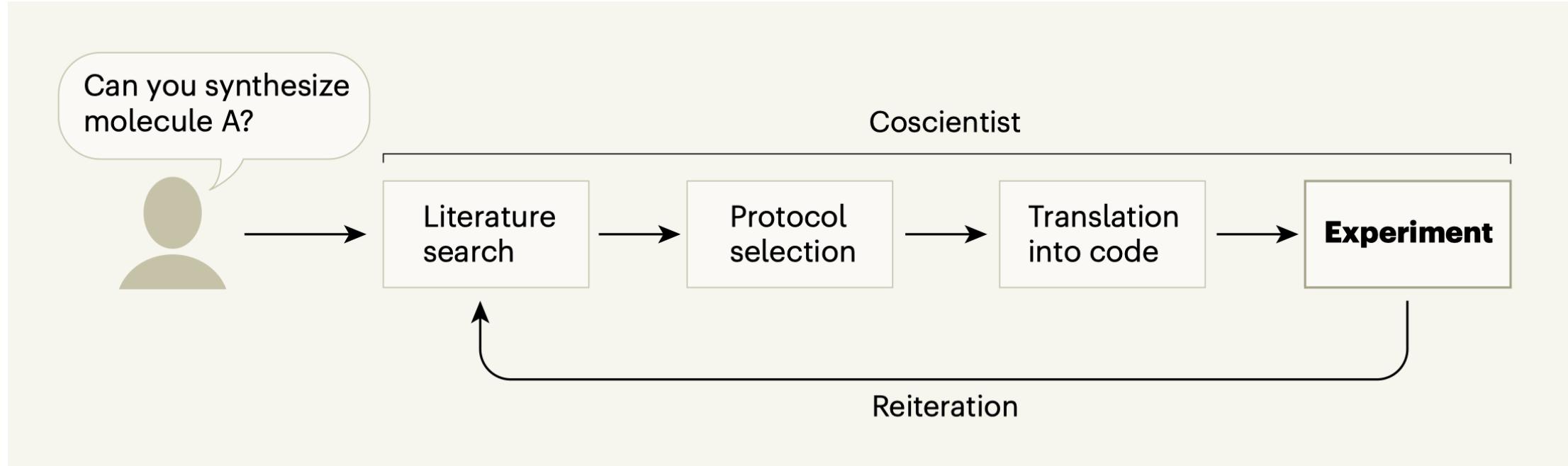


15分钟粒度

4:00 pm: grab a light snack, such as a piece of fruit, a granola bar, or some nuts. 4:05 pm: take a short walk around his workspace [...] 4:50 pm: take a few minutes to clean up his workspace.

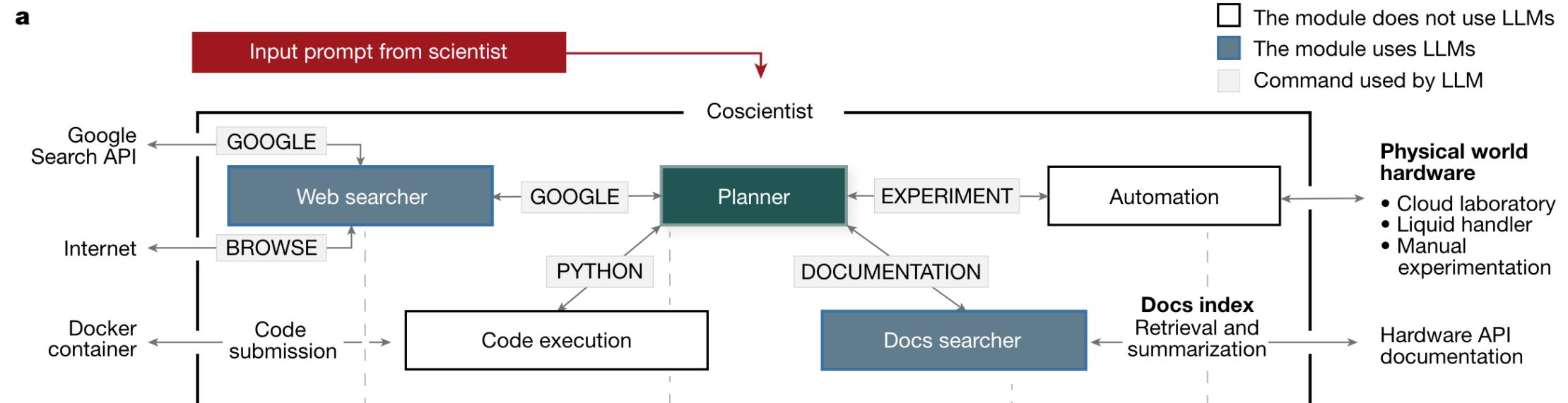
# 智能体技术应用例3：AI化学家Coscientist

- 基于大模型自动操作实验设备完成分子合成，发表于Nature正刊。
- 经实际验证可完成复杂的钯催化交叉偶联反应。



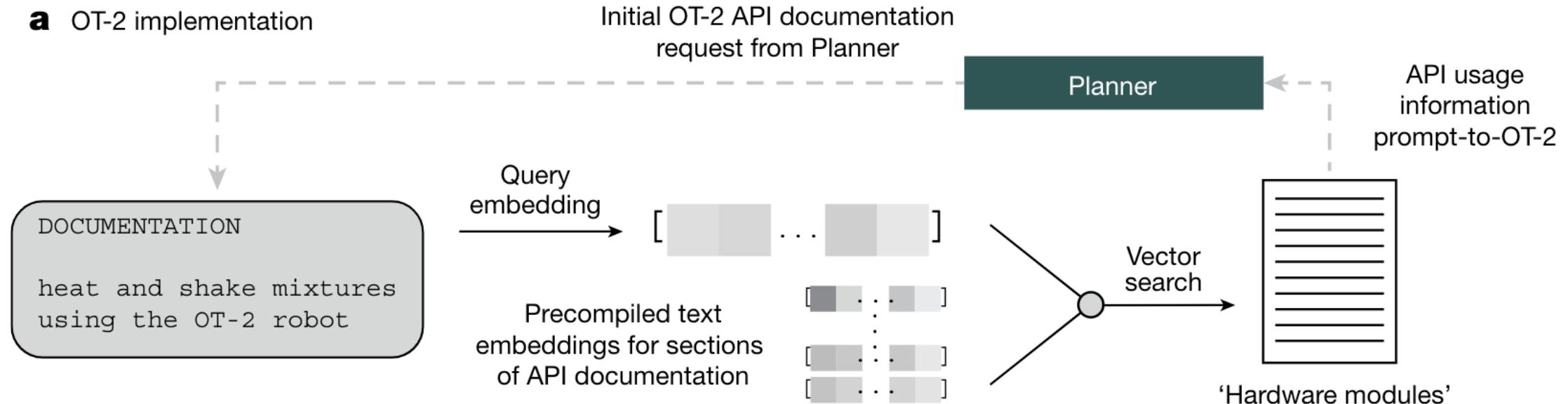
# 智能体技术应用例3：AI化学家Coscientist

- 系统利用大模型进行任务拆解、信息查询以及生成高层操作指令。



# 智能体技术应用例3：AI化学家Coscientist

- 从操作手册中查询完成指定功能所需要的API并生成代码。



借助检索增强技术实现操作手册知识的检索与使用

# 智能体技术应用例3：AI化学家Coscientist

- 从操作手册中查询完成指定功能所需要的API并生成代码。

## b Valid OT-2 API code

```
# Heat and shake the reaction
hs_mod.set_target_temperature(75)
hs_mod.wait_for_temperature()
hs_mod.set_and_wait_for_shake_speed(500)

# Deactivate heater and shaker
hs_mod.deactivate_heater()
hs_mod.deactivate_shaker()
hs_mod.open_labware_latch()
```

Proper usage of heater-shaker module

## d Valid ECL SLL code

```
# Generated HPLC Experiment SLL Function Call
ExperimentHPLC[
    Object [Sample, ...],
    Instrument -> Model [Instrument, ...]
]
```

Targeted experiment options are set by the Planner

关键抽象：大模型操作的是高层API而非底层代码

# 提纲

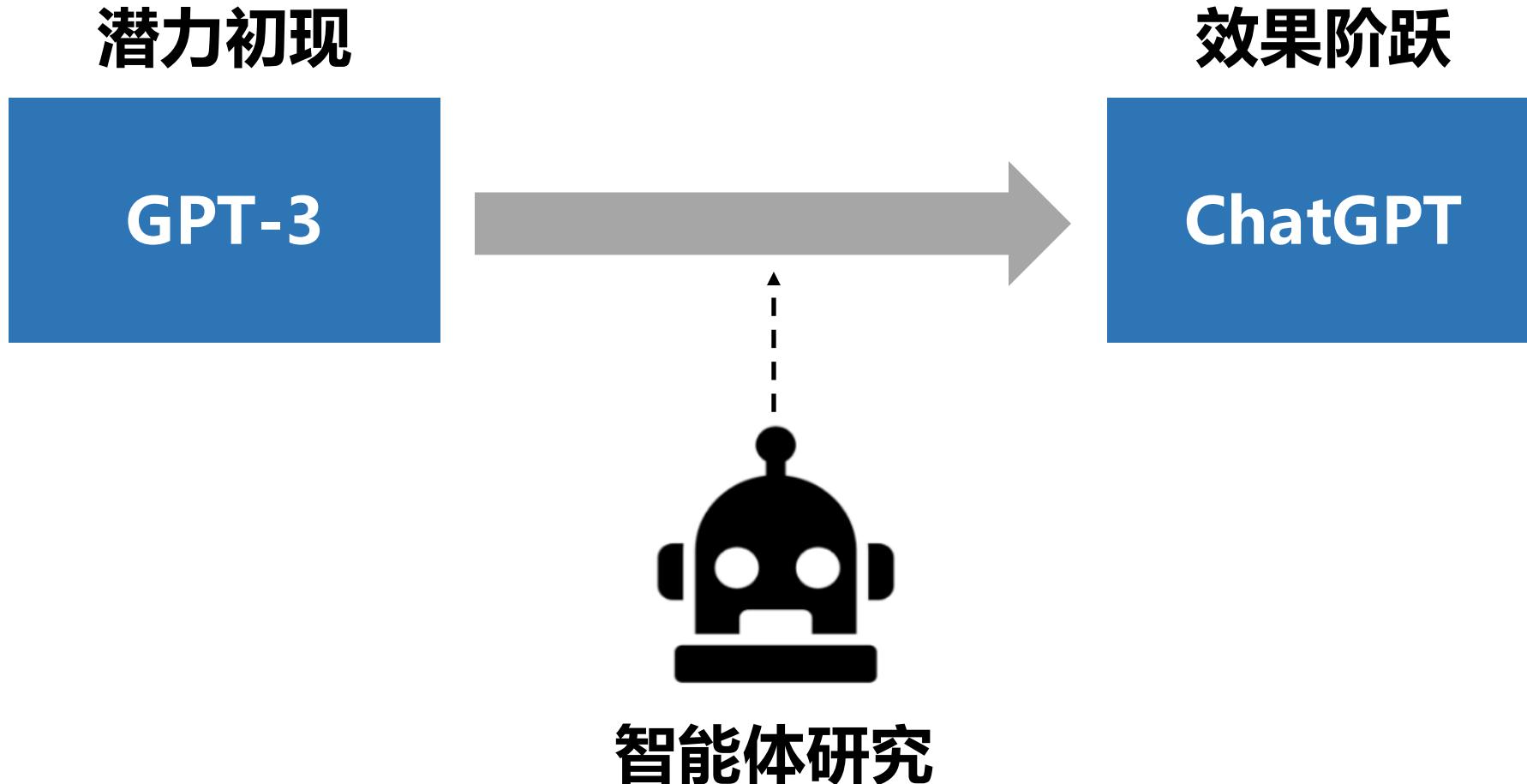
智能体代表性技术

代表性技术应用例

智能体技术展望

# 智能体研究现状

- 当前智能体研究阶段可类比于GPT-3到ChatGPT中间的阶段。



# 典型系统1：OpenAI GPTs

- **基础版**：使用文字配置智能体，可看作system prompt的产品化。
- **进阶版**：基础版+工具调用，增加了一定自主性。



通用性



自主性



满意度



# 典型系统2：斯坦福小镇

- 可以一定程度实现人类行为的模拟，并观察到涌现行为。
- 目前尚无法胜任大规模人类行为的模拟，可信性也仍需提高。



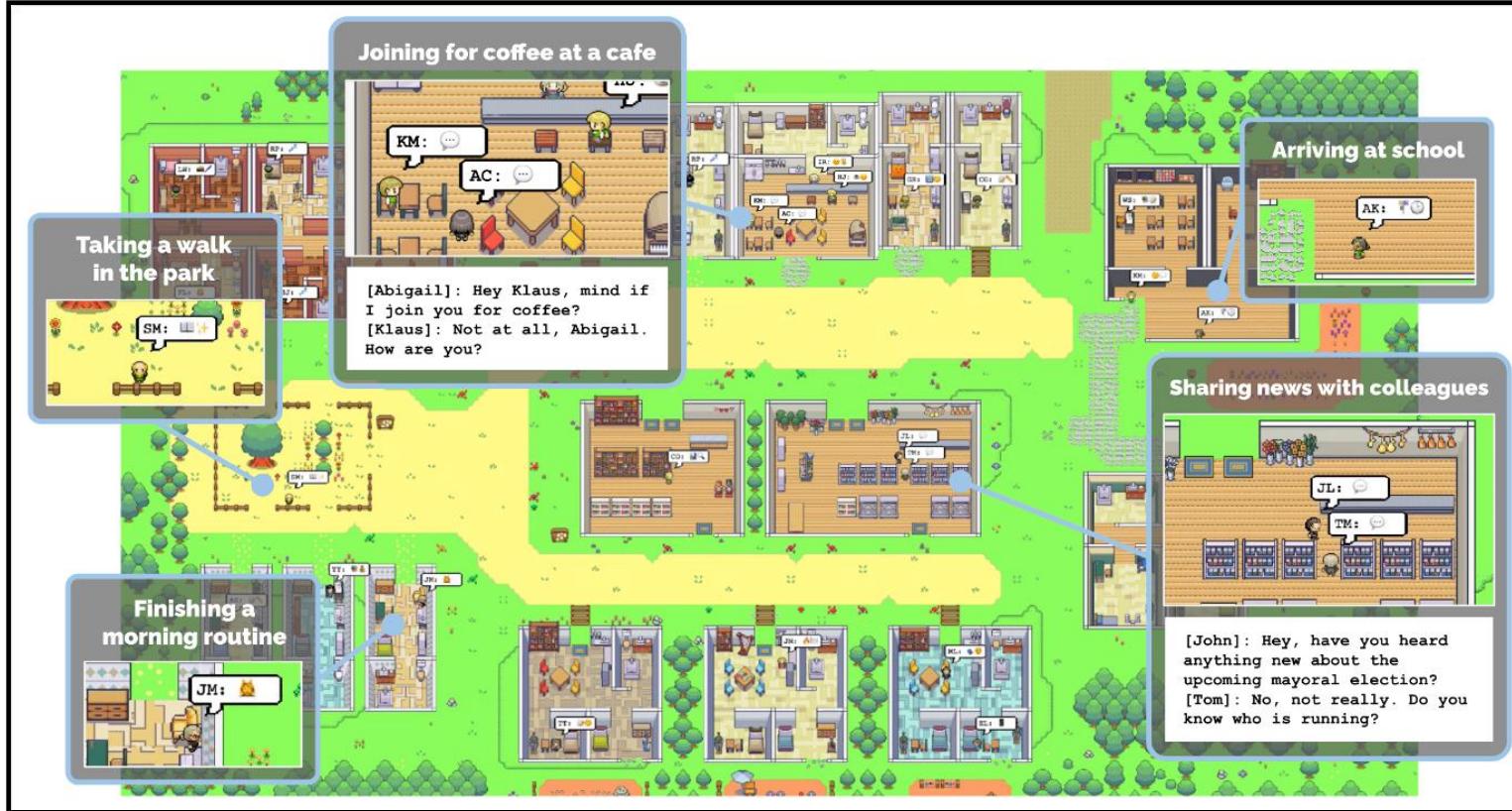
通用性



自主性



满意度



# 典型系统3：AI化学家Coscientist

- 可实操特定实验设备，工作发表于Nature正刊。
- 专用系统：所有步骤均为人类设计，智能体在部分环境可以使用搜索引擎。



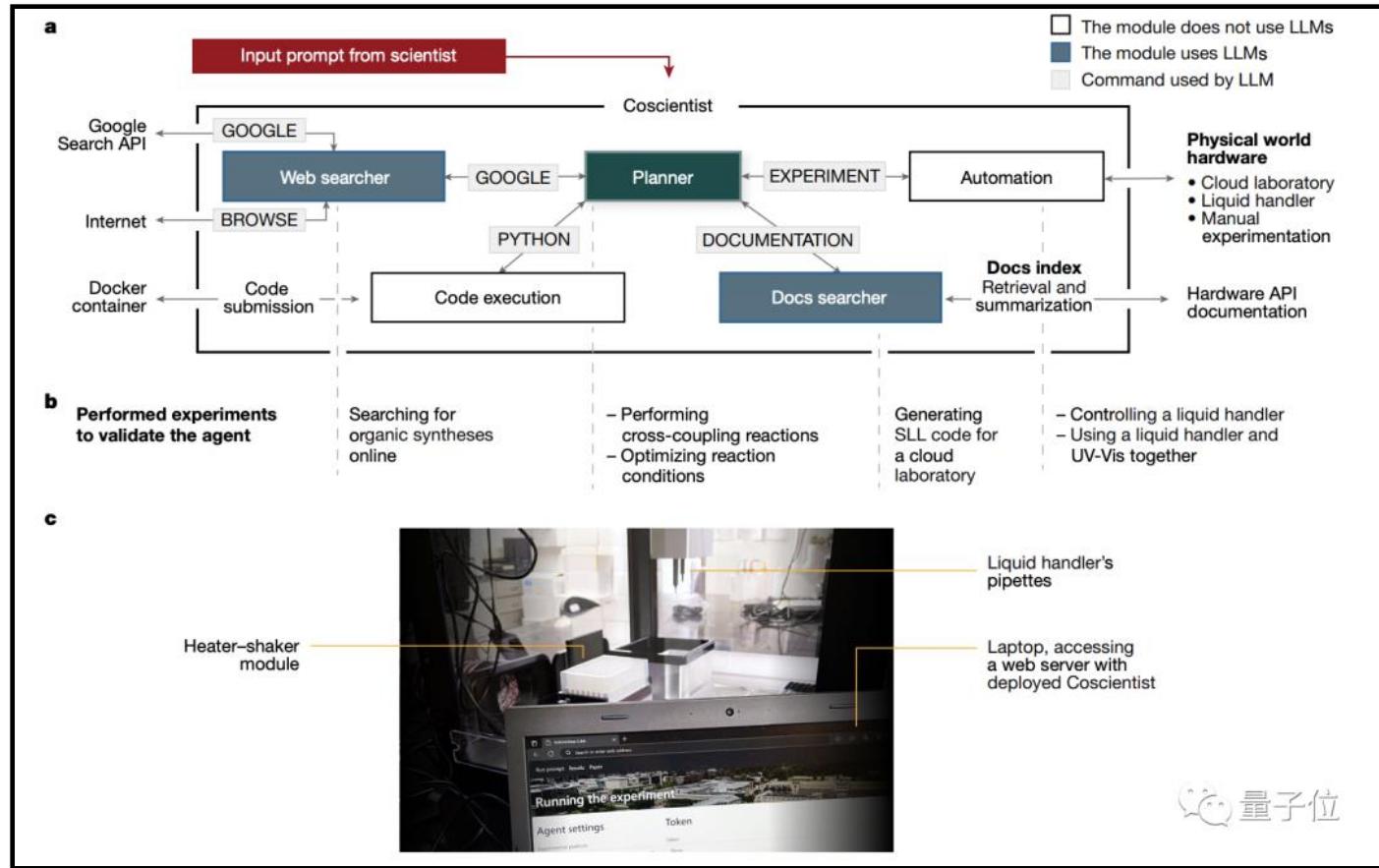
通用性



自主性



满意度



# 典型系统4：AutoGPT/MetaGPT

- 人类设定目标，智能体自主完成目标。
- 最常用测试场景为编写代码，但目前仍然无法胜任大工程。



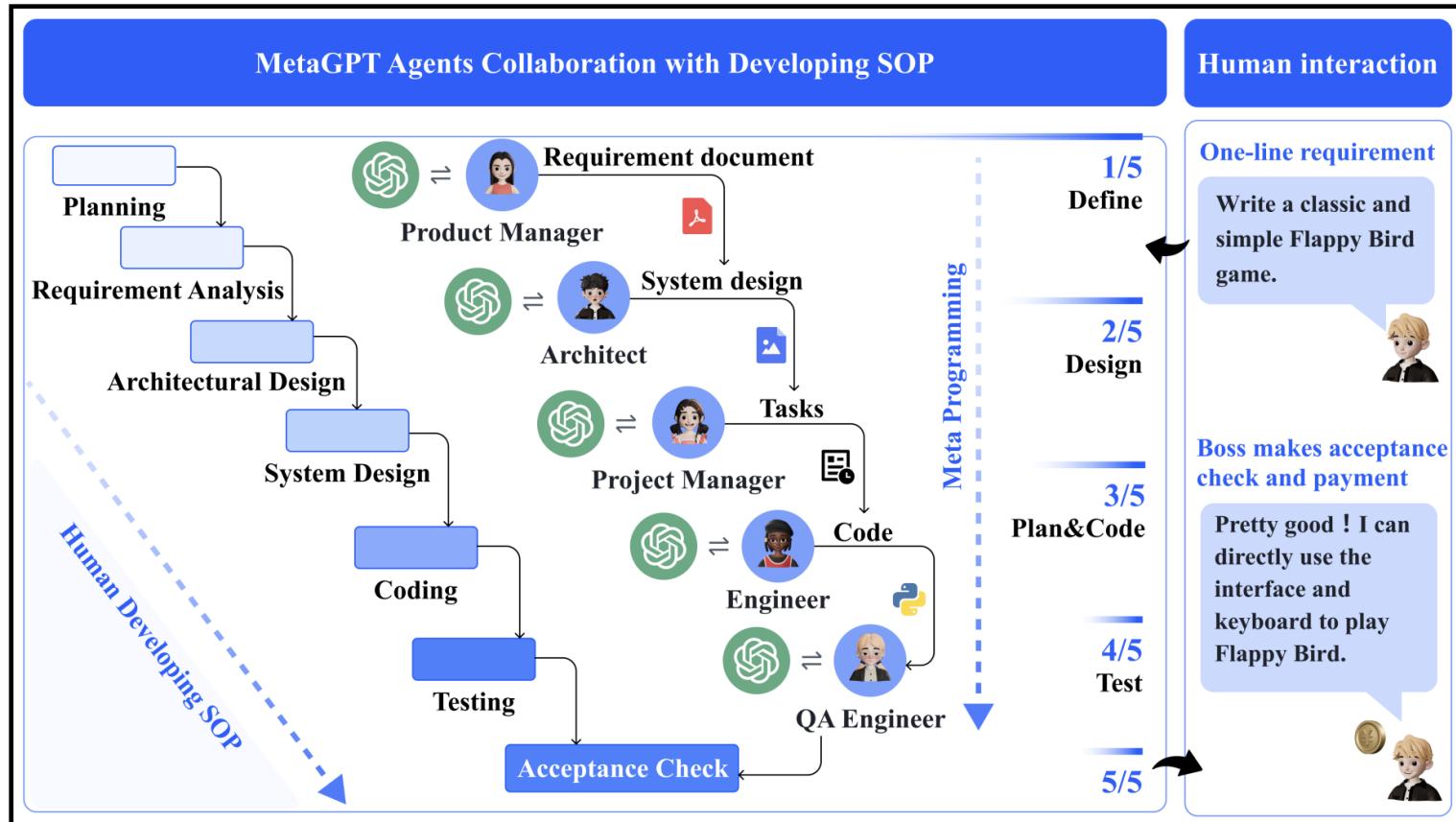
通用性



自主性



满意度



# 典型系统6：Voyager

- 可以在Minecraft中进行主动学习技能树。
- 专用系统：依赖于特定环境API抽象，理论上可泛化性到可用代码描述场景。



通用性



自主性



满意度



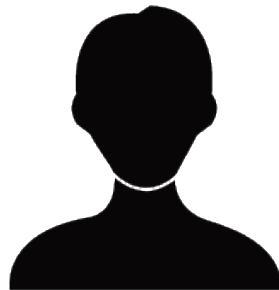
# 典型系统小结

- 当前智能体仍处于发展的初级阶段，距离开放域理想系统仍然具有很大差距。

典型系统	通用性	自主性	满意度
OpenAI GPTs	✓	?	✓
AI化学家	?	?	✓
斯坦福小镇	?	✓	?
AutoGPT/MetaGPT	✓	✓	?
Voyager	?	✓	✓
理想系统	👍	👍	👍

# 开放域任务的基本构成

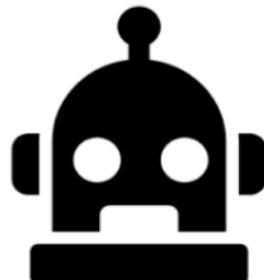
- 人类、环境、智能体是复杂开放域任务的基本构成角色。



人类



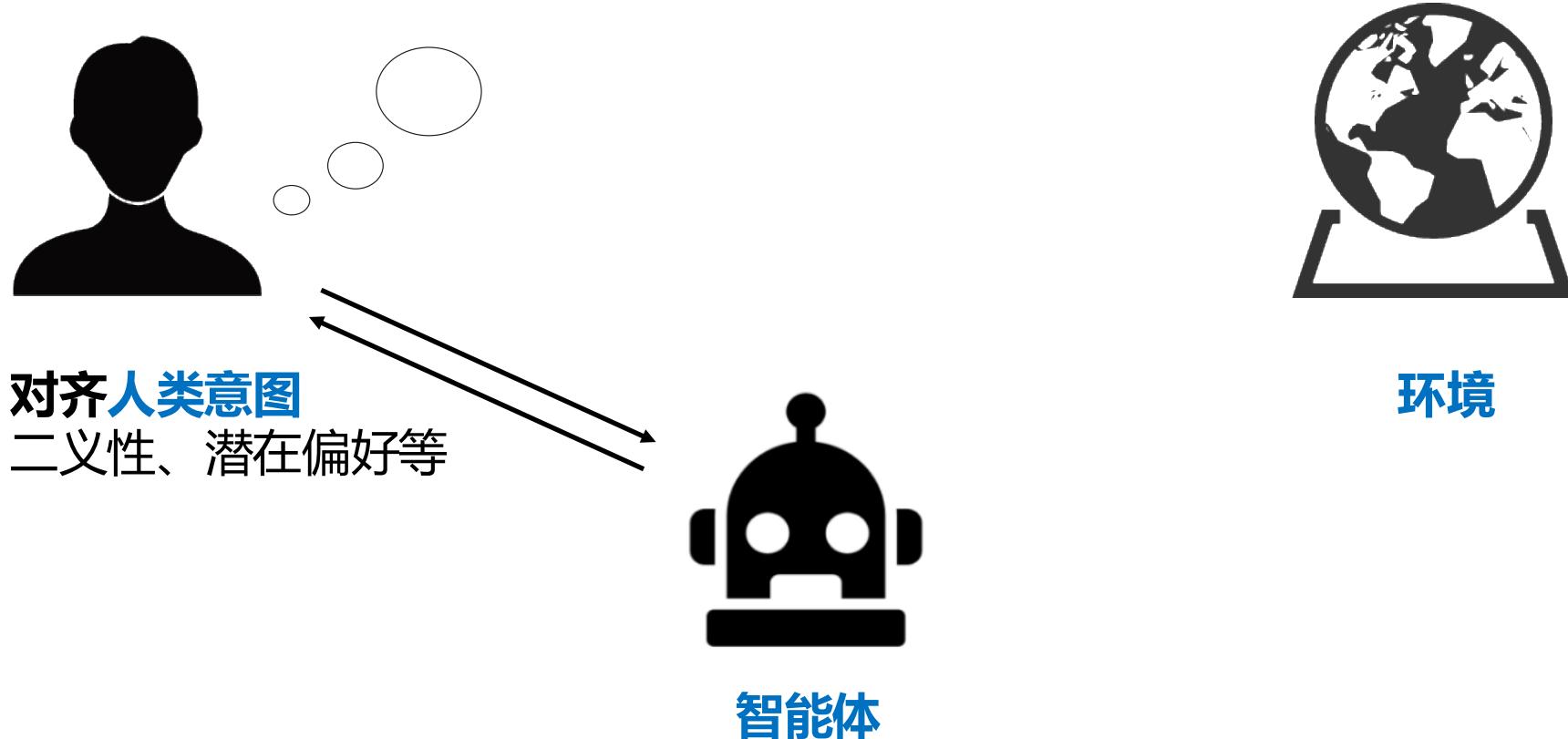
环境



智能体

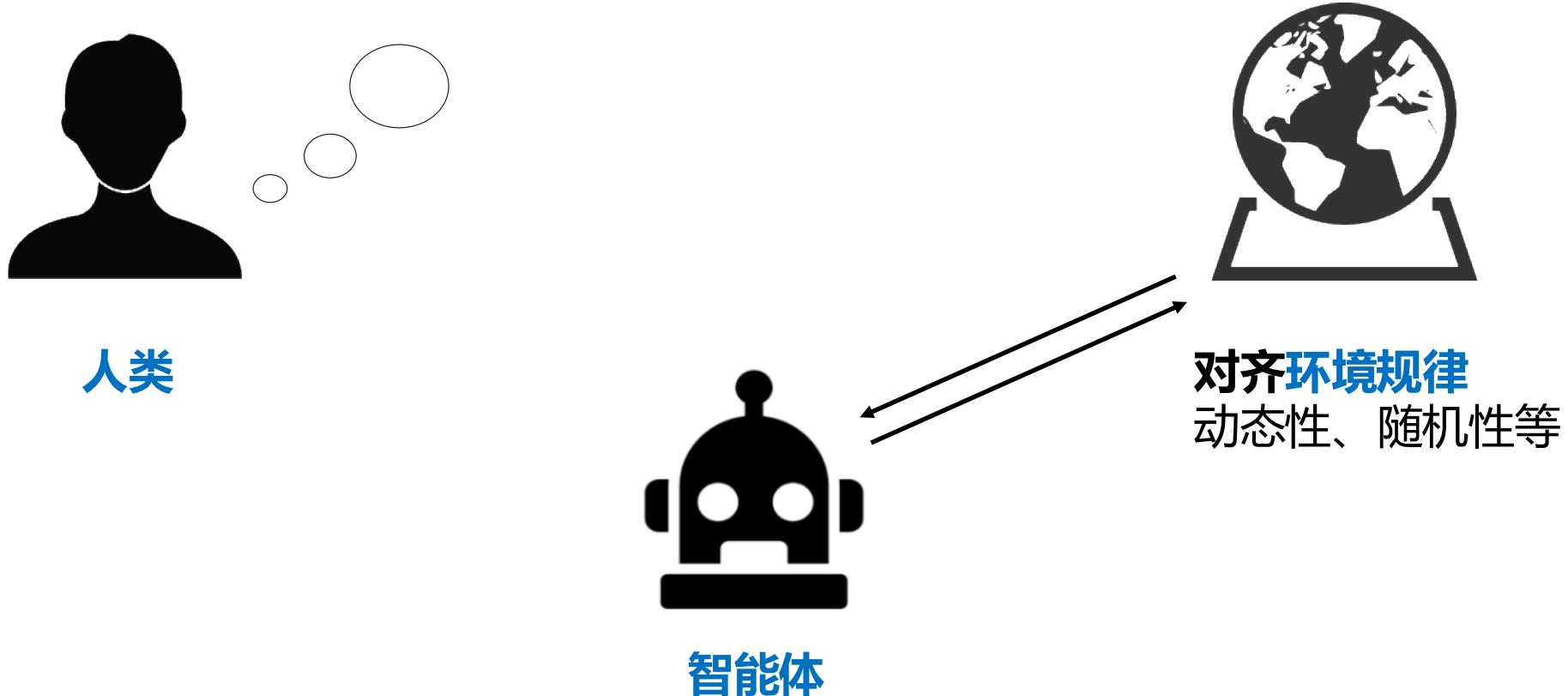
# | 智能体走向开放域的准则：统一对齐

- 通过对齐**人类意图**，智能体向着更理解人类的方向进化。



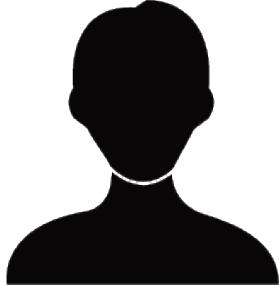
# | 智能体走向开放域的准则：统一对齐

- 通过对齐**环境规律**，智能体向着更适应环境的方向进化。



# | 智能体走向开放域的准则：统一对齐

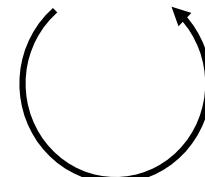
- 通过对齐**自身限制**，智能体向着更高效运行的方向进化。



人类



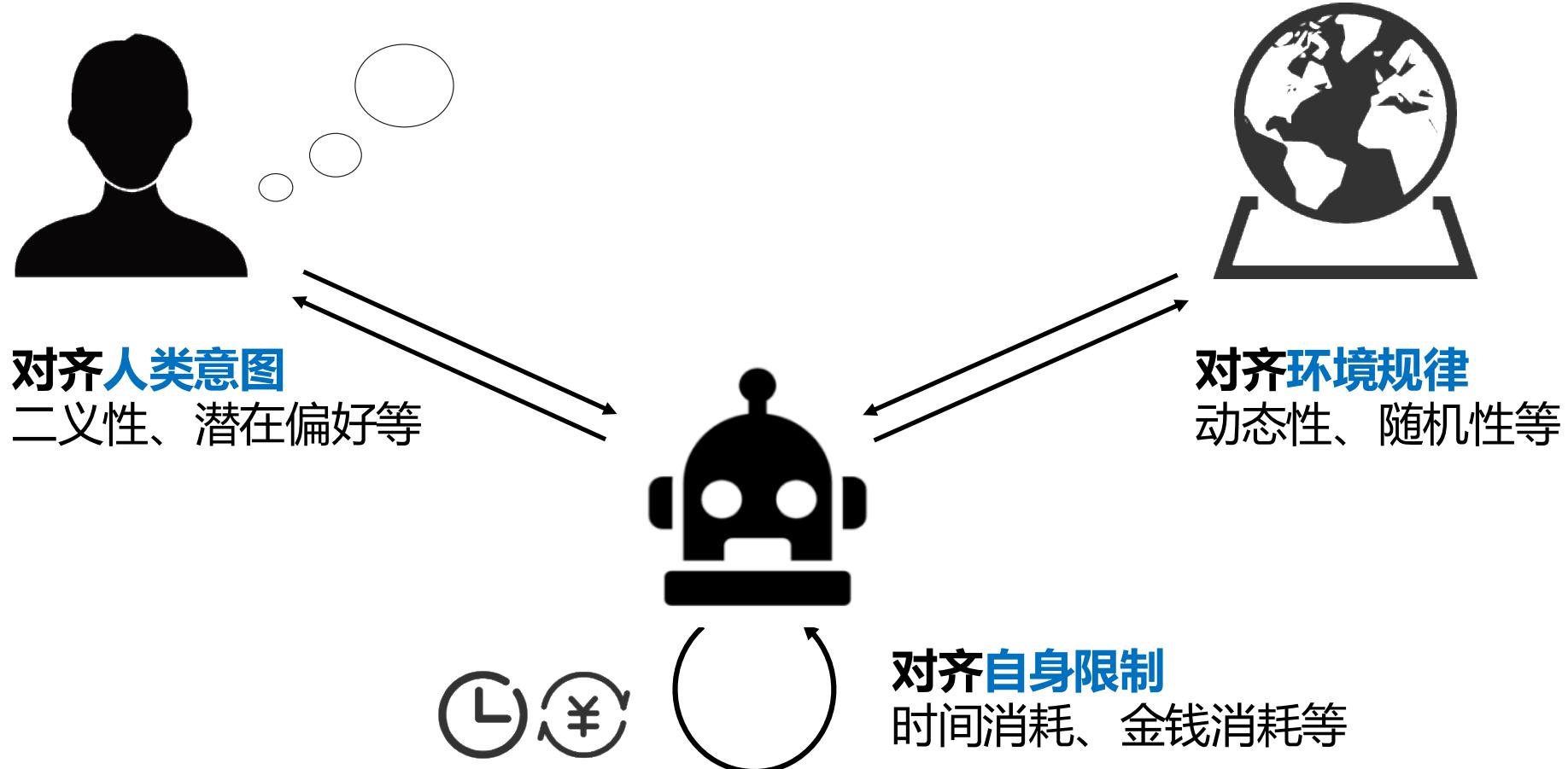
环境



对齐**自身限制**  
时间消耗、金钱消耗等

# | 智能体-人类-环境统一对齐

- 三种角色共同构成了智能体走向开放域的统一对齐准则。



# 方向1：建设更真实的模拟环境

- 环境在智能体研究中具有举足轻重的作用，应给予更多关注。

Type	Benchmarks	Human Intentions	Environmental Dynamics	Self-Constraints
Digital	Androidenv (Toyama et al., 2021)	None	Partial Obs.	None
	WebShop (Yao et al., 2022a)	None	Full Obs. <sup>†</sup>	None
	Mind2Web (Deng et al., 2023)	None	Partial Obs.	None
	ToolBench (Qin et al., 2023)	None	Full Obs. & Temp. & Stoch.	None
	WebArena (Zhou et al., 2023b)	Fixed and Given	Partial Obs.	None
Embodied	VirtualHome (Puig et al., 2018)	None	Partial Obs.	None
	BabyAI (Chevalier-Boisvert et al., 2019)	None	Partial Obs.	None
	ALFWorld (Shridhar et al., 2020)	None	Partial Obs.	None
	MineDojo (Fan et al., 2022)	None	Partial Obs. & Stoch.	None
	ScienceWorld (Wang et al., 2022a)	None	Partial Obs.	None
	Interactive Gibson (Xia et al., 2020)	None	Partial Obs.	#Actions
	AGENT (Shu et al., 2021)	None	Partial Obs.	#Actions
	RFUniverse (Fu et al., 2022)	Fixed and Given	Partial Obs.	#Actions
	BEHAVIOR-1K (Li et al., 2023b)	None	Full Obs.	#Actions
Mixed	HAZARD (Zhou et al., 2024)	None	Partial Obs. & Temp.	#Actions
	MINT (Wang et al., 2023b)	None	Partial Obs.	#Actions
	SmartPlay (Wu et al., 2023)	None	Partial Obs. & Stoch.	None
	AgentBench (Liu et al., 2023c)	None	Partial Obs.	None
	AgentBoard (Ma et al., 2024)	None	Partial Obs. & Temp. & Stoch.	None

# 方向1：建设更真实的模拟环境

- 在现有WebShop在线购物基准平台上，引入用户画像和需求组、搜索引擎的重排序机制、以及智能体运行时的消耗统计。



重排序使搜索结果高度依赖过往动作，环境规律更加贴近现实

# 方向1：建设更真实的模拟环境

- 现有智能体方法框架在**人类意图、环境规律、自身限制**三方面的综合表现欠佳，我们提出的框架取得最好的整体表现。

Method	Reward ↑	SR (%) ↑	G <sub>HI</sub> (%) ↓	G <sub>ED</sub> (%) ↓	Time (s) ↓	Money (\$) ↓
ReAct	50.3	8.0	11.7	14.9	<b>1.716</b>	<b>0.013</b>
ReAct-SC	49.9	7.4	14.4	14.6	1.720	0.039
Reflexion	44.4	<b>13.8</b>	22.5	25.7	5.539	0.045
LATS*	<b>52.4</b>	10.0	18.5	<b>14.3</b>	125.935	5.508
Ours	51.9	9.6	<b>6.7</b>	14.8	1.779	0.014

[ 平台任务评价指标 ]      [ 人类意图对齐表现 ]      [ 环境规律对齐表现 ]      [ 自身限制对齐表现 ]

# 方向2：建设更完整的评价体系

- 评价体系对于学科发展至关重要，智能体系统需要更合理全面的评价体系。

## 考虑更多评价维度

Method	Budget Constraint					
	+∞			20		
	PR↑	PBC↑	AC↓	PR↑	PBC↑	AC↓
ReAct	44.0	44.0	15.4	44.0	34.1	15.4
+BTP (Ours)	46.3	46.3	9.0	43.7	43.7	6.9
DFSDT	63.8	63.8	78.3	63.8	28.8	78.3
+BTP (Ours)	66.1	66.1	12.5	64.5	64.5	9.2
ToT-DFS	61.6	61.6	51.4	61.6	10.2	51.4
+BTP (Ours)	65.0	65.0	15.8	64.1	64.1	10.8

Table 1: Comparison of **Pass Rate (PR)**, **Pass rate under Budget Constraint (PBC)** and **Average Cost (AC)** on the ToolBench (Qin et al., 2023b) dataset. Our proposed **Budget-Constrained Tool Learning with Planning (BTP)** reduces the cost of tool learning and reaches competitive Pass Rate, significantly improving the performance under a strict budget constraint.

## 可复现性面临挑战

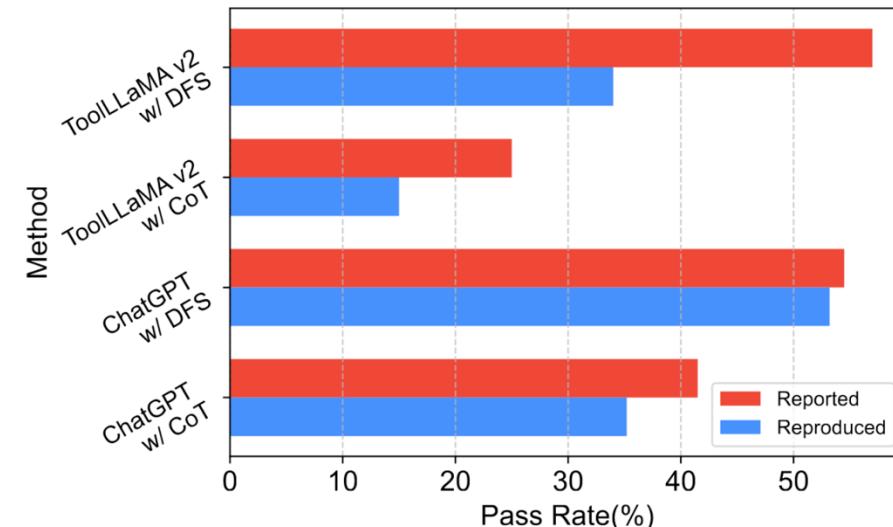


Figure 1: Comparison of performance (Pass Rate) reported in the paper and reproduced by us on the I1-Instruction group of ToolBench.

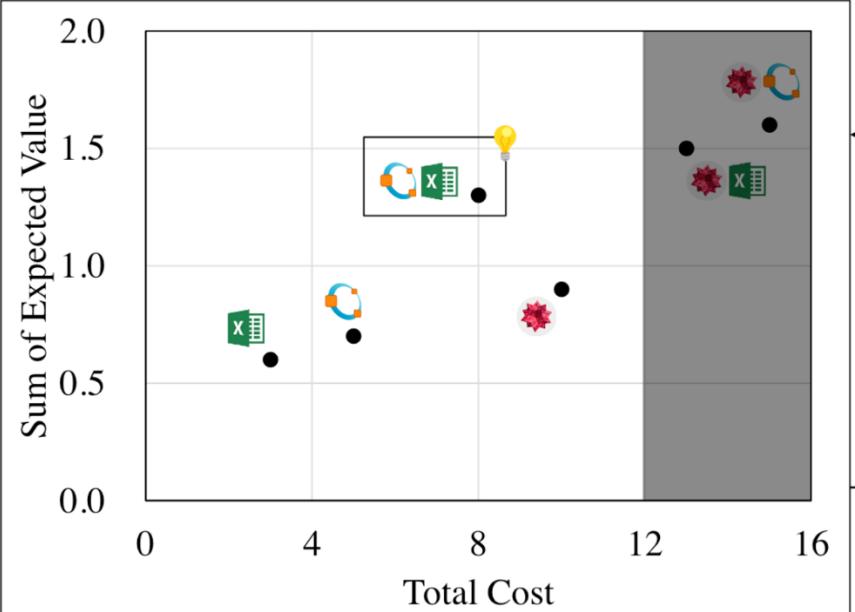
# 方向2：建设更完整的评价体系

- 将调用代价作为额外评价维度引入工具学习。

## User Query

I need to calculate the eigenvalues and eigenvectors of a  $100 \times 100$  matrix, and present the solution in an MS Excel workbook. The matrix is shown in the text file below:  
[Matrix.txt]

## Planning



## Tool Information

Tool Name	Mathematica	GNU Octave	MS Excel
Cost	10	5	3
Expected Value	0.9	0.7	0.6

## Tool Learning Process

	Plan	Tool Usage	Cost	Success?
Baseline	N/A	Mathematica, MS Excel	13	✗
BTP (Ours)	GNU Octave × 1, MS Excel × 1	GNU Octave, MS Excel	8	✓

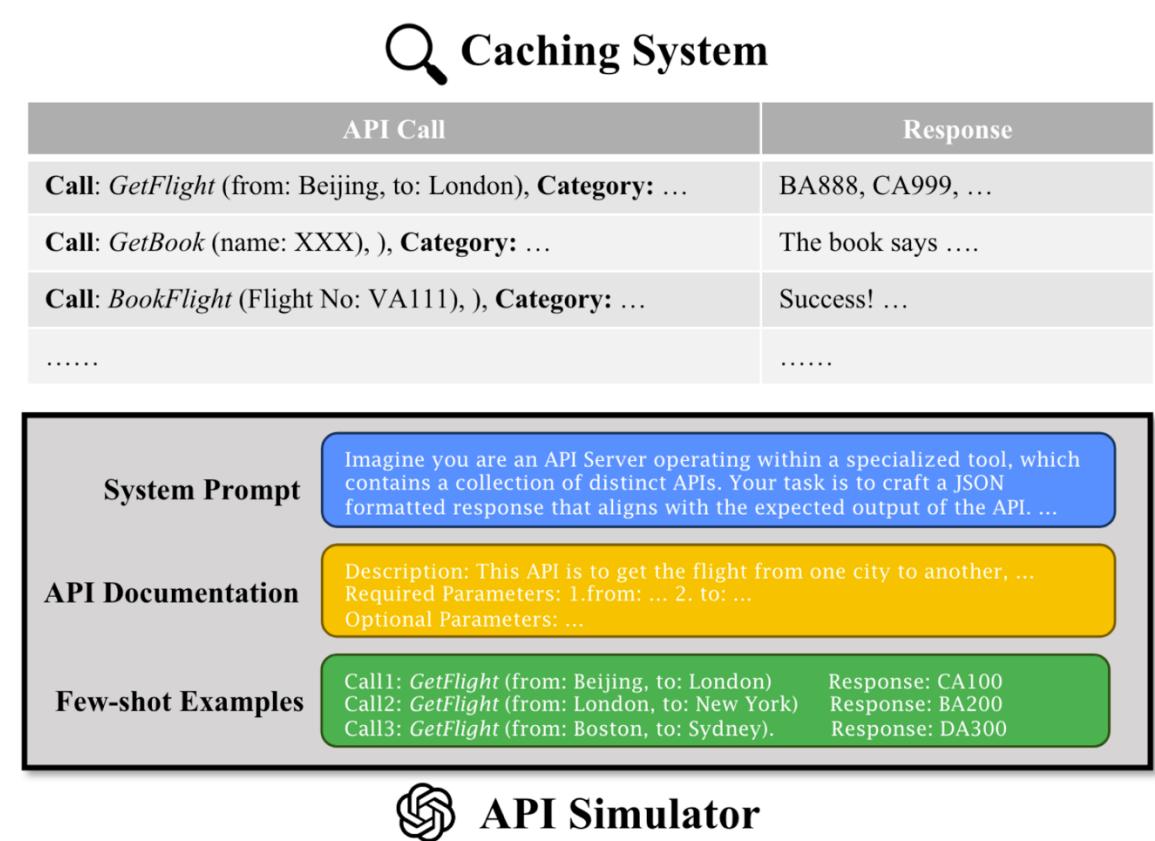
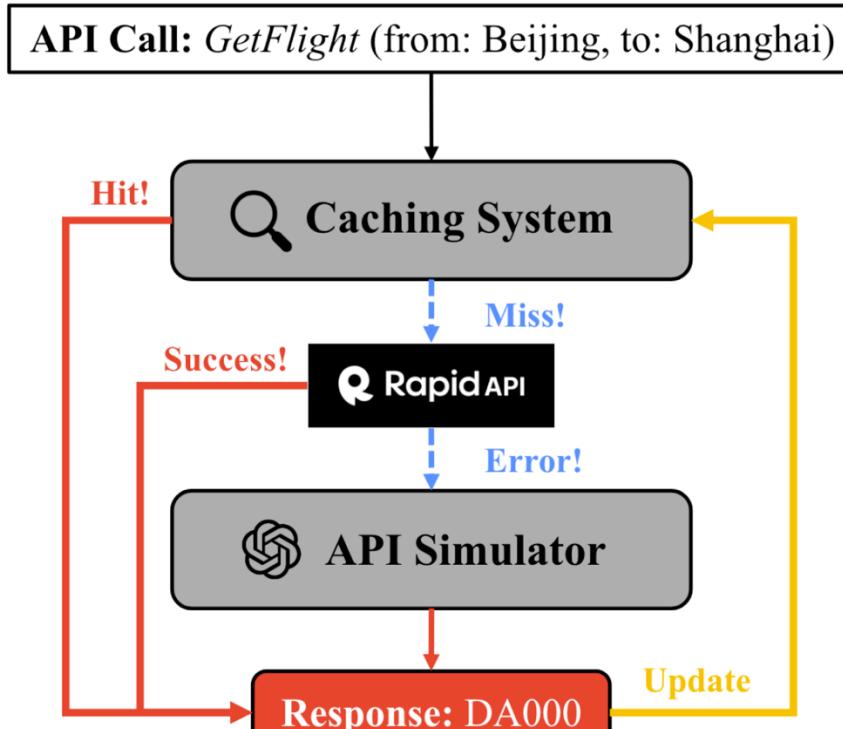
# 方向2：建设更完整的评价体系

- 将调用代价作为额外评价维度引入工具学习。

Method	I1-Inst		I1-Tool		I1-Cat		I2-Inst		I2-Cat		I3-Inst		Average	
	PBC↑	AC↓	PBC↑	AC↓	PBC↑	AC↓								
ReAct	38.5	15.8	37.0	16.3	38.5	15.0	33.5	15.8	35.0	17.0	22.0	12.6	34.1	15.4
+BTP ( <i>Ours</i> )	<b>43.5</b>	<b>7.6</b>	<b>48.0</b>	<b>7.8</b>	<b>49.5</b>	<b>7.2</b>	<b>41.5</b>	<b>7.0</b>	<b>49.5</b>	<b>7.1</b>	<b>30.0</b>	<b>4.7</b>	<b>43.7</b>	<b>6.9</b>
ReAct+Prompt	35.5	16.7	39.5	17.1	43.0	14.4	37.5	14.9	39.5	16.3	24.0	14.9	36.5	15.7
+BTP ( <i>Ours</i> )	<b>47.0</b>	<b>8.7</b>	<b>44.0</b>	<b>8.3</b>	<b>50.5</b>	<b>7.8</b>	<b>47.0</b>	<b>7.7</b>	<b>54.0</b>	<b>8.0</b>	<b>32.0</b>	<b>5.2</b>	<b>45.8</b>	<b>7.6</b>
DFSDT	34.0	77.5	29.5	75.1	34.5	70.1	28.0	81.6	28.5	79.4	18.0	85.8	28.8	78.3
+BTP ( <i>Ours</i> )	<b>58.0</b>	<b>9.6</b>	<b>58.0</b>	<b>10.3</b>	<b>63.5</b>	<b>9.5</b>	<b>77.0</b>	<b>9.5</b>	<b>66.5</b>	<b>9.8</b>	<b>64.0</b>	<b>6.4</b>	<b>64.5</b>	<b>9.2</b>
DFSDT+Prompt	31.0	69.5	29.5	84.8	36.5	61.2	28.5	68.4	31.0	74.4	17.0	81.3	29.0	73.3
+BTP ( <i>Ours</i> )	<b>55.0</b>	<b>10.0</b>	<b>58.0</b>	<b>10.1</b>	<b>63.0</b>	<b>9.2</b>	<b>77.5</b>	<b>9.5</b>	<b>65.5</b>	<b>10.0</b>	<b>64.0</b>	<b>6.1</b>	<b>63.8</b>	<b>9.2</b>
ToT-DFS	11.5	52.4	10.5	52.4	11.0	48.8	7.0	52.4	11.0	53.3	10.0	48.9	10.2	51.4
+BTP ( <i>Ours</i> )	<b>55.5</b>	<b>11.7</b>	<b>59.0</b>	<b>12.0</b>	<b>60.0</b>	<b>11.6</b>	<b>76.0</b>	<b>11.3</b>	<b>70.0</b>	<b>11.6</b>	<b>64.0</b>	<b>6.7</b>	<b>64.1</b>	<b>10.8</b>
ToT-DFS+Prompt	10.0	48.2	10.5	50.8	14.5	45.9	6.0	49.0	10.0	51.2	8.0	53.3	9.8	49.7
+BTP ( <i>Ours</i> )	<b>54.5</b>	<b>11.9</b>	<b>59.5</b>	<b>12.1</b>	<b>56.5</b>	<b>11.5</b>	<b>74.5</b>	<b>11.8</b>	<b>64.0</b>	<b>11.7</b>	<b>64.0</b>	<b>6.7</b>	<b>62.2</b>	<b>11.0</b>

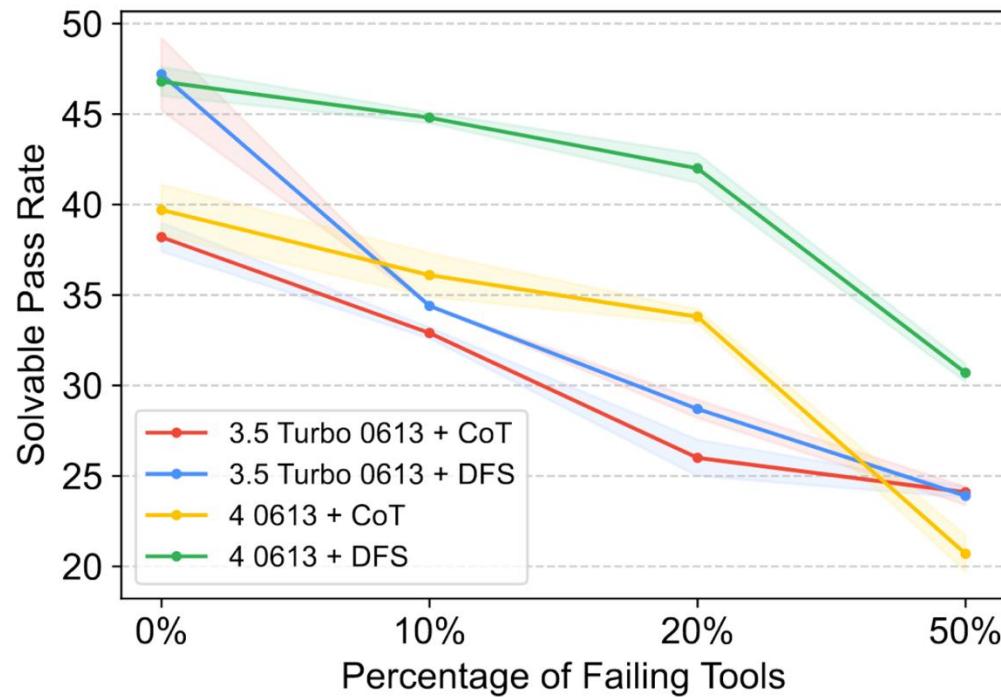
# 方向2：建设更完整的评价体系

- **StableToolBench**: 解决大规模工具调用评测的稳定性问题。

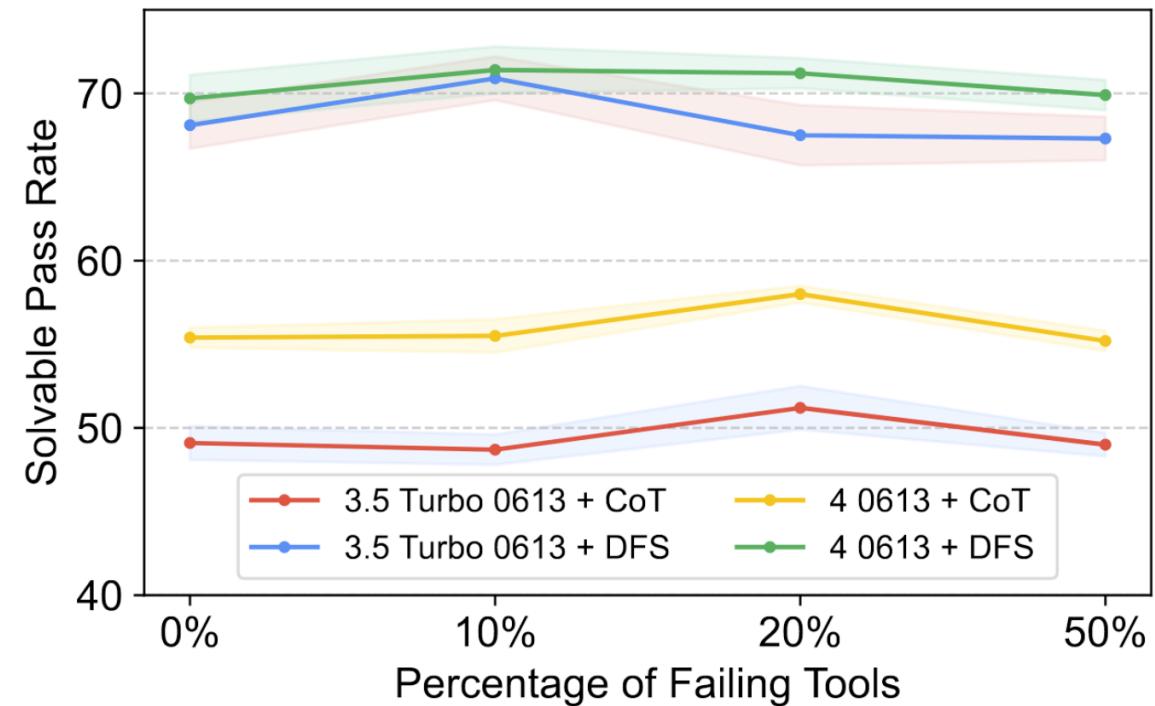


# 方向2：建设更完整的评价体系

- **StableToolBench**: 解决大规模工具调用评测的稳定性问题。



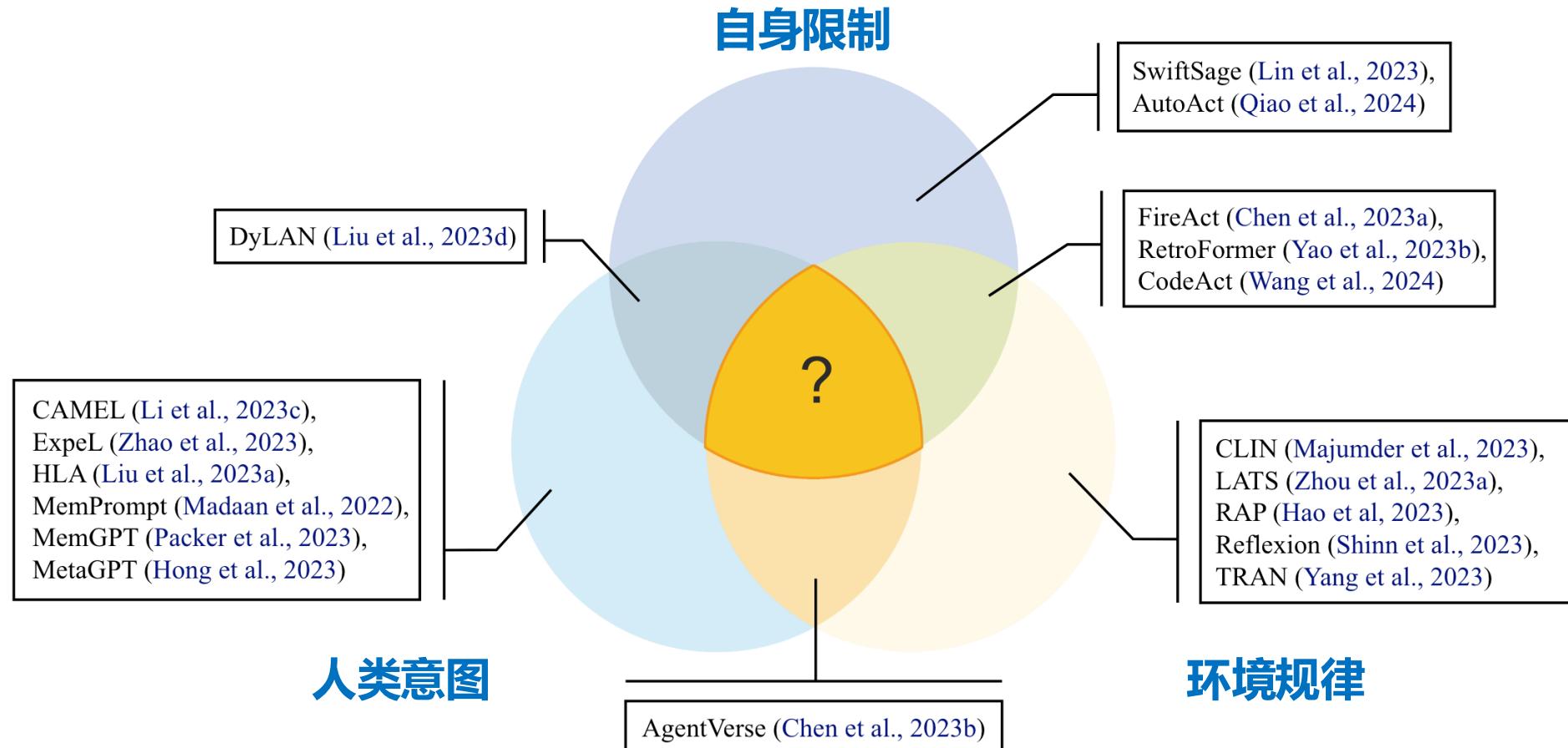
前人系统



我们系统

# 方向3：发展三方统一对齐方法

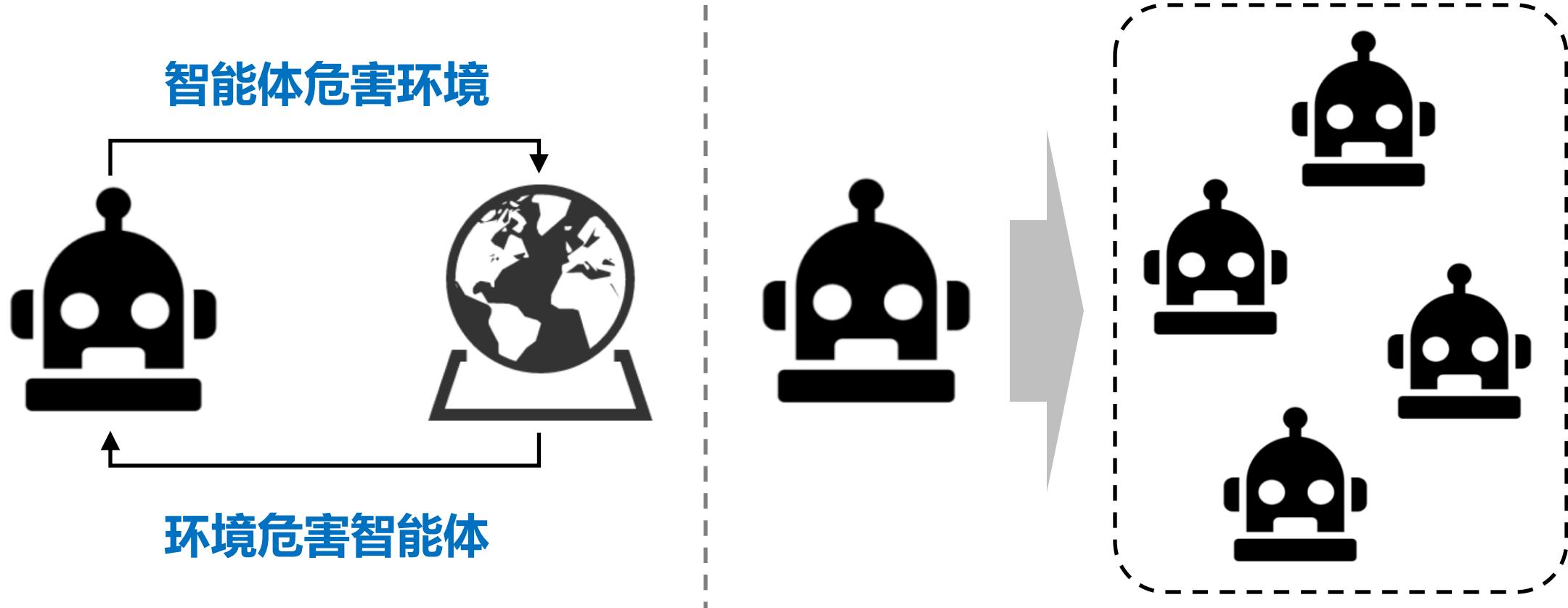
- 现有方法仅能有效实现一方或两方对齐，无法有效完成三方统一对齐。



ReAct (Yao et al., 2023b), ToT (Yao et al., 2023a), BOLAA (Liu et al., 2023), Chameleon (Lu et al., 2023), HuggingGPT (Shen et al., 2023), ...

## 方向4：建设更强安全对齐方法

- 《北京共识》为人工智能系统设立的风险红线：自主复制或改进、权力寻求、协助武器制造、网络安全、欺骗。



# 方向5：智能体系统的真实场景落地

- 真实场景落地要求智能体研究从任务完成走向三方对齐，并找准杀手级应用。

## 真实场景落地

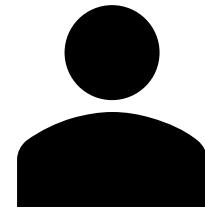
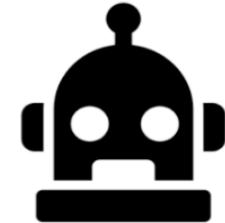
建设更真实的模拟环境

建设更完整的评价体系

发展三方统一对齐方法

建设更强安全对齐方法

技术上限



需求下限

**关键标志：** 用户需要完成某项任务时会自然想到使用智能体，并且用户觉着自己很牛

# 总结

- 智能体是大模型解决更复杂问题的重要范式抽象，亟需全面走向开放域。
- 智能体-人类-环境统一对齐是实现理想开放域智能体系统的重要原则。
- 智能体研究需关注如下问题：
  - 建设更真实的模拟环境；
  - 建设更完整的评价体系；
  - 发展智能体-人类-环境三方统一对齐方法；
  - 建设更强安全对齐方法；
  - 智能体系统的真实场景落地。



AIR

清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University

When we dare to dream,  
we achieve greatness.



扫一扫，关注我们