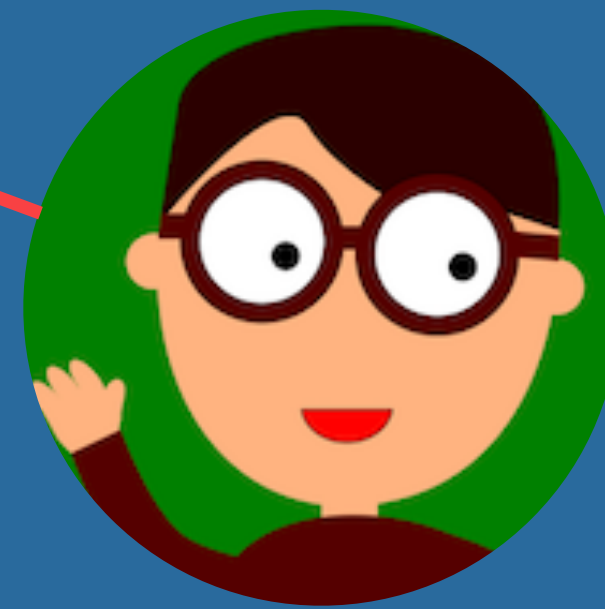


生成式 AI：文字與圖像生成的原理與實務

09.

## AI Agents：打造你專屬的超級代理人



蔡炎龍

政治大學應用數學系





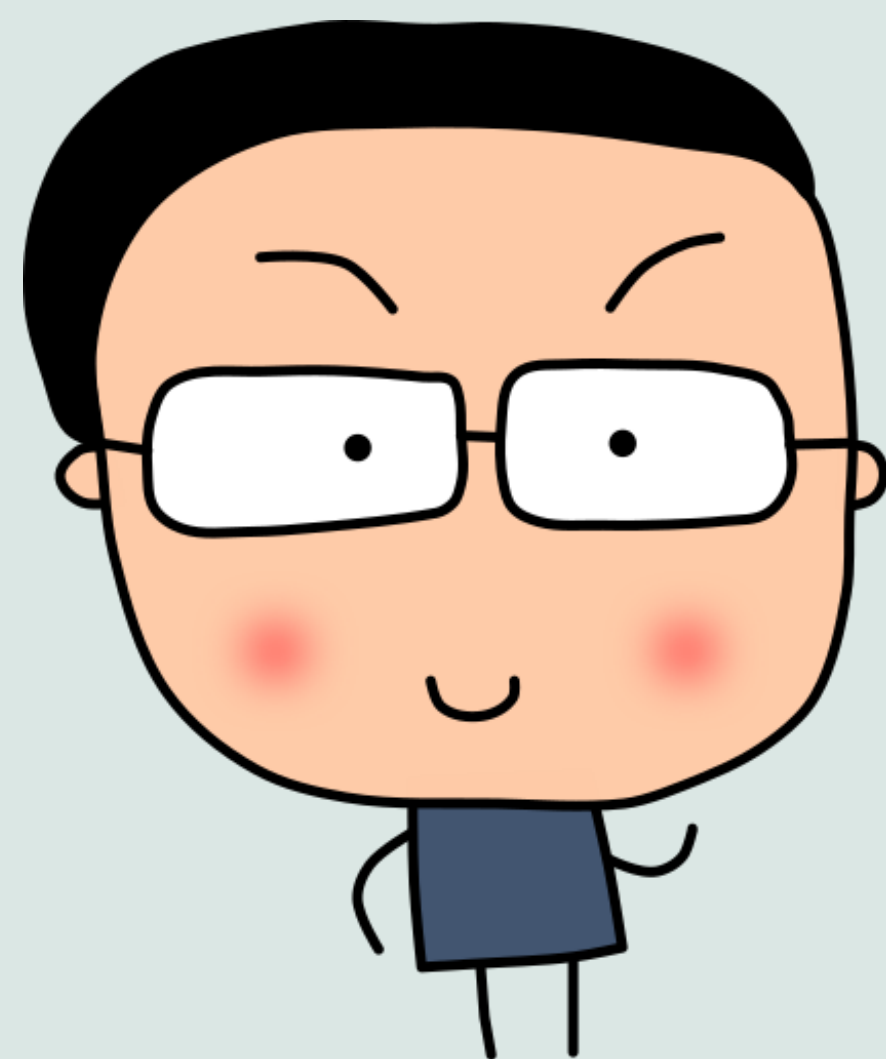


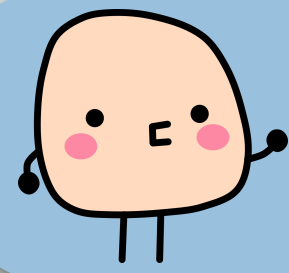
01.

AI Agents 是什麼呢?



很多人說 2025 是 AI 元年





# 非常多的討論

遠見雜誌

## AI Agent應用與優缺點真實揭露：哪些企業正導入？

當這些企業紛紛提出，AI Agent (AI代理) 以及衍伸的應用，就是下一波關鍵趨勢。顧能 (Gartner) 也



www.tca.org.tw

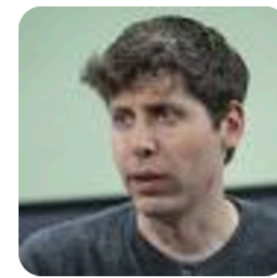
## [翱騰國際] AI代理人 (AI Agent) 重磅登場，創新金融業智慧客服體驗

AI代理人 (AI Agent) 具備多項核心功能，能協助企業處理及多樣化資料來源整合，協助企業

TechNews 科技新報

## 代替用戶自動執行任務，OpenAI 傳打造 AI 代理工具「Operator」

《彭博社》引述2名知情人士消息指出，OpenAI 準備推出一款代號「Operator」新型AI代理 (AI agent)，代表用戶採取行動，例如編寫程式碼或預訂出遊行程等...



ITmedia

## AIエージェント (AI Agent) とは？：AI・機械学習の用語辞典

用語「AIエージェント」について解説。特定の目標を達成するために自律的に作成し、計画的に各タスクを実行するAIシステムのこと。



CIO

## 세일즈포스 마크 베니오프 "AI 에이전트, 클라우드·소셜·모바일 수준의 혁신"

AI 에이전트가 비즈니스 환경에서 '혁명'이 될 것이라고 마크 베니오프가 평가했다.



數位時代

## AI Agent是什麼？邁向AI界聖杯，AI代理為何是關鍵一步？

AI Agent中文是AI代理，它被視為實現通用型人工智慧 (AGI) 的一小步，最大不同就是擁有完成任務的「彈性」、「感知」與「調整」。

브런치스토리

## 본격 AI가 떠먹여 주는 시대, AI Agent가 온다

시리아"라고 부르면 뭐든 다 해주는 시대, 우리는 지금 어디쯤 있을까? | 안녕하세요. 카카오벤처스 투자팀입니다. 투자팀은 늘 창업 초기 단계에 있는 스타트업들과...



電子工程專輯

## 「前瞻 AI Agents，顛覆未來想像」論壇-圓桌討論：探索AI Agent未來的應用新場景

當AI Agent 正在顛覆各行各業的運營方式，企業如何在這波新浪潮中創造新的商業模式？在「『Agentic』顛覆未來」焦點座談中，台杉投資總經理翁嘉盛與AI...



PR TIMES

## ジーニーのグループ会社 JAPAN AI株式会社『日本企業のための最も実用的なAIエージェントへ』自律型AIエージェント「JAPAN AI AGENT」を提供開始

株式会社ジーニーのプレスリリース (2024年11月14日 15時30分) ジーニーのグループ会社 JAPAN AI株式会社『日本企業のための最も実用的なAIエージェント...



DIGITIMES

## 前瞻 AI Agents 顛覆未來想像

本活動將由主辦單位進行出席資格審核，與主題及屬性符合者為優先考量。通過審核後，將以電子郵件方式寄發含有報到編號/QR Code的「報到通知...

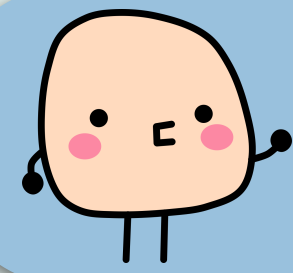


## 速覽Binance AI Agent報告：大器初顯，常若玩物

在曼谷Devcon 的火熱和街道的煙火霓虹中，AI Meme 們迎來了自己的高光時刻。從Binance 光速上線ACT，到GOAT 再破新高，一切的注意力可能都始於山羊背後...





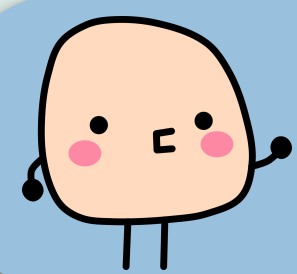


## 吳恩達力推 AI 最重要的方向



What's next for AI agentic workflows ft. Andrew Ng of AI Fund

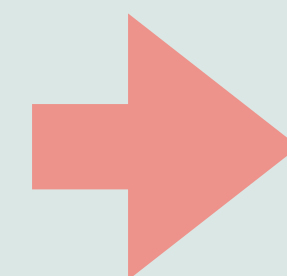
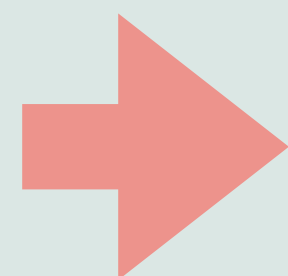




以前我們用 LLM, 是下了 prompt, 就要 LLM 回答我們



Prompt



outputs





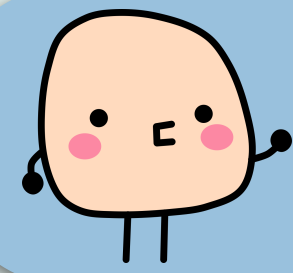
## 吳恩達老師說



這就好比要一個人，寫一篇文章，但不可以用刪除鍵、不可以修改，一路寫下去。

人很難做到這樣...





## 寫程式也是這樣

```
In [ ]: from langchain.embeddings import HuggingFaceEmbeddings

class CustomE5Embedding(HuggingFaceEmbeddings):
    def embed_documents(self, texts):
        texts = [f"passage: {t}" for t in texts]
        return super().embed_documents(texts)

    def embed_query(self, text):
        return super().embed_documents([text])
```

我們喜歡一邊寫、一邊  
試，一邊改。







## 我們用 LLM 也常常要做不少事

資訊

可能要查一下資訊

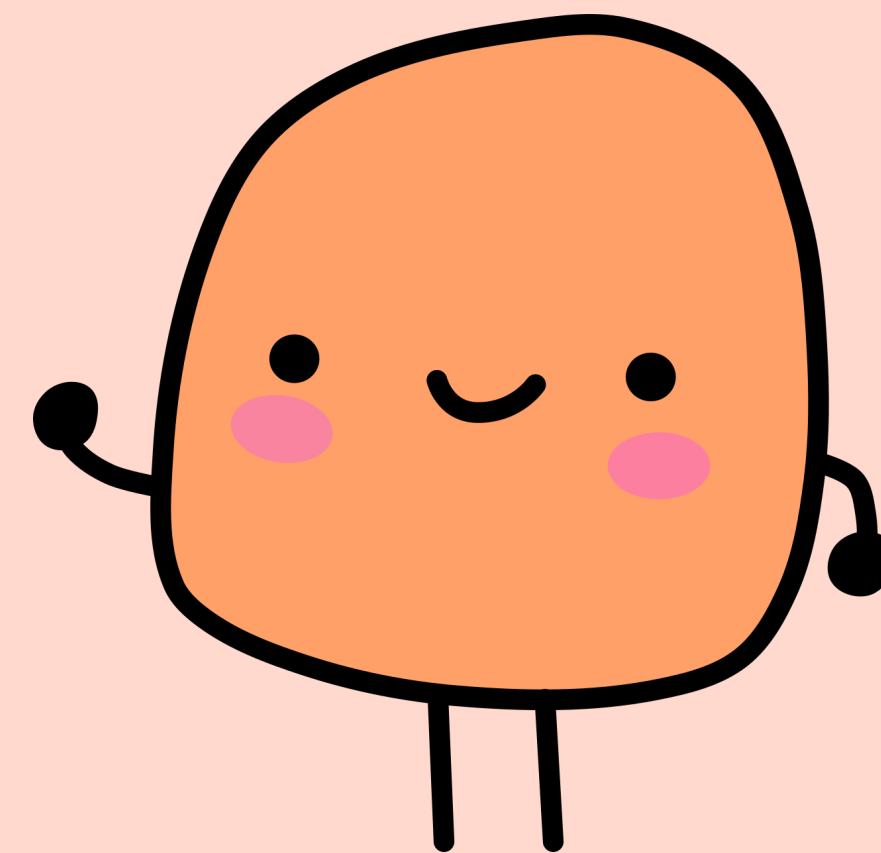
提供需要的正確資訊。

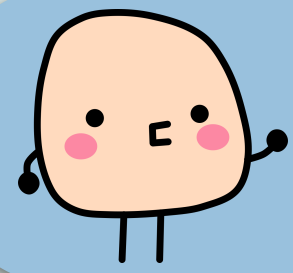
可能要按一下計算機

可能要上網搜尋

清楚的指引

例如, 以上面的資訊, 用什麼樣的格式、風格, 來回答使用者的問題。





也常常是來來回回要 LLM 修改

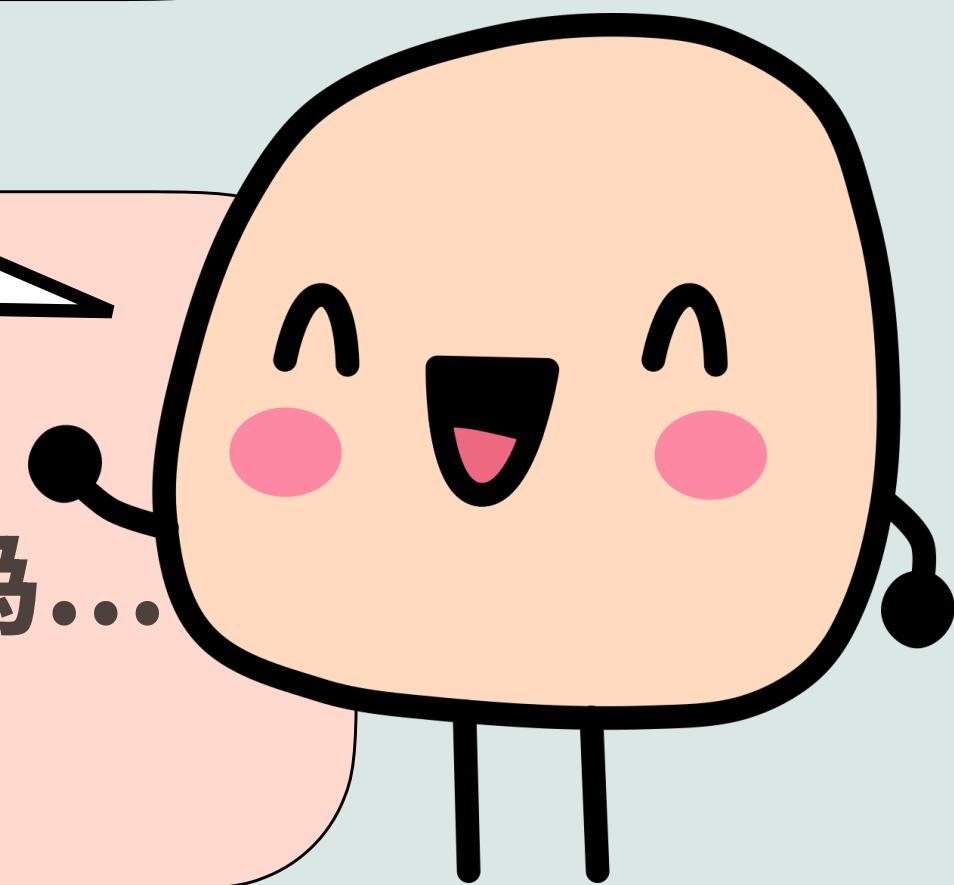
**Assistant**

我說了一堆, 可能有些地方不錯, 有些地方還可以修正...

**User**

某某地方, 實際上的情況是... 另外, 某某地方是否可以改為...

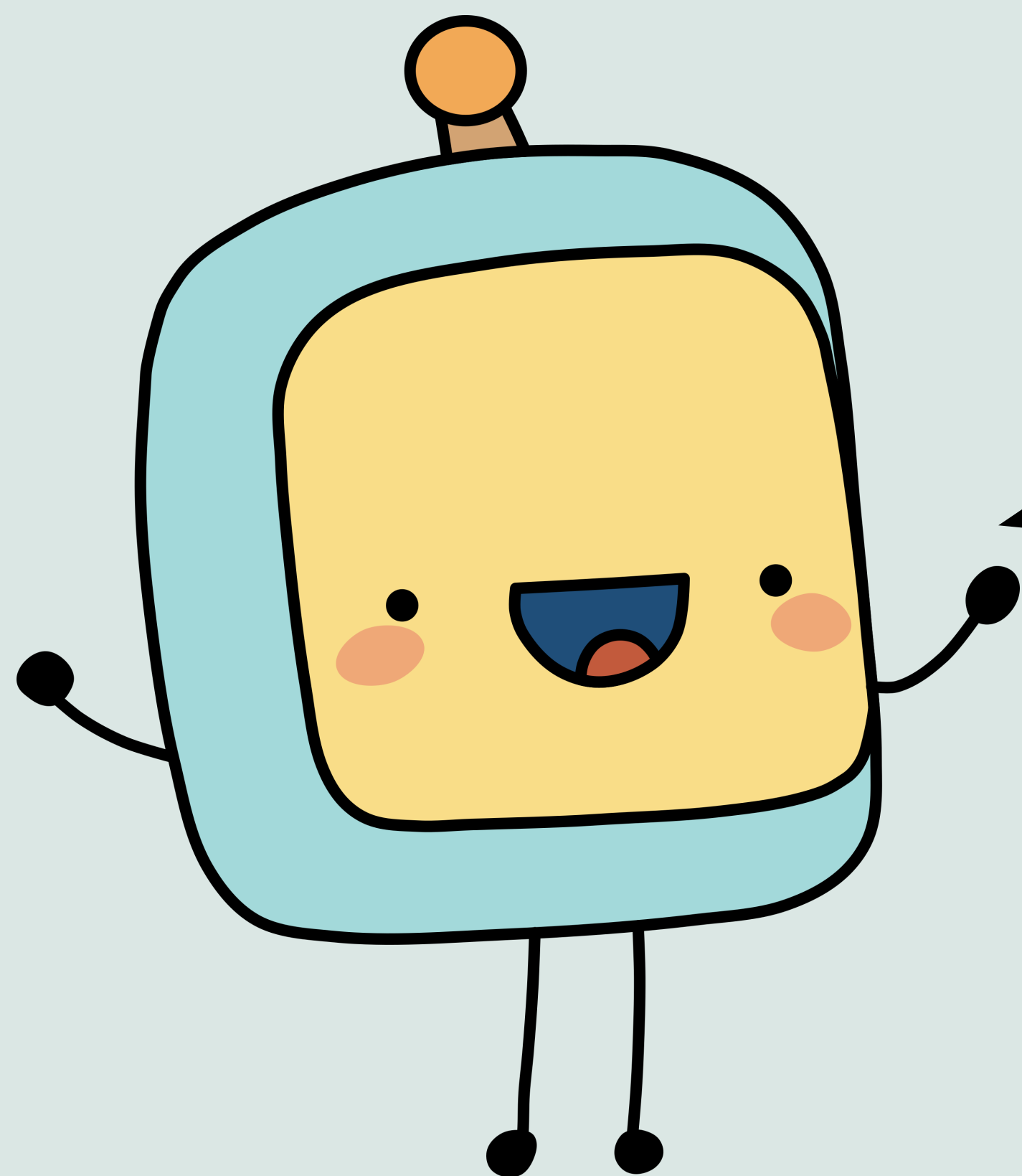
如果 AI 自動做這些事, 豈不美哉?







AI Agent: 本來你要做的, AI 自動幫你做完!



舉凡規畫、收集資料、  
不斷迭代嚐試等等, 都  
由 AI 來負責做!



## RAG 其實也是一種 AI Agent

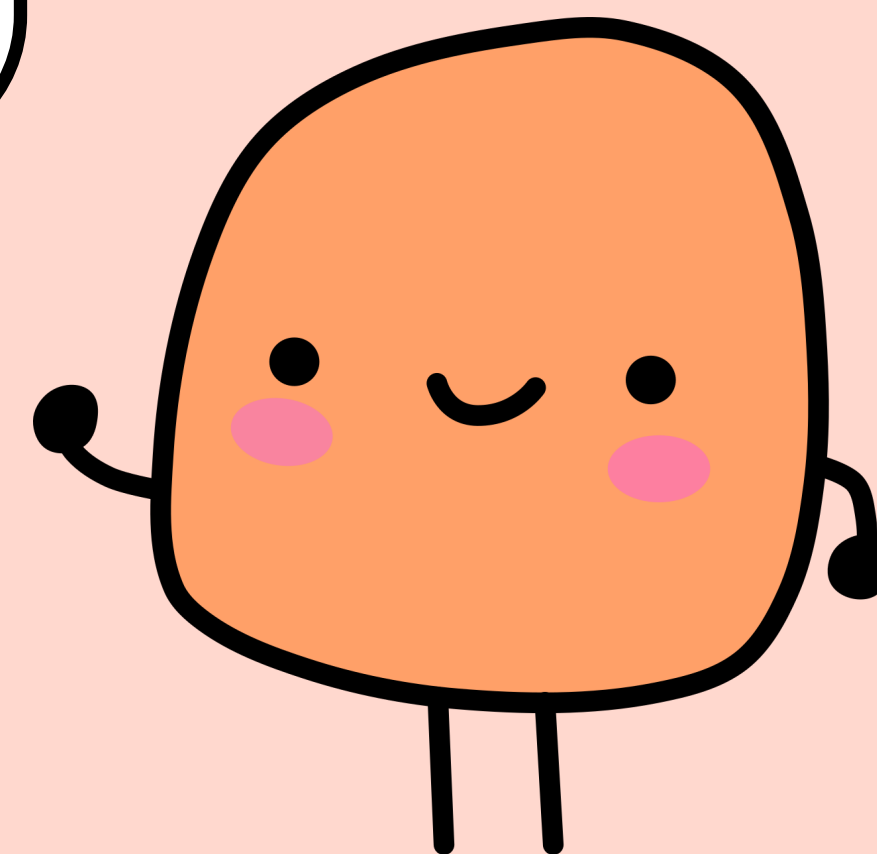
### 資訊

提供需要的正確資訊。

這裡由電腦**自動**從  
資料庫中尋找。

### 清楚的指引

例如, 以上面的資訊, 用什麼樣的格式、風格, 來回答使用者的問題。







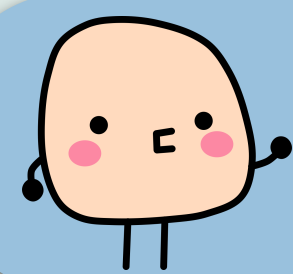
## 現在更多幫忙找足夠資訊的

我來用 RAG 找  
內部相關資訊。

我幫忙上網  
收集資料。

寫出來的東西讓  
我檢查一下。

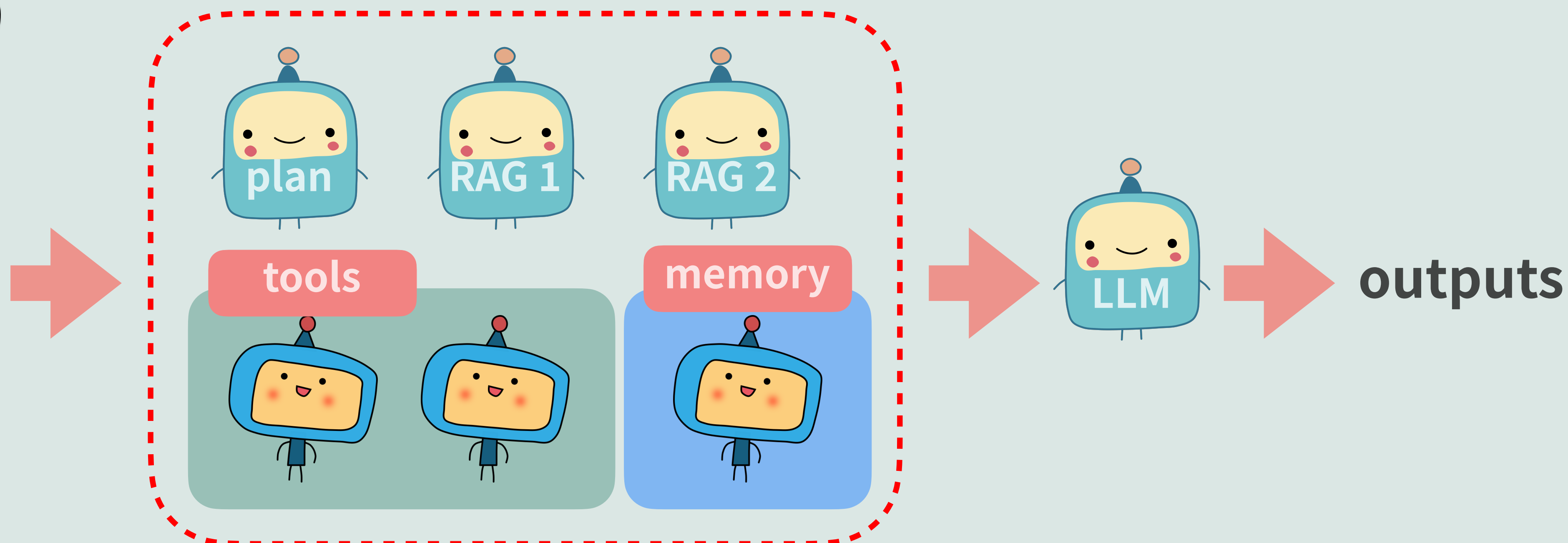
我按計算機檢  
查一下數字。



現在老闆 (你) 一句話, 員工 (AI Agents) 動起來!



Prompt

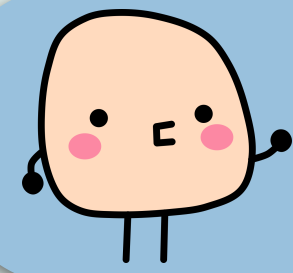






02.

# AI Agent 的設計模式



## 吳恩達 AI Agents 的四個設計模式 (Design Patterns)



Reflection

Tool Use

Planning

Multiagent Collaboration

<https://youtu.be/sa178ACtGTc>



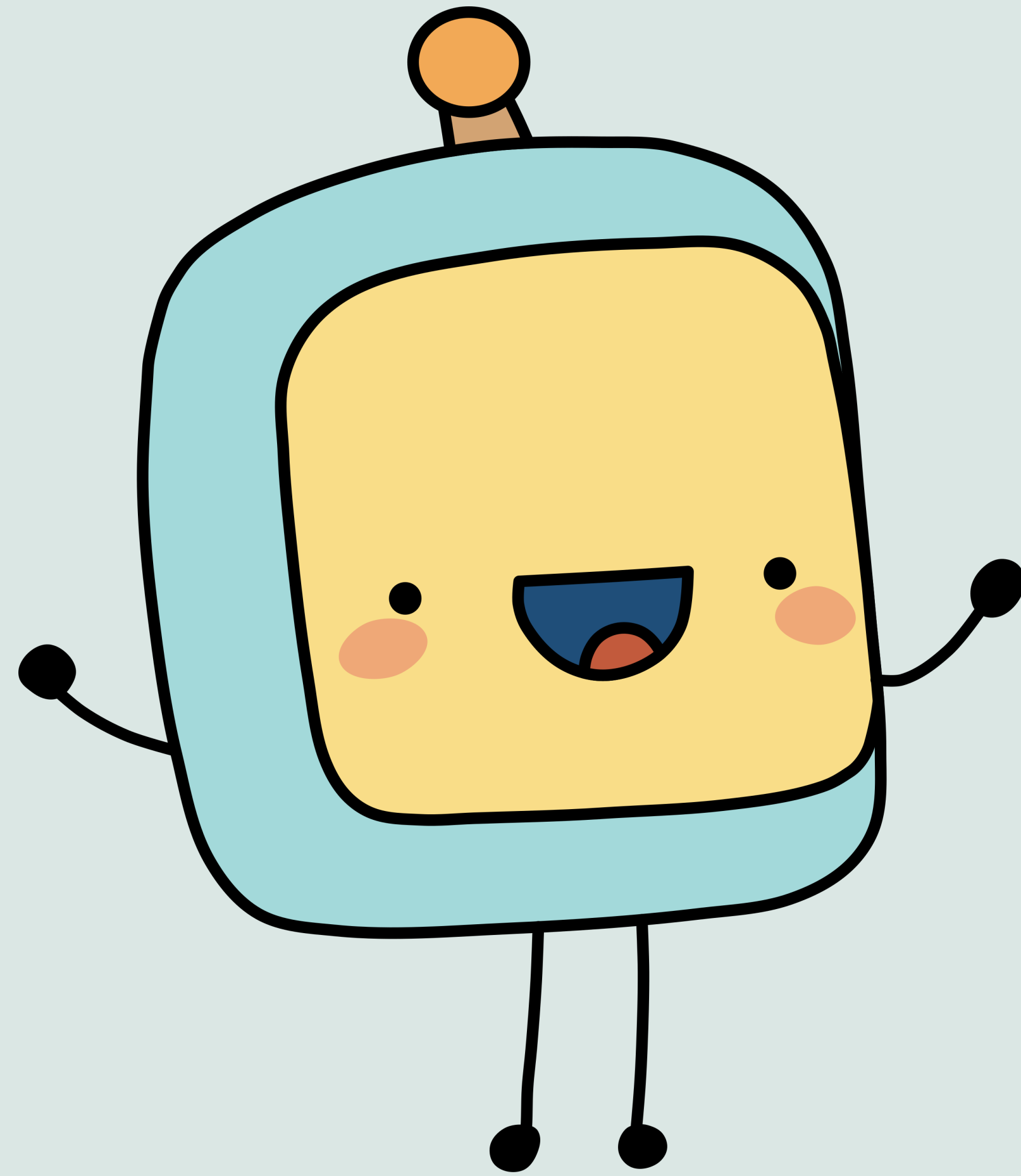


今天會用吳恩達老師的 AISuite 來實作

# AISuite

- \*統一的使用方式
- \*輕鬆切換模型
- \*同時使用不同供應商模型
- \*安裝容易

<https://yenlung.me/AI07>



# Reflection

## 反思

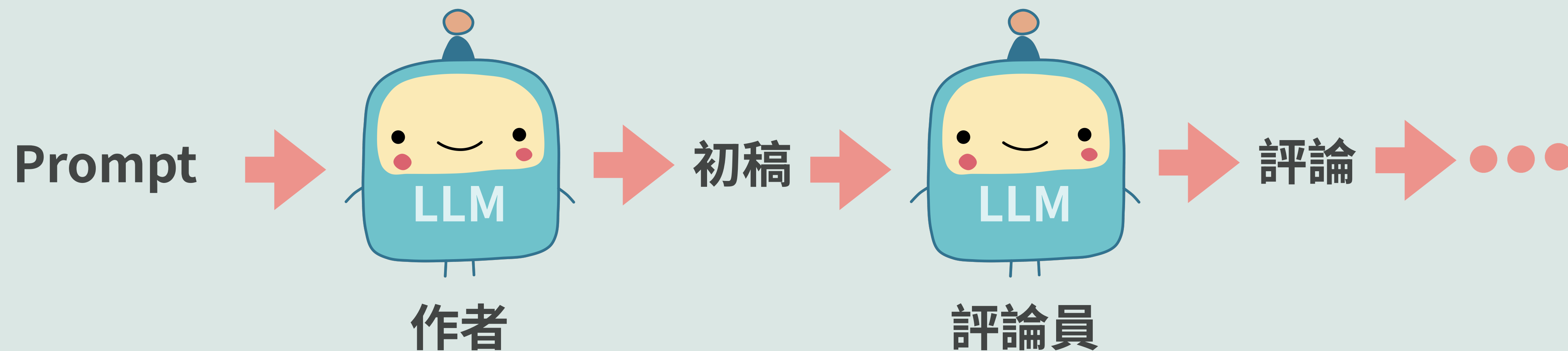
第一個 LLM 產生初稿, 第二個 LLM 進行評估, 提供改善建議, 然後第一個 LLM 進行修改。可重覆數次。

<https://yenlung.me/AI07a>



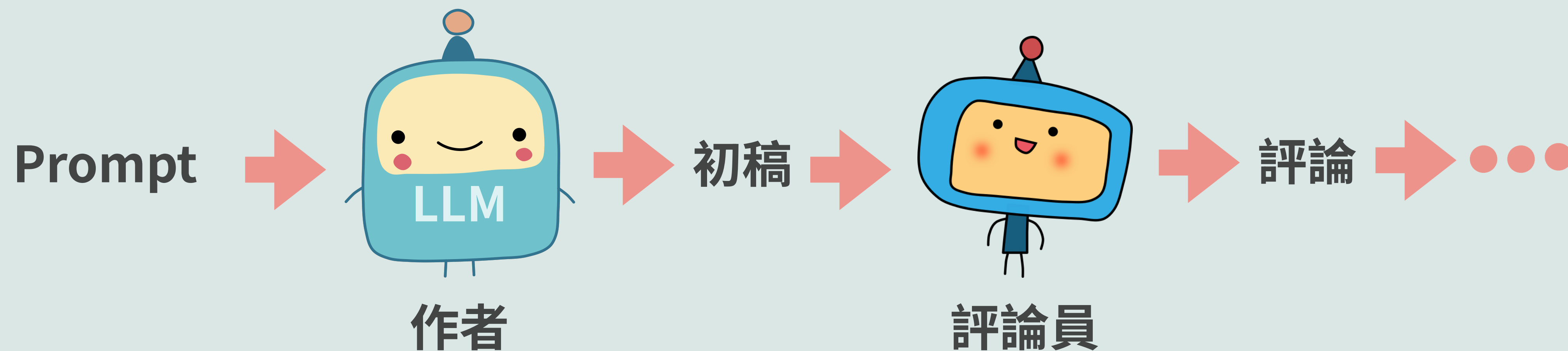


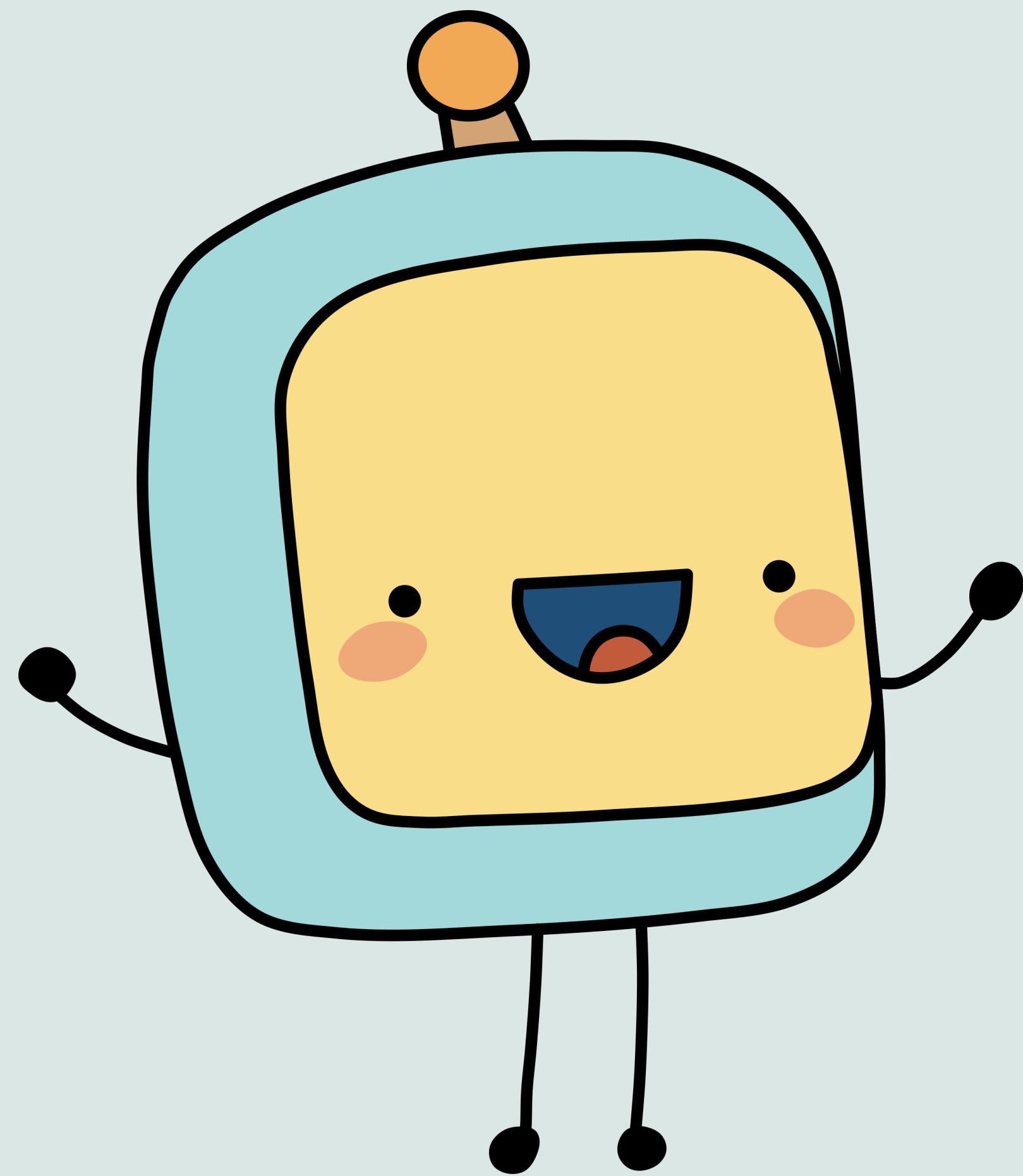
## 【Reflection】





## 【Reflection】當然「評論員」可用不同模型



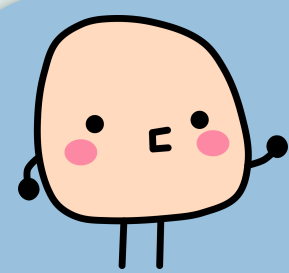


# Tool Use

## 使用工具

不是 LLM 擅長的, 或者不會的, 呼叫工具來幫忙。

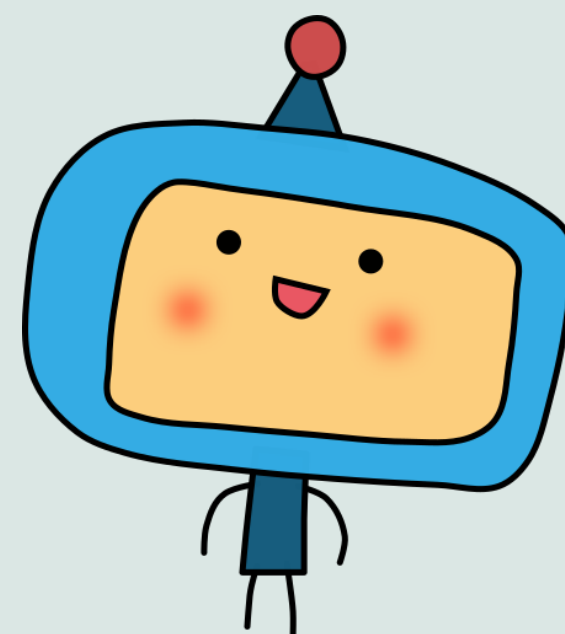




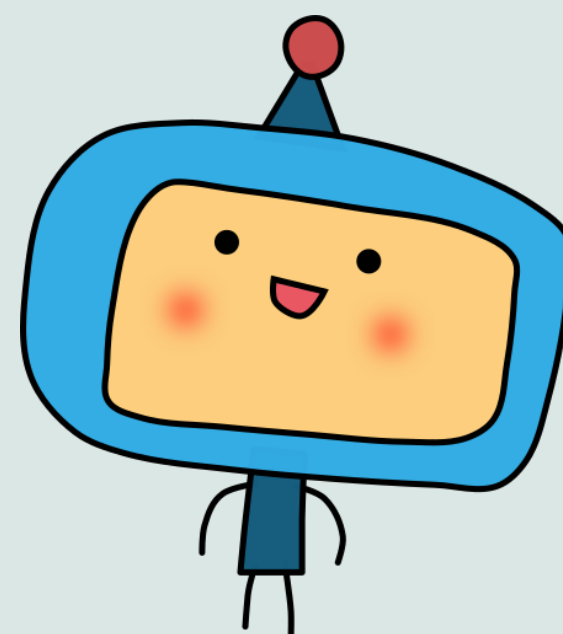
## 【Tool】工具呼叫



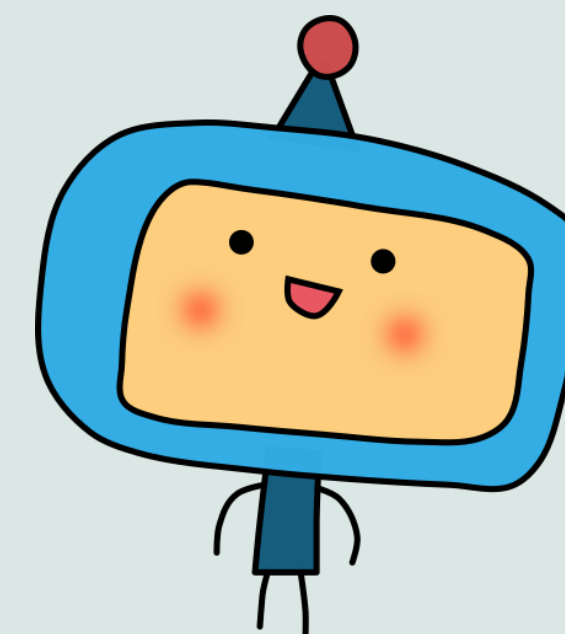
LLM 不太會做的事，  
就不要再逼他了...



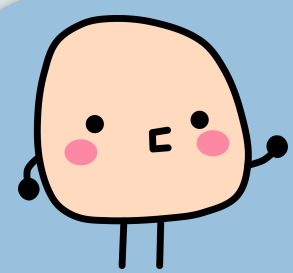
計算機



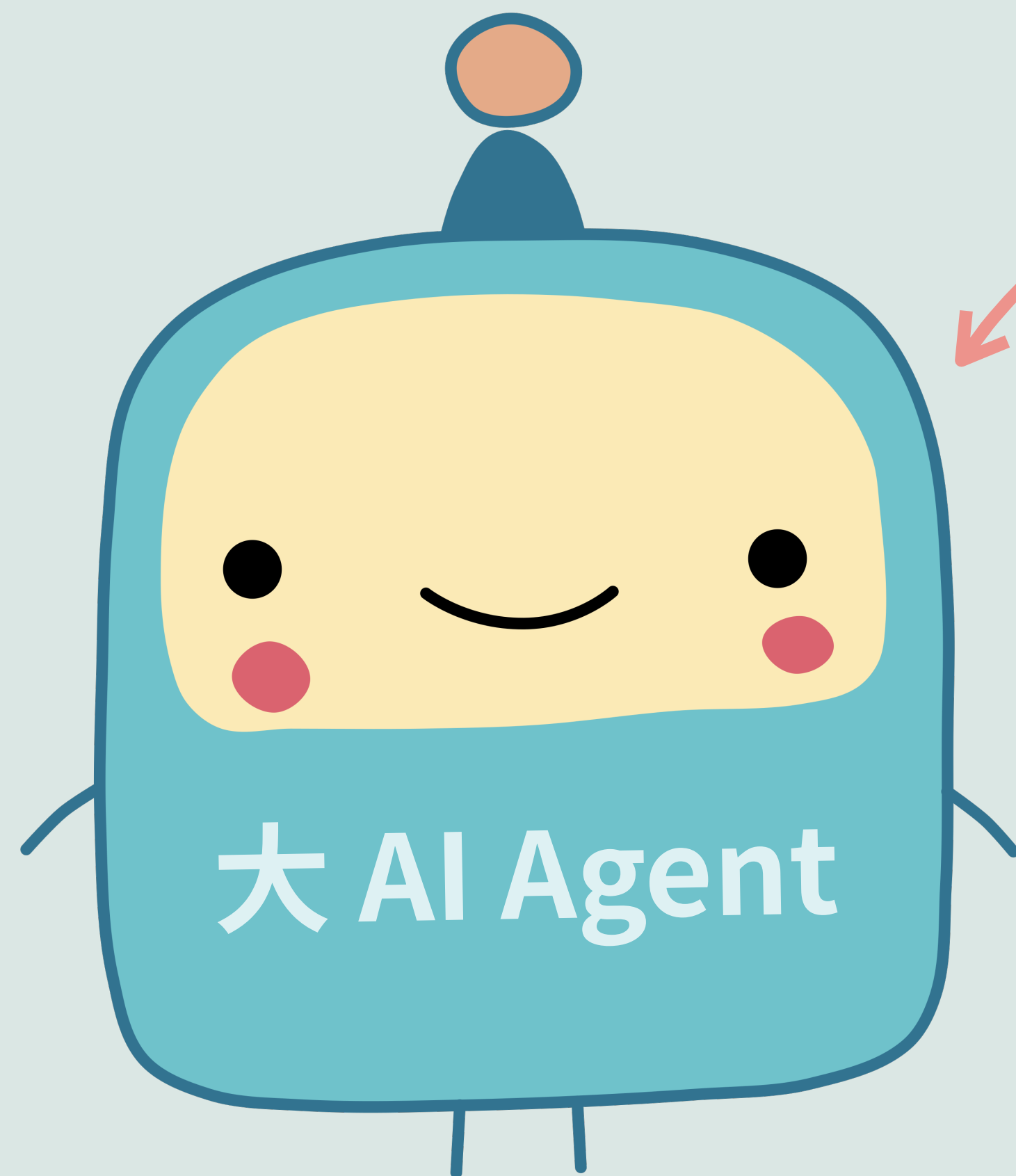
數據搜尋



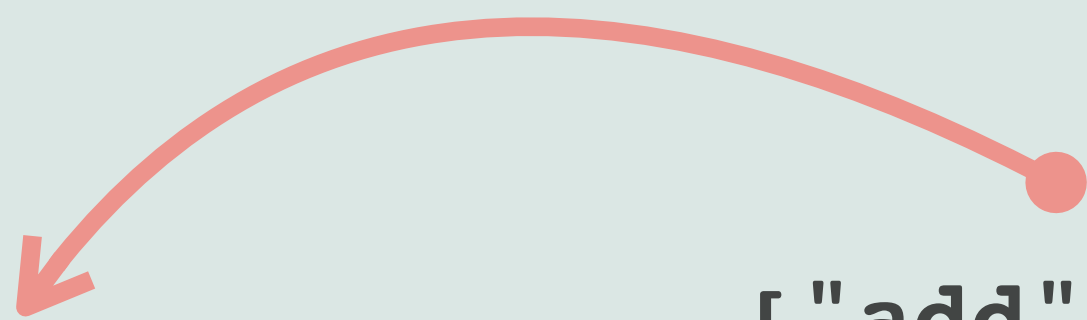
網路搜尋



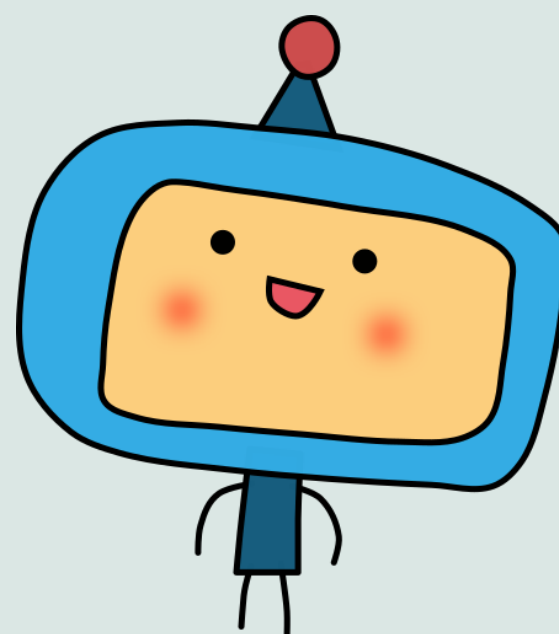
## 【Tool】先告訴我們的 AI Agent, 有哪些工具



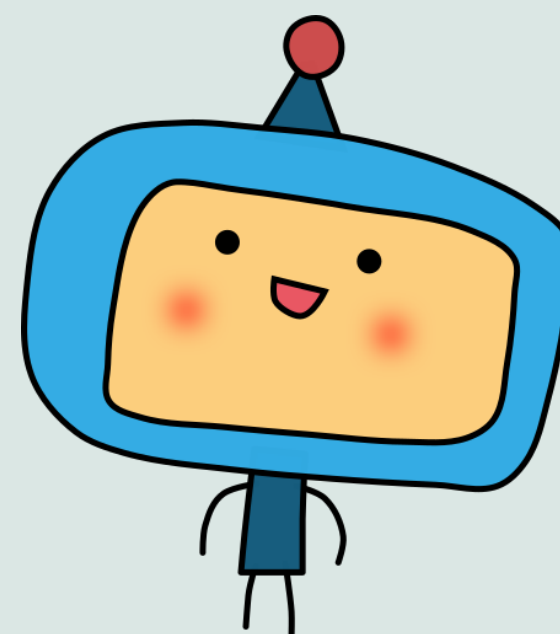
**bind\_tools**



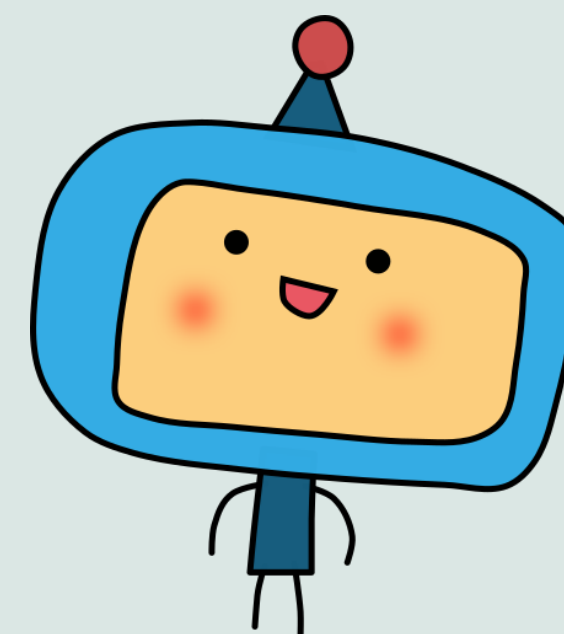
`["add", "get_stock_info", "get_stock_news"]`



`add`



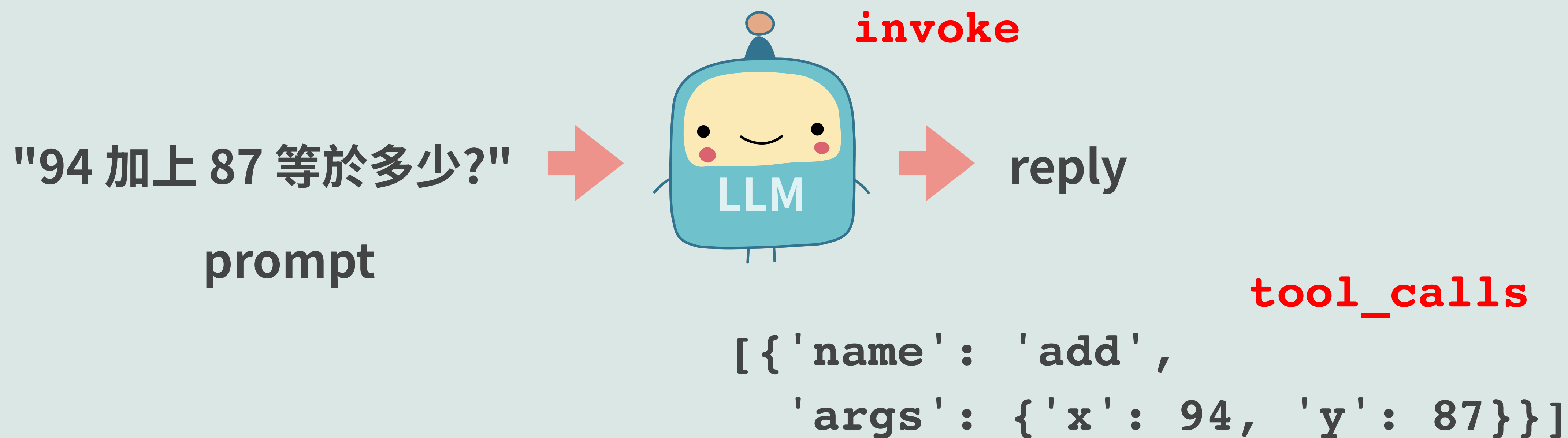
`get_stock_info`



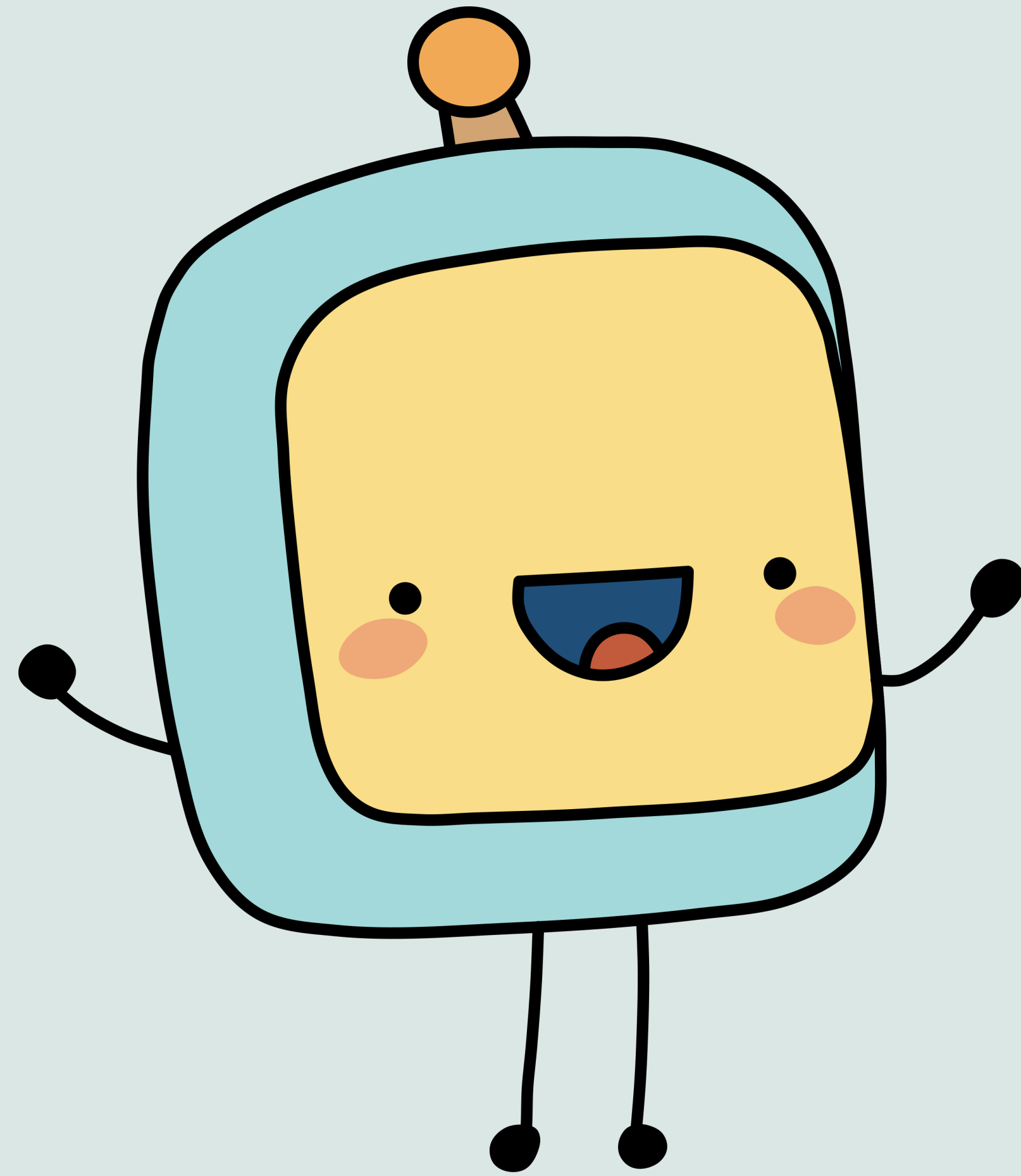
`get_stock_news`



## 【Tool】接著從 prompt 中找到可以用 tools 算的地方





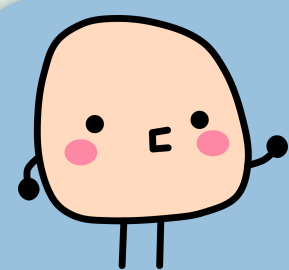


# Planning

## 計畫

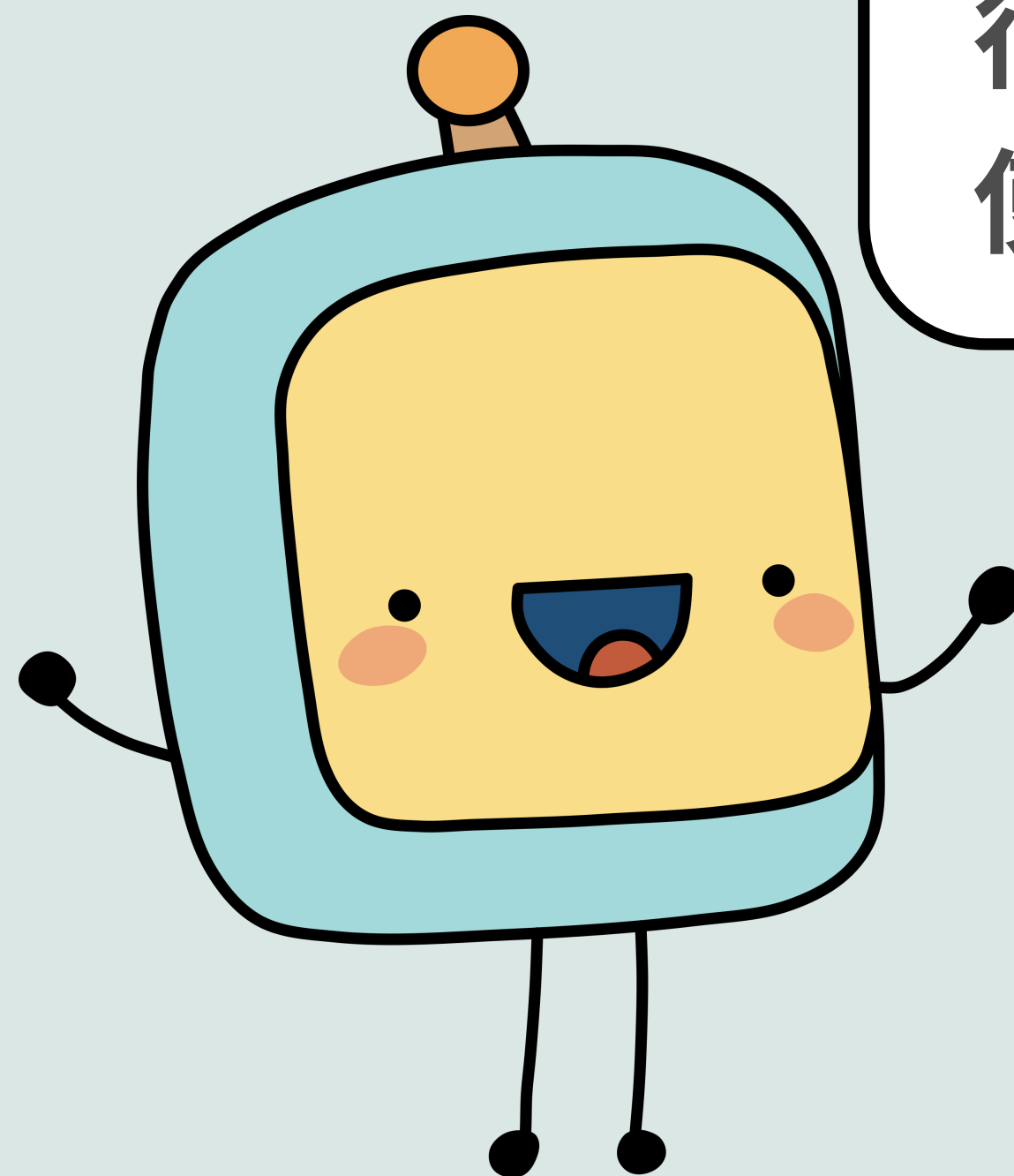
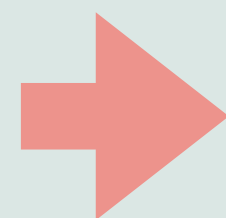
先請 LLM 「想一想」、「打草稿」再正式回應。

<https://yenlung.me/AI07c>



## 【Planning】

prompt

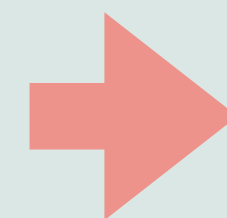


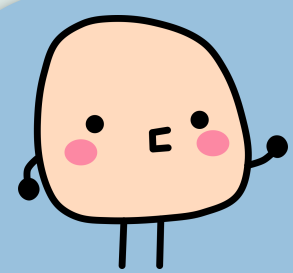
我先來規畫要怎麼進行, 比較好完整回覆使用者需求。

1. ....

2. ....

3. ....





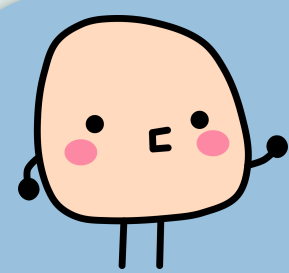
## 【Planning】例如 CoT



**CoT** (Chain-of-Thought) 是一種有名的方法。

- 有系統地拆解複雜議題
- 找出議題間的關聯性
- 更完整地涵蓋各個面向
- 容易看出優先順序





# CoT 是讓 DeepSeek-R1 厲害的關鍵

先產生「思考」用  
`<think>` 和 `</think>`  
tag 包起來。

## Assistant

好的，我需要幫助使用者了解唯識學，但他提到它很多名相很難學，所以我要用淺顯易懂的中文來解釋。

首先，要簡單明了地定義唯識學。它的核心是「一切唯心造」，也就是所有現象都是意識產生的。我可以比喻成夢境，因為夢中的世界看似真實，卻只是大腦的創造，這樣比較容易理解。

接下來，要講如何學習唯識學。使用者可能希望有步驟或方法，所以應該分點介紹，比如從基本概念開始，再逐漸深入，參考經典和現代解釋等。還要強調實踐的重要性，因為佛法不是只停留在理論上。

關於生命連結，要聯繫到人生意義和自我認知。唯識學可以幫助人們理解自己對世界的感知是主觀的，從而改變負面思維，提升生活品質。這樣就能讓使用者看到實用價值，而不僅僅是理論。

需要注意避免使用佛教術語或過於專業的名相，比如直接翻譯英文詞可能會有困難，所以用更常見的比喻和例子來解釋。例如，用鏡子的光影來形容心識的作用，或者將心識與媒體類比，讓人容易聯想到現有的知識。

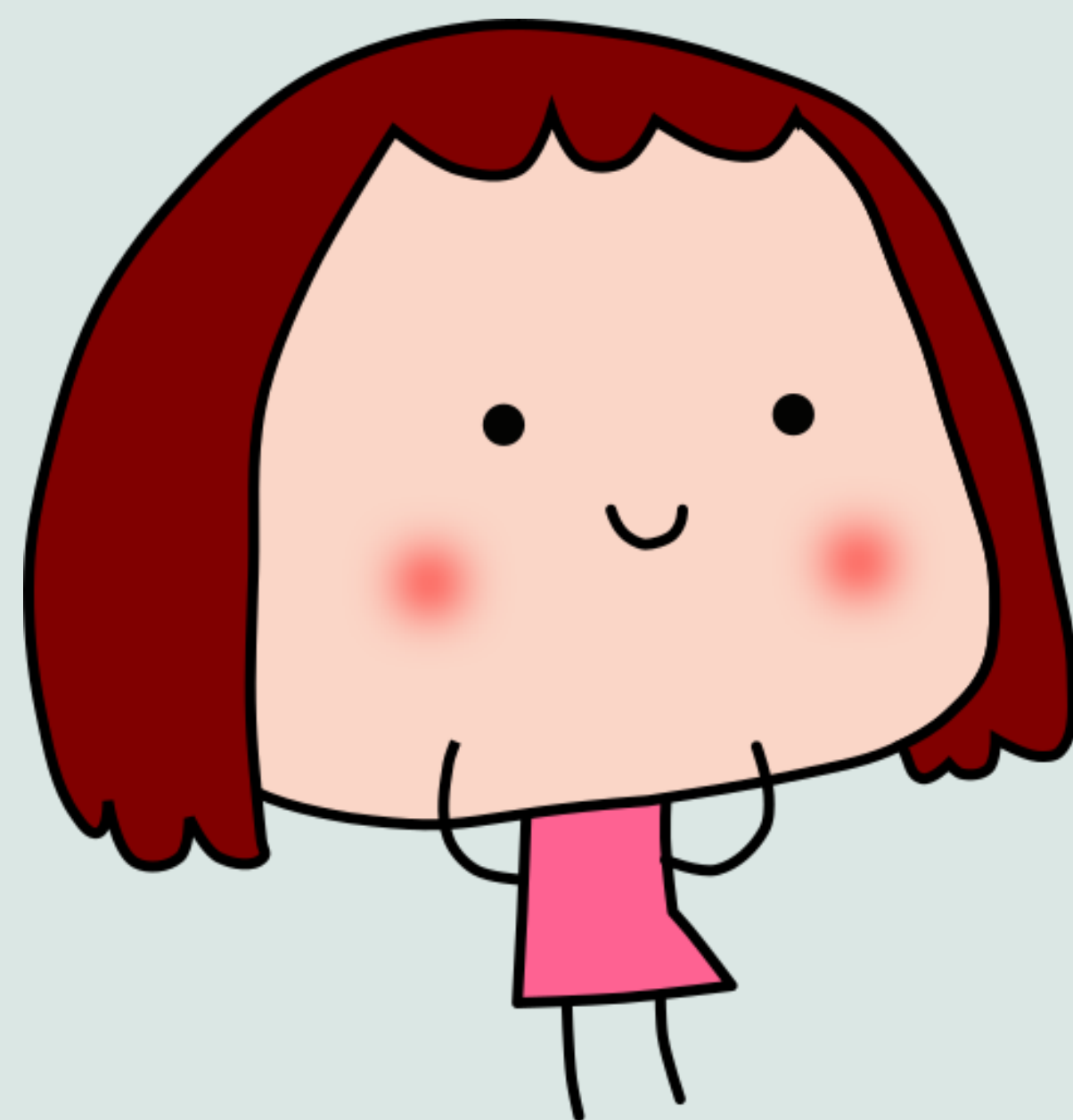
最後，要確保回答結構清晰，有定義、學習方法和生命連結三部分，每一部分都簡潔明了，不會讓讀者感到負擔。還要檢查是否有錯誤或不夠淺顯的地方，確保每個解釋都能幫助使用者逐步理解唯識學的核心思想。

## 唯識學到底是什麼？

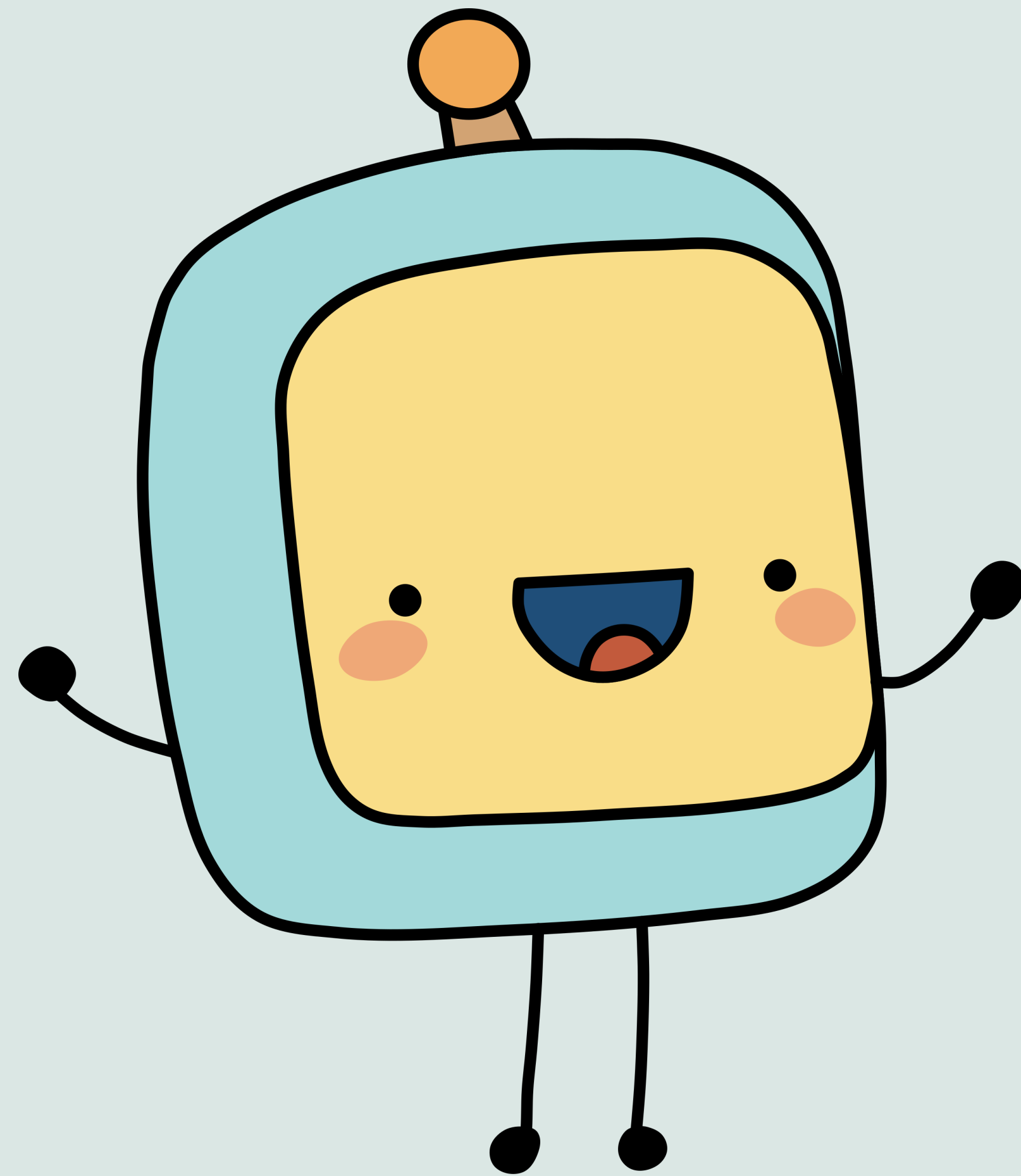
簡單來說，「唯識學」可以理解為「一切現象都是心造的」。佛教裡常用一個比喻來形容：我們就像躺在床上睡夢中的人，將手臂當作蟒蛇、把風聲當作妖怪，其實這些恐懼都只是「心識」投射出來的幻象。唯識學的核心就是告訴你：世界、我、生命……一切看似固定的現實，都源於你的意識在「分割」與「詮釋」的過程中創造的結果。



或許我們可以做得更好!



像 DeepSeek-R1 這種模型, 目標是任何問題都能產生好的思維。如果我們只是單一型的任務, 當然有可能打造更好的「思考」。

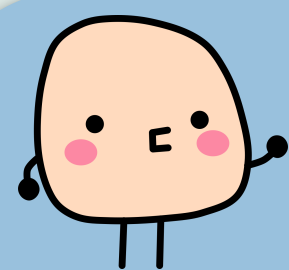


# Multiagent Collaboration

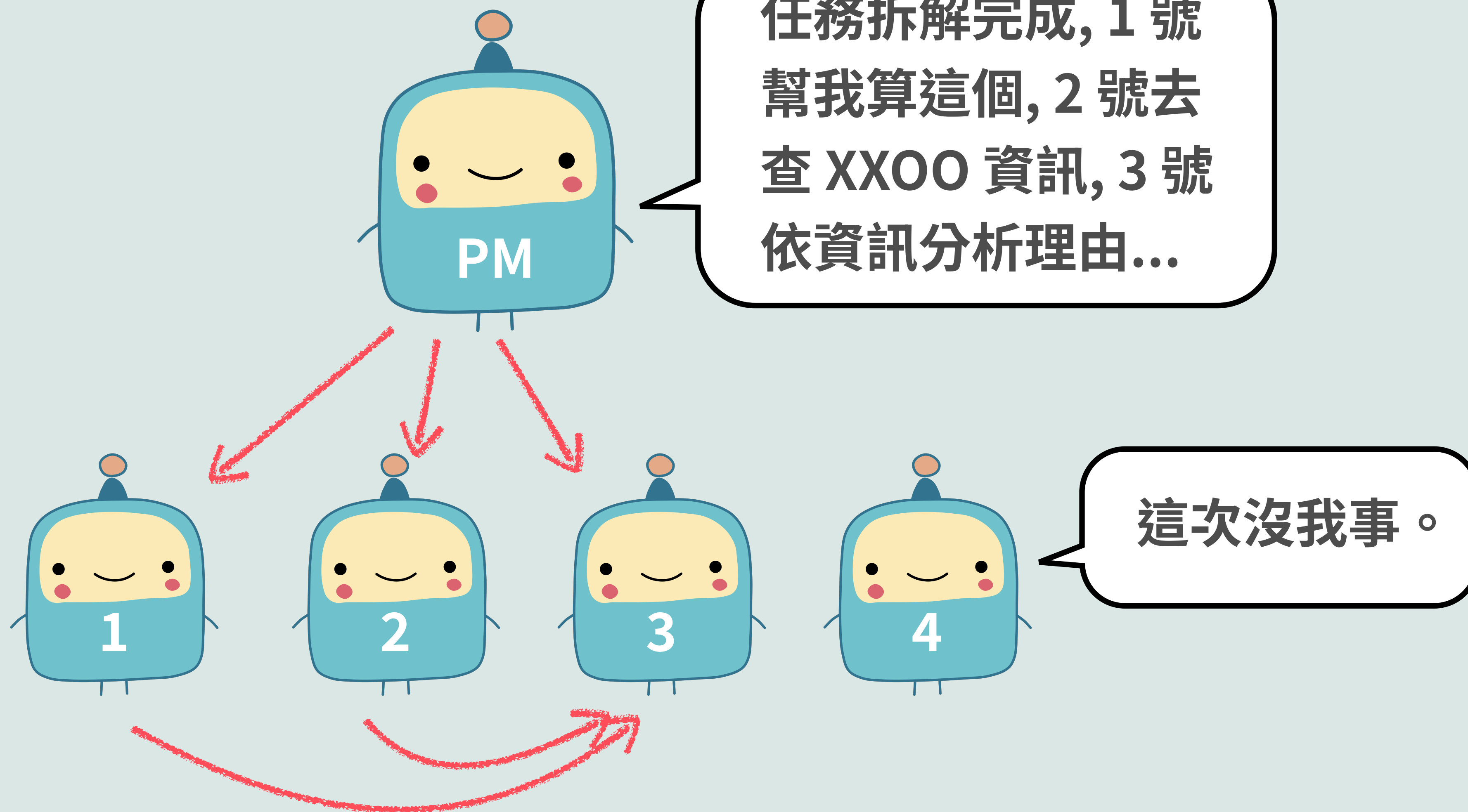
多代理合作

多個 AI Agent 一起工作。



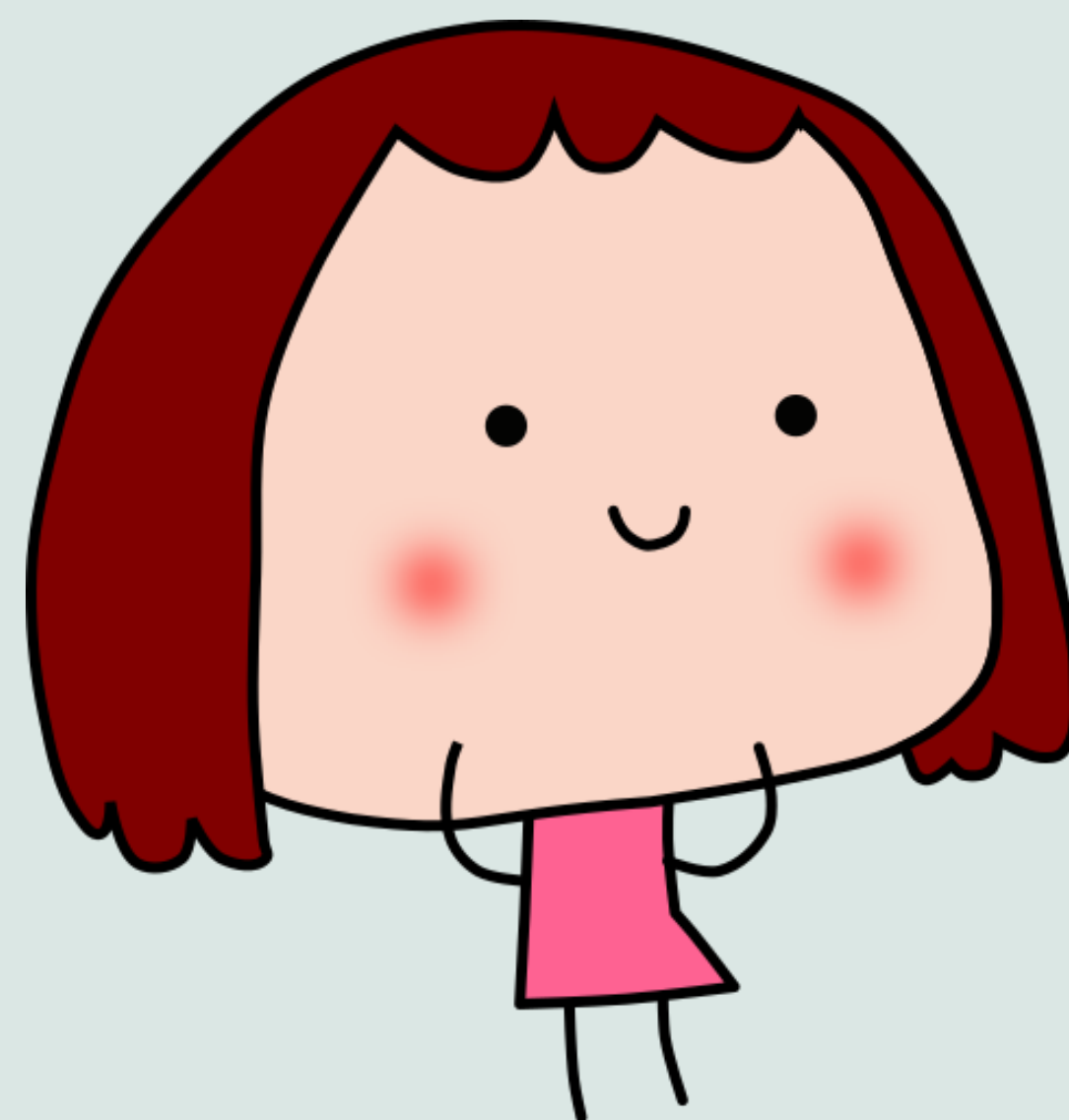


## 【Planning】常常有位如 PM 角色的 LLM 分配工作

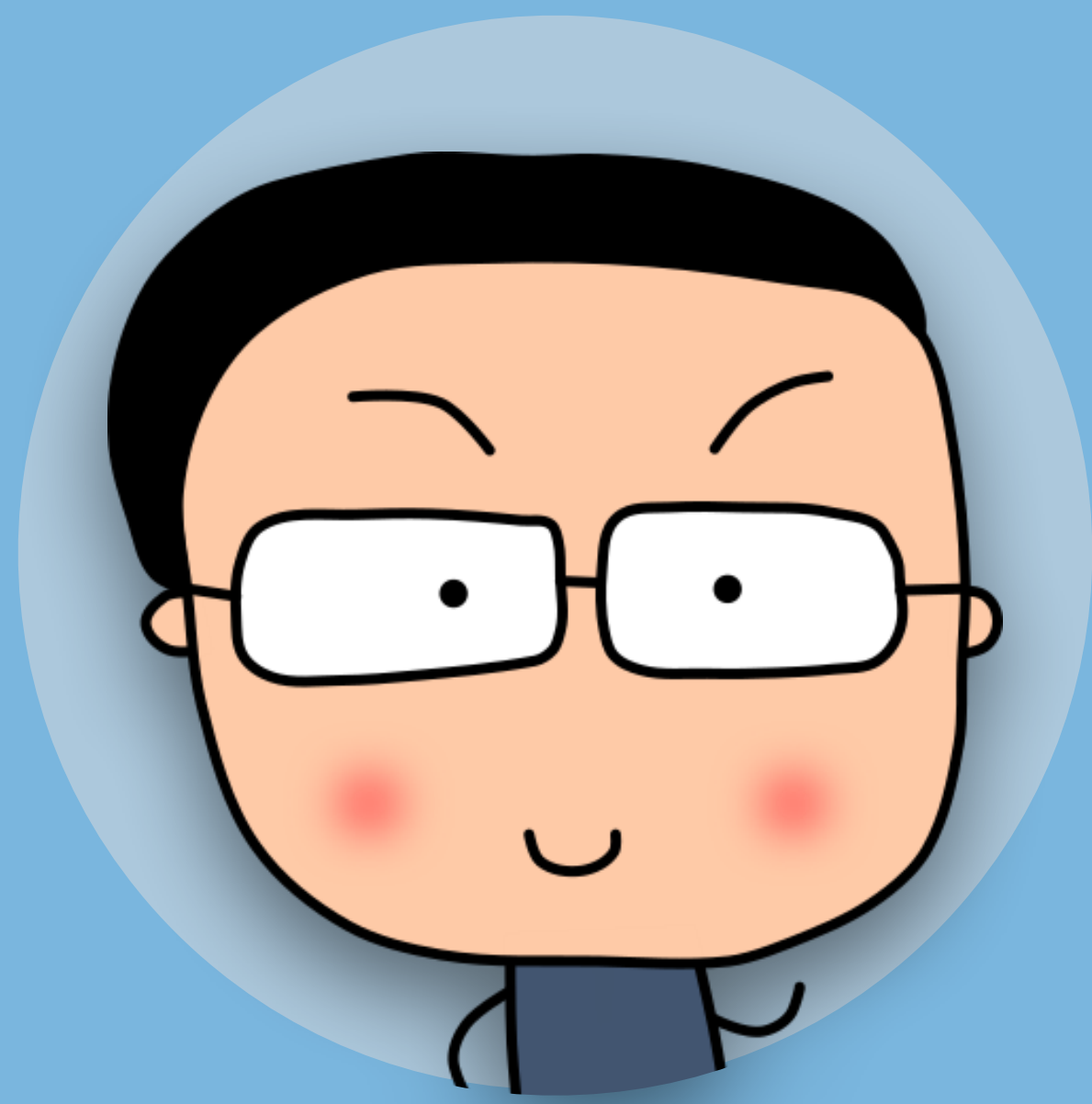




# 2025 年是 AI Agent 元年

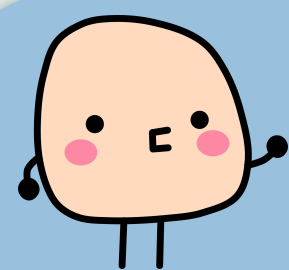


Karpathy 說我們還要  
幾年, 才能真正的理解  
怎麼好好做 AI Agent 。



03.

進階學習



## 值得注意的學習方向

| 類型     | 工具/框架                       | 適合對象 | 用途與學習重點   |
|--------|-----------------------------|------|---|
| ✓ 入門實作 | AlSuite                     | 所有學生 | 中文化、簡單易用的 Agent 系統，適合理解 Agent 架構與創作屬於自己的角色      |
| ✓ 入門互動 | Gradio                      | 所有學生 | 快速建立介面，展示你的 AI Agent 網頁互動版本                     |
| 🌱 進階學習 | LangChain                   | 進階學生 | 建構 RAG + Agent 系統，支援記憶、工具呼叫與流程編排                |
| 🌱 進階實作 | AutoGen (Microsoft)         | 進階學生 | MCP 設計，能構建多 Agent 系統進行任務規劃與協作                   |
| 🌱 進階設計 | CrewAI                      | 進階學生 | 強調角色分工與流程控制，適合建構 AI 團隊協作流程                      |
| 🌱 外部整合 | FAISS / Chroma / Weaviate   | 進階學生 | 建立自有知識庫，用於 Retrieval-Augmented Generation (RAG) |
| 🌱 嵌入模型 | BGE / E5 / OpenAI Embedding | 進階學生 | 將文字轉為向量，用於相似查詢與任務規劃                             |





## 讓 LLM 能和外界溝通的標準

# MCP

Anthropic

Model Context Protocol

與外部世界溝通

MCP 關注的是單一 Agent 如何規劃和使用外部工具的能力, 是在標準化 LLM 與外部工具、資料來源或環境的互動方式。

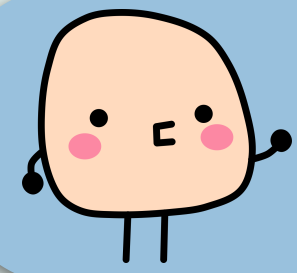
# A2A

Google

Agent-to-Agent

代理之間溝通

A2A 著重 Agent 之間的溝通, 定義了代理之間交換資訊、協作和協調的方式。



## Perplexity & Felo



Perplexity

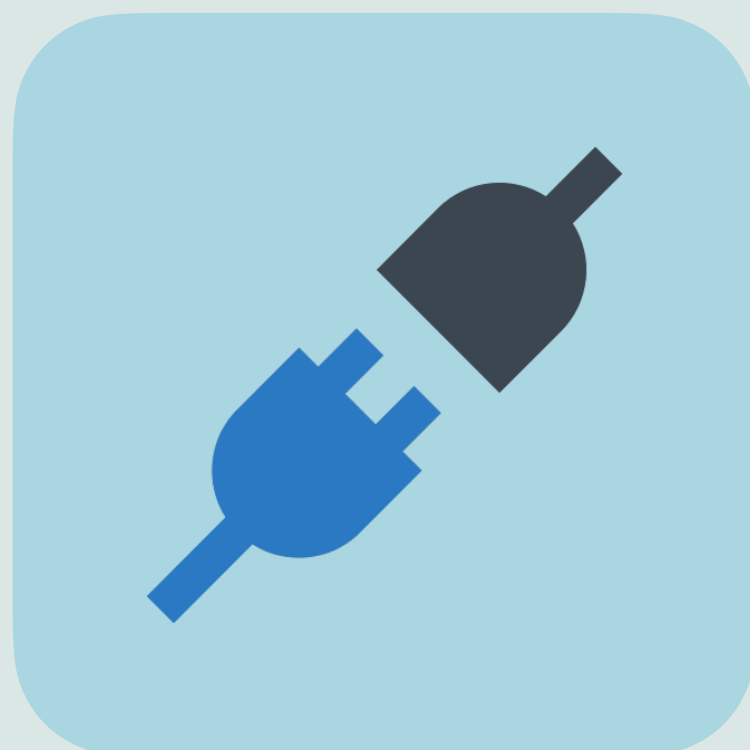


Felo

都可以算是一種 AI Agent



## OpenAI 出了不少個 AI Agent



**Connectors**

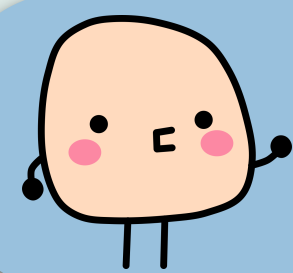


**Search**



**Deep Research**

**還有超多很少人弄得懂差異的模型**



## ChatGPT 5: 導向單一模型

# ChatGPT 5

自動決定要用哪個 AI Agent 的新模型。





# Andrej Karpathy — AGI is Still a Decade Away



<https://www.dwarkesh.com/p/andrej-karpathy>



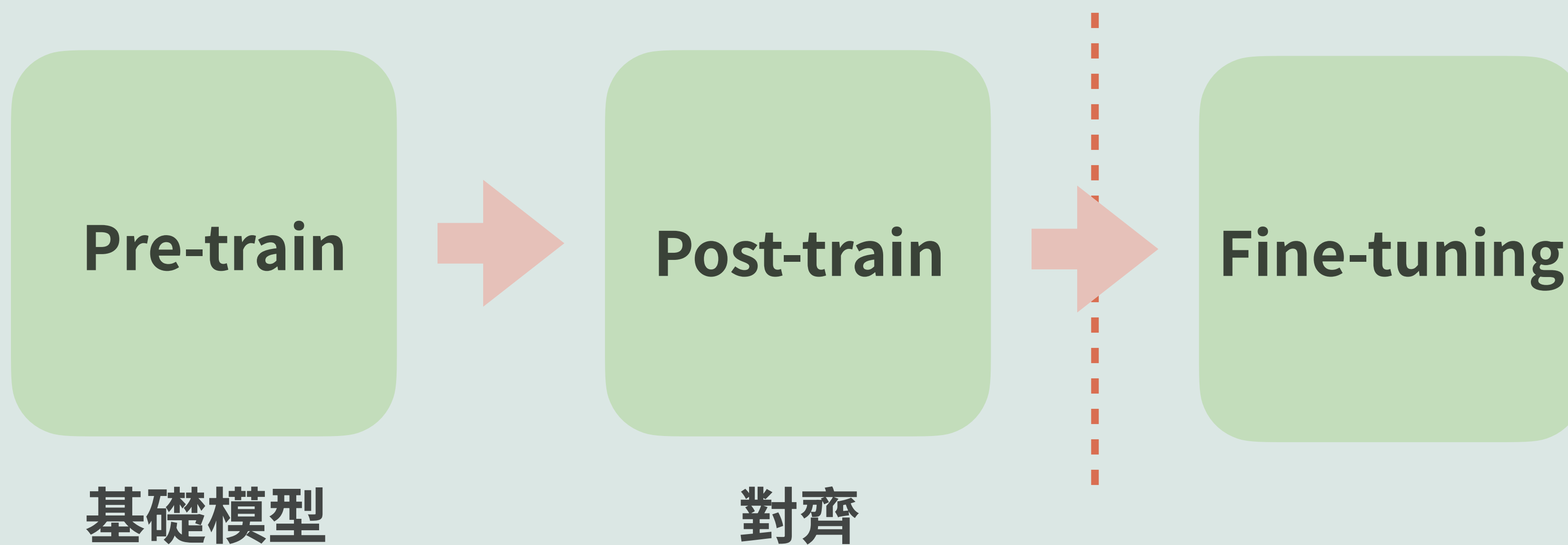
## Karpathy 的重要想法



- AGI 距離現在還有十年: 現在的 AI 還不像真正人類實習生
- 人類學習不等於強化學習 (RL) —— RL 很糟糕
- 模仿人類的網路 (LLM) 「知識壓縮」障礙



## LLM 的訓練三個階段







再一次, 很多人說 Karpathy 對 Vibe Coding 有意見是誤讀

AI 時代三種寫程式的方式:

- 完全不用 LLM (純手工、Karpathy 認為不合理)
- 人類主導 + 自動補全
- Vibe Coding：把需求用自然語言丟給 LLM 生成整段程式



而且這「不是」本次訪談重點





# Hinton: Will AI Outsmart Human Intelligence?



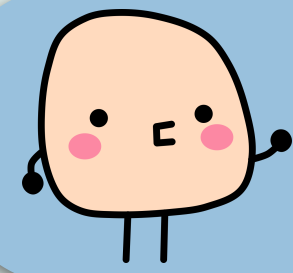




## 人工智慧的學派

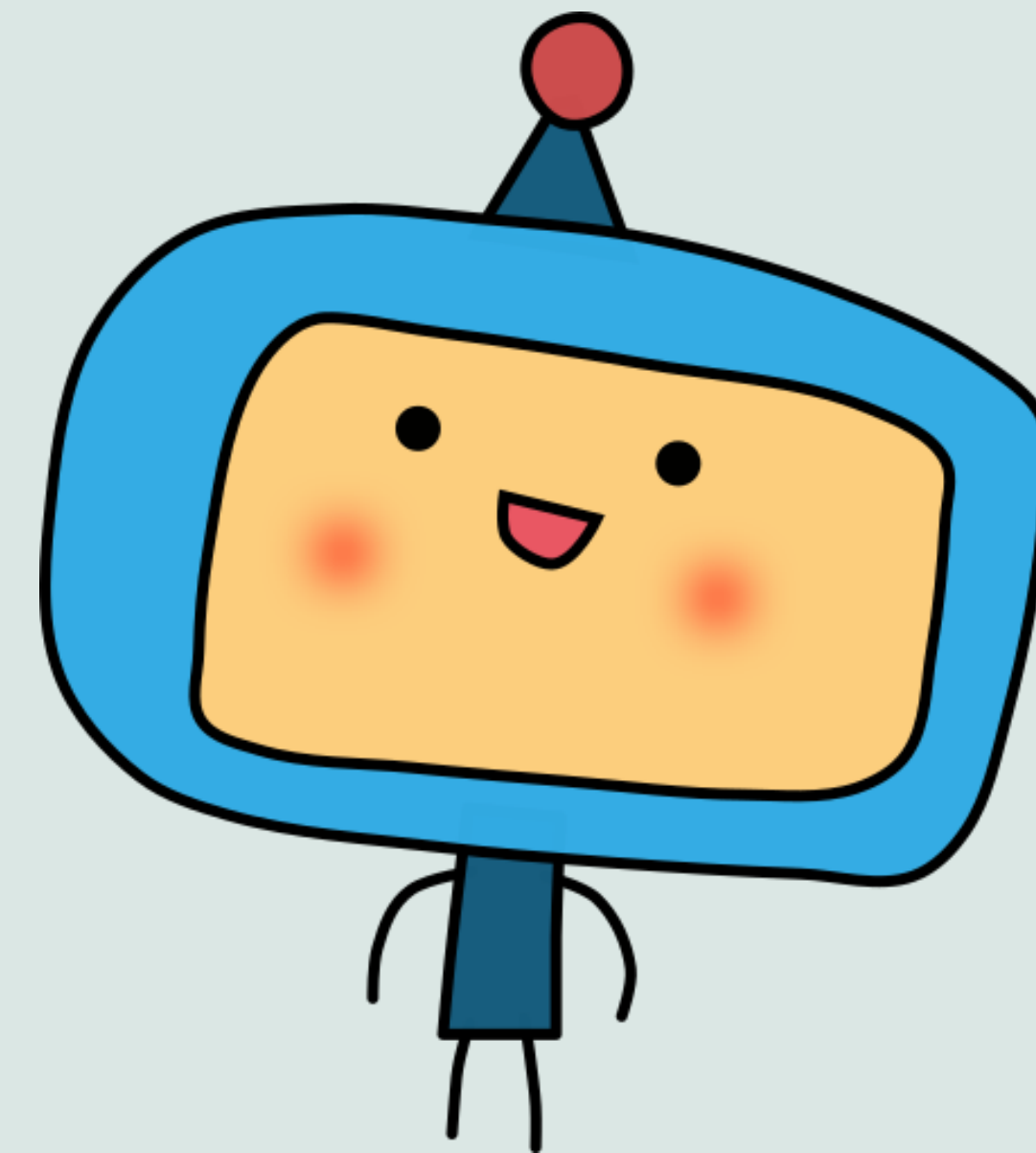


- 符號式邏輯學派 (logic-inspired AI)
- 生物學式學派 (biologically-inspired AI)



## AI 將會比人類聰明

- AI 已經在以與人類相同的方式理解世界
- AI 將會比我們聰明，且想要控制
- AI 甚至可能有主觀經驗 (subjective experience)





Q & A



有問題嗎？