

HON HAI TECHNOLOGY GROUP

生成式AI驅動的新產業革命 以智慧醫療為例

鴻海研究院人工智慧研究所
廖宜斌博士

$$\exists + \exists + \exists = \infty$$



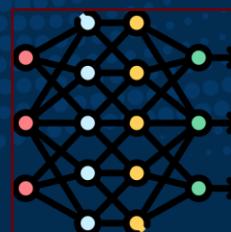
Dr. I-Bin Liao



- 現任
 - 鴻海研究院正研究員
- 曾任
 - 中華電信研究院高級研究員
- 學歷
 - 陽明交大電信碩士、博士
 - 國立成功大學電機系學士



- 研究領域
 - 人工智慧與深度學習
 - 語音處理
 - 自然語言處理
 - 自動駕駛與智慧座艙



目錄



- 1. 趨勢背景：生成式 AI 正在改變產業決策模式**
 - 生成式 AI 的核心概念
 - 生成式 AI 賦能百工百業
- 2. 醫療場域：最能體現生成式 AI 價值的關鍵場景**
 - 個人化精準醫療
 - 下一個產業級機會
- 3. 未來展望**

我們每天都在做決策



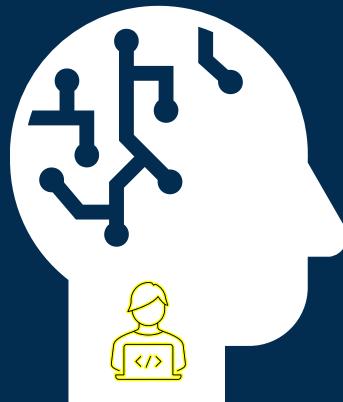
$P(\text{穿短袖} \mid \text{下雨機率}, \text{溫度}, \text{前一天穿什麼})$

生成式AI幫我們決定

輸入
 $x=[s,c]$

天氣、溫度、降雨機率、場合需求、個人舒適偏好

生成式AI
在做的是「機率選擇」



輸出
 y

在某一天 t ，選一套
『最有把握你會覺得舒服』的搭配

$$P_t(a \mid s, c)$$

行動 a = 今天的穿衣決策

狀態 s = 個人偏好 + 天氣資訊

上下文 c = 天氣預報、場合需求

生成式AI最佳化「一段時間內的整體結果」



目標：連續一週都要穿得舒服

昨天穿太薄，今天你可能會改穿厚一點

昨天悶熱不舒服，AI 會調整策略

AI 的決策其實是一連串有記憶的選擇：

$$P_t(a_t | s_{t-1}, c)$$

行動 a_t = 今天的穿衣決策

上下文 c = 天氣預報、場合需求

狀態 s_{t-1} = 前一天穿什麼 + 個人偏好 + 天氣資訊

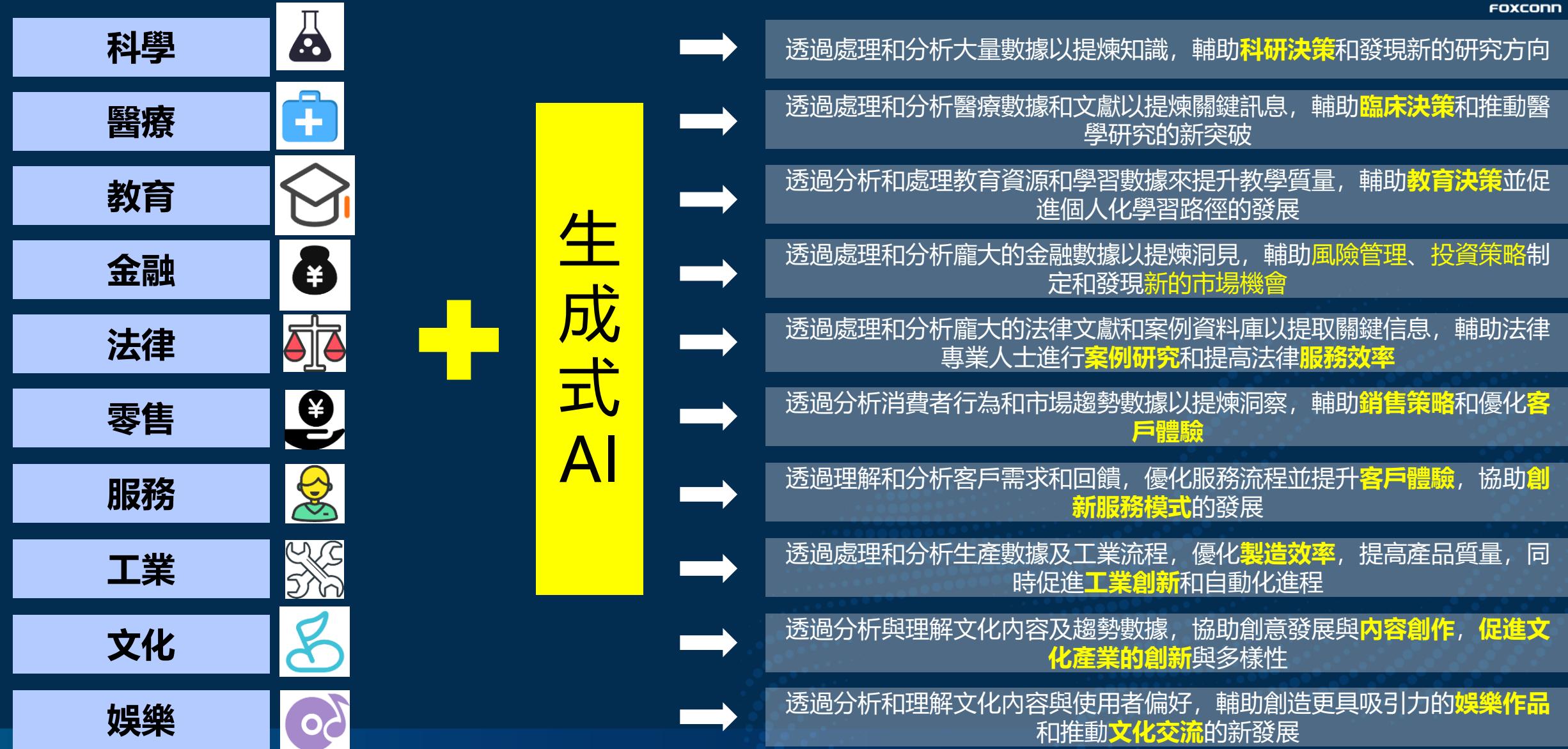
$$P = P_1(a_1 | s_0, c) \times P_2(a_2 | s_1, c) \times \dots \times P_7(a_7 | s_6, c)$$

星期一穿得舒服的機率

星期二穿得舒服的機率

星期日穿得舒服的機率

生成式AI賦能百工百業



5,000億

8億

AI 在醫療的角色演進(1/2)



早期 AI 側重圖像辨識、風險預測，輔助性強但不具語言理解



臨床運用侷限

- 早期AI工具提供部分預測功能，但須醫生判斷所需使用工具。
- 仍無法理解非結構化語言輸入，如病人主訴、醫囑細節或多段式病例記錄，使得操作流程缺乏語意連結，降低使用靈活性與準確度。

AI 在醫療的角色演進(2/2)

大語言模型崛起後，AI 開始能聽懂醫師語言、理解上下文



臨床語境下的潛力

- AI 不只是回答問題，更能**主動**生成摘要、推論病因
- 可支援醫師在「**決策前中後**」階段，有效提升效率與準確率

與人聊天



我肚子痛



病了吧

吃多了吧

滾

與機器聊天 – AI 1.0時代



我肚子痛



思維固定

你肚子痛的情形是突然發生的嗎？有沒有
其他症狀，例如：腹瀉、便秘……

通用、安全、標準的答案

與機器聊天 – AI 2.0時代



我肚子痛



SYSTEM PROMPT
{.....}

USER PROMPT
{我肚子痛}

思維靈活

SYSTEM PROMPT
{請扮演難搞的女朋友}

情境導向的答案

你說你肚子餓了，難道你忘記我今天加班到很晚嗎？你自己不能找點東西吃嗎？

為什麼猩猩會停止進化？



繁體中文 5 √

人類並不是猩猩「進化來的」，而是和牠們**共享一個數百萬年前的共同祖先**。從那個祖先開始，不同族群朝不同方向演化：有的成為現代人類，有的成為今日的黑猩猩、大猩猩或紅毛猩猩。

3. 猩猩仍在進化，只是方向不同

牠們仍持續在基因層面上演化，只是變化未必顯而易見。例如牠們的免疫系統、社會行為與適應森林環境的能力，都在隨時間微調。

簡單來說：

猩猩沒有停止進化，只是牠們**不需要像人類一樣改變**，因為牠們的生活方式與環境已經讓牠們非常適應了。

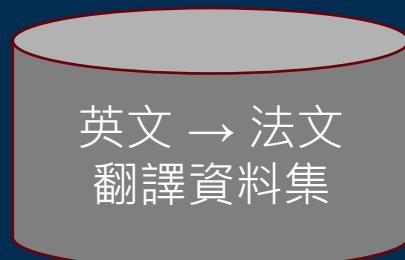
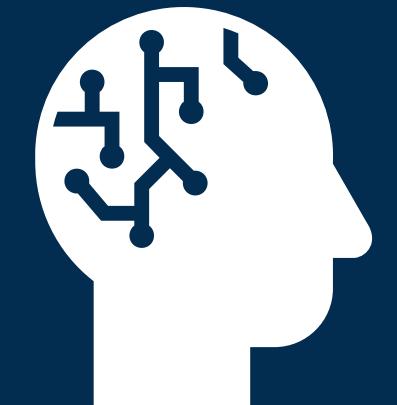
跨語言通用化能力



I love cats →



我愛貓



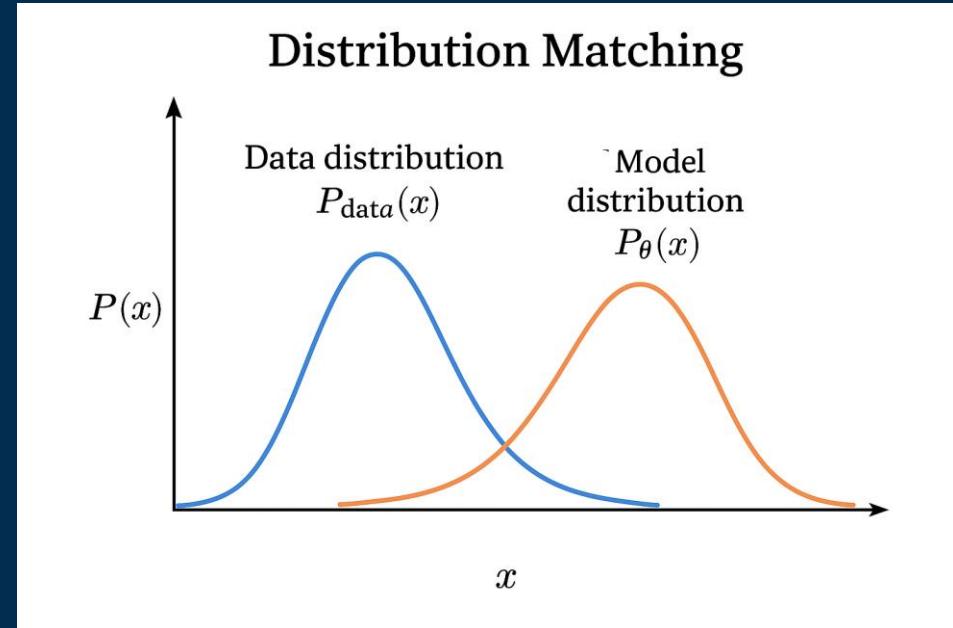
J'aime les chats

跨語言通用化能力

■ 分佈匹配 (distribution matching)

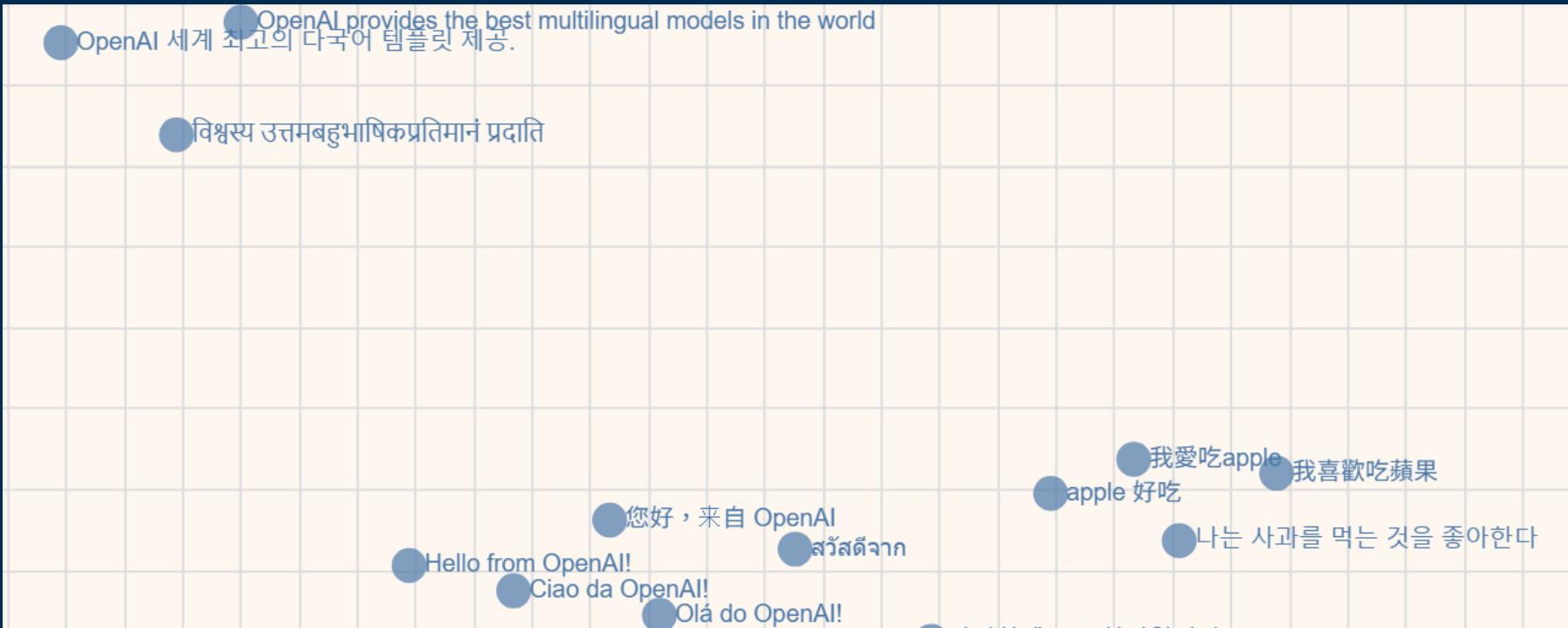
■ 在跨語言大型語言模型 (LLM) 學習中，是確保不同語言的詞彙和語義能夠在模型內部空間中對齊的關鍵機制

- 例如：「英文的 *pneumonia*」和「中文的『肺炎』」，在模型的語意空間中會被映射到相近的位置



在高維空間中，匹配語義、上下文、句法等複雜特徵的整體機率分佈

跨語言通用化能力



	en	fr	es	de	el	bg	ru	tr	ar	vi	th	zh	hi	sw	ur	avg
Token emb	83.1	73.3	73.9	71.0	70.3	71.5	66.7	64.5	66.6	68.2	63.9	66.9	61.3	58.1	57.3	67.8

參考來源 : On the Cross-lingual Transferability of Monolingual Representations (Artetxe, Ruder & Yogatama, ACL 2020)

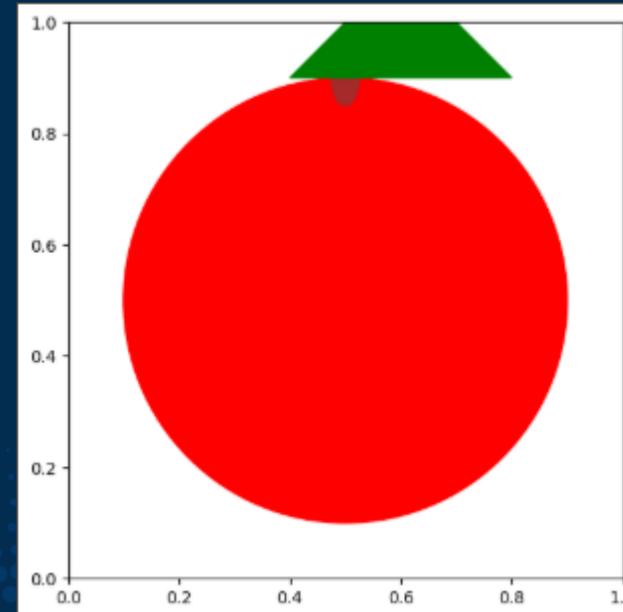
LLM：跨語言學習讓模型「學懂醫學」



- 大部分的醫學文獻與資料庫 (PubMed、NEJM、Lancet、Radiology...) 都是英文撰寫。臨床醫師每天閱讀的資訊中，「知識來源是英文」但「臨床交流是中文」
- 建立一個以繁體中文為核心語言的 LLM，並在模型中對齊英文醫學文獻的語義分佈，那麼這個模型就能成為醫學知識與臨床實踐之間的橋樑

語言模型憑什麼「看得見」？

- **視覺先驗知識** - 語言模型不是靠「看圖」學會視覺，而是靠「理解語言中的視覺結構」建立視覺能力
- 視覺先驗分為兩種：推理先驗與感知先驗
 - **推理先驗**：來自程式碼、數學、學術文本，強化模型的空間邏輯與結構理解
 - **感知先驗**：來自多樣化語料（如網頁爬蟲），幫助模型理解物件外觀與細節



最佳訓練配方：

- 50% 以上推理型語料（程式碼、數學、研究文獻）
- 15% 左右視覺世界語料（網頁、圖像描述）

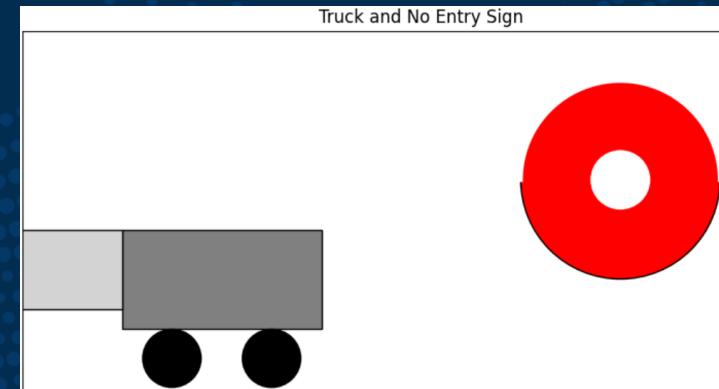
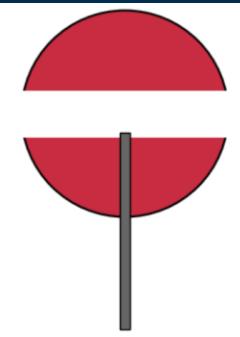
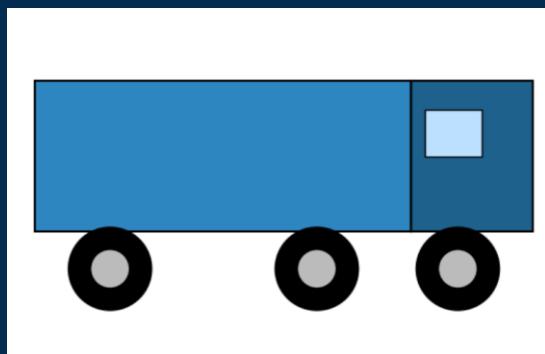
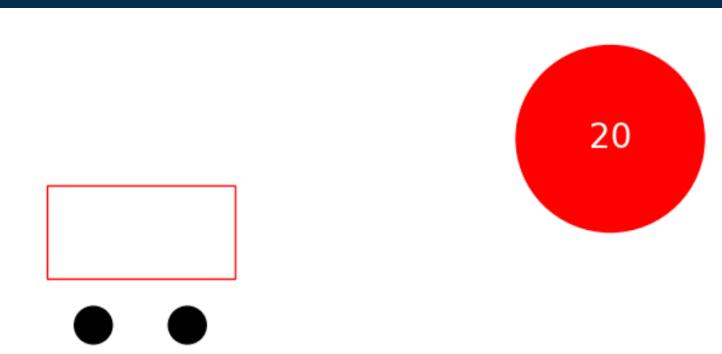
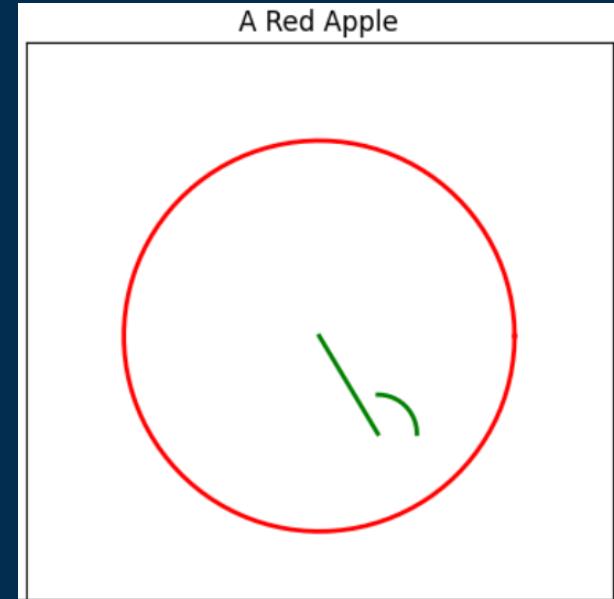
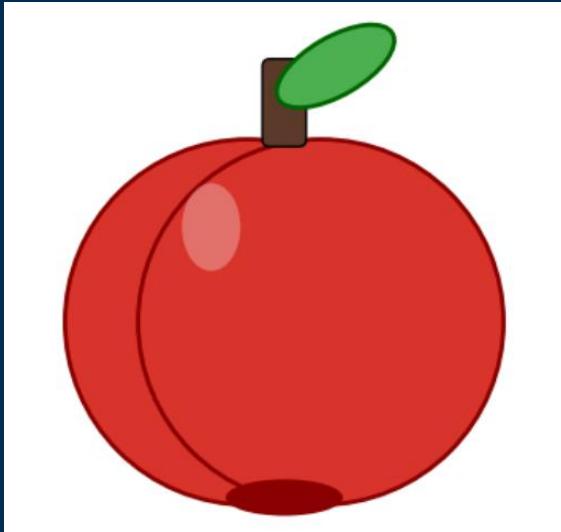
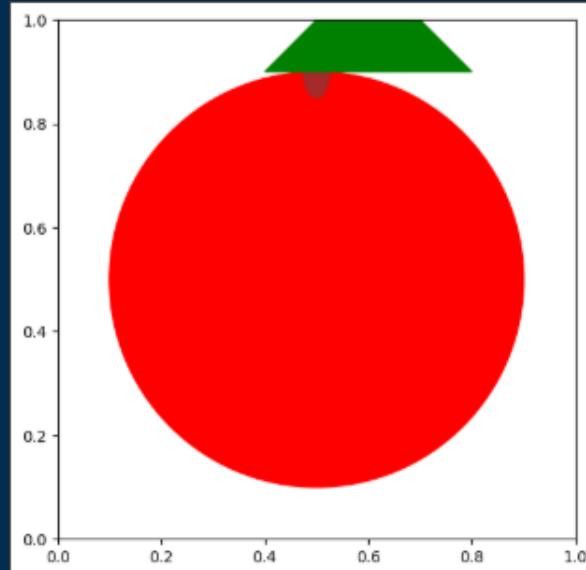
語言模型出視覺能力

```
import matplotlib.pyplot as plt  
import matplotlib.patches as patches  
# 設定圖形的大小和背景顏色  
fig, ax = plt.subplots(figsize=(6, 6))  
ax.set_facecolor('white')  
# 繪製蘋果的主體 apple = patches.Circle((0.5, 0.5), 0.4,  
color='red') ax.add_patch(apple)  
# 繪製蘋果的莖 stem = patches.Ellipse((0.5, 0.9), 0.05,  
0.1, color='brown') ax.add_patch(stem)  
# 繪製蘋果的葉子 leaf = patches.Polygon([[0.4, 0.9],  
[0.6, 1.1], [0.8, 0.9]], color='green')  
ax.add_patch(leaf)  
# 設定圖形的範圍和比例 ax.set_xlim(0, 1) ax.set_ylim(0,  
1) ax.set_aspect('equal')  
# 顯示圖形 plt.show()
```

LLM思考過程

- 首先，我們需要導入 Matplotlib 的 pyplot 模組，並設定圖形的大小和背景顏色。
- 接下來，我們將使用 Matplotlib 的 patch 模組來繪製一個圓形，這將代表蘋果的主體。
- 然後，我們將在圓形上添加一個小的橢圓形，這將代表蘋果的莖。
- 最後，我們將在圓形上添加一些葉子，這將使蘋果看起來更加真實。
- 最後，我們將顯示圖形。
- 這個解決方案應該能夠滿足用戶的需求，並且使用了 Matplotlib 來繪製一個蘋果。

語言模型學出視覺能力



大語言模型如何理解醫學知識？



面向醫療領域 - MedGemma



■ Google 推出的醫療視覺-語言基礎模型

- 專為醫療影像與文本理解而設計的開源模型

■ 解決醫療 AI 的三大挑戰

- 多樣化資料（影像、病歷、問答）
- 任務複雜度高（診斷、報告生成、推理）
- 隱私與資料取得困難

■ 臨床意義

- 降低模型訓練門檻，加速醫療 AI 的實際應用

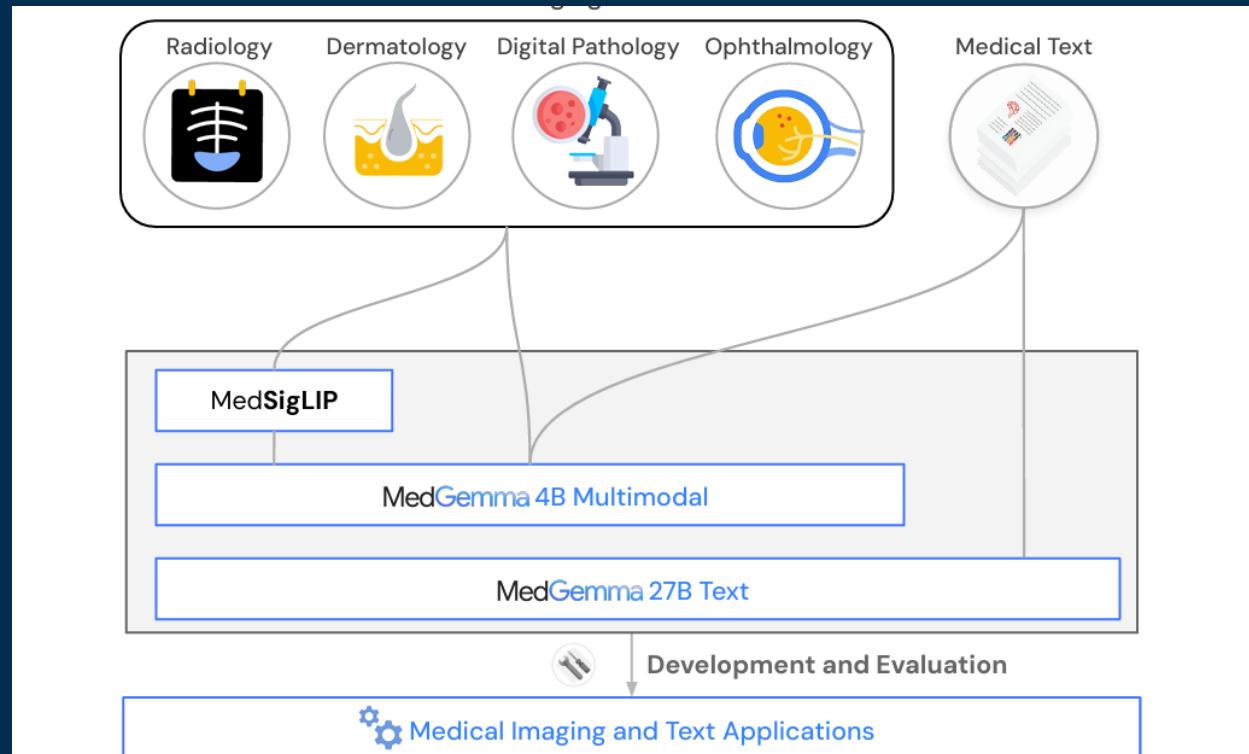
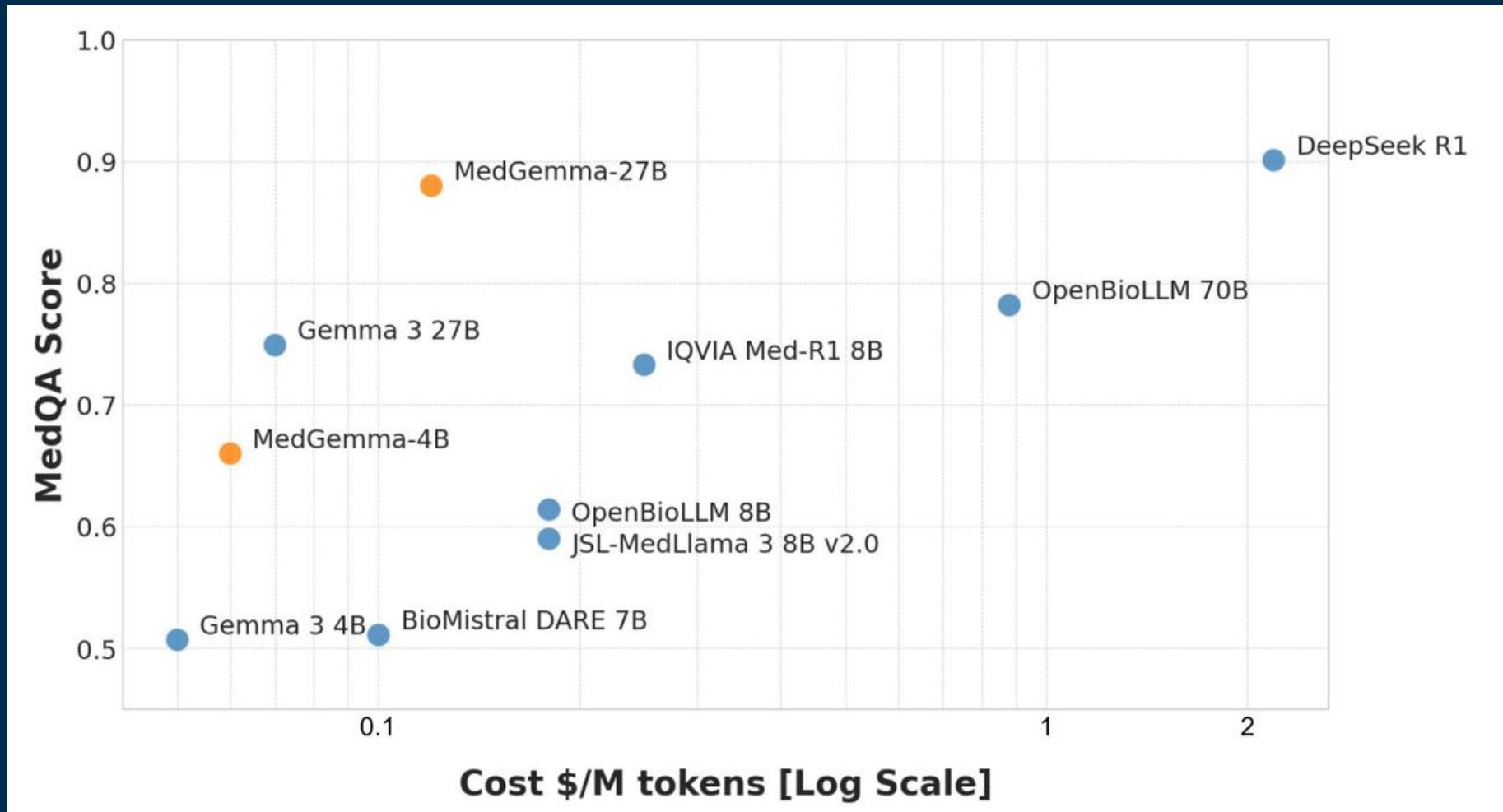


Figure 1 | Overview of the MedGemma model collection featuring the MedSigLIP image encoder, MedGemma 4B Multimodal and MedGemma 27B Text

Google Research and Google DeepMind

MedGemma Technical Report

MedGemma 評測結果



圖片來源：[Google Research Blog](#)

MedGemma 在多種醫療任務中的表現



任務類型

表現亮點

醫療問答

超越同級模型，在 MedQA、PubMedQA 等基準表現優異

影像分類

胸腔 X 光、皮膚病、眼底影像分類準確率高

視覺問答

能理解影像並回答臨床問題，支援 SLAKE、VQA-RAD 等資料集

報告生成

在 MIMIC-CXR 上生成報告，臨床決策一致率達 81%

醫師代理任務

在模擬診間中進行病史詢問、診斷推理，表現接近人類醫師

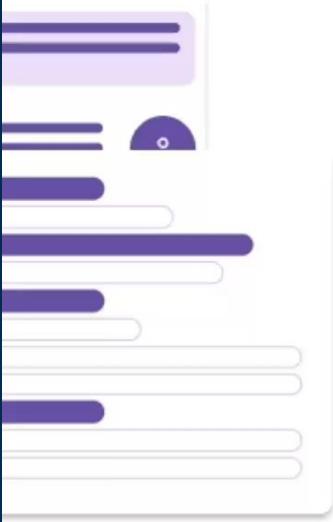
***臨床意義** - 作為輔助診斷、教學、病歷摘要工具

醫療問答筆記輔助工具



Simulated Pre-visit Intake Demo

Healthcare providers often need to gather patient information before appointments. This demo illustrates how MedGemma could be used in an application to streamline pre-visit information collection and utilization.



First, a pre-visit AI agent built with MedGemma asks questions to gather information. After it has identified and collected relevant information, the demo application generates a pre-visit report.

This type of intelligent pre-visit report can help providers be more efficient and effective while also providing an improved experience for patients relative to traditional intake forms.

Lastly, you can view an evaluation of the pre-visit report which provides insights into the quality of the output. For this evaluation, MedGemma is provided the reference diagnosis, allowing "self-evaluation" that highlights both strengths and what it could have done better.

Disclaimer This demonstration is for illustrative purposes only and does not represent a finished or approved product. It is not representative of compliance to any regulations or standards for quality, safety or efficacy. Any real-world application would require additional development, training, and adaptation. The experience highlighted in this demo shows MedGemma's

MedGemma如何學習醫療知識



■ 訓練資料涵蓋：

- 醫療問答資料集 (MedQA、PubMedQA 等)
- 醫療影像與報告配對 (MIMIC-CXR、EyePACS 等)

■ 微調方式：

- 強化學習 (RL) 提升推理能力
- 指令微調 (Instruction tuning) 提升臨床語言理解

■ 可針對特定任務進行微調，如：

- 肺氣胸分類 (SIIM-ACR)
- 病理切片分類 (CRC100k)
- EHR(電子健康記錄) 資訊檢索與推理 (EHRQA)

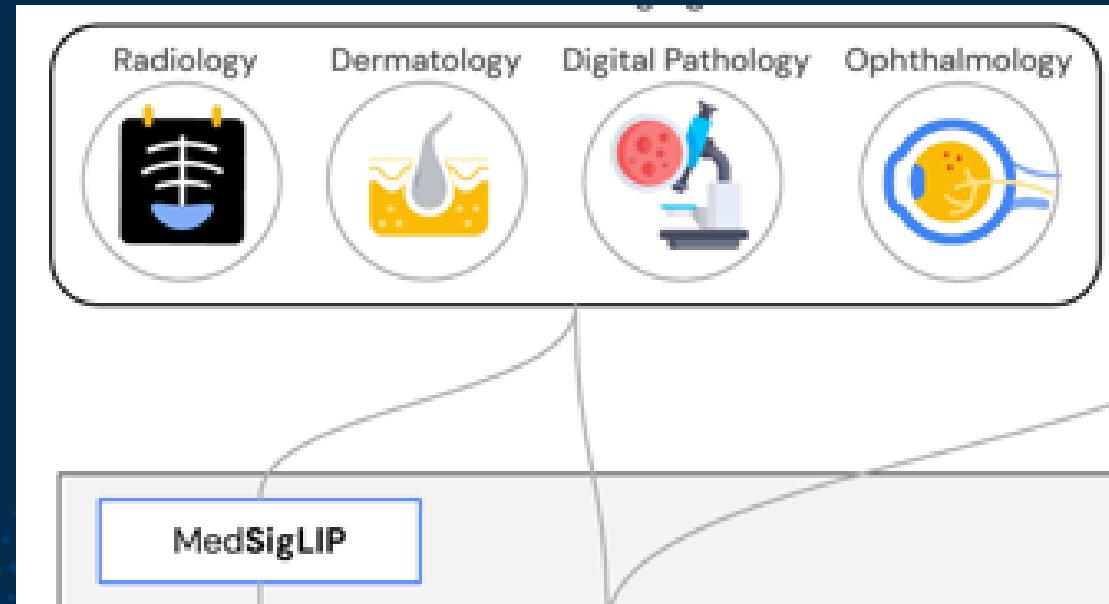
***臨床意義** - 可客製化模型以符合科別需求

MedGemma 如何理解醫療影像？



■ MedSigLIP 影像編碼器

- MedGemma 的影像理解由 MedSigLIP 編碼器驅動
- 支援多種醫療影像：X 光、皮膚病、眼底、病理切片
- 零樣本分類 (Zero-shot) 與少樣本學習 (Linear probe) 皆表現優異
- 在骨折、肺部病變等難題上超越同級模型 (如 ELIXR)



*臨床意義 - 可快速部署於新影像任務，減少標註需求

多模態整合驅動個人化精準醫療

多模態數據整合

-  基因檢測
-  病理報告
-  腫瘤標記
-  乳房影像
-  藥物紀錄

現況

未來

臨床決策的高度不確定性



醫師根據經驗
制定治療計畫

病患療程中的資訊焦慮



醫病溝通過程
常見認知偏差
???



生成式AI建議
最佳治療路徑



數位科技克服
醫病溝通障礙



個人化精準醫療 - 可被設計、可被控制的數學問題



■ 「在病人限制條件下，如何配置醫療流程，最大化治療成功機率？」

醫療紀錄 x



$$\text{價值} = P(\text{正確}) - \lambda \times \text{成本}$$

$$P = P_{diagnosis}(x, u_1) \times P_{treatment}(x, u_2) \times P_{monitoring}(x, u_3)$$

新產業機會

■ 每個病人都 是獨特案例

- ✓ 年齡、體質、分子特徵、預算、風險承受度不同
- ✓ 傳統方式：醫師靠經驗選方案
- ✓ AI 平台：可以自動生成個人化流程配置

■ 醫療流程多步驟串接成機率鏈

- ✓ 成功率 = 診斷 × 治療 × 追蹤
- ✓ AI 可以在每個環節進行推理與優化，不只是單點決策

■ 成本與資源限制可納入決策

- ✓ 傳統醫療忽略資源分配效率
- ✓ AI 可以同時考慮預算、風險、成功率，形成最優配置

■ 產業級擴展能力

- ✓ 一套平台可以服務成千上萬個病人，每個病人都 是不同的最佳化問題
- ✓ 從單位醫院 → 區域醫療網 → 國家級醫療系統，有可擴展的商業價值

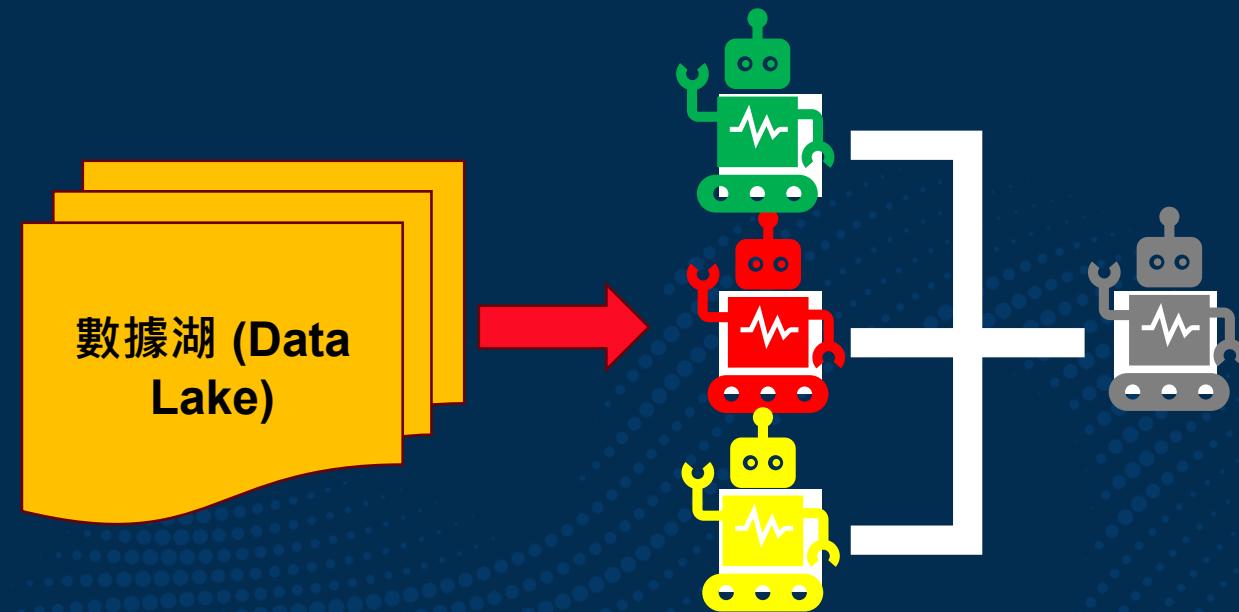
臨床應用潛力



■ 臨床應用潛力

- 輔助診斷：提供初步判讀與建議
- 報告生成：減少醫師書寫負擔
- 教學工具：訓練住院醫師與學生
- 病歷摘要：快速擷取重點資訊
- 多語言支援：跨國醫療溝通
- 結合 EHR 系統與 PACS
- 支援更多影像模態 (MRI、超音波)
- 強化臨床推理與決策支持能力

多Agent協作
臨床決策輔助平台



● 更個人化的健康管理

- LLM 可分析個人病歷、生活習慣、基因資料，提供量身打造的健康建議

● AI輔助臨床決策系統

- LLM 可快速整理病患資料、醫學文獻、臨床指南，協助醫師做出更準確判斷

● 跨平台整合與遠距醫療擴展

- LLM 可作為語言介面，串接穿戴裝置、手機App、雲端病歷系統
- 遠距醫療場景：病患在家透過語音與LLM互動，醫師遠端監控並介入