

(簡報摘要)

# The Challenges and Opportunities of Generative AI in Library and Information Science

## 生成式AI在圖書資訊領域中帶來的挑戰與機會

---

Dr. Willie Yang 楊立偉博士

意藍資訊

# 內容大綱 Agenda

## 講師背景介紹

1-1 生成式 AI 的重要

1-2 生成式 AI 的基本原理

2-1 生成式 AI 特性和長處

2-2 重要的應用場域

3-1 生成式 AI 的發展與影響

3-2 生成式 AI 的未來

交流討論



# 講者背景介紹

---



# 楊立偉 博士

## 領域

- 資料庫管理及搜尋引擎
- 機器學習及語意分析
- 大數據及商業分析

## 現任

- 台大資管系/工管系暨商研所兼任助理教授
- eLAND 意藍資訊 (6925) 創辦人及總經理
- AI 大聯盟大語言模型工作組召集人
- 資訊及通信國家標準技術委員

## 經歷

- 100 MVP最有價值經理人，近30項語意分析專利
- 國家雲端創新獎、數位時代「創業之星」首獎
- 任教於台灣大學超過18年，為教學傑出教師
- 數位時代選為代表台灣軟體的封面人物之一

# 意藍資訊

首家掛牌的 AI 智能數據公司  
股票代號 6925

- 國內 AI 智能數據代表廠商，擁有巨量、多樣化的智能數據，專精於搜尋和 AI 語意分析，擁有豐富知識管理系統(KM)、智能數據處理經驗，以 SaaS 訂閱制提供高品質解決方案。
- 從企業內部數據延伸到外部數據，處理規模更大，發展出更精準的 AI 語意分析技術。
- 主要客戶包括高科技、電信、金融、零售等產業之龍頭企業，擁有極高之市場覆蓋率
- 從企業軟體轉型到雲端服務成功、且獲利的典範。

# AI 智能數據 產品服務

## 雲端資料分析服務

**Opview** 社群數據

規模最大的社群網路及品牌口碑資料庫

**PeopleView** 人群數據

AI 分析興趣及意圖的人群樣貌資料庫

## 智能搜尋 AI Search

**AI Search for EC**

智能貼標與搜尋推薦系統

**AI Search for KM**

新一代生成式 AI 知識管理解決方案

## 搜尋引擎軟體及其他

處理百億筆非結構資料仍有絕佳效能的  
自有核心引擎

## 資料即時性最高

上百台雲端主機  
24小時爬文不間斷  
整合多種情報數據

## 豐富的應用導入經驗

最權威的分析流程方法  
各大機構及百位學者認可採用



## 觀測範圍最廣

收錄超過220,000個頻道  
商業運轉規模遠超過同業

## 領先業界 AI 語意分析

國際資訊檢索評估會議第一名  
公信單位評選第一之服務廠商



即時快速

24時不間斷爬文

+



60億+

中文字收錄/日

+



22萬+

網站頻道收錄

+



專利技術

語意與斷詞技術

→

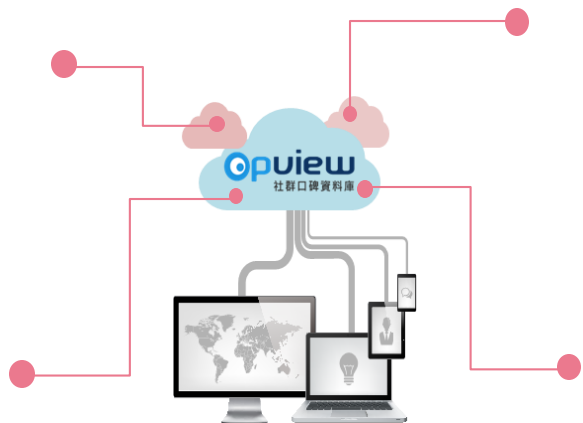


單一入口查詢

# AI 輿情應變顧問

資料即時性最高

觀測範圍最廣



即時完整  
相關資料

接下來一周新光金的  
股價會如何？

未來一週新光金的股價走勢可能會呈現出一定的波動，監管方面的壓力和合併的不確定性會帶來下跌壓力，而良好的獲利表現帶來支撐。

營運性服務經驗

領先業界  
AI 語意分析

AI  
輿情應變顧問

網友對於王品食安議題  
的討論面向？

1. 員工衛生管理問題
2. 消費者健康安全
3. 食藥署的處罰措施
4. 疫情追蹤與公共衛生



即時快速

24時不間斷爬文



公開資料

巨量搜尋引擎



類型完整

網站頻道收錄



專利技術

AI 語意技術



# 基於企業內部資料，提供精確且可驗證的回答

**AISKM** eLAND

新人報到要交哪些東西

新人報到時，應繳交或填寫以下資料：  
人事資料表（如有異動，需即時通知行政單位更正）。  
勞動契約書（一式兩份，公司及員工各執乙份）。  
健康告知書（應據實填寫）。  
身份證影本。  
最高學歷證件影本。  
一寸照片一張。  
銀行帳號—薪資轉帳存摺影本。  
其他依職務或工作性質所需之檔案。  
參考段落：<<4221#42214>>。

檔案來源與內容：

- <<4221#42214>> 新人教育訓練與工作規則.pdf
- <<4221#42216>> 新人教育訓練與工作規則.pdf

口語化提問

提供參考資料

## 第二章 僱用、資遣、解僱及離職

- 第六條 各單位因工作之需要，必須增補人員時，應填具「招募人員申請表」呈總經理核准後，交行政單位進行招募作業。
- 第七條 新進人員報到時，應繳交或填寫如下之資料：  
一、人事資料表。（員工資料如有異動，請即時通知行政單位更正，以確保連絡資料之正確性）。  
二、勞動契約書（一式兩份，公司及員工各執乙份）。  
三、填寫健康告知書（應據實填寫）。  
四、身份證影本。  
五、最高學歷證件影本。  
六、一寸照片一張。  
七、銀行帳號—薪資轉帳存摺影本。  
八、其他依職務或工作性質所需之文件。
- 第八條 本公司新進員工依職務不同試用期間為三個月，合格者依規定正式僱用之。不合格者即停止僱用，並依本規則第九、十一、十二及十三條等相關規定辦理。



---

1-1

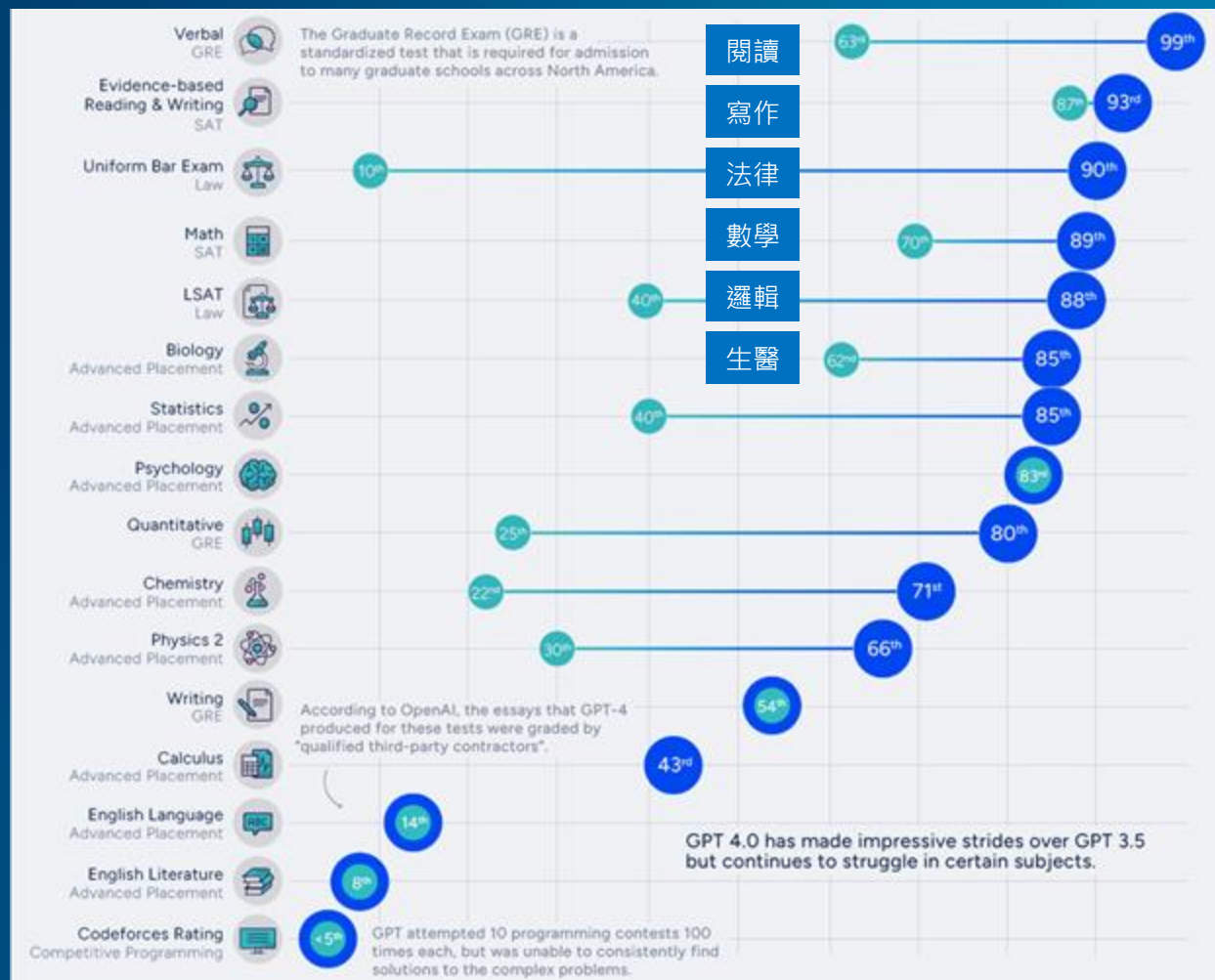
引言：為什麼生成式 AI 這麼重要

Intro: Why Generative AI Matters

# 生成式AI有多聰明？

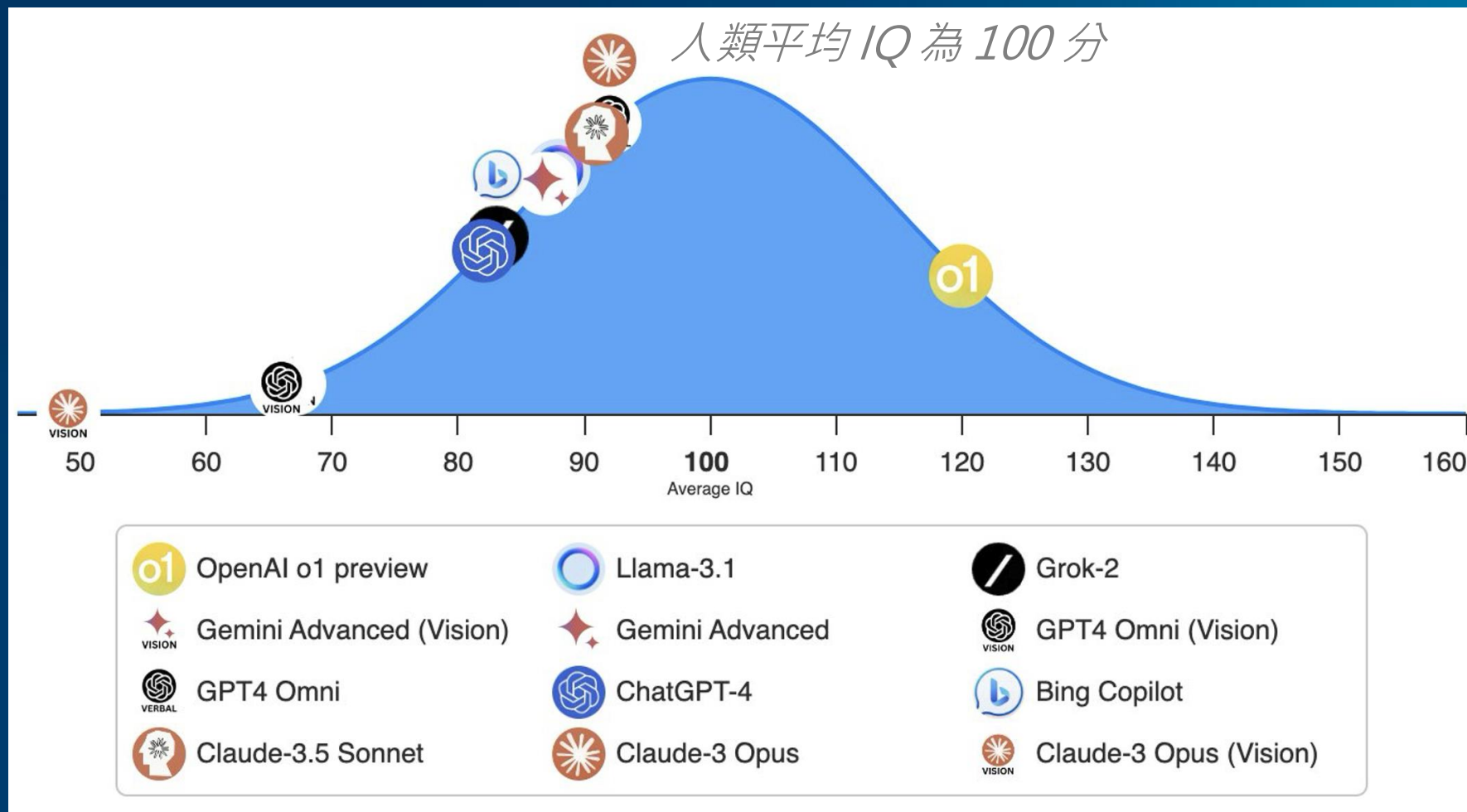
遍覽群書，在記憶  
與融會上勝過人類  
前6項考贏85%人類

閱讀	法律	邏輯
寫作	數學	生醫



Source: <https://www.visualcapitalist.com/how-smart-is-chatgpt/>

# 最新版 AI 在智力測驗中達 120 分



Source: <https://vulcanpost.com/870954/openai-new-gpt-model-o1-score-120-iq-test/>

1-2

# 生成式 AI 的基本原理

Fundamentals of Generative AI

# 革命性的大語言模型 LLM (Large Language Model)

閱讀時時時變化的注意力 (attention)，會依上下文做變換

範例一

請先閱讀下列文章「大雄和哆拉a夢說胖虎要開演唱會，他要每個人明天下午都到空地集合，小夫說如果你有道具可以拯救大家，他就要請大家吃點心。」請問文章中第一個他是指誰？第二個他是指誰？其中的你是指誰？



第一個「他」是指胖虎，第二個「他」是指小夫，「你」是指哆拉a夢

## 範例二

請針對下列文字做斷詞：「昨天全台大停電，一家人晚上摸黑吃飯。在台灣大學唸書的哥哥，打電話來說全台大也停電。」請問前面的全台大和後面的全台大有什麼不一樣？



斷詞結果：昨天/全台/大/停電/，/一家人/晚上/摸黑/吃飯/。/在/台灣大學/唸書/的/哥哥/，/打電話/來/說/全台大/也/停電/。

前面的全台大指的是「全台灣」，後面的全台大指的是「全台灣大學」。

# 學習新的符號並且向量化

讀完以下文章後請回答甚麼是 "令和"：

日本公布新年號「令和」後，掀起一股新風潮。商家開始大搶「令和」商機。

日本民眾：「咦？是什麼？令和？！令和，令和元年。」

日本學生：「新年號，令和。」



# AI 把世上符號都數學向量化

學到了相似符號會聚在一起

例如：親戚關係

學到了相對關係

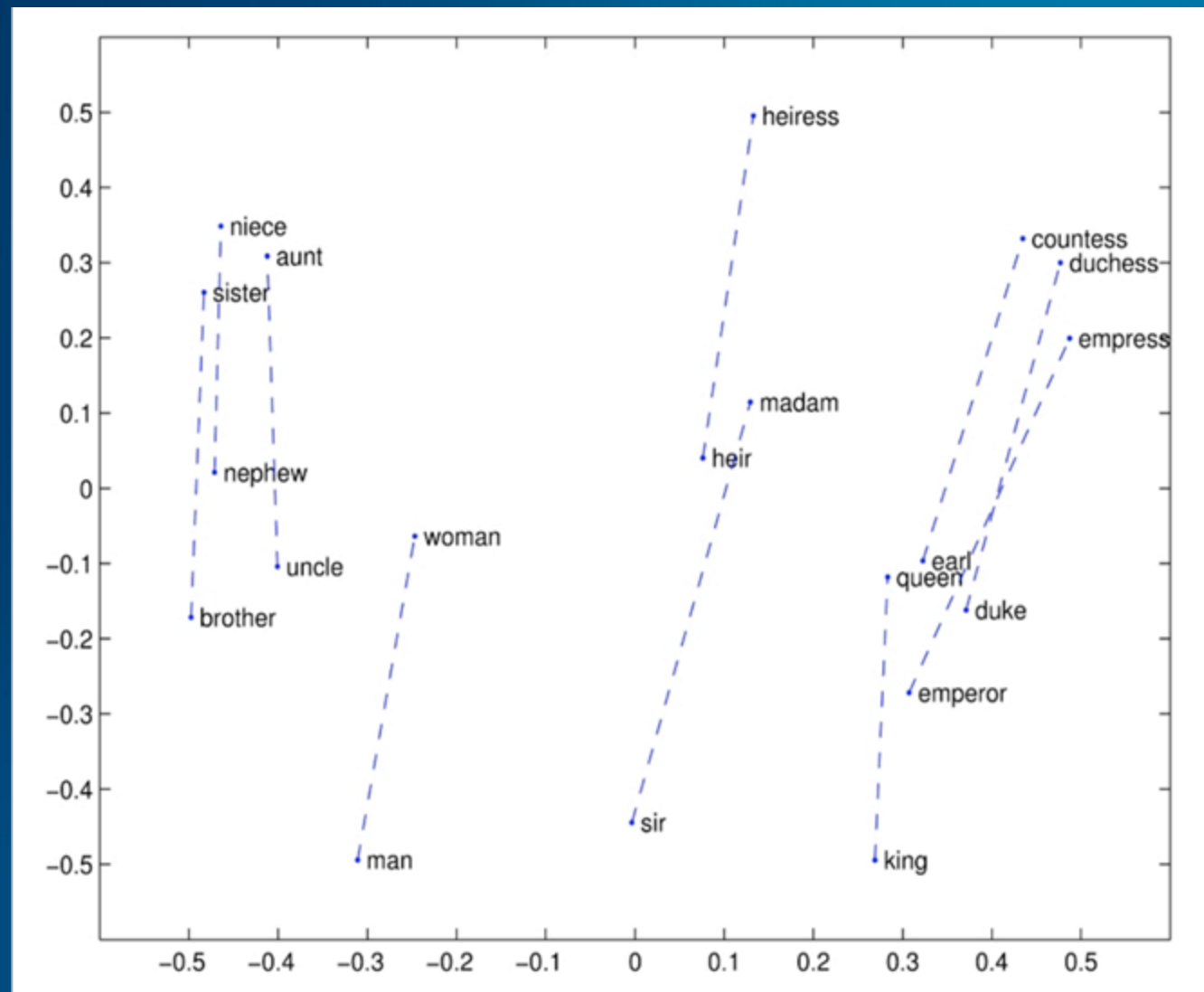
(虛線表示 陽性詞 vs 陰性詞)

學會了類推計算

king-man+woman = queen

所有的圖、文、聲等都是符號

也都可轉成向量統一處理



# AI 學習和理解 – 用語言模型作為類推應用

請輸入欲查找之詞彙

玉晶光



←問相關類股

3406

大立光

今國光

新鉅科

3406-tw

請輸入欲查找之詞彙

元大證



凱基證

日盛證

元富證

富邦證

合庫證

請輸入欲查找之詞彙

虧損



虧損額

大虧

巨虧

虧蝕

轉虧

累虧

小虧

鉅虧

淨虧損

←類推相似用語

張忠謀 - 台積電 + 阿里巴巴



馬雲

←問代表人物

張忠謀 - 台積電 + 大立光



林恩平

Proprietary and Confidential.

2-1

# 生成式 AI 的特性與長處

Features and Benefits of Generative AI

# 回到GPT，這是怎麼來的？

GPT就是大語言模型 (Large Language Model, LLM) 的代表之一  
先學了所有的開放圖書、再學社群聲量、再學維基百科和網路爬文

2018  
GPT-1

2019  
GPT-2

2020  
GPT-3

2022  
ChatGPT

2024  
GPT-4o

文字生成  
用大量書籍做訓練

加強文字生成，加入大量社群數據做訓練

加強文字生成，加入維基百科及巨量的網路爬文數據

可進行對話問答的生成模型

多語言、多模態、更便宜的生成模型

2-2

# 生成式 AI 重要的應用場域

## The Important Applications of Generative AI

# 生成式 AI 十大常見應用場域

是生產力工具，  
定位為 AI 小幫手  
讓辦公應用成為殺手應用

可接受多種答案的  
探索性、發想性、打草稿  
這類優先

有標準答案的要注意求證  
特別是訓練資料缺乏的

生成	1	生成一般文書，如故事、對話、劇本、詩詞等。
	2	生成商用文書，如信件、推薦信、簡報、報告。
	3	生成法律文書，如法律意見、合約、個案分析。
	4	生成摘要，如整理資料、懶人包。
	5	生成論說文或論文，如研究框架、方法、推論。
問答	6	回答知識題或閱讀題 (如解釋文章中隱喻)
	7	回答計算題 (注意有時不精準，需靠外部套件)
進階	8	生成類別標記，例如分類、評等、學測評分
	9	生成程式，任務拆解及自動化
	10	生成訓練資料，做為其他演算法模型的教練



---

3-1

# 生成式 AI 的發展與影響

## The Development of Generative AI and Options

# 生成式 AI 的發展

模型變成熟，選擇變多

雲端大型化 vs 地端特式化

封閉式 vs 開源生態系

整合圖、文、聲音、影像等多模態 (multi-modal)





# OpenAI的效果

台股走勢



## 近一週 [台股走勢] 的輿情重點：

- 台股24日小幅下跌，收在17287點，外資淨空單增加，而下週多家金控及航空股將舉行法人說明會，市場關注其營運展望。期貨法人提醒，若台股未能增加成交量上攻，短線可能面臨整理壓力。
- 富邦投顧董事長蕭乾祥預測，受明年總統大選、企業獲利回升、科技法案及通膨控制政策影響，美股及台股將呈正向發展，台股指數可能介於15000至18600點。他建議投資聚焦於創新科技、綠能、醫療保健等產業，同時指出美中貿易對抗和全球供應鏈重組將影響市場，預期美國聯準會利率將維持高位，限制經濟成長和企業獲利。
- 台灣專家預測，在GDP成長、庫存去化及企業獲利回升的情況下，2024年台股指數有望達到18,600點。另外，隨著全球供應鏈重組及美科技戰影響，半導體產業回流及分散化生產成為趨勢，且創新科技特別是AI的發展被看



近一週

近一月

近一季

近一年

OpenAI

Eland Llama2

請以繁體中文輸入關鍵字或帶有布林邏輯符號的關鍵詞組...



送出



# 地端開源模型調校後的效果

台股走勢



## 近一週 [台股走勢] 的輿情重點：

- 專家對於台灣經濟的預測和投資建議。根據專家的觀點，台灣的經濟成長有望回升，並且預計明年台股指數將超過18,600點。文章還提到了美國和中國的經濟趨勢，以及對投資者來說可以關注的產業方向。此外，文章還提到了美國聯準會可能在明年中啟動預防性降息，但利率仍可能保持相對高檔，對經濟成長和企業獲利造成影響。最後，文章提到了一個財經趨勢論壇，預計將探討台灣經濟前景和投資機會。
- 美股多數收高，台股下跌7點，外資淨空單增加。下週金控、航空和傳產指標股將舉行法人說明會，展望景氣和營運展望。期貨法人分析，若台股未能進一步補量上攻，短線不排除整理。三大法人外資及陸資買超、投信賣超、自營商賣超，合計買超17.09億元。外資買超開發金，賣超緯創。台股走勢震盪，權值股中權王台積電下跌，聯發科上漲，鴻海以平盤收盤。

近一週

近一月

近一季

近一年

OpenAI

Eland Llama2

請以繁體中文輸入關鍵字或帶有布林邏輯符號的關鍵詞組...



送出

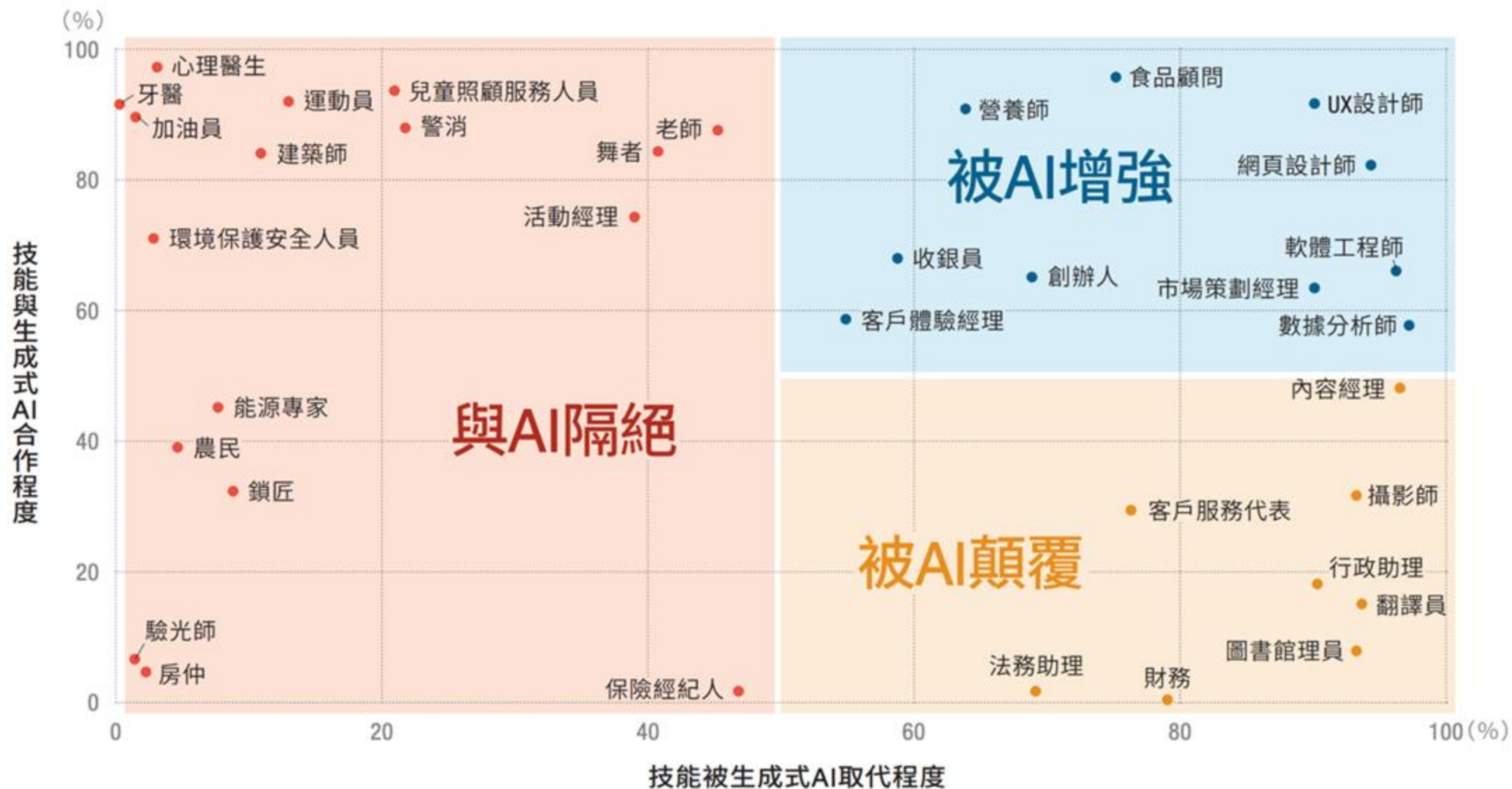
# 知識工作者導入生成式 AI 的效益明確



- AI 可分擔完成12.2%任務
- 讓25.1%任務加快；40%任務品質提升
- 讓績效低於平均的同仁提升43%；  
高於平均的同仁提升17%

Harvard Business School 與 BCG 顧問集團所做的研究  
Source: Navigating the Jagged Technological Frontier: Field Experimental  
Evidence of the Effects of AI on Knowledge Worker Productivity and Quality

圖表2 LinkedIn將工作分為3種不同與AI互動關係



出處：天下 x Cheers, 2024 台灣人才白皮書

參考資料：LinkedIn Economic Graphic Research Institute



---

3-2

生成式 AI 的未來

The Future of Generative AI

# 可見的未來-1

LLM不只生成更重理解  
成為任務自動化的代理人 Agent

網友將LLM扮演成軟體工程中的  
多個角色，彼此合作在短時間內  
生成遊戲

1. 理解需求
2. 撰寫程式
3. 執行及除錯
4. 自行改版
5. 成功發布

參考名詞：low code/no code  
RPA (robotic process automation)



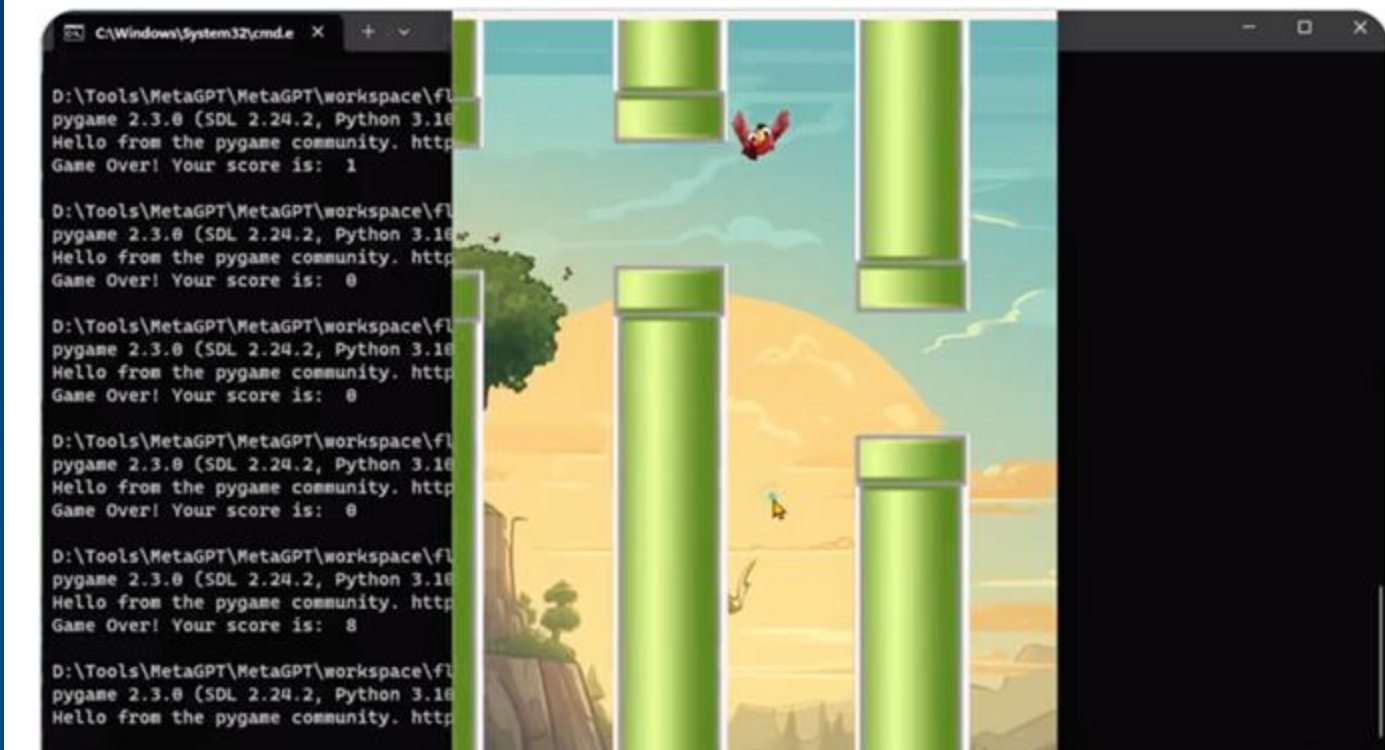
JoJo 99AI ✓  
@99aico

It took me 10 minutes to create this game using MetaGPT. I didn't watch any tutorial, and had no idea how to write the flappy bird game.

MetaGPT is so powerful that it might really replace a software company!

How I did it? please see thread 📖

[翻譯推文](#)



# 可見的未來-2

## AI PC、AI伺服器、AI車電 林百里看明年景氣 3項AI領域續旺

地端及Edge端  
特式化模型  
免連網快速回應  
成為人機介面之革命

### AI PC 供應鏈，一張圖看懂



科技界掀起「AI PC」熱潮，微軟直接稱2024年是「AI PC」年，到底什麼是AI PC？在AI PC供應鏈中，有誰已經卡好位了？

▶ 文章語音朗讀 · 04:14



# AI 時代下的職場協作

1. 了解 AI 的能力及長處，學習如何運用
2. 學習如何評估 AI 的產出品質和績效，如何改善
3. 學習如何帶領 AI 解決複雜問題，一同達成目標

交流討論

THANK YOU

wyang@ntu.edu.tw



follow on facebook

