書報討論課程之心得報告與延伸探討

<生成式 AI 技術解析與異質平台整合之反思>

國立勤益科技大學資工系系主任 智慧網路最佳化實驗室

楊振坤教授 講座教授

2025/11/11

壹、 前言

本報告聚焦於一場主題為「生成式人工智慧與異質平台整合系統」的專題演講，並嘗試將演講內容重新整理與深化。近年來，生成式 AI 的發展速度近乎翻頁般迅速，不論是模型的能力、應用範圍，或對產業運作的衝擊，都已成為科技領域討論的核心。本次演講從技術基礎談到最新的系統整合案例，也延伸到治理與倫理等較少被細談的面向，使我得以從更完整的角度理解這項技術的現況與限制。

貳、 生成式 AI 的技術演進與核心架構

若要真正看懂生成式 AI 的現況,就得先回到它的技術脈絡來思考。理解這些底層的發展，不只是為了分辨各種模型的差異，更能看出它們能力的邊界與形成原因。本章將從資料處理方式的演進談起，再逐步介紹深度生成模型的主要類型。之後，我會以目前仍相當具代表性的生成對抗網路（GAN）為例，說明它的核心概念與運作方式，作為理解後續技術的基礎。

2.1 資料處理的典範轉移

演講中提到，現代 AI 能走到今天，其實和電腦處理資料的方法改變有很大關係。這段歷史不是單純的技術更新，而更像是整個思維方式被翻新了一遍。

- 1950 年代初期-完全靠人寫規則的時代：那時的電腦什麼都得靠人一步步告訴它。所有可能的情況、應該採取的動作，都要先寫進程式裡。電腦不會自己找答案，只會照指令走。這種做法雖然清楚，但遇到稍微複雜的情境就很難處理。
- 1980 年代末至今-改用資料說話：後來資料量暴增，加上運算能力變強，做法開始出現根本性的變動。電腦不再只是執行規則，而是從一堆資料裡自己找出規律。統計與機率模型慢慢成了主角，像「語言是怎麼被理解的」、「什麼特徵有意義」這種問題，開始交給模型自己學。
- 這種「規則換成經驗」的做法，就是深度學習與生成式模型後來能起飛的原因。如果把這段變化拉在一起看，可以發現 AI 的發展並不是突然冒出來，而是整個資料處理方式改寫後，長期累積的結果。

2.2 深度生成模型（Deep Generative Models）概覽

演講中提到的模型分類圖，把深度生成模型的整體樣貌整理得相當清楚。這類模型的核心想法都圍繞在同一件事上：理解資料本身的分佈，然後依據這種理解去產生新的內容。多數生成模型的分類方式，都和「最大概似估計」這種常見的學習方法有關。若按照它們處理機率密度的方式來看，大致可以分成兩條主要的技術路線：

- 顯性密度（Explicit Density）
  這一類模型會直接給出一個機率密度函數，因此能算出某個資料點出現的機率。不過它們的做法又可以再細分：
  ✓ 易處理密度（Tractable Density）：模型本身就能直接算密度，步驟相對明確。
- 隱性密度（Implicit Density）
  這類模型沒有直接寫出密度函數，而是用某種「能夠抽樣」的過程，讓模型從資料裡抓住分佈的形狀。
  ✓ 直接生成（Direct）：生成對抗網路（GAN）屬於這一類。
  ✓ 馬可夫鏈生成（Markov Chain）：例如 GSN，透過一連串的轉移過程來生成樣本。

2.3 案例解析：生成對抗網路（GAN）

GAN 是理解生成式 AI 運作原理的絕佳案例。其核心機制是一種巧妙的「對抗性學習」，由兩個相互競爭的神經網路組成：

- 生成器（Generator）：其任務是從一組隨機雜訊（noise）中，生成看起來足以以假亂真的資料（例如：偽鈔）。它的目標是盡可能地欺騙判別器。
- 判別器（Discriminator）：其任務是判斷輸入的資料是真實的（例如：真鈔）還是由生成器偽造的。它的目標是盡可能地準確分辨真偽。

這兩個網路在訓練過程中不斷進行博弈：生成器努力提升偽造技巧，而判別器則努力提升辨識能力。最終，當判別器再也無法輕易分辨真偽時，代表生成器已經學會了如何產生高度逼真的資料。

根據演講總結，GAN 的兩大主要優勢在於：

1. 針對已有部份資料來產生不存在的資料：例如，在資料集不完整時，可以用來填補缺失的部分。
2. 擴充原有功能：例如，將低解析度圖片提升為高解析度，或為黑白照片上色。

本章節從資料處理的歷史演進，到深度生成模型的分類，再到 GAN 的具體運作，勾勒出 AI 技術的基礎樣貌。然而，驅動這些先進模型實現驚人能力的背後，是令人咋舌的龐大資源投入。

參、 現代 AI 的驅動力：成本與算力的軍備競賽

演講中關於訓練成本的揭露，深刻地指出了當前尖端 AI 技術發展的另一面：這不僅是一場演算法的突破，更是一場由巨量資本與頂級硬體所支撐的資源投入競賽。AI 模型能力的躍進，與其背後訓練成本的指數級增長密不可分。本章節將量化分析 GPT 模型的演進成本，並探討其背後由高階運算單元（GPU）所引領的全球算力佈局。

3.1 GPT 模型的演進與代價

透過整合演講中的「GPT Evolution」與「Training Cost」兩張圖表，我們可以建構一個更完整的發展時間軸，清晰地看到從 GPT-1 到 GPT-5 的演進軌跡及其伴隨的巨大資源消耗。

模型版本	發布/啟用時間	關鍵發展里程碑	訓練資源與時間成本
GPT-1	2018/06/11	使用簡單的網路技術，以 7000 本文學著作為訓練基礎。	使用 8 個 P600 運算單元，訓練 30 天。
GPT-2	2019/11/15	以 GPT-1 的架構調整為適合口語的資料判讀。	未知。
GPT-3	2020/05/28	以 GPT-2 的架構為基礎，調整為適合更大範圍的資料。	估計使用 1000 張 A100 運算單元，訓練 30 天 + 調校 60 天。
GPT-3.5	2022/03/15	可以生成文本，具備對話能力。	未提供。
ChatGPT	2022 年底	ChatGPT 橫空出世，引爆全球 AI 熱潮。	未提供。
GPT-4	2023/04/14	接受以圖片作為資訊輸入的管道，具備多模態能力。	使用 8192 個 H100 運算單元，訓練 90-100 天。
GPT-5	2025/08/07 (預計)	最佳化各項細節，包含幻覺降低、擬真回覆、深度分析等。	OpenAI 預計需要 5 萬張 H100 進行訓練。

從上表可見，從 GPT-1 到 GPT-5，模型不僅在功能上（從純文本到多模態、從生成到對話）實現了飛躍，其訓練所需的算力與時間成本也呈現指數級的增長。這清晰地反映了模型複雜度與能力提升背後的巨大代價。

3.2 AI 訓練的硬體核心：A100 與 H100

這場算力競賽的核心，是 NVIDIA 等公司推出的高階 AI 晶片。演講資料揭示了其驚人的價格與效能：

- NVIDIA A100：單價高達 60 萬 新台幣。
- NVIDIA H100：單價至少 100 萬 新台幣。其效能提升極為顯著，演講中提到一個驚人的案例：「若使用 3584 張 H100 訓練 GPT-3，僅需 11 分鐘」。

這種硬體的迭代不僅是價格的提升，更是對訓練效率的顛覆性改變。因此，掌握這些尖端硬體成為了科技巨頭在 AI 領域競爭的關鍵。目前全球的算力佈局高度集中：

- Microsoft Azure：至少擁有 5 萬張 H100。
- OpenAI：為訓練 GPT-5，需要 5 萬張 H100。
- Google：手上大概有 3 萬張 H100。
- Meta：則是 2.5 萬張 H100。
- Inflection：需求 2.2 萬張 H100。
- Oracle：大概有 2 萬張左右的 H100。
- 其他：特斯拉、亞馬遜也至少各有 1 萬張左右。

這些數字凸顯了當前全球 AI 發展的算力集中化現象，只有少數頂級企業能夠負擔得起參與這場競賽的「入場券」。

高昂的成本與資源門檻已成為定義當代 AI 發展的重要特徵。正是基於如此巨大的投入，生成式 AI 才得以實現多元化的應用，但同時也引發了一系列新的問題與挑戰。

肆、 生成式 AI 的應用、挑戰與侷限

在龐大資源的支持下，生成式 AI 已在各個領域展現出驚人的應用潛力，從創意內容的生成到企業流程的優化。然而，隨著應用的深入，其內在的技術挑戰與潛在風險也日益浮現。本章節將探討其多元應用場景，並深入剖析其在可靠性、 logique 性及倫理層面所面臨的關鍵問題。

4.1 應用光譜：從創意生成到企業整合

生成式 AI 的應用範圍極其廣泛，演講中提及了以下幾個代表性方向：

- 創意生成：AI 已具備生成影片、製作歌曲、設計圖片等能力，為內容創作領域帶來了新的可能性。
- 企業流程整合：AI 不僅是創意的工具，更是提升企業營運效率的核心引擎。以生產排程為例，演講對比了傳統與 AI 賦能後的流程：
  - ✓ 現況（As-Is）：生管人員需手動開啟排班系統，一步步找出可行的完工時間、物料準備狀態、換線時間等，過程繁瑣且耗時。
  - ✓ 未來（To-Be）：生管人員僅需向具備 GAI 的先進規劃與排程系統（APS）下達指令，APS 便能主動介入協調工單完成時間、物料狀態、換線時間等，大幅提升反應速度與準確性。
- 智慧化辦公：以 Microsoft 365 Copilot 為例，其架構清晰地展示了 GAI 如何與日常辦公軟體深度整合。Copilot 將三大核心要素串聯起來：
  1. 大型語言模型（LLM）：提供強大的自然語言理解與生成能力。
  2. Microsoft Graph：存取使用者個人的數據，如電子郵件（emails）、檔案（files）、會議（meetings）、日曆（calendar）等。
  3. Microsoft 365 Apps：將 AI 能力嵌入 Word、Excel、PowerPoint 等應用程式中。 透過此架構，Copilot 能夠理解使用者的情境與需求，提供高度個人化的智慧協助。

4.2 可靠性與邏輯性的挑戰

儘管應用前景廣闊，但生成式 AI 的輸出並非永遠可靠。演講中列舉了幾個關鍵挑戰：

- 邏輯謬誤：AI 生成的內容可能違背基本常識。例如，演講中展示了一張同時出現盛開的櫻花與厚重的雪景的圖片，這種「不合邏輯的生成結果」顯示 AI 在理解物理世界與時序關係上仍有缺陷。
- 幻覺（Hallucination）現象：這是目前大型語言模型最嚴重的問題之一。在「收集相關文獻」的案例中，當研究人員要求 AI 列出主題相關且引用數最高的論文時，AI 雖然提供了一份看似專業的清單，但經過查證後發現：
  - ✓ AI 聲稱某篇論文「引用數 40」，但實際查核後發現資訊不實。
  - ✓ 另一篇聲稱「引用數 35」的論文，同樣存在引用數錯誤的問題。 這種憑空捏造或提供不實資訊的行為，嚴重影響了 AI 作為可靠資訊來源的可信度，是當前技術上的一大缺陷。

4.3 不同 GAI 服務的特性比較

市場上不同的 GAI 服務在設計理念與能力上各有側重。演講中比較了 GPTs 與 Gemini 兩大主流服務的核心差異：

特性	GPTs	Gemini
處理模糊問題	偏好於生成該問題可能的虛擬內容。	僅生成信心值較高的答案；若信心值不足或超出範圍，則不會生成結果。
付費方案	可 Fine-tune；可存取所有 GPT 模型；128,000 tokens。	可 Fine-tune；可使用所有 Google Assistant 和 2TB Drive；可以任何 Android 設備存取；1,000,000 tokens。
推論效能	常識推論 (Commonsense Reasoning) 高達 95.3%，優於 Gemini 的 87.8%。	大規模多任務語言理解 (MMLU) 能力為 90%，優於 GPT 的 86.4%。
網路存取範圍	可存取 Internet 資源，但不確定其存取範圍。	可以正確地存取所有 Google 資源，而 Internet 資源也可存取，但不確定結果的正確性。

此比較顯示，不同的模型在推論策略、生態整合與處理不確定性方面採取了不同的路徑，使用者需根據具體需求選擇合適的工具。

總結而言，儘管 AI 的應用前景令人振奮，但其內在的不可靠性與潛在風險，促使我們必須嚴肅地思考其背後的倫理與治理問題，這也為下一章節的個人反思提供了重要的切入點。

伍、 學習心得與批判性反思

在系統性梳理了演講內容後，本章節將從一名科技管理研究生的視角，提出個人的綜合性心得與批判性反思。生成式 AI 的發展不僅是純粹的技術議題，更涉及商業策略、資源配置與社會倫理的複雜交織。我將探討生成式 AI 的雙面性、成功導入的關鍵要素，以及其所引發的深層次倫理議題。

5.1 AI 的雙面刃：缺點與潛在切入點

演講中的「Recall」圖表，巧妙地將生成式 AI 的缺點與其對應的潛在機會並陳，揭示了其「雙面刃」的特性。我認為，理解此一對應關係，是找到 AI 最佳應用場景的關鍵。

缺點 (Flaws)	潛在的切入點 (Potential Entry Points)
不是很準確	能夠有些容錯的空間
需要高品質資料	應用於延伸或擴展的情境
需要有人輔助	讓危險的工作由 AI 代替
無法在意外發生時負責	與專家負責人員協同合作

基於上述分析，我的觀點是，與其追求一個全知全能、零錯誤的通用 AI，現階段更務實的策略是「化缺點為機會」。例如，正因為 AI 不夠準確，它最適合應用於那些允許犯錯、且後果可控的創意發想或初步草案生成場景，而非高風險的決策領域。正因為它無法負責，它就應該被定位為「專家的得力助手」，在人類的監督下執行任務，而非完全取代人類。這種認知上的轉變，有助於我們在實際應用中，為 AI 找到最能發揮其價值的定位。

5.2 成功導入 GAI 應用的三大基石

演講中的 Venn 圖揭示了成功建構一個 GAI 應用的三大核心要素，我將其歸納為「三大基石」。這三者相輔相成，缺一不可，共同構成了企業導入 AI 應用的核心護城河。

1. 訓練平台 (Training Platform)：這不僅指涉 Google Colab、Azure、AWS 等雲端平台，更包含了高效能的軟硬體整合能力，以及針對特定任務設計有限時間的執行演算法。這是 AI 應用的基礎設施。
2. 訓練資料 (Training Data)：資料是 AI 的燃料。成功的關鍵在於取得「有用」的資料，即標註了特定領域知識（domain knowledge）的高品質資料。沒有好的資料，再強大的平台也無法產出有價值的模型。
3. 調校知識 (Tuning Knowledge)：這是最容易被忽視，卻也最關鍵的一環。它包含了對領域知識相當豐富的專業知識與經驗，以及對 GAI 執行細節的精準掌握。唯有具備深厚專業的團隊，才能將平台與資料的潛力最大化，進行有效的模型調校與優化。

這三大基石提醒我們，導入 AI 並非單純的技術採購，而是一項涉及技術、資料與人才的系統性工程。

5.3 倫理與治理的深層思考

演講最後提出了一個極具挑戰性的情境問題：「當生成式 AI 所產生的決策發生醫療糾紛時，該如何解決？」這個問題直指 AI 發展的核心困境——責任歸屬。

對此，我認為解決方案必須是多層次的：

- 責任歸屬的釐清：在法律與規範層面，必須建立清晰的責任鏈條。責任應由誰承擔？是 AI 開發者、模型訓練者、醫療機構，還是最終下達指令的醫生？這需要跨領域的專家共同探討，制定新的法律框架。
- 演算法透明度與可解釋性 (Explainable AI, XAI)：當 AI 做出錯誤決策時，我們必須有能力追溯其決策過程。黑箱模型在醫療、金融等高風險領域是不可接受的。推動可解釋性 AI 的研究與應用，是建立信任的基礎。
- 健全的監管機制：政府與行業協會應設立監管機構，對高風險 AI 應用進行審核、認證與持續監督，確保其符合倫理與安全標準。

演講中的「Trustable Service」圖表，列出了建構「可信賴 AI」的多個關鍵要素，包括**公平性（Fairness）、隱私（Privacy）、安全（Security）、倫理（Ethics）**等。這再次強調，建構可信賴的 AI 不僅是一項技術挑戰，更是一項需要全社會共同參與的社會性責任。

總結而言，我認為技術的進步最終是為了服務人類。因此，在我們驚嘆於生成式 AI 強大能力的同時，更應將人文關懷與倫理框架置於其發展的核心，確保技術的航向始終與人類的長遠福祉保持一致。

陸、 結論

本次「生成式人工智慧與異質平台整合系統」的專題演講，提供了一場兼具廣度與深度的知識饗宴。從技術的歷史演進、核心模型的運作原理，到驅動其發展的巨大成本、多元的應用場景，再到其所面臨的可靠性挑戰與深刻的倫理困境，演講勾勒出了一幅生成式 AI 發展的全景圖。

總結而言，本次演講的核心價值在於揭示了生成式 AI 不僅是一次單純的技術革新。其背後高昂的算力成本，決定了這是一場由少數科技巨頭主導的競賽；其應用的廣泛性與內在的不可靠性，使其成為一柄需要謹慎使用的雙面刃；而其對社會決策與責任歸屬帶來的衝擊，更使其成為一個亟需跨學科對話的社會性議題。

作為一名資訊領域的研究生，這次演講加深了我對「技術始終鑲嵌於社會」的理解。未來，我們不僅需要關注演算法的精進，更需要研究如何建構一個可信賴、負責任且具備倫理框架的 AI 生態系統。展望未來，生成式 AI 與人類社會的協同演化將是一個充滿機遇與挑戰的動態過程，值得我們持續關注與深入研究。