依107年6月22日輔綜字第1070050960號函修訂

出國報告（出國類別：開會）

2025年美國波士頓麻省理工學院電腦暨人工智慧實驗室(MIT CSAIL)年會

服務機關：臺中榮民總醫院人工智慧科
姓名職稱：李佳霖科主任
出國期間：2025年04月27日至2025年05月02日
報告日期：2025年05月05日

摘要

本次參與麻省理工學院電腦與人工智慧實驗室（MIT CSAIL）2025 Alliances 年會，主要目的是掌握 AI 最新技術於臨床醫療應用的趨勢，並探索其在台灣實務落地的可行性。期間深入了解多項關鍵技術，包括 DeepSeek 的低成本高效能訓練策略、Mantis 的語意導向多模態資料整合平台，以及 Liquid AI 所提出之液態神經元架構，具備處理時間動態與低功耗應用的潛力。

會中亦與駐點 MIT 的緯創、台達電工程師交流，探討以人為橋樑引介國際技術的新策略。整體而言，本次出訪除拓展國際視野，也促進對生成式 AI、視覺語言模型（VLM）與混合式運算架構的系統性認識。

本報告建議：一、重新思考 LLM 在臨床決策中的角色；二、鬆綁研究經費制度；三、同步部署高階與中階 GPU 建構 AI 算力基礎；四、邀請 MIT 歸國工程師演講分享。唯有以 startup 式思維與行動，在體制內創造 AI 新創，方能掌握醫療 AI 發展的主動權。

關鍵字：麻省理工學院(MIT)、電腦暨人工智慧實驗室(CSAIL)、生成式人工智慧、Mantis、跨領域創新、液態神經網路（Liquid AI）

目次

一、 目的

1. 了解AI前沿技術於醫療應用之現況與挑戰，特別針對生成式AI平台的於醫院應用之可行性。
2. 拓展跨領域合作關係，與MIT CSAIL重點研究人員進行討論交流，探討技術應用的可行性與落地策略。
3. 評估中榮AI部署之可行策略方向，主要包含「成熟技術部署」及「醫療/行政人員需求導向應用」。

二、 過程

本次參訪與趙文震主任結伴參訪，行程主要包括幾項重要內容：

(一) DeepSeek 模型技術

在第一天，我們特別去聽了 Yoon Kim 教授主講的「DeepSeek Deciphered」並探討其對大型語言模型（LLM）技術生態的影響。LLM 的訓練過程分為兩大階段：

1. **預訓練階段（Pre-training）**：透過大量語料進行自監督學習，學習語言結構與知識關聯性。
2. **後訓練階段（Post-training）**：透過人類回饋（RLHF）與指令微調進行模型對齊，提升安全性與應用可控性。

DeepSeek 技術創新亮點

1. **稀疏專家架構（Sparse Mixture-of-Experts, MoE）**
   - 每層由多位「專家」組成，但僅啟用少數幾位處理每筆輸入
2. **低精度訓練（8-bit Low-Precision Training）**
   - 採用 8-bit 精度訓練替代傳統 16-bit FP，顯著降低記憶體與能源消耗。
3. **多項訓練效率優化策略**
   - **專家負載平衡（Expert Load Balancing）**：避免特定專家被過度使用，提升模型泛化穩定性。
   - **多潛變注意力（Multi-Latent Attention）與多 token 預測**：提升預訓練效率與下游任務適應力。

訓練成本與性能表現

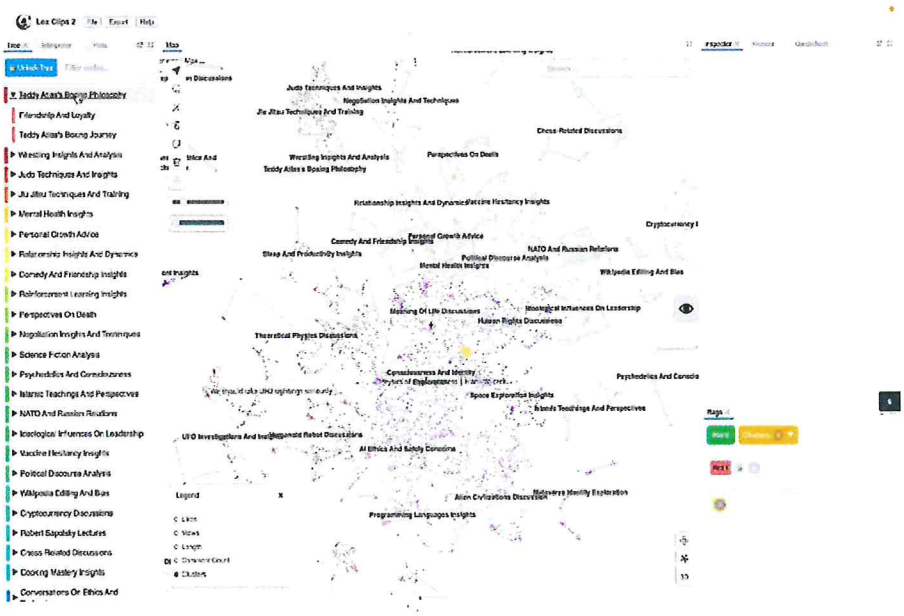
成本比較（每百萬 tokens 推論成本）：GPT-4o 需要 10 美金，但 DeepSeek-v3 只需要 0.28 美元，**雖然 deep seek 在訓練資效率有了重要進步，但** Yoon Kim 教授不斷強調這並**不代表** deep seek 有了 breakthrough 的發展，這只是在速度的突破，對 MIT 這類追求創新的學者來說，他們追求的是全新架構的突破，比方說取代 transformer，這點在後續的 Liquid AI 也是如此，整體來說 DeepSeek 的成功展示了在低成本條件下訓練高效能 LLM 的可行性，但仍需面對合規與倫理挑戰，其架構與策略可作為全球 AI 開發者的重要參考典範。

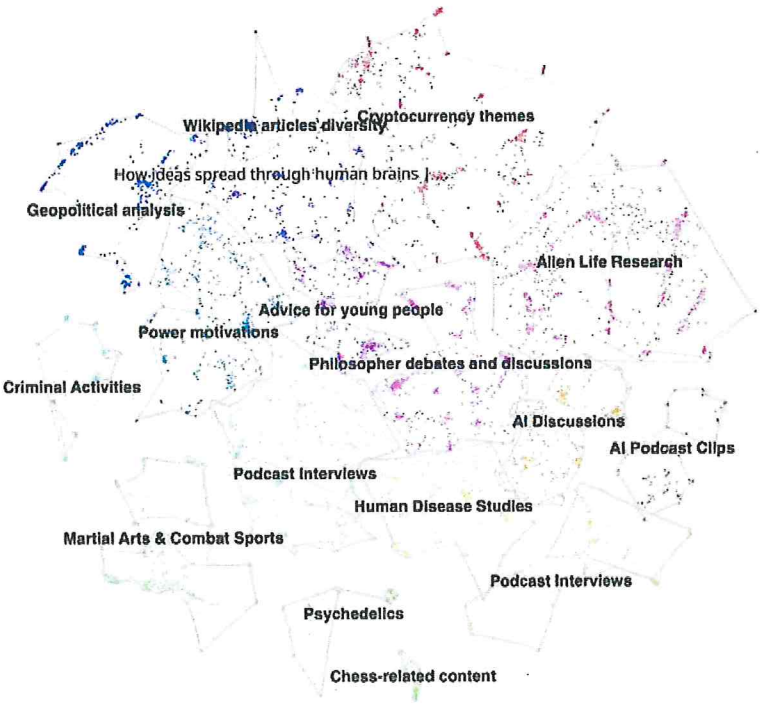
(二) Mantis 模型技術

Mantis 是 Manolis Kellis 團隊最新研發的技術平台，在演講開始前，Manolis Kellis 就在布置會場，我們剛好偶遇，教授一如往常的拿起了手機要自拍，同時回憶起了一起吃龍蝦的時光（但印象中我真的覺得台灣龍蝦比波士頓好吃…），教授想起兩年前吃龍蝦的時光後，立刻跟我說他們這幾年又多了很多的成果，後來確認應該是 Mantis 的 platform。

Mantis 的 platform 實在很難解釋，當場聽我只是模糊的知道這是一種 latent space cluster

的視覺化技術，好在後來他的博士生有 present poster，在他們的同意下，我要到了相關的 slide，原始檔案很大，我擷取幾張圖片當範例：





坦白說，大神的 slide 的確是直接看是看不懂的，因此我試著理解 Mantis 的核心，其出發點主要是現代資料分析面臨日益複雜的結構，尤其在醫療、科研與產業應用中，資料特徵具有**多模態（Multi-modal）**：資料來源包含結構化數值、自由文字、影像、時序等類型

- **高維度（High-dimensional）**：單一病患或樣本可包含上百種變項
- **異質性（Heterogeneous）**：來自不同資料庫、設備、時間點與觀察尺度的資料難以整合等特質，傳統分析工具（如統計平台或單一視覺化工具）往往無法同時兼顧有效整合異質性資料、建立直覺式探索與語意導向的使用者界面和支援即時、協作式的人機互動決策流程。

因此 Manolis Kellis 團隊提出了一個解法，所有的資料都可以投影到高維的 latent space,而只要到 latent spalce，所有的資料就有可能被對齊、整合、甚至降維，因此 **MANTIS** 是一個 AI 驅動核心概念為「**認知地圖 Cognitive Cartography**」。使用者可以在資料的高維度空間中**進行導航、探索與推論**，就像在思維地圖中定位每筆資料所屬的意義與潛在脈絡。其中用到的技術架構很多，大致上有：

1. **語意嵌入與階層式聚類**
   - 使用 **UMAP** 或 **t-SNE** 將高維資料投影至可視化語意空間
   - 搭配**階層式聚類**（如 K-means, HDBSCAN）形成語意區塊
   - 支援動態語意地圖，讓使用者以視覺化方式理解資料分布與內在結構
2. **語意搜尋與互動式篩選**
   - 搜尋結果直接高亮於地圖中，便於即時比對與聚焦
3. **多視角探索模式**
   - 使用者可從不同面向（病人子群、變項空間、時間軸）觀察資料特性
4. **多模態整合**
   - 將**文字、影像、數值、時間序列等異質資料**嵌入共同語意潛在空間
   - 允許跨模態推論與相似性比對（如病例 × 醫囑 × 化驗結果）
5. **圖神經網路與 RAG 架構**
   - 將資料點間的邏輯/時間關聯建構成 **graph 結構，並**搭配 RAG（Retrieval-Augmented

Generation），實現以類比資料為基礎的自然語言敘述與查詢補強

6. **Human-in-the-loop + RL 強化互動學習**

- 使用者互動（篩選、標記、指派任務）將反饋至模型訓練流程

會後我有向 Manolis Kellis 教授提到隱私與付費的問題，此 platform 是地端運行，但需要他們幫忙佈署客戶端的 server，同時這是需要付費的，但價格尚未訂出來。

(三) Liquid AI 模型技術

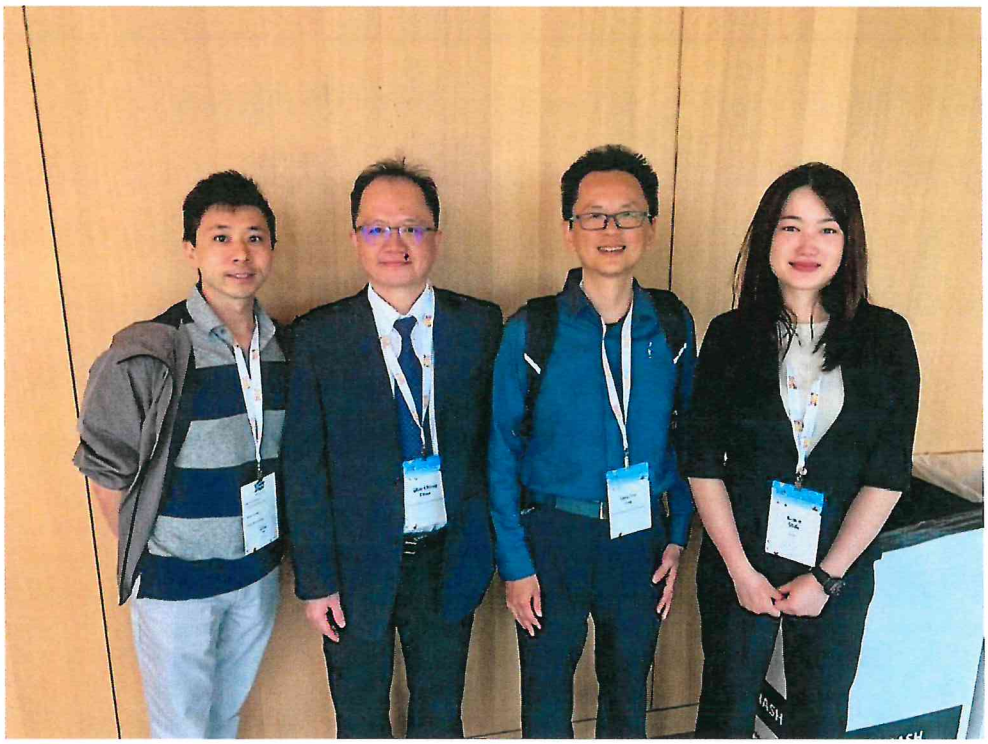
Daniela Rus 教授是 MIT CSAIL 的主任，上次來 CSAIL 時，她幫我們上了一堂空間幾何學，因為聽得很頭痛，所以印象很深刻，但時那時候我有跟她提出個人化營養的 project,因為她也是個好媽媽，要幫小孩準備午餐，所以她還記得我，但她對台灣人最為人知的應該是陳文茜的節目曾數次訪問，Liquid AI 在台灣的訪談節目已經數次被提及，因此這次要來之前已經鎖定是要優先理解的技術

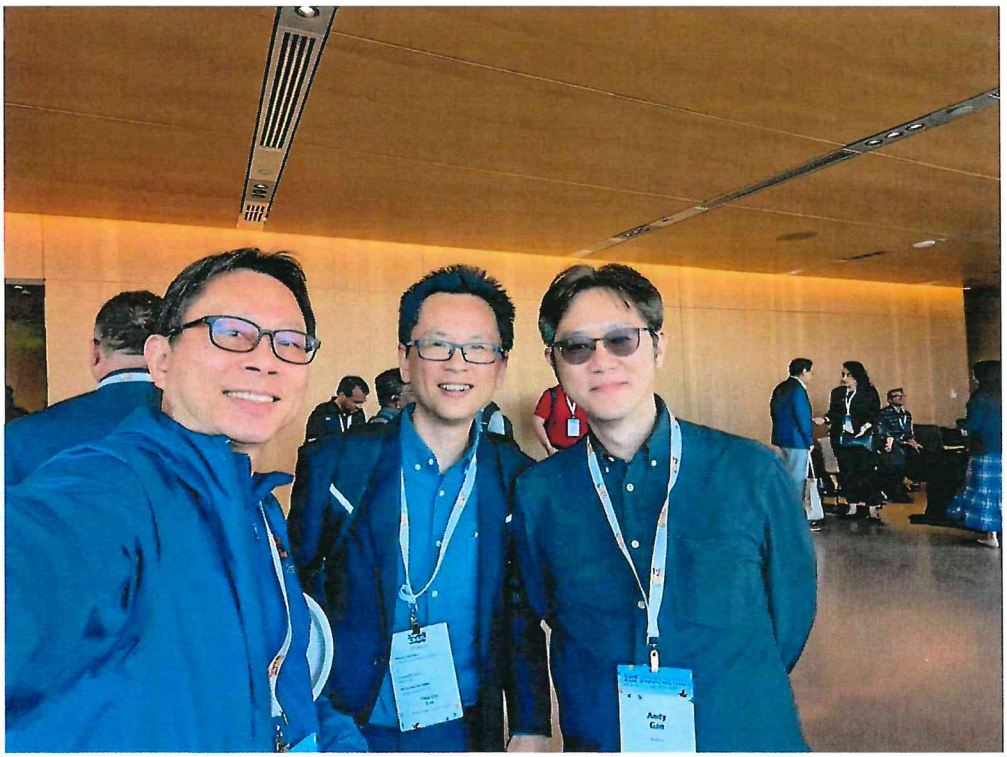
Daniela Rus 教授一開始引用秀麗隱桿線蟲（C. elegans）這個僅有 302 個神經元的模式生物，引出「簡約神經網絡仍可產生複雜行為」的概念。這為現代深度學習模型的精簡提供生物啟發式的支持，傳統模型壓縮策略：剪枝（Pruning）」與「量化（Quantization）雖然能透

過刪減不必要的神經元與降低精度來大幅減少模型參數與運算成本，但這部分需要人力去手動刪減參數，對大規模佈署仍不見得實用，而 Daniela Rus 教授從線蟲獲取靈感，(順帶一提，她上次的花朵型機器手臂靈感來自於聖誕禮物…)，線蟲僅由 302 個神經元構成，但能展現複雜的行為。因此，他們提出以少量神經元模擬動態決策過程，並用動態微分方程取代靜態激活函數，使模型能更自然地捕捉時間序列與因果變化。因為傳統神經元使用權重加總後經激活函數（0 或 1）輸出結果。Liquid 神經元則是基於時間動態的微分方程，所以能更好地處理時間變化的輸入訊號和真實世界的動態任務。實驗的結果，Liquid Network 相對 Transformer 提升 123% 表現（square error 減少），圖像處理：+46%，長距離軌跡重建：+96%，且記憶體使用量大幅減少，在處理 1M token 時仍可在 16GB 內運行，然而 Liquid 基於時間動態的微分方程，因此，此模型對靜態的處理不如 transformer based, 然而高效、低功耗、高度泛化、在未曾見過的環境條件下仍能穩定運作等種種特性，讓 Liquid AI 有成為下一代 foundation model 的潛力。

(四) 與台灣科技廠的交流

在經過艱難的技術洗禮後，終於有一些輕鬆的交流場合，晚宴的時候遇到台灣緯創團隊外派的工程師團隊，他們負責駐點在 MIT 一年，也認識了台達電的研發長(前任 HTC 副總裁與 google 地區總監)，交流了相同技術但在不同場域的應用，巧合的是，緯創外派團隊恰好在 Dina Katabi 教授以及 Polina Golland 教授 Lab，這兩位教授都是上次有幫我們上課的教授，Dina Katabi 教授主要是用 Wifi 波進行 vital sign detection and parkinsonism prediction，Polina Golland 教授則是長期研究醫學影像，這次雖然沒有見到兩位教授，但得知緯創 8 月份會邀請她們來台演講。




Andy
Gan

三、 心得

(一) 這是一個 medical AI 的黃金時代，特別是對台灣醫界，從 Deep- seek, Liquid AI 等發展，顯示 AI 發展往節能省電、地端（邊緣）佈署、創新發展的方向發展，因此 GPU 的軍部競賽雖然仍存在，但應用端就可以靠想像力來突圍，這對原本我們是硬體弱勢但 domain data 優勢的我們，多加了好幾分。然而我個人的看法是，AI 在發展上存在陷阱，雖然未來百工百業需要生成式 AI，但他的重點在於" reshape future" ，而不是僅僅加速效率的考量，AI 重點不是在 release 人類雙手，而是讓人類的大腦更有時間去創造。AI 的發展不是要讓人類更輕鬆，而是要讓人類更聰明。AI 應該要用來創造價值，而不是停在減少勞動力，如同當年的瓦特發明蒸汽機，線性的想法是用火車取代馬車，然而指數的想法是利用蒸氣機導入紡織業，因而引其工業革命，現在的 AI 如同當年的蒸汽機，我們不該只是讓 AI 做人類本來做的事而已，應該要利用想像力，讓 AI 做原本人類做不好的事(比方說在醫療上，是不是 AI 能比人更早看出疾病風險潛勢)，人類應該是雙手放開悠閒地喝下午茶，然後發揮想像力，不斷想出新奇困難的任務讓 AI 去解決，這樣才能 reshape future。雖然在短期內，AI 勢必仍有助於提升效率，但長期來看，真正的價值創造將來自它如何擴張我們的認知疆界，以下這段話是 AI 總結我這段話的範例，證明了 AI 的作文完全比我好： ***AI 的真正意義，不是讓人類脫離勞動，而是釋放出我們更大的思考能量；如同蒸汽機未必是更好的馬，而是整個文明的轉折點。未來的 AI 應該是智識的引擎，而不是勞動的替身***。

(二) 這個世界的連結：在當初 google 翻譯使用 RNN 取得巨大進步的時候，有人用世界的巴比倫塔倒塌了來形容，因為世界的語言終將被統一，後來 LLM 盛行了，這個說法被改成 LLM 串聯起全世界，因為人類跟機器人也能經過 LLM 來溝通，但其實這個說法並不精確，我覺得這世界連結方式最後只有一種，就是高維度的向量(vector)空間，而 LLM 只是這種高維度向量空間降維成語言方式表現，讓人類可以理解 vector 的流動，而 Mantis，則是降維成圖形，讓人類可以理解，本質並無差別，差別的只是降維的方式，以下仍然引用 AI 的總結：***如果我們把 AI 看成住在***

高維宇宙的外星人，向量是 AI 外星語的母語，LLM 是翻譯器，Mantis 是地圖。

(三) AI 的變動太快了，幾乎是用每星期在變動來計算，因此我們應該考用 **think like startup company** instead of hospital management, **move like pirate** instead of Navy. 這樣才能有機會跟上每個 AI 潮流的轉折.我們需要在體制內劃出 AI sandbox，讓少數團隊可以『像 新創』思考，『像 海盜』行動，否則我們會錯過每一波浪頭。

四、 建議事項

(一) 應全面重新思考 LLM 在臨床實務中的角色。相較於僅用於撰寫 weekly summary 或語音轉文字，LLM 的應用應進一步升級為協助整合結構化與非結構化資料（如病歷、檢查報告與影像），進而提升疾病進展預測的準確性，並於臨床介入決策中，提出符合醫療實證的第二選擇。我們科別也將根據此策略，加速部署 LLM 與 VLM 模型，整合進臨床決策支持系統（CDSS），提升病患照護品質。。

(二) 為加速臨床新創轉型，建議本院可考慮調整現行研究部的經費運用彈性。以目前「月支／年支型研究獎勵金制度」為例，該制度設計良好，對於有科技部計畫的主持人，提供額外獎金以供研究，同時也能支付主持人參與短期國際會議，但在實務執行上若能進一步放寬用途限制，不僅限於短期國際會議，只要主持人研究帳戶經費許可，也能支持主持人赴海外進行中長期進修（如臨床 AI 計畫、國際合作實驗室等），此舉亦可補足目前出國進修獎勵金額不足的問題，特別是在全球通膨、旅費攀升的當下，更具政策急迫性。

(三) 在未來 1－2 年內，本院需持續加碼 GPU 投資，並採取高低搭配的策略配置：一方面建議及早佈建高端 GPU（如 H200），以因應大型模型訓練需求；另一方面，也應預先布局中階 GPU（如 RTX 5090 或 A6000），分散於多節點本地端伺服器。此類資源不僅可加速模型訓練與 prompt 測試，也有助於建立靈活的 hybrid AI 平台架構，降低對雲端計算的依賴，並因應多樣化的臨床 AI 工作負載需求。

(四) 因為與 MIT 一線教授直接建立研究合作的門檻較高，建議可考慮另一條策略路徑：與曾在 MIT 實驗室工作的科技業工程師(比方說像緯創工程師)建立溝通管道。這些工程師在美期間曾參與頂尖 AI 或醫療技術開發，他們返國後，邀請他們至本院演

講，請他們基於其所學技術，建議本院合適開發的題目，可開拓本院的視野。

五、 附錄

略 (相關圖片皆檢附於上述報告)。