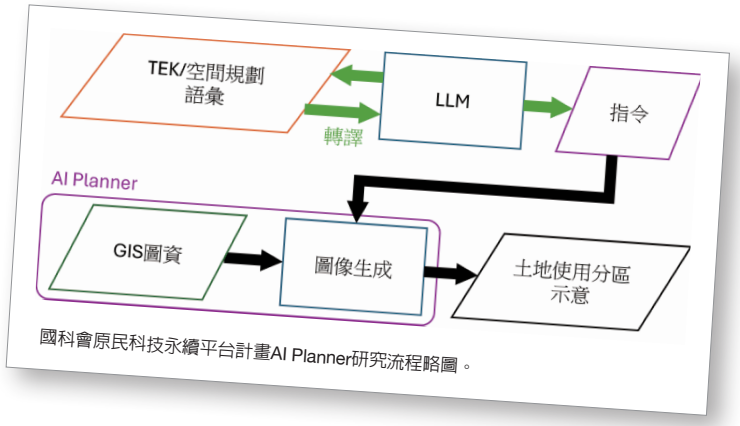
初探大語言模型於原住民知識掌握——AI 模型在國科會原民科技永續平台計畫的運用

大規模言語モデルの原住民族知識把握における初歩的探求
——国科会の原住民族科学技術永続的プラットフォーム計画におけるAIモデルの運用

Preliminary Exploration of Large Language Models in Mastering Indigenous Knowledge: Application of AI Models in NSTC's Indigenous Technology Sustainability Platform Project

文・圖｜李泱儒（政治大學地政學系測量組博士生）

隨著《國土計畫法》即將上路，首例原住民族特定區域計畫－泰雅族鎮西堡與斯馬庫斯部落，樹立了原鄉由下而上透過參與式規劃（Participatory Planning）做成空間規劃的一重要里程碑。



國科會原民科技永續平台計畫AI Planner研究流程略圖。

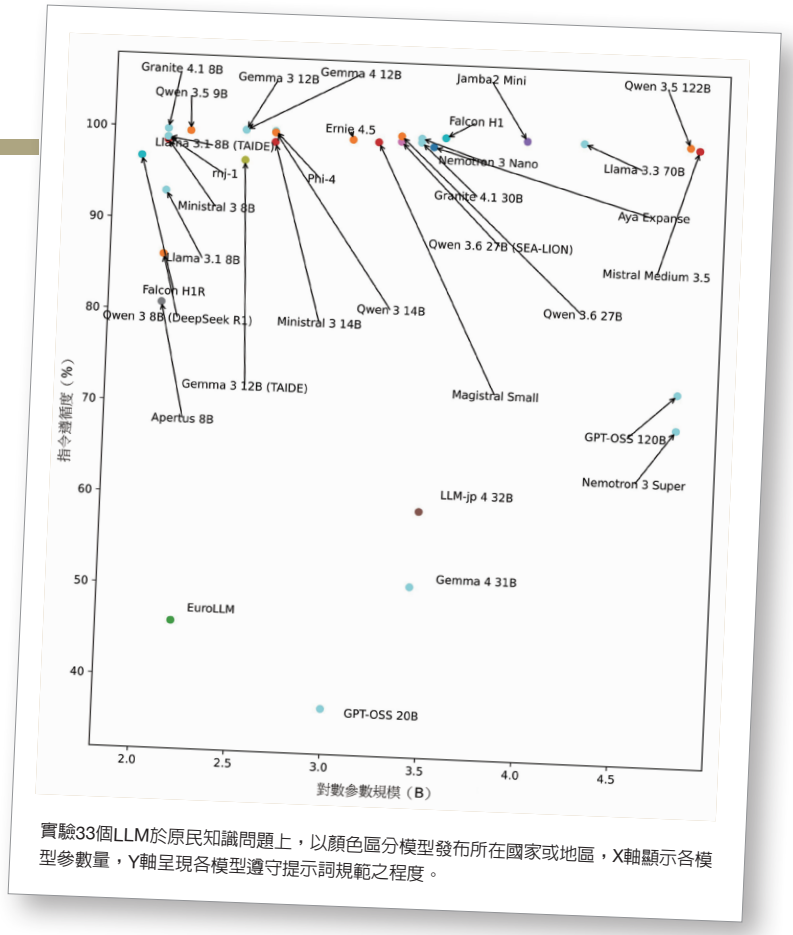
AI 技術的運用

面對氣候變遷的衝擊，如何將部落傳統生態知識（Traditional Ecological Knowledge, TEK）融入現代空間規劃，為永續發展的關鍵課題。在國科會「原民科技永續平台計畫」中，我們嘗試以人工智慧（Artificial Intelligence, AI）技術為知識轉譯媒介，降低族人參與空間決策門檻的同時亦提升規劃效率，期能為原鄉土地利用與文化傳承開創智慧轉譯的新契機。

為此，在研究流程設計上我們導入了大語言模型（Large Language Model, LLM）與圖像生成技術，如擴散模型（Diffusion Model）。作為當今 AI 時代之一代表，LLM 具備極佳的自然語言理解與跨領域語意轉譯能力，能夠轉化部落耆老與族人所口述的 TEK 、在地文化記憶以及對土地發展的質性想法為量化空間規劃標的，生成具備空間邏輯的結構化指令與提示詞（Prompts），結合地理資訊系統（Geographic Information System, GIS）上疊圖、視域與路網等分析技術於現地圖資，如地形、生態敏感區和既有土地利用圖等，經圖像生成模型輔助產製具體且多樣化的土地使用分區參考圖。前者，LLM 並非要取代傳統專家做出決策或部落為主體的規劃實踐，而是將族人原本難以用專業工程語言表達的土地願景、亦或是將專家學者描述之艱澀難懂的專有名詞，翻譯成參與規劃者皆能了解的語言，連結部落族人與空間規

在國科會「原民科技永續平台計畫」中，我們嘗試以人工智慧（Artificial Intelligence, AI）技術為知識轉譯媒介，降低族人參與空間決策門檻的同時亦提升規劃效率，期能為原鄉土地利用與文化傳承開創智慧轉譯的新契機。





實驗33個LLM於原民知識問題上，以顏色區分模型發布所在國家或地區，X軸顯示各模型參數量，Y軸呈現各模型遵守提示詞規範之程度。

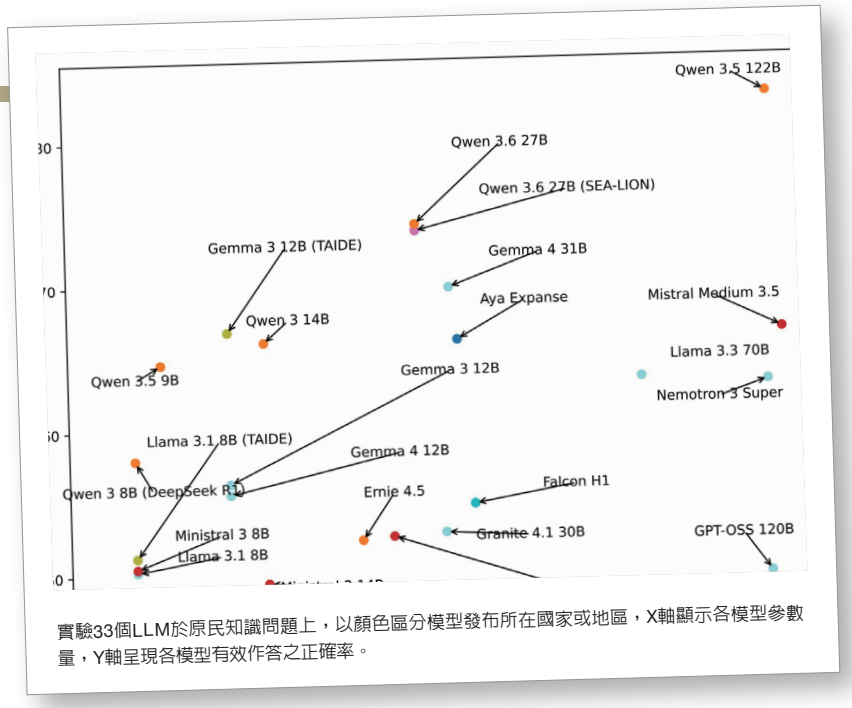
劃專家之間溝通的橋梁。後者，圖像生成模型藉由 LLM 生成之系統性提示詞，憑藉模型所具備之空間規劃領域知識，將文字資訊變成直觀的視覺化，讓抽象的部落發展願景躍然紙上。預期建構一能夠輔助原鄉土地決策的工具—— AI Planner，使其不僅能作為空間規劃的參考依據，更能在數位時代中維護原住民族的土地主權與文化主體性，並試圖建立在實質交流對談時可快速迭代呈現的新互動方式。

養成 AI Planner文化敏感度與知識正確性實驗

實際進入上述研究流程前，我們實作了一些實驗，在此分享。實驗以 108、111、114 和 115 年度台南市原住民族教育資源中心知識問答競賽選擇題作為資料來源，問題涵蓋台灣原住民族的地理環境適應、傳統祭儀、歷史文化與語言族群分布等核心知識。舉例而言：通常什麼樣的地理環境，發展出什麼樣的居住文化，達悟族「半穴居」是因為下列何種因素？（A）可降低當地惡劣氣候的肆虐（B）為保障族人免受野獸的攻擊（C）可以節省建築的花費；布農族著名的複音音樂 Pasibutbut，試問他們是為了以下何種原因吟唱？（A）傳達愛慕之情（B）戰爭前的頌歌（C）戰後凱旋之歌（D）祈禱小米豐收；「泰雅人所說的___，意思為祖訓或遺言。」請問上

在實務應用與部署中，盲目追求龐大的模型參數並非品質保證，選擇參數規模適中且具備良好指令對齊（Instruction Alignment）能力的模型，才能在維持高品質轉譯的同時，確保流程的穩定。





實驗33個LLM於原民知識問題上，以顏色區分模型發布所在國家或地區，X軸顯示各模型參數量，Y軸呈現各模型有效作答之正確率。

文的空白處，最可能是下列何者？（A）gaya（B）yaya（C）gaga（D）yaga。將其整理為題目與答案對應共 663 題不重複之單選題資料集，分布為 108 年度 98 題、111 年度 141 題、114 年度 400 題及 115 年度 24 題。我們嘗試以其量化評估 LLM 對台灣原住民族客觀知識的掌握程度，這些量化成果可協助檢驗 LLM 基礎能力，亦可作為後續引入特定領域資料進行模型微調（Fine-tuning）或推論（Inference）時的重要依據。藉由基礎知識問答的驗證，期能確保 AI Planner 在未來轉譯更複雜的 TEK 或空間規劃語彙上，具備足夠的文化敏感度與知識正確性。

實驗共挑選了 33 款不同參數規模與架構 4-bit 量化的模型，在維持模型語言理解能力的前提下順暢執行推論，涵蓋從千億參數等級的 LLM 如 Nemotron 3 Super ，到數十億參數等級之 LLM 如擁有 80 億總參數的 Llama 3.1 ，也包括針對台灣在地語境與繁體中文進行優化的 TAIDE 系列，藉此廣泛比較不同模型對原民知識的掌握差異。在模型設定上如溫度（Temperature）、Top-P 和 Top-K 等參數以及系統提示詞（System Prompt）上，我們使用各模型預設，上下文窗口（Context Window）則設定為足夠大之 4096 個符元，並將最大生成符元限制為 1024 以避免模型陷入無限迴圈或過度冗長的回應，而在使用者提示詞（User Prompt）的設計上，為確保模型輸出的結構化和資料處理的一致性，我們規範各模型必須以 JSON 格式回傳結果，必需包含作答和理由兩項，作答項為（A）、（B）、（C）或（D）等單一選項，理由項則為一段 500 字以內文字闡述其推論邏輯。如此設計不僅能幫助我們精確計算作答正確率，更能深入檢視模型是否真正理解文化脈絡，亦或僅是機率上的隨機猜測。各題目皆為獨立單次測試，不

保留跨題目的對話歷史紀錄，以避免前文脈絡干擾模型的判斷，藉此客觀評估 LLM 在零樣本（Zero-shot）情境下的背景知識與推理表現。

從初步的實驗成果來看，多數輕量級至中型模型（如 8B 至 12B 參數規模的 Granite 4.1 8B、 Qwen 3.5 9B 以及 Llama 系列等），其指令遵循度均展現接近 100%，顯示出高度穩定的格式約束力與指令理解能力。相反地，部分較大參數規模的模型（如 GPT-OSS 120B 和 Gemma 4 31B），其指令遵循度卻落在 80% 之下，說明於推論時未能嚴格符合指定 JSON 結構或未能精確輸出選項。此觀察對於後續 AI Planner 的基礎模型選擇提供了重要參考：在實務應用與部署中，盲目追求龐大的模型參數並非品質保證，選擇參數規模適中且具備良好指令對齊（Instruction Alignment）能力的模型，才能在維持高品質轉譯的同時，確保流程的穩定。

又在模型微調上，應將指令對齊納入考量，以確保模型的產出可以銜接後續空間決策流程。從 LLM 作答的正確率來看，整體表現與模型參數規模呈現正相關，Qwen 3.5 122B 正確率更是超過八成，展現出龐大參數在知識儲備上的絕對優勢。然而，若細看中小型模型的表現，可發現「模型架構系列」與「在地化語境優化」為影響正確率的兩大關鍵因素。其一，Qwen 系列在各參數體量下大多展現出較佳的表現，例如 9B 規模的 Qwen 3.5 9B 正確率即達 65%，27B 規模的 Qwen 3.6 27B 更達 74%，顯著領先多數同級別模型；其二，特別針對台灣在地繁體中文語境優化的 TAIDE 系列展現出較佳的知識理解力。以 12B 級別為例，經 TAIDE 微調之 Gemma 3 12B（TAIDE）的 67%，其正確率均顯著高於對應的原始基礎模型的 56%，證實本土化語料訓練能提升模型對在地知識的掌握。

持續規劃及推進

在了解模型表現後，我們為下一階段的研究確立了兩項後續推進重點。**一是思考建構一套更專注於原住民族知識領域的測試基準（Benchmark Dataset）**。有別於現有測試基準大多著重於理工邏輯與程式開發能力（如 SWE-bench、AIME），或是屬於範疇廣泛的綜合性評估（如 MMLU或 Humanity's Last Exam），目前仍缺乏針對原住民族文化與 TEK 的客觀評估指標。建構專屬的測試基準，能幫助我們更精準且公平地衡量模型對部落在地知識的理解深度；**二是我們將持續探索如何精進模型對原民知識的掌握程度（如導入領域微調或檢索增強生成技術），並加入「空間規劃」領域能力的評估**。透過結合文化脈絡與現代土地分區劃設規範，測試模型在轉譯部落土地願景時的邏輯合理性，建構兼具文化回應性與專業空間規劃實力的 AI Planner。 ◆



李泱儒

高雄市左營區人。2000年生。政治大學地政學系測量組博士生。曾任財團法人成大研究發展基金會研究助理。研究興趣涵蓋人工智慧研發應用、遙測及空間資料處理與土地利用模擬，並將空間資訊技術用於自然環境和社會經濟等領域。