

## 本土語言 AI 工具發展機會與挑戰



國立陽明交通大學廖元甫教授分享本土語言 AI 工具（攝影：蘇進榮）

### 【語文教育及編譯研究中心 吳鑑城】

語言承載著獨特的文化內涵，是人類重要且寶貴的無形文化遺產，然而現今不少語種正面臨著被同化與消失的嚴峻考驗。現今人工智慧蓬勃發展並大量應用，有鑑於此，本院語文教育及編譯研究中心於 2024 年 5 月 31 日邀請國立陽明交通大學智能系統研究所廖元甫教授兼所長主講「本土語言 AI 工具發展機會與挑戰」。廖教授在人工智慧領域擁有深厚學識與豐富經驗，專精於語音訊號處理、自然語言處理、機器學習及多模態基礎模型等研究。

廖教授分享多項進行中的重要研究工作與語料收集經驗，包括收集閩南語以及四縣腔和海陸腔的客家語語音語料，這些寶貴語料為語音識別和語音合成研究奠定堅實基礎。他們開發的語音技術應用範疇廣泛，例如疫情期間為指揮中心與總統新聞記者會提供即時字幕轉譯服務；與科技巨擘 Meta 合作，提供閩南語語料訓練台英即時口譯系統，以及在電影《流麻溝十五號》的製作中，透過語音合成重現已故前總統蔣經國的聲音。

接著，廖教授介紹生成式 AI 的最新進展，包括 Deepfake、ChatGPT-4V 與 Solar 等技術，並展示這些技術在語音識別及影像理解方面的應用，但也同時憂心地表示，在人工智慧時代，主流語言在技術發展上佔盡優勢，主流語言使用者與人工智慧的互動也較為順暢；相形之下，弱勢語言的使用者卻面臨重重阻礙，高科技引發的語言隔離加劇，迫使更多人轉向使用主流語言，進一步邊緣化了弱勢語言的地位。訓練大型人工智慧模型需要龐大的訓練資料，而這正是許多弱勢語言所缺乏的。

他強調台語和客語語言復振的急迫性，並提出利用 AI 技術進行本土語言復振的構想。當與會者詢及 AI 技術可否用於原住民族語，廖教授抱持樂觀態度，並回應技術本身可以適用於各種語言，然而，高質量語料的數量往往是決定人工智慧技術效能的關鍵，而閩客乃至於原住民語目前仍面臨語料不足的挑戰。

最後，談到 AI 應用於語言復振的挑戰與未來展望，廖教授指出語料問題是主要挑戰，語言模型需要更多台語和客語語料及更強大算力以達到主流語言模型的水準。而跨領域合作亦至關重要，需結合工程領域及人文社會科學領域的專業知識。他期待開發出更完善的 AI 工具，促進本土語的學習與應用，進而促進語言復振。此次演講在熱烈的掌聲中圓滿結束，本院將繼續舉辦更多類似學術活動，促進語言教育與跨領域研究之發展。