「生成式AI諮詢問答系統服務」館員實務分享座談會紀實

陳俐卉

成功大學圖書館知識服務組

壹、前言

隨著生成式AI技術迅速發展，圖書館界亦積極探索AI於讀者服務上的應用，以提升服務效能與使用體驗。本次實務座談分享會邀請來自不同類型的圖書館館員，包括：成功大學圖書館、明志科技大學圖書館，以及國立公共資訊圖書館，針對生成式AI諮詢問答系統的規劃與應用進行分享與交流，探討內容涵蓋圖書館如何建置與導入AI系統、相關資源與服務整合、後續功能改善，以及系統上線後的優化調整等面向，藉由各館間的實務交流，共同思考生成式AI如何成為圖書館服務與館員工作的重要助力。

貳、從概念到實踐的圖書館AI參考館員之路

成大圖書館系統資訊組張凱雅副管理師分享如何運用Dify軟體，透過Docker於Windows系統部署成大圖書館AI參考館員。這套系統整合了大型語言模型（LLM）與檢索增強生成（RAG）技術，運用歷年累積的參考諮詢問題集進行訓練，打造出以AI技術為核心的圖書館參考諮詢系統（圖一），為有意導入生成式AI的圖書館同道提供了實務參考。

除了介紹系統的整體規劃與開發歷程外，講者也分享建置過程中所遭遇的問題與解決

方法，並提出系統正式上線前可進一步優化的功能面向。最後，講者深入分析實用的提示語（Prompt）設計技巧與開發經驗，並展望AI客服系統的未來發展趨勢，包括：模型上下文協議（MCP）、原生代理型AI技術及AI代理人等應用的可能性，也讓與會者對未來AI參考館員的發展有了更深的期待。



圖一　張凱雅副管理師展示畫面

參、自建圖書館生成式AI聊天機器人（客服機器人）經驗分享

明志科大圖書館諮詢推廣組林農偉組長首先分析傳統客服系統的限制，並說明導入生成式AI可提升客服系統的語意理解和回應彈性，打造更自然、人性化的客服互動體驗。明志科大圖書館開發的客服機器人是以檢索增強生成（RAG）架構為核心，搭配各種開源軟體及語言模型，如：Python、Gemma 3、LangChain框架與Ollama管理平臺等，並整合MySQL資料庫與Hugging Face向量嵌入技術，建構出可在單一GPU上運行的AI聊天機器人問答系統（圖二）。

講者進一步介紹全封閉式與部分封閉式兩類回應方式：前者以館內常見Q&A資料庫為基礎，像是借書規則、館藏查詢、設備使用等；後者則應用於主題書籍的推薦，讓模型發揮語意理解能力來延伸對話。在講者精彩的實作展示中，可以看到系統不僅能記憶讀者的提問歷程，也能以中英文雙語進行查詢與回應，充分展現其回應邏輯與內容的多元性。

為驗證系統效能，講者也展示變體問句自動生成與壓力測試的流程，以及使用程式模擬多用戶同時提問、蒐集準確率與回應延遲時間等資訊，作為系統優化的依據。此外，也針對硬體與經費提出可行的發展策略，如：館際共用GPU或租用虛擬伺服器，以降低建置與維運成本。



圖二　明志科大圖書館諮詢推廣組林農偉組長展示畫面

肆、AI館員的秘辛：從曉書談建置、維運到優化

本場次以導入「AI館員——曉書」實際經驗為主軸，國立公共資訊圖書館閱覽諮詢科陳筠凡科員深入介紹「曉書」的建置歷程與維運挑戰。講者以「不是工程師也能做」為出發點，說明從0到1推動AI服務的關鍵步驟，從資料清理、語意分析、語音辨識、語言切換到動畫設計，各項技術整合相互配合，逐步打造出兼具互動性與語言理解能力的虛擬館員（圖三）。

「曉書」同樣採用大型語言模型（LLM）與檢索增強生成（RAG）技術建置。系統上線後，館員必須每月檢視與分析逾兩千筆的對話紀錄，找出錯誤回應、語音辨識失準或資料擷取不完整等問題，透過補充知識庫或優化提示語（Prompt）的方式進行持續修正與

微調。講者指出，常見錯誤來源包括：語音辨識失敗、資料分類不精確、語境前後不連貫，以及語言切換異常等。雖然AI館員可協助處理參考服務臺的常見諮詢，但其角色更像是「須要訓練的工讀生」，仍然須要仰賴館員持續投入人力進行維護與調整，才能穩定運作並確保服務品質。

在AI模型的訓練與調整方面，講者分析不同手法的成本與適用性：提示語（Prompt）可以調整語氣與格式、RAG提供即時查詢能力，而Fine-tuning雖然能改變模型的行為，但成本過高，不建議優先採用。

展望未來，講者提出理想AI館員應具備的特質，包括：理解讀者的意圖、整合多元資源、回饋學習能力與互動體驗升級等，並提醒圖書館應審慎考量系統建置的彈性、讀者習慣與成本效益，才能在快速變動的AI技術浪潮中穩健前行。



圖三　國立公共資訊圖書館閱覽諮詢科陳筠凡科員展示畫面

伍、綜合座談及交流

在綜合座談交流階段，與會者針對系統建置、使用情境與維運的細節提出各種面向的問題，反映出圖書館界對生成式AI實作的高度關注與實務考量。其中許多提問聚焦在系

統效能與維運上的挑戰，包括：詢問明志科大如何維護RAG資料的更新頻率與處理方式、在使用量高峰時是否會出現存取限制，以及是否須人工檢視正確性等；亦有與會者詢問國資圖是否有錄音機制回溯檢測以提升語音辨識的準確率，聚焦於語音系統建置時可能產生的挑戰。另有多位與會者探詢知識庫與資料來源的問題，例如：書摘與小說資訊是否來自原書目資料、是否整合其他校內資料、如何處理季節性或大量資料庫擴增的維護需求等。

部分提問則關注在實體與虛擬服務的整合，如：生成式AI是否影響讀者對實體參考館員的依賴、曉書是否限館內使用、是否可於閉館期間提供服務，以及使用者在進行推薦書與一般問答時是否需主動切換模式等。在模型與技術選用方面，講者們也回應了關於LLM模型選擇困難、成本投入（含人力、硬體與雲端流量）的考量，以及提示語設計中角色、任務、格式與限制四要素的應用心得。

整體而言，與會者的提問涵蓋了生成式AI服務從理論導入到實務操作、從技術框架到使用者體驗等各種面向，顯示大家對智慧參考服務的好奇與期待，也讓本次活動的討論更加深入全面。

陸、結語

館員實務分享座談會為館員跨校交流與知識共享的重要平臺。本次座談會吸引來自全國126所大學及技專校院圖書館共450位館員以視訊方式參與，充分顯示圖書館界對生成式AI議題的高度重視。

三場分享分別代表了AI參考館員建置的不同路徑與發展階段：成大圖書館展示了以開源工具自主部署的技術深度；明志科大圖書館則呈現了以RAG架構打造雙語客服機器人的完整流程，並提出降低硬體成本的可行策略；國資圖則以「曉書」的實際維運經驗說明，AI館員的持續優化須要投入相當的人力成本，並非一勞永逸的解決方案。

三位講者不約而同地強調：提示語（Prompt）設計是影響AI表現的關鍵，RAG架構是目前圖書館環境中最具可行性的技術選擇，而無論採用何種技術，持續的人工維護與資料更新仍是確保系統品質的根本。線上與會者反應熱烈，交流討論十分踴躍，為圖書館界在AI應用的探索之路上，提供了豐富而務實的參考方向。