

數位人文，從地土到雲端

文——林佩蓉 研究典藏組

臺灣文學研究走到數位化、走進科技環境的今天，有多大的能量、多豐富的風景，就在我們眼前展開？臺灣文學館的數位人文工作，希望兼顧研究與推廣，擁有速度及能量。正在進行的「數位全臺詩」計畫則是一座新的里程碑。

千里馬的真正能量

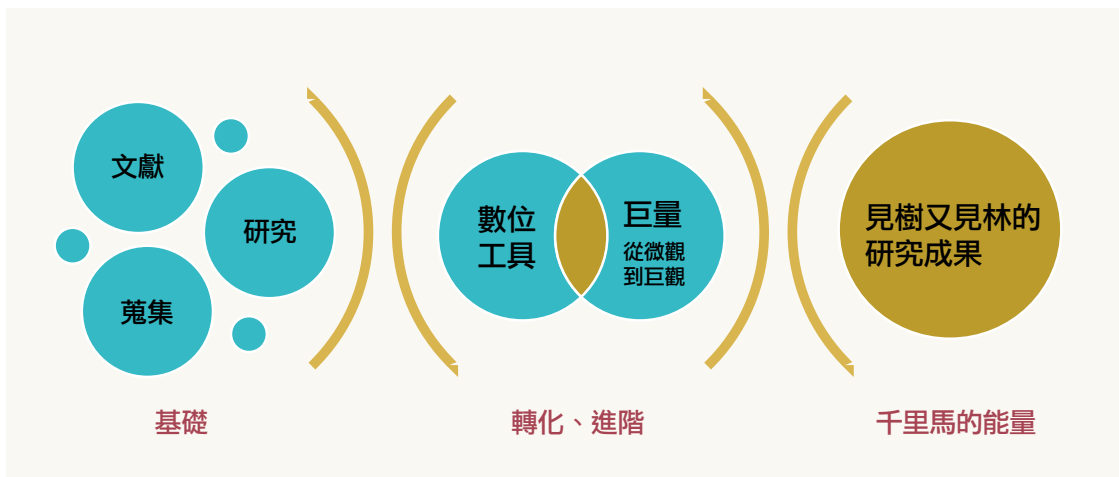
進入21世紀的「滑世代」，臺灣文學研究方法，在2010年代起進入科技數位時代，不只是透過資料庫查找所想要的資料，接著或許剪下貼上，放在EXCEL、WORD格式裡，可能還會再給予排序、計算作品發表的先後順序、作家參與活動的次數等，這半人工方式介入資料庫，是以電腦作為書寫工具的我們，最基本的能力。只是這樣的方式，就好像讓一匹千里馬，只跑了一趟千里，而牠真正的能力則是來回千里，千里馬沒能發揮真正能量，資料庫的發展很快走到了一種既定生存的模組，它的能量，從設計的起初就被限制，雖然它的價值還是必須被肯定，至少減少影印資料、降低翻閱古老原件的機會，只是我們必須要問，既然已經搭上千里馬，怎能不一探牠的千里能力？臺灣文學研究走到數位化、走進科技環境的今天，有多大的能量、多豐富的風景，就在我們眼前展開？

擁有16個資料庫、1個整合查詢系統、2個作家主題部落格（持續增加中）的國立臺灣文學館（以下簡稱臺文館）持續在想，當備齊所有查考功能的資料庫之後，開始使用中英對照的數位

平臺推廣臺灣文學、開始面對要不要重新盤點現有的資料庫，以及探求其開展性。於是邀請了幾年前開始投入數位人文研究的幾位學者，包括邱貴芬教授、黃美娥教授，他們分別就數位平臺的國際發展以及資料庫研究的延展性進行思考及研究，此外臺文館也透過獲得「科技部專題補助計畫」——「數位全臺詩」進行全臺詩知識庫的改造，透過現行出版的40冊《全臺詩》計800多位詩人，對其詩作進行元素類別分析、作家行跡考究、社群的脈絡等探究。這個計畫揭示了臺文館的文學研究進入「大數據（Big data）」的方法領域。面對這樣的時代，臺文館做了「數位人文前期研究」，委請國立臺灣大學黃美娥教授、佛光大學林以衡兼任助理教授兩位學者，統整、概述當今文學研究大數據的發展概況。

數位人文計畫：從微觀到巨量、見樹也見林的可能性

根據項潔教授的研究（項潔、涂豐恩〈導論——什麼是數位人文〉，《從保存到創造：開啟數位人文研究》，2011年），國外大學很早就已設立不少數位研究的單位，例如倫敦國王學院「數位人文學系」、史丹佛大學「人文實驗室」、加州大學洛杉磯分校「數位人文中心」、哈佛大學「數位人文先導計畫」、日本立命館大學「日本藝術與文化數位人文中心」，以及國內臺灣大學的「數位人文研究中心」，證明國內外對數位研究的重視。而自2006年起，世界舉辦「數位人文年會」，並有跨國機構如數位人



文研究聯盟（The Alliance of Digital Humanities Organizations, ADHO），其下設有專職推動全球數位人文發展之研究小組。數位研究蔚為風潮，也讓我們開始思索，數位研究與人文學科結合的意義。

現行的成果

數位人文近十年來在臺灣學界倍受關注，臺灣各大研究機構或中心皆投入不少資源拓展此議題。臺灣目前對數位人文研究起步較早、積極思考的學術單位，以中研院、臺灣大學與科技部為最，部分大學如逢甲大學、東吳大學近年來亦注意數位研究的重要性。臺灣數位人文發展概況與資料庫分析如下：

1. 中央研究院數位文化中心

中研院各所處中心長期致力於運用資訊科技典藏該院珍藏文化遺產、歷史檔案、考古文物、生物標本與各式攝影圖像，以及致力於數位相關資訊與網路技術的研發，為臺灣數位典藏機構起帶領與示範的作用。

2. 科技部

長久以來，臺灣科技部引領臺灣學術取向，為臺灣厚植不少豐碩研究成果，可見科技部對臺灣學術發展的重要性。也因此，在科技部的推動與鼓勵下，不少研究者以數位研究作為探究重點，開發不少數位研究的可能。這其中，也包括數位人文方面的探究，從「數位人文前期研究」報告書的第五項附錄，即可知科技部這幾年來，獎助不少以數位人文研究為主題的計畫案，積極提升臺灣在數位人文研究方面的能力。

3. 臺灣大學數位人文研究中心

臺灣大學作為全臺大學的龍頭，對於數位人文的研究和推廣，也是不遺餘力。臺灣大學數位人文研究中心，前身為2007年成立之「數位典藏研究發展中心」。自2012年起，為加速推動數位人文研究，更名為「數位人文研究中心」。此中心重點工作之一在典藏品數位化與資料庫建置，除持續針對臺大所擁有重要性、唯一性、即時性、脆弱性的珍品進行數位典藏，並由內而外地延伸到國內外的重要文化資產。目前已陸續建立33個大型數位典藏資料庫，內含逾500萬筆詮釋

資料 (metadata records)，近3,000萬幅影像，近4億字的全文，以及數百小時影音資料。內容包括臺灣歷史、政治、社會、法律、經濟等各方面，資料所屬年代橫貫4百年臺灣歷史，且多為第一手史料，不僅是國內數位史料累積最豐富的數位典藏機構之一，亦是目前世界上累積最獨特、大量的臺灣研究資料。

這些研究資料庫所收錄的史料均有獨特、重要、唯一、即時等特性，並對外提供研究者使用，已在文史學術界及古契書研究領域產生震撼並被廣泛運用。結合「檢索後分類」為主的檢索方法與資訊探勘分析工具，更突顯出檢索結果 (query result) 的整體性，並用後分類及統計方法分析呈現其中的關聯性與脈絡。同時開發新的資訊取得及資訊分析方法，目前已開發20種以上知識探勘及呈現的工具，藉此提升人文研究的品質與效率，促進數位人文的發展。

除持續進行典藏數位化、建置資料庫及數位研究工具外，此中心打破使用者使用資料庫時，單純檢索搜尋的習慣用法，藉由資訊技術，提供使用者觀察資料後進一步勾勒資料彼此間的可能脈絡，以此作為開發「數位人文」研究方法和學科領域的基礎思維。

4. 逢甲大學「地理資訊系統研究中心」和東吳大學「巨量資料管理學院」

這兩所大學的研究發展除了提供學生的教育與培訓外，因為研究與產業越來越密不可分，並預計與業界進行合作，一方面可以幫助企業解決目前所面臨龐大而複雜資料內容的難題，也可從與企業的合作中，協助學生獲得更多的就業機

會；另外一方面，亦可發展成為提供業界解決問題之研發中心，建立長期合作的夥伴關係。

千里上路

經由上述梳理和分析後，計畫成果歸結出近年來臺灣對於數位人文研究的幾個趨向如下，以供文學館未來對臺灣文學與數位人文相輔相成時為參考。

1. 以GIS為主要的研究途徑

GIS (Geographic Information System) 地理資訊系統，是目前數位人文研究的重點之一。它透過地理學和地圖學的結合，運用數位科技輸入、儲存、查詢、分析和顯示的功能，幫助研究者對欲查詢資料建立地圖式的輪廓，達成資料整合與學科跨界運用的目的。目前臺灣以GIS為研究取向的數位研究，以中央研究院人社中心的地理資訊科學專題研究中心為最，它結合中央研究院數位典藏資源網開發「臺灣百年歷史地圖：臺灣堡圖相關圖層」，主要運用該中心「地圖與遙測數位典藏計畫」所累積大量的臺灣地圖資料，結合Google地圖介面，為研究者建立一個全新的網站服務，對於認識臺灣和鄉土教育有極大功用。前述逢甲大學「地理資訊系統研究中心」亦是以此為發展特色。

2. 以「觀念史」為主要的研究途徑

在金觀濤、劉青峰兩位教授協助與規畫後，政治大學「歷史與思想數位人文實驗室」在中文系鄭文惠教授的主持下，已成為目前臺、中、港三地重要的思想與人文研究資源。本實驗室以

「中國近現代思想及文學史數據庫」為基礎，利用數據庫的數據挖掘和重要關鍵詞意義的統計分析，揭示中國近現代觀念如何在西方衝擊下起源和形成之過程。通過近十年的研究，完成了「理性」、「社會」、「共和」、「權利」、「個人」、「經濟」、「科學」、「革命」等重大現代觀念形成以及相應近百個重要術語演變研究；具體成果已展現在《觀念史研究：中國現代重要政治術語的形成》（香港：香港中文大學當代中國文化研究中心，2008）一書，並獲得臺、中、日、韓等地學界熱烈迴響。

在此基礎下，數據庫辦公室在科技部與學校的支持下，拓展其對數位人文的研究面向，目前主要在中國近現代觀念形成與演變研究上，勾劃出中國近現代國族觀念形塑與轉變的大歷史結構，並運用數位方法尋找音譯詞，找到突破口，以解決各式問題，並試圖在未來使編寫中國近現代人名地名大字典成為可能。此外，本實驗室另發展數位方法尋找核心關鍵詞，為數位方法實現中文歷史文獻分析建立基礎，並用數位方法突出觀念、事件、行動緊密關聯、彼此作用的複雜關係與脈絡化轉向。

而與學校其他數位單位合作的情況，也足供臺文館參考，此實驗室與同校如未來力實驗室、數位資料稜鏡實驗室、數位足跡探索實驗室、數位民主治理實驗室、資料科學跨領域實驗室、計算行為社會科學實驗室、態度與決策巨量資料實驗室、語言與認知實驗室與創新行動金融商務科技實驗室進行整合，成為政治大

學近年來數位研究的大型團隊，對於數位研究的發展有重要影響。

3. 文本標記與文獻閱讀

按魏希德教授的解釋，所謂「文本標記」就是：「數位工具對於人文學的研究，不只能進行巨量的統計、宏觀趨勢的分析，也能細緻地記錄過去文件的樣貌、研究者的成果和心得，讓數位人文學者可以任意地在宏觀與微觀之間自由移動，進行各式各樣的探索。要做到這點，其中一種方法就是文本標記；亦即，把文件中的重要資訊利用標記的方式以利於電腦的辨識，再在這樣的基礎上進行種種的利用。然而，大量的文本要如何進行標記，卻是一個大問題。如果利用人工進行不但曠日廢時，而且常常很難取得一致的標記成果，因此如何讓電腦進行自動標記也就成了資訊學者的一大課題。」目前中研院數位人文研究室已舉辦過多場關於文本標記與文獻閱讀的演講。而清華大學中文系祝平次教授近年來亦以發展文本標記與文獻閱讀為重點，此方向對於人文學科的研究，有極大助益。

從地土到雲端——數位臺灣文學館

走在地土上、計量在雲端，是臺文館面對臺灣文學的研究態度與學習的技術，我們期待走在臺灣文學研究者的旁邊、臺文系所的前方，整合眼前的資源式的資料庫、發展推廣具國際使用規格的平臺、運用微觀與巨觀的研究技術，是本館走上數位研究型機關的必經之路。☒