

台灣文學文獻數位化的回顧與前瞻

2004年漢學研究國際學術研討會，2004年10月29、30日，
雲林科技大學漢學資料整理研究所主辦。

羅鳳珠*

共同作者

陳萬益、何淑華、蔡宛純、林亮君、黃欣慧、林淑鈴、林宜萱、黃詩媛

摘要

從1984年7月中央研究院「史籍自動化計畫」開始，中文資訊技術即有計畫的應用在中文文獻資料的處理上，迄今將近二十年，台灣在漢學數位化累積的數量已達幾億字，種類也已經從文字資料擴充到文物、影像、聲音等各種不同媒材的資料，投入數位化工作的人員，也從當年學者自發性的研究到如今政府以國家型計畫作政策性的推動。

二十年來，資訊環境進步的快速以及影響的深遠，恐非當時所能想像。台灣文學數位化的起始時間晚於其他漢學，雖說起步較晚，但相較於二十年前中央研究院的「史籍自動化計畫」工作，台灣文學的數位化已經有比較成熟的資訊環境，漢學數位化工作所累積的經驗，也正可使台灣文學數位化的工作避免很多摸索的時期。

近年台灣文學系所的成立及台灣文學研究不再成為禁忌，台灣文學喚起了台灣人的感情，2003年國家台灣文學館成立，意義更是不同，但是面對起步較晚的現實環境，如何藉助資訊科技在網路環境裡建立台灣文學數位博物館，結合典藏、研究、教學、推廣等多元功能，從文學擴充到文學與語言文化的結合，讓台灣文學藉著網路伸向全世界，這是台灣文學數位化可行的方向。

關鍵詞：台灣文學、文學數位化、數位博物館、詮釋資料、賴和、鍾理

* 元智大學中語系講師

壹、前言

從 1984 年 7 月中央研究院「史籍自動化計畫」開始，中文資訊技術即有計畫的應用在中文文獻資料的處理上，文獻媒體種類由文字資料擴充到文物、影像、聲音等各種不同媒材的資料，投入數位化工作的人員，也從當年學者自發性的研究到如今政府以國家型計畫作政策性的推動。

二十年來，資訊環境進步的快速以及影響的深遠，使得引用資訊科技作為文學的研究工具已是必然的趨勢，台灣文學數位化也是刻不容緩的工作。相較於二十年前中央研究院開始執行史籍自動化的情況，台灣文學的數位化工作，已經有比較成熟的資訊環境，前人文獻數位化工作所累積的經驗，也正可使台灣文學數位化的工作避免很多摸索的時期。

近年來台灣文學研究逐漸受到重視，原典文獻、文物的蒐藏，受限於人力、物力、自然環境、時間、空間等因素，除無法讓更多人使用外，文獻及文物的保存狀況，也因種種主客觀條件而異。因此，如何藉助資訊科技在網路環境裡建立台灣文學數位博物館，結合典藏、研究、教學、推廣等多元功能，從文學擴充到文學與語言文化的結合，讓台灣文學藉著網路伸向全世界，這是台灣文學數位化可行的方向。

本文將以「賴和數位博物館」與「鍾理和數位博物館」為例，藉以探討「文學數位博物館」的建置模式與設計方法，並進一步探討「以作家為單位」及「以文體為單位」的文學網站建置方式，以及在文學研究方面的應用、在其他層面擴充性的應用與研究。

貳、台灣文學文獻數位化的回顧

近年來台灣文學文獻數位化的工作陸續進行，已完成的網站如客家文化委員會策劃的「台灣客家文學館」（第一年：<http://cls.hs.yzu.edu.tw:88/hakka/default.htm>；第二年：<http://lit.hakka.gov.tw/index.htm>）、「客家研究圖書資料庫」（<http://140.115.130.214/hakka/use.htm>）；國家台灣文學館策劃的「作家手稿」（<http://nmtl.ncrpcp.gov.tw/literaturetotal.htm>）、「全臺詩電子文庫」（<http://www.wordpedia.com/twpoem/index.asp>）；行政院文化建設委員會策劃的「當代文學史料影像全文系統」（<http://lit.ncl.edu.tw/>）、「當代台灣文學系統作品英譯本索引」（http://www.cca.gov.tw/imfor/new/work_11/）、「2001 台灣文學年鑑」（

<http://www4.cca.gov.tw/2001/index.html>)、「2002 台灣文學年鑑」(<http://www4.cca.gov.tw/2002/>)；行政院文化建設委員會國家文化資料庫策劃的「賴和數位博物館」(<http://cls.hs.yzu.edu.tw:88/LAIHE/>)、「鍾理和數位博物館」(<http://cls.hs.yzu.edu.tw:88/ZHONGLIHE/>)等。看到這些琳瑯滿目的成果，著實令人振奮，本土意識的提升，台灣文學漸漸受各界重視，越來越多的人力、精力投入台灣文學的領域，台灣文學文獻數位化也在進步中求成長。回顧早期台灣文學數位化的工作，如何將文學與資訊領域結合，將文獻資料妥善的分類整理，建構成網路資訊系統，是大不易的事。文學領域的專家學者對資訊領域的陌生，不知如何引用資訊科技幫助文學的典藏、研究與教學；相對的，資訊領域的學者對文學內容與文學數位化的需求也不熟悉。如何將二者結合，對多數人文學背景的人而言，仍是一知半解，因此無法有效的利用資訊科技，結合多種文本進行資料的分析、標記(taging)，將單一資料提供多元的加值運用；而資訊領域的專家學者，因對文學內容的不熟悉，無法將文獻資料妥善的分析、整理，只能表面的在網頁的設計上求變化，無法做到完善的典藏功用。

由以上列舉的網站進一步比較，如客委會策劃的「台灣客家文學館」¹與其後的文建會國家文化資料庫策劃的「賴和數位博物館」²、「鍾理和數位博物館」³，前者僅提供作家代表作品、少數客語詞彙庫、作家即時資訊、電子報、少數作家的照片、手稿、影音資料；少數作家文學資料檢索、生平大事年表、作品繫年等資料。後者典藏文獻資料倍增，除在內容廣泛蒐羅文學家的典藏文獻及研究資料，針對作家的個別差異特別獨立架構外，在系統功能及資料庫的設計，建立共同平台以及跨資料庫的檢索，更完整、精確的典藏文獻資料，並凸顯其特色。

文學資料庫的建置，如客委會策劃的「客家研究圖書資料庫」(<http://140.115.130.214/hakka/use.htm>)與文建會的「當代文學史料影像全文系統」(<http://lit.ncl.edu.tw/>)比較，前者以客家族群作家為範圍提供研究書目、博碩士論文、期刊文獻、會議論文等資料目錄之資料庫的檢索；後者則全面的網羅五十餘年來台灣地區當代文學作家約兩千位之基本資料、及其生平傳記、手稿、照片、著作年表、作品目錄、評論文獻、翻譯文獻、名句及歷屆文學獎得獎記錄，建置完整而豐富的檢索資料庫。

¹ 「台灣客家文學館」由客委會策劃，苗栗聯合大學承辦，第一年計畫由清華大學陳萬益主持，建置作家包括龍瑛宗、吳濁流、鍾理和、鍾肇政、李喬、鍾鐵民等六位作家，網站分別安裝於元智大學之網路展書讀網站：<http://cls.hs.yzu.edu.tw:88/hakka/default.htm>，由羅鳳珠負責網站架構規劃及建置，另一份安裝於客委會網站，經過聯合大學修改，已非原貌：<http://literature.ihakka.net/hakka/default.htm>，建置時間是 2002 年。第二年計畫由聯合大學承辦，靜宜大學彭瑞金主持，延續第一年網站架構，建置作家增加林柏燕、曾貴海、鍾延豪、杜潘芳格、謝霜天、利玉芳，網站安裝於客委會：<http://lit.hakka.gov.tw/index.htm>，建置時間是 2003 年。

² 「賴和數位博物館」：<http://cls.hs.yzu.edu.tw:88/LAIHE/>、「鍾理和數位博物館」：<http://cls.hs.yzu.edu.tw:88/ZHONGLIHE/>，由文建會策劃，屬於國家文化資料庫之計畫，由陳萬益、羅鳳珠、林瑞明、鍾怡彥共同主持，鍾彩焱協助手稿文物拍攝，建置時間是 2003 年。

³ 同前註。

參、進行中的台灣文學文獻數位化工作

台灣文學文獻數位化工作在台灣文學愛好者及政府努力推動下，台灣文學網站如雨後春筍，一個個冒出頭來，目前進行中的數位化工作如：「閩南語典藏：歷史語言與分布變遷資料庫」⁴，內含「閩南語歌仔冊」、「客家歌本」（客語）單元的建置，針對台灣戲劇文學進行文本點校、考證以及語言整理。其他如台灣文學館「全臺詩電子文庫」計畫⁵、清華大學葉榮鐘網路紀念館⁶持續進行之中，文建會國家文化資料庫台灣文學主題網站的推動、客委會客家文學作家網站的策劃等，期待有進一步的創新與研發。

肆、賴和、鍾理和數位博物館建置模式與設計方法

一、賴和、鍾理和數位博物館的製作目的與網站定位

賴和、鍾理和數位博物館的建置是為落實文建會國家文化資料庫規劃之構想，所以將網站定位於「數位博物館」的形式。

國家文化資料庫建置計畫乃文建會「網路文化建設發展計畫」之下的基礎建設系列子計畫之一，其目的是要系統性的、計畫性的執行文化藝術資源之蒐集、整理和保存，並藉由資訊科技將其數位化典藏，留下文化資源的長久記錄。亦即蒐集、調查全國之藝文資源，包括：美術、音樂、舞蹈、文學、戲劇、建築、電影、古文書與老照片類等，並進行建檔（實體文物）、拍攝、建檔（數位文物）、數位化等工作，以文字、聲音、影像及視訊媒體，以適當的知識架構與資訊表達架構呈現，使分散各地的藝文資源，得以透過網路，讓民眾、研究者整合性的查詢利用。

國家文化資料庫提供了一個全民參與文化保存的機制，整合全國的文化資料，並透過各種檢索機制開放給全民共享，如此一來資料將可經由不同使用者的創新發展，產生活化與再利用效果，更可以經由網際網路的強大傳播功能，達到全民文化的提昇與文化對外傳播的目的。

⁴ 「閩南語典藏：歷史語言與分布變遷資料庫」是數位典藏國家型科技計畫，中央研究院語言研究所承辦，鄭錦全主持，楊秀芳、魏培泉、羅鳳珠共同主持，網址：<http://cls.hs.yzu.edu.tw/southernmin/index.htm>，計畫時間 2002 年至 2006 年。

⁵ 「全臺詩電子文庫計畫」，網址：<http://www.wordpedia.com/twpoem/>，是國家文學館的長期計畫，由成功大學台灣文學所施懿琳教授主持。

⁶ 「葉榮鐘網路紀念館」，網址：<http://lib032.lib.nthu.edu.tw/jcyeh/>，是國科會數位加值計畫，由清華大學主持。

在此目的下，賴和、鍾理和數位博物館定位於三方面：其一為典藏：將所有文物數位化以方便典藏；其二為推廣：文物及文學館實體空間數位化之後，以適當的知識結構呈現，藉著網路無遠弗屆的特性，提升推廣的寬廣度與速度；其三為加值應用：提供研究、教學之加值應用。

二、內容媒介分析與資料說明

網站資料內容以原典文獻手稿、文物，基金會藝文活動成果為主，以相關的研究資料為輔。此外，作家會因時代背景、生長經歷、地理環境的異同，影響文學作品的內容與風格，產生不同的創作靈感與素材，故這部分的資料也包含其中。舉例說明：彰化為台灣中部地區文化、政經與交通的重鎮，文教風氣鼎盛，人才輩出，堪稱獨步全台。陳肇興、吳德功、洪月樵等人傳承漢詩文，新文學之父賴和以及陳虛谷、楊守愚、王白淵、謝春木等濟濟群英，奠立台灣新文學的基礎，因此在「賴和數位博物館」的內容知識結構設計，將「彰化區作家」獨立架構，期由賴和個人資料的典藏開始，以日治時期彰化地區的作家為核心，逐步擴充為「日治時期台灣文學」的數位紀念館。目前網站已將館藏的楊守愚和賴賢穎的手稿、文物數位化，置入研究區中。以「鍾理和數位博物館」而言，重新點校鍾理和全集時，將作品中所描寫之美濃人文訊息建立索引檔，作為建置「美濃文學文化地圖」的資料，結合美濃客家鄉鎮之文學、風俗、民情、文化，以文學為起點，以文學為軸心，呈現美濃地區特殊的客家人文風貌。

網站執行製作需蒐集內容資料，並分析內容媒介，資料內容之使用媒體包含文字、圖片、照片、影像、虛擬導覽、文物、聲音等不同形式。詳述如下：

- (一) 文字資料：文學原典資料，如三校本賴和全集、鍾理和全集、校定本鍾理和作品集等；相關研究資料，如藏書目錄、研究論文、評論雜文、生平年表、寫作年表、學術活動資料、賴和平反資料等；其他文字，如生平介紹、紀念館館史、專家導覽紀錄、基金會介紹、語料資料、基金會活動簡介、文學導讀、音標資料、彰化縣文學史年表、作家紀念亭文句及生平簡介等。
- (二) 圖片資料：手稿、刊稿、筆記、剪報、賴和紀念碑、生平掛圖、學術活動資料、賴和平反資料、友人往來書信、祭悼輓詞、出版品、紀念館館藏文物、活動海報、選入教材作品等。
- (三) 照片資料：賴和相關照片、學術活動照片、楊守愚相關照片、基金會活動照片、八卦山文學步道等。
- (四) 影像資料：專家導覽影像、賴和身影、學術活動資料、賴和平反資料、基金會活動影音資料等。
- (五) 虛擬導覽：紀念館場景網路虛擬導覽，館藏內容專家導覽。

- (六) 文物資料：賴和紀念碑、醫療器材、生前遺物、八卦山文學步道等。
- (七) 聲音資料：專家導覽、虛擬導覽、賴和身影、賴和餘音、文學活動資料、賴和平反資料、基金會活動影音資料、背景音樂等。

三、設計內容之知識結構與資訊表達架構

賴和、鍾理和數位博物館的建置，無論是系統架構或內容知識架構，均已將未來擴充建置台灣文學家數位博物館的需求納入規劃及建置之考慮，建立共同平台、跨館檢索資料庫功能。

內容之知識結構與資訊表達架構的設計，以紀念館導覽及索引檢索區為兩大主體結構，進而由「紀念館導覽」衍生出作家生平、專家導覽、虛擬導覽、紀念館館史等知識架構；由「索引檢索區」延伸為綜合索引區、文學檢索區、文學研究區與文教推廣區等四項主題知識架構。以下詳細說明整體知識架構設計內容：

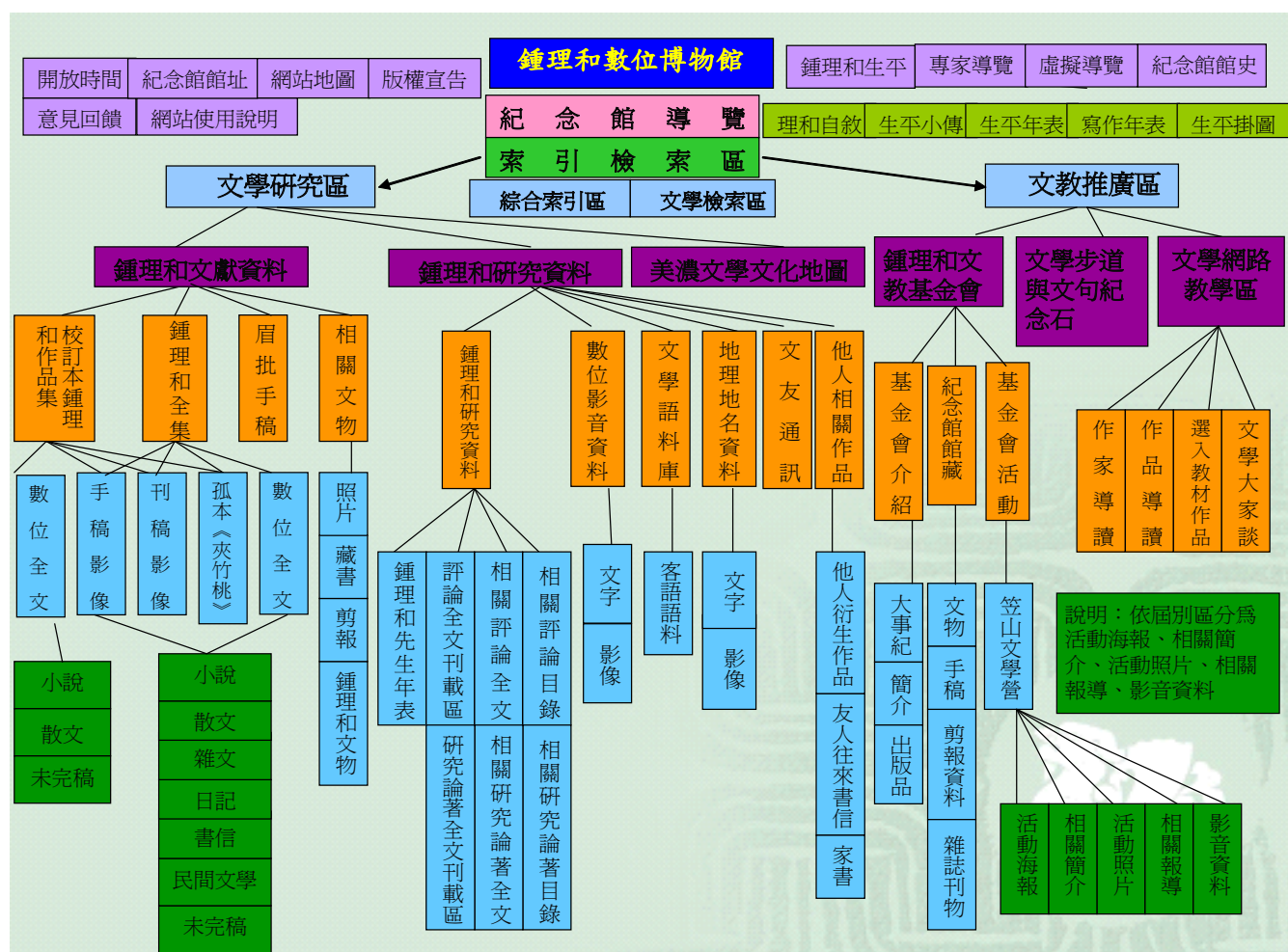
- (一) 紀念館導覽：包含四部分，其一為作家生平：蒐集各版本之生平資料，如生平小傳、生平年表、寫作年表、生平掛圖等，藉此能更親近瞭解賴和先生的文學生命與現實人生，如靜宜大學陳建忠老師所提供的〈賴和生平介紹目錄〉與〈賴和生平與創作年譜〉。其二為專家導覽：由不同領域的專家、學者以及和作家有相當淵源的人士，為作家及紀念館提供多元的深度導覽，藉由史學家的眼光檢視作家、藉由文學家的眼光透視作家，或是藉由作家們的後輩，親近其生活與生命，如鍾鐵民先生談其為父親所創建的鍾理和紀念館，由賴悅顏先生談賴和紀念館，由楊洽人先生談楊守愚。其三為虛擬導覽：實地環場拍攝現場實物，透過虛擬環場感受紀念館的實體空間，提供不受時空限制的服務。其四為紀念館館史：說明賴和紀念館的發展沿革及其歷史。
- (二) 索引檢索區：包含四部分，其一為綜合索引區：依據網站架構整理出索引表，綜合索引網站資料。其二為文學檢索區：提供網站作家文學作品之全文檢索，例如：二校本賴和全集、三校本賴和全集、鍾理和全集等。目前提供檢索的作家作品只有一人，但仍提供作者欄位，檢索出來的資料，包含「作者」的欄位，是預期之後會擴充到其他作家，連帶也可以檢索到其他作家的資料，亦即架構共同平台，串連相同的數位博物館。其三為文學研究區：包含二部分，其一為文獻資料：重新點校手稿、刊稿以及全集，將各版校訂本文字資料輸入成電子檔，全數翻拍手稿、刊稿為數位檔案，提供數位文字稿與手稿、刊稿交互參照閱讀的功能，並提供手稿線上翻閱的功能，讓現實生活中的閱讀機制，搬到螢幕上，使用者可以深刻感受到資訊近乎無所不能的展現，並寫定校勘記，以供研究考證；相關文物含照片、藏書、剪報、生前遺物。其二為研究資料：依資料內容主題區分成相關研究資料、數位影音資料、文學語料庫、他人相關作品。相關研究資料收錄各版

本賴和生平年表、寫作年表，以及相關研究論著、評論雜文，提供檢索目錄，瀏覽全文及使用者可以自行上傳研究資料等功能；數位影音資料收錄與作家相關的數位影音資料，如賴和烈士復祀忠烈祠實況錄影、鍾理和之妻鍾台妹談鍾理和。文學語料庫提供詞彙隨選隨譯的功能，此作法是先建立方言詞彙、方言音標、語意解釋的資料庫，再利用 JavaScript 的程式搭配所產生的效果。這裡所謂的「隨選隨譯」意思是指當滑鼠拖曳至方言詞彙的地方，會立即浮現該詞彙之音標和語意解釋，而非選取一段文字，就會出現整段翻譯。將來希望可以進一步做到語音的紀錄，希望能保存完整的語音資料。其四為文教推廣區：此單元分別為文教基金會、文學網路教學區，介紹基金會及其館藏手稿、文物資料，基金會相關的藝文活動、文學步道、作家及其作品導讀、選入教材作品、文學討論園地。文獻資料內容的部分，介紹基金會大事紀、簡介、出版品、紀念館館藏手稿、文物、基金會活動海報、相關簡介、活動照片、相關報導、影音資料等成果。

- (三) 其他：多元動態導覽，依使用者與內容媒體作為不同類別的導覽選擇，如一般閱聽者、教學研究者，為兩種不同使用需求人，則呈現不同取向的導覽內容；內容媒體包含文字、圖片、照片、影像、虛擬導覽、文物、聲音等不同形式的資料，可選擇各式媒體類型的導覽。意見回饋，提供研究與教學使用。靜態網站使用說明，網站資料內容繁多，使用功能也依不同的需要而有所區別，故此處依各主題單元，清楚描述整體網站的使用說明及步驟。

作家們的原典文獻資料分析歸納後，除為建立共同平台而設計的內容之知識結構與資訊表達架構，另外可強調作家的個別差異性，以嚴謹的研究為基礎，典藏正確、完整的文獻資料。如鍾理和生前唯一出版的小說集《夾竹桃》，內含〈夾竹桃〉、〈游絲〉、〈薄芒〉、〈新生〉等四篇作品，應是作者自費，由北平馬德增書店出版印行。網站呈現的線裝本《夾竹桃》是僅存的孤本，由鍾家珍藏，內有鍾理和先生親筆眉批，本網站從鍾理和的文學作品中，將孤本《夾竹桃》獨立成主要架構元素之一，整本翻拍，重新輸入、點校，以供原稿與輸入稿作圖文對照瀏覽。又如鍾理和藏書有眉批的部分，以數位影像檔獨立呈現，從這批資料，可以知道鍾理和生前閱讀那些書以及在書上留下那些眉批。鍾理和所珍藏的《文友通訊》，是戰後台灣作家第一份聯誼性的通訊，一九五七年四月由鍾肇政發起，一九五八年九月九日刊出最後一次通訊為止，前後共計維持了一年四個月之久，共發行十六次。參加的作家有：陳火泉、鍾理和、李榮春、施翠峰、鍾肇政、廖清秀、許炳成、許山木、楊紫江等九人。主要通訊內容是報導各文友動態，包括生活的近況，和寫作成果、得獎、出版消息、互相交換作品觀摩與發表對傳閱作品的見解。本網站將這份珍藏數位化，可詳見《文友通訊》的原稿以及文字記錄。或從鍾理和文學作品中，找尋美濃客家農村之文學、風俗、民情、文化足跡，建立美濃文學文化地圖，呈現美濃地區特殊的客家人文風貌。其他如賴和學生時代的醫學筆記跟其他筆記，彰化區作家楊守愚、賴賢穎的文獻資料以及相較於鍾理和基金會活動的笠山文學營

鍾理和數位博物館架構圖



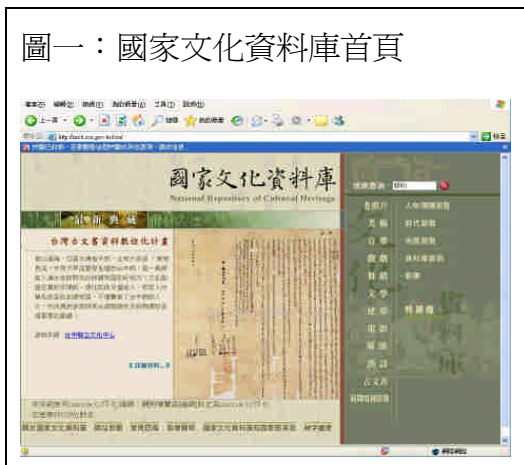
四、建立詮釋資料

文建會國家文化資料庫提供了一個全民參與文化保存的機制，整合全國的文化資料，並透過各種檢索機制開放給全民共享，如此一來資料將可經由不同使用者的創新發展，產生活化與再利用效果，更可以經由網際網路的強大傳播功能，達到全民文化提昇與文化對外傳播的目的。為了方便資料的交換與整合，所有資料都逐筆建立詮釋資料（Metadata）。

詮釋資料簡言之，是對資訊資源屬性的一組描述，目的在促進資料系統中對資料之檢索、管理與分析，具有傳統目錄之「著錄」功能，目的在使資料的管理維護者及使用者，可透過詮釋資料瞭解並辨識資料，進而去利用和管理資料。主要功能有二，一為提供後端詮釋資料輸入及系統管理之相關功能，一為於前端網站上提供民眾瀏覽查詢功能。前端網站上瀏覽查詢功能包含：使用者可依作品類別、人物與團體、時代、地區、及資料庫等分類作關鍵字邏輯運算查詢，查詢結果採圖文式排

列，依名稱或年代排序顯示。以「國家文化資料庫」(http://nrch.cca.gov.tw/cca/) 為例，利用所建置的詮釋資料誇資料庫查詢「賴和」，查詢的資料庫類別包括：老照片、美術、音樂、戲劇、舞蹈、文學、建築、電影、報紙、漢詩、古文書、新聞電視影像等十二類，例如以「賴和」查詢，結果如下：

圖一：國家文化資料庫首頁



圖二：「賴和」查詢結果清單



圖三：閱覽詳細詮釋資料



圖四：從「文學」類查詢「賴和」



國家文化資料庫詮釋資料格式依據都柏林核心集 (Dublin Code) 為基礎，再加以延伸，目前依主題、形式分為：音樂、美術、戲劇、舞蹈、文學、電影、建築、古文書、老照片、新聞、編年史等十一類。九十三年度行政院文化建設委員會國家文化資料庫詮釋資料規定之必備著錄欄位範例說明如下⁷：

⁷ 參考國家文化資料庫知識管理系統 (http://km.cca.gov.tw/) 詮釋資料格式相關規範。

欄位名稱		範例說明	
type 資源類型	worksType 作品類型	建築作品	
	Original Surrogate 原件與否	重製	
	Aggregation Level 藏品層次	單件	
Format+ 資料格式	Medium 媒體類型	照片	
	Digital Scale+ 數位化規格	Digital Category 數位化類別	影像（圖片）
		fileType 檔案類型	jpg
		File Name 檔案名稱	cca300001-ar-hpoe001-0001-w.jpg
		File Desc 檔案描述	永久保存兼網路下載
		resolution 解析度	300dpi 為單位。
title+作品名稱		玻璃詩之一	
subject 主題與關鍵 詞	Subject Matter+ 內容主題	〈十日春霖〉詩三首錄一	
	Keywords+ 關鍵詞	賴和 十日春霖 漢詩 玻璃詩	
description 描述	abstract 摘要	毛玻璃，是杜忠誥老師所選出的詩句去雕刻。杜老師也是彰化人，特別挑選賴和詩句，雕刻在毛玻璃上。詩句如下：「心情俗化久無詩，墜落雖深卻不悲。要向民間親走去，街頭日作走方醫。」摘錄於賴和〈十日春霖〉詩。	
	conditions 保存狀況	良好	

	Acquire Method 收藏取得方式	陳朝林先生拍攝		
creator+ 創作者	Creator Role 角色/頭銜	攝影師		
	Creator Name+ 姓名	陳朝林		
date+ 日期時間	Date Type 日期類型	拍攝日期		
	dates+日期	2003-11-05		
identifier 識別資料	Local Number 地方識別號碼	Local File Name+ 檔案名稱	cca300001-ar-hpoe001-0001-w.jpg	
	Metadata ID 詮釋資料識別碼（具唯一性）		848.1-093-210004-ar00001	
coverage 時空涵蓋範圍	spatial+ 空間類別	Place Type 地點類型	拍攝地點	
	spatial+ 空間類別	place+ 地點	賴和紀念館	
rights 權限範圍	Copyright Restrictions 著作權/使用 限制	Rights Statement 著作權授權狀態	{ 20031105, }	
		Access Restrictions+ 使用限制	網際網路公開展示與檢索 非營利、教育及研究使用	
	owner 典藏者	Owner Country 典藏單位國家	中華民國	
		Owner Name 典藏單位	本數位作品由元智大學羅鳳珠主持之網路展書讀網站及行政院文化建設委員會國家文化資料庫分別典藏。	

各類別詮釋資料內容及規範，請參見行政院文化建設委員會，國家文化資料庫知識管理系統網址：<http://km.cca.gov.tw/>。

「賴和數位博物館」與「鍾理和數位博物館」是行政院文化建設委員會國家文化資料庫的計畫之一，所以在執行典藏品詮釋著錄時，必須依循文建會各類詮釋資料標準建置詮釋資料，以便未來資料之交換共享可以順利進行。

詮釋資料格式依據不同主題、形式設計其欄位，藉由交換方式或是知識分類方式，達到跨資料庫的檢索，容許不同研究領域的資源共享的功能，例如由資料庫檢索「笠山」兩個字，即可顯示文學作品類別中的鍾理和長篇小說《笠山農場》、美術類別中畫家曾文忠筆下的「笠山」水彩畫、老照片類別中的「笠山風情照」等資料。又如檢索「美濃」二字，即可查詢到音樂類的美濃交工樂團所演奏之「美濃印象」，老照片類的「美濃風情照」，古地圖類的「美濃文學文化地圖」等資料。

伍、預期的加值應用與研究

引用資訊科技作為文學研究的工具已是必然的趨勢，因此，如何藉助資訊科技在網路環境裡建立台灣文學數位資料中心，結合典藏、研究、教學、推廣等多元功能，內容是成敗的關鍵，應由內容專家主導，並與資訊及媒體專家充分合作。由內容專家負責提出需求，從使用者的角度去思考網站的呈現方式，資訊專家協助滿足需求，所以在網站整體的規劃上，程式設計應以滿足使用需求為前提，而非要求使用需求去遷就程式設計。十多年前中央研究院張仲陶教授曾經說過：「只要使用者提得出的需求，在現今資訊環境下，應該沒有做不到的事。」十多年後的今天，資訊環境的進步一日千里，在漢學研究數位化累積的基礎下，台灣文學的數位化工作，已經有更成熟的資訊環境，因此，只要使用者提得出的需求，技術層面都足以滿足。

如何將台灣文學文獻數位化，並將多種媒體資料結合，或將單一資料提供多元加值利用，以開展新的研究議題，擴充研究領域，是文學引用資訊科技需要思考的方向，本文將文學資料建置為數位博物館，其研究與加值應用的方向略述如下：

- 一、網站知識結構與系統架構需預留成長與擴充的空間：以鍾理和數位博物館為例，設計知識結構與系統架構時，利用開放式的架構建立共同平台、跨資料庫檢索功能，節省程式開發的成本，以期將來可以建立更多台灣文學家的數位博物館。所以在設計資訊表達結構時，將「文學步道與文句紀念石」的架構獨立，以文句紀念石之作家及文學資料作為繼續擴充的延伸點，從美濃地區文人擴充至高雄地區，再擴充至全省。
- 二、隨時從使用者的角度思考改善的空間：以鍾理和數位博物館典藏三十五幅「施明正素描」畫為例，在鍾理和紀念館成立之初，原本預計要繪製一百幅台灣文學家

鉛筆素描，贈送給紀念館典藏，可惜施明正先生尚未完成此願望便過世了。目前鍾理和數位博物館將這批珍貴資料掃描放在網站上，所陳列的台灣文學家素描畫的詩人有：管管、陳千武……；評論家：何欣、尉天聰……等人，對使用者而言，也許不認識這些詩人、評論家，於是在網站內容設計上，加入每一位詩人、評論家的簡介，介紹其生平、作品等，讓使用者一邊欣賞文學家素描畫像的同時，可以進一步了解詳細資料。又如紀念館虛擬實境導覽，目前網站上提供三百六十度的環場導覽，事實上還有七百二十度的環場和其他比較理想的呈現方式，如將虛擬實境設計成一般直向行走的導覽方式等。

- 三、如何與新觀念結合：以本體知識為基礎的語意概念檢索，將全文檢索從字形搜尋提升到智慧型代理人的層次。現今的全文檢索只能分辨字形，依輸入的字型檢索出結果，但是我們知道，文學作品中有許多通同義字詞，譬如要檢索「月亮」，若作品使用「嬋娟」，便檢索不到這筆資料。學界很多人正在進行「以本體知識為基礎的語意概念檢索」的相關研究，輸入「月亮」，除了可以檢索到含有「月亮」詞彙的資料之外，還可以檢索出「玉盤」、「玉輪」、「玉蟾」、「冰輪」、「桂魄」等其他屬於「月亮」語意概念的資料，如此便能將全文檢索從字形搜尋提升到智慧型代理人的層次。
- 四、如何與新工具結合：與地理資訊系統結合，全面建立文化地理資訊系統；與 PDA 結合，提供隨時隨地上網的機制。譬如：整合台灣各地的文學文化資料庫，與地理資訊系統結合，全面建立文化地理資訊系統，建立整個台灣文化地理資訊，從文化地圖上搜尋「苗栗」的文學資訊，即會出現「李喬」、「《寒夜》」；搜尋「桃園」，即會出現「鍾肇政」、「《魯冰花》」的資訊；或者說與 PDA 結合，提供隨時隨地上網的機制，如任何時候在 PDA 都可獲得鍾肇政長篇小說《魯冰花》電影的拍攝地點，或是李喬大河小說《寒夜》電視劇拍攝的地點，讓使用者可以循其足跡，感受電影戲劇小說人物的生活環境。
- 五、詮釋資料如何與網站結合，提供更多元的知識加值利用：如目前文建會國家文化資料庫建立一個完整的詮釋資料系統，也許是因為計畫執行之初，還來不及兼顧，除了文建會的「國家文化資料庫」網站依詮釋資料提供跨資料庫查詢功能之外，只見其他子計畫建立很多詮釋資料，但卻無法有效的運用在網站的資訊表達上，造成詮釋資料與網站各自獨立的狀態，失去了建立詮釋資料的意義，殊為可惜，期待往後兩者可以結合起來提供更多元的知識加值利用。
- 六、如何提供更方便的使用管道：網路版與光碟版同時製作，除了 PDA 之外，建議網路版與光碟版同時製作發行，以方便網路連線速度緩慢的偏遠地區使用，跨資料庫查詢等無法在光碟版呈現的功能，保留在網路版。
- 七、語言學研究與教學的加值利用：以語言來說，逐步建立每一個文學家文學作品的

語料，針對其時代、區域性等主題，做進一步的語言學比較與研究，成為台灣文學語料與研究資料庫。語言學研究還包含歷史語言學，如筆者曾經執行製作一個以閩南語撰寫的《荔鏡記》的計畫⁸，《荔鏡記》紀錄的是十六世紀的閩南語，我們將十六世紀的閩南語與現在的閩南語詞彙作比較，可以發現因為文字演變比較慢，語言演變比較快，經過幾世紀，語言已經有一些差別了，建立語言的時空座標，可以做語言學研究與教學的加值利用。

八、其他如文建會國家文化資料庫中各類型的網站，能彼此互相連結利用。譬如鍾理和網站中談到飲食文學部分，除了這個網站資料以外，可以連結到國家文化資料庫裡的其他相關飲食專業網站，或如鍾理和網站中，整理客語詞彙資料，建立客語自動標示讀音以及客語語音合成，將來在客語語言的教學上，也是一種加值利用，站在使用者的角度去設想，做延伸性的計畫。

陸、未來的發展

文獻數位化的工作，雖然已經執行許多年了，但各計畫執行似乎都是各自獨立運作，分散在國家台灣文學館、各級學校圖書館、各縣市文化中心、各研究機構圖書館、各地方文史工作室、私人蒐藏文獻資料、私人圖書館、個人建置之網站等，缺乏資料交換與經驗分享的機制，如台灣文學館的「作家手稿網站」、「全臺詩電子文庫」；個人建置的網站，如：「陳虛谷文集網站」、「台灣古典漢詩網站」等，目前尚無最完整、最正確、最權威的整合性台灣文學網站，因此，台灣文學文獻數位化仍有很大的空間可以做更細緻、更宏觀的規劃。

台灣文學文獻數位化在前人的努力下，已有為數不少的成果，如何在前人的基礎上，建立更宏觀、細緻的規劃，筆者提供未來研究發展的方向如下：

一、以作家為單元的台灣文學網站：

為了兼顧台灣文學網站的整體性，並保留每一位作家的獨立性，以及未來陸續增加其他作家的擴充彈性，在以作家為單元的台灣文學網站的規劃，系統方面宜採用共通欄位的開放性架構，內容及網頁設計方面，採用可延伸性的架構。以鍾理和、賴和數位博物館為例，共同性的內容架構包含：一、作家生平資料：生平小傳、生平年表、寫作年表；二、作品文獻資料；三、研究論著資料等三類。每一位作家的個別網站都包含這三類的資料，當由兩個以上的作家整合而成的台

⁸ 參見「荔鏡姻，河洛源：閩南語第一名著《荔鏡記》多媒體數位教學網站」
 (http://cls.hs.yzu.edu.tw/LM/Lm_home.asp) 及「閩南語典藏：歷史語言與分布變遷資料庫」
 (<http://cls.hs.yzu.edu.tw/southernmin/index.htm>)。

灣文學網站，如能讓使用者從整合網站的介面，從各個使用需求的角度看到所有作家的整合資料，將可提高使用的效益。以作家生平資料為例，從作家生平資料可以查詢個別作家某一年的寫作資料，也可以查詢所有作家在某一年的所有寫作資料；以作品文獻資料為例，查詢範圍可以限定在某一位作家、某幾位作家或所有作家；研究論著資料的查詢亦同。要在各作家個別網站與全部作家整合網站達到上述功能，資料建立體例的統一非常重要，例如年代資料的建立、地名資料的建立都需要有統一的標準。⁹

二、依作品體裁類別建置：

目前學術網路上已有的文學資料庫，依作品體裁類別建置的網站有「全唐詩」、「唐宋詞」、「宋詩」等等。「全臺詩電子文庫」由國家台灣文學館策劃，成功大學施懿琳教授主持，以明鄭至日治時期（1661—1945）近三百年間台灣文學中的古典詩為範圍，「結合國內多位古典文學研究者，共同進行台灣古典詩的蒐集、校對、編纂，俾能提供一部資料齊全、校勘精確的《全臺詩》，做為國人閱讀、研究之用。」¹⁰編輯體例採取「以人繫詩」的原則，逐一撰寫「作者提要」、「作品」、「註」資料，網站上提供「按作者姓氏字首筆畫瀏覽」，「不能區分欄位的全文檢索」二種功能。如就主持人所期待的，希望《全臺詩》的編纂出版，提供研究使用，未來可以與資訊科技作更好的結合，以開拓研究的領域，提高研究者使用資料的方便性。《全臺詩》的內容包含作者、詩題、詩序、詩文、註、參考書目資料、作品出處資料等，除了全文檢索的欄位應予以區分之外，以現在的資訊環境，可以開發更多功能的智慧型《全臺詩》電子資料庫，內容部分可以與前述之台灣文學研究資料中心整合，詩人之生平資料、作品之地理資訊資料可以整合為台灣文學地理資訊系統。

三、與地理資訊結合的台灣文學地理資訊系統。藝術創作的風格會受到地理景觀環境的影響，音樂如此、繪畫如此，文學也如此。地理景觀對文學的主要影響是作家利用自然景觀為意象所形成的文學內涵與風格，其次是文學的語言與聲韻的變化。對作家而言，不同的地理景觀給予不同的創作靈感與素材；對地方而言，文人的流動對當地的文風帶來影響，因此文學地理時空資訊系統的建立，除了可以提供文學與地理景觀交互影響的研究、文學風格意象的研究，還因為文學的內涵所呈現的民情風俗，文人的遷徙流動所帶動的文化氣息，可以作為文化盛衰變遷的研究。如：「以 XML 可延伸式標註語言建立文章標誌系統研究—以蘇軾詩為範圍」網站¹¹，以詩的內容標誌方法建立蘇軾詩的時間、空間標誌，再以此時空內容標

⁹ 參見筆者與陳萬益共同發表之拙文，〈挑戰 2008：如何建立一個兼具客家語言復甦及傳播功能的客家文學網站〉，2002 年 10 月 30-31 日，客家文化學術研討會、行政院客家委員會委託、中央大學客家研究中心主辦。

¹⁰ 施懿琳，〈全臺詩電子文庫編序〉，<http://www.wordpedia.com/twpoem/about.asp>。

¹¹ 羅鳳珠、李錫捷，「以 XML(eXtensible Markup Language)可延伸式標註語言建立文章標誌(Content Markup)

誌的基礎建立蘇軾詩的文學時空地理資訊系統，藉以探討文學時空地理資訊系統的設計方法以及在文學研究方面的應用、其他文化層面擴充性的應用與研究。¹²台灣文學地理資訊系統可以此構想繪製屬於台灣地區特殊的地理人文風貌藍圖，建置以台灣文學為起點，以台灣文學為軸心的「台灣文學地理資訊系統」，內容包含台灣文人地理資訊資料與台灣文學作品地理資訊資料。

柒、結語

台灣文學文獻資料的典藏，通常透過兩種方式，其一是文學家個人蒐藏，如文壇耆老鍾肇政先生；其二是建立紀念館典藏，成立文教基金會，如賴和紀念館、鍾理和紀念館。文獻資料的管理及典藏，若是由文學家私人典藏，資料較易發生流失的情形，而且因為沒有完善、充足的典藏空間及經費，導致手稿、文物褪色、泛黃、腐蝕，使珍貴資料損壞。若是由紀念館統籌典藏與推廣，文學資料的蒐集將更加完整、豐富，進而舉辦各項藝文活動推廣。如賴和文教基金會及賴和紀念館成立後，秉持賴和精神，發揚台灣文化。除館藏有完整之賴和遺物、藏書、字畫、手稿及相關文獻資料，並陸續蒐藏、展示彰化地區作家之手稿文物。為了讓文化紮根、落實本土教育，幾年來一直持續不斷地積極推動舉辦各類相關活動，頒發賴和醫療服務獎與文學獎、頒發台灣文學博碩士論文獎學金、舉辦各種文學講座與文學活動、獎助優良台灣文化圖書雜誌出版、其他相關活動等，希望透過這些文藝活動，能夠讓文學種子回歸到孕育生長的土地上，發芽並且茁壯。讓民眾重新認識台灣的人文之美，重視自己的鄉土，關心社會周遭變動事物，進而對台灣文學、文化有更深入的瞭解，正確的去認識台灣文學的文化美感。

文獻資料數位化的工作，有賴於資料的完整典藏，數位典藏的資料需建構在嚴謹的學術性知識體系及邏輯性的資訊分類基礎上，才能提供便捷的、符合使用者邏輯思維的使用介面，如「台灣客家文學館」網站，其中所收錄的文學家，文獻資料屬於個人蒐藏，未經適當分類，「賴和數位博物館」網站與「鍾理和數位博物館」網站，資料已經由紀念館作初步分類，由紀念館提供典藏的文學家作品、文物，鍾理和紀念館在高雄縣政府的資助下，建置文學步道與台灣文學家紀念石，資料完整而豐富，典藏文物已有系統性分類，相對的，取自文學家個人蒐藏而建置的數位博物館，內容便相對的顯得貧乏與沒有系統。

因此，如何將文學家文獻資料有效的作系統管理與典藏，加速文獻資料數位化的工

系統研究－以蘇軾詩詞為範圍」，網址：<http://cls.hs.yzu.edu.tw/cm/>，國科會，2002年7月1日至2003年6月30日。

¹² 參見拙文，〈文學地理時空資訊系統設計與應用：以蘇軾詩為例〉，第二屆數位地球國際研討會，2004年5月27-28日，中央研究院、國家實驗研究院、中國文化大學主辦。

作，是未來需要思考的方向。筆者認為台灣文學文獻數位化的工作，宜優先處理已有相當研究成果的資料，吸引有研究成果的學者參與；在建立網站的過程中，以大格局與寬闊的胸襟，建立著作所有權人對執行團隊的信任，以製作的品質及網站的權威性吸引著作權人以著作能被收入台灣文學網站為榮；參與的團隊定期舉辦研討會，彼此交換心得，培養人才。執行先後順序原則如下：

- 一、以資料性質區分：已建置的網站先作整合，如「賴和數位博物館」、「鍾理和數位博物館」、「陳虛谷文集網站」、「葉榮鐘網站」……等；已數位化的素材先重新建置網站，如台灣文學館的「全臺詩電子文庫」；已整理好的資料先做，如胡萬川教授蒐集整理的「民間文學」；已取得授權的資料先做，如「楊守愚文集」。
- 二、以典藏區分：私藏優於公藏；已整理好優於未整理；已數位化優於未數位化。
- 三、以作家區分：已過世的作家優於在世的作家；影響層面廣的作家優於影響層面小的作家。
- 四、以重要性區分：對台灣文壇影響層面較廣的作家優先，如鍾肇政、林海音等。

台灣文學在很多前輩的努力之下，直到這幾年，隨著大學院校之台灣文學系所成立，台灣文學館設立，研究與親炙台灣文學不再是禁忌，在台灣文學受到壓抑的年代裡，台灣文學環境誠然艱困，當此之際，引用資訊科技做為台灣文學典藏、研究、教學、推廣的工具，透過無遠弗屆的學術網路，讓台灣成為台灣文學的資源中心，是將台灣文學推向國際舞台最有效的方法，政府機構如能支持有心人，這個構想將可逐步落實。

感謝與說明：

- 一、本研究之執行，感謝文建會提供經費，鍾理和紀念館鍾鐵民先生、賴和紀念館賴悅顏先生協助，使研究得以順利進行；感謝鍾彩焱老師、許介彥老師協助文物拍攝工作以及許榮達校長、楊洽人先生義務幫忙，使研究得以順利進行。本文由羅鳳珠、何淑華執筆，參與研究的成員有陳萬益老師、林瑞明老師、鍾怡彥老師，蔡宛純、林亮君、黃欣慧負責網站程式設計，何淑華、林淑鈴、林宜萱、蔡明諺、黃詩媛、吳姿慧、周馥儀、張綵芳協助資料整理建置。也感謝提供研究及著作資料授權的所有單位及個人。
- 二、本文完稿於 2004 年 10 月，文中所引資料以 2004 年 10 月為準。