

# 中國文史哲工具書編製的 幾點思考

劉寧慧 臺北大學中文系文獻學研究所助理教授

【摘要】臺灣地區數十年來中文文史哲工具書的編製在質與量上均有相當的發展成就，近年在資訊化潮流下，數位型態的工具資料庫也在開發。然而數位化的初期階段，以資料庫的遷就檢索功能，卻形成利用上的若干障礙，乃至部分失卻紙本型態可提供的效能。本文闡述於此紙本工具資料庫發展過渡階段，在未來可能出現的趨勢之前，提出以長期紙本編輯經驗的思考與成就，注入資料庫的設計與建置，希望工具性資料庫的發展能夠更為理想地繼承紙本成就，並真正利用資訊優勢，活化工具書的利用功能。

關鍵詞：工具書；文史哲研究；數位化；資料庫；目錄編輯；索引

## 一、前言

筆者所處的世代跨越紙本工具書時期以及轉變為資料庫建置以後，前一個階段還完全仰賴人力編輯，以抄寫及排列卡片為主要的編製方式；後一個時期則已經是電腦輸入、線上公用檢索，以及講究後設資料等的資訊管理形式。在這形式、媒體的巨大轉變中，文史哲工具書的編製也歷經了內容與形式上的諸多變異與演化。這些年來，我們一直在嘗試新的工具書形式，也做了許多適應，基本上是逐步脫離了對紙本工具書的依賴，而轉向資料庫的檢索搜尋。習慣在改變，方

法在更新，編輯思考的重點也在轉移。然而追逐了多年以後，我們真的感到更進步與完善了嗎？除了便利、快捷，我們的資料查詢需求完全被數位化所滿足了嗎？現在是不是一個時機對這股潮流所形成的景況乃至問題作一些反省與思考？是不是應該更深層並細膩地考察在這個形式轉變中，我們因過度偏重某種檢索功能而帶來了盲點？這個問題不僅應直接檢視資料庫的製作，還應包括對紙本型態的借鑑。以一個在兩階段分別有不同使用習慣及製作經驗的工具書關切者來說，有些疑惑實在不得不提出來。筆者的領域在中國文史哲，特別是文獻方面的研究，工具書與



資訊資源的編輯與製作都是主要關切內容，近年也參與過一些工具性資料庫的建置，因此本文將略談面對這幾年轉變的使用與觀察心得，並提出幾點想法就教於各前輩先進。以下的討論將分內容編製與呈現製作兩方面來談，它們的對象包含紙本與資料庫形式的工具書。

## 二、內容的編製——以區域性總結為主的思考

### (一) 臺灣地區典藏資料及研究成果的總結介紹

臺灣典藏有為數眾多並珍貴稀罕的古籍，但自 1960 年代後期初步整理迄今，還沒有一部公藏古籍的分類聯合目錄，也沒有形式、內容完整且精詳的聯合資料庫。近年來，大陸、香港、日韓及北美等地的漢籍古籍典藏單位，都紛紛推出紙本目錄或聯合資料庫，展現該館、該地區典藏的情況，而重要典藏地的臺灣，卻一直未能完成這項工作。臺灣目前的紙本古籍目錄仍停留在 1970 年代初期編輯的各單位簡要書目，三十餘年來的單位裁撤、改易及收藏的增減等變化，都未獲有新穎的反映。而資料庫的建置又受限於各單位分別製作，內容的呈現詳簡不一，古籍相關訊息的提取也還未能在網路上全面滿足。這些情況不僅對在臺灣從事古籍研究的學術人是一大障礙，對彰顯臺灣地區的典藏也是憾事。又如臺灣地區典藏的豐富臺灣史料，像中央圖書館臺灣分館的珍稀臺灣史料等，是海內外所殷切期待進一步整理公布的珍貴資源。

此外，臺灣地區的學術研究成果，六十餘年來一直是世界漢學圈中受到重視的一環，專書與

期刊論文等資料很早就開始有紙本目錄的蒐集，發展為資料庫檢索也已相當完備，不過學術資料中仍缺乏一項內容的開發，即論文集篇目的分類析理。臺北大學古典文獻學研究所與國家圖書館合作建置有「館藏臺灣近六十年文史哲論文集篇目資料庫（1945-2005）」，雖然收錄不過三千餘部論文集，近六萬篇篇目，但因都是臺灣地區所編印出版的內容，且過去還沒有過類似的整理，據日前到大陸參加圖書館會議的國圖人員表示，大陸的圖書館人員便注意到了這個資料庫。以在地的優勢，呈現地區的總結性成果，不僅以紙本形式表現典藏的詳細內容與結構，也發展成線上資料庫，便捷查詢檢索，乃至全文圖檔的掃描呈現，相信將可以提升國內相關研究水準並引起海外重視。

### (二) 海外漢學研究成果的中文工具書編製

在臺灣蒐集學術資料或者編製專題、專科論著目錄，臺灣本地及大陸地區的資料還有為數不少的索引目錄可資查找，但海外地區像日、韓、美、歐等，能利用的資源便極為有限。日本的《東洋學文獻類目》編輯出書較緩；韓國的漢學研究目錄也多屬個別專題成果，且涵蓋的時間與內容範圍都不廣；西方歐美成果則常見到選目或是以英語等西文為主的資料庫內容，查詢時往往有不知如何下關鍵字搜尋的問題，而瀏覽也因動輒數千筆有相當的困難，是以這些海外成果在收錄上總有無法完善及完整的困擾。如果臺灣能與海外各國漢學單位合作，編製這些東亞及西方漢學研究成果的中文版漢學研究論著分類目錄，廣為採取漢學各門類的研究成果，以符合中文資料查找習慣的形式編排，語文方面當然是中外對照，



相信對臺灣學術界深度了解海外成果以及作為專科工具書編製採取利用，都有莫大的助益。今天我們編輯中文內容為主的索引目錄，也實應補強過去所一直忽略的外語漢學研究成果資料。

### 三、形式與方法——資料庫非萬能，紙本功能仍值借鏡

#### (一) 資訊化潮流下的反省

資訊化以來，圖書單位的管理也隨之大幅調整。多數的館藏圖書目錄系統主要在提供檢索，不能瀏覽，失去了過往卡片的功能。期刊篇目資料庫主要也僅提供關鍵詞查詢，分類瀏覽的設計未盡理想。在資訊化的發展下，學術資料的結構性認識趨於模糊，僅存提取式的資料擷取。又如國家圖書館漢學研究中心過去十幾年曾將所藏大陸期刊每期目次頁影印下來存放資料夾，但至2002年已全面停止影印工作，這些主要都被線上公用目錄、檢索資料庫的發展所「取代」。然而它們的功能真的完全被替代了嗎？事實是，關鍵詞沒有辦法查全所需的資料，即使交叉使用多個關鍵詞或進行全欄位的搜索，也不一定能保證蒐羅完全。例如古籍叢書研究成果，僅鍵入「叢書」一詞，會出現大量訊息，但多數是非需求的雜訊，真正需求的內容則因沒有「叢書」兩字而無法提取。資料庫固然有強大的全文搜尋功能，但建置者除了考慮明確資料的查詢外，也應考量非特定名詞資料的查找。尤其對於像專題論著目錄的編製，除了既定詞彙以外，還需要判斷過濾許多相關類別與內容的資料，在紙本式全國圖書、期刊論文目錄早已全然停止編印的情況下，資料庫設

計必須重視分類瀏覽的功能便日益凸顯。另如筆者十餘年前開始進行漢學研究中心所藏期刊篇目選錄工作，藉以編輯專題論著目錄及掌握最新學術發表成果，但至2002年後幾近停滯，相關研究訊息的更新也自此零落緩慢。因期刊無法全面開架，資料庫收錄也不全，索引目錄編刊（指大陸上海圖書館《全國報刊索引》之類刊物）又有時間落差，傳統資料搜尋方式遭遇不小的挫敗。資料庫的問題不僅在關鍵詞的查全與查準率，還有內容本身的限制。「中國期刊網」、「世紀期刊」的收錄雖然號稱擁有中國大陸數千種核心刊物，但也並非所有領域的重要期刊都已全備，像大學學報及諸多圖書館學刊物等都還未盡收錄，而收錄的內容中又有資料項目不全的問題，像「中國期刊網」的論文條目不表現起迄頁碼與發表年月等。上述情況使得近幾年來的學術資料反而不容易彰顯，也欠缺結構性的認識，專題論著目的資料蒐羅更形成一種無奈的困境。

#### (二) 資料庫檢索形式應結合紙本工具書功能

資料庫的強項在檢索，它的壯大也正以這項優勢而來。如果資料庫還要逐筆瀏覽內容，似乎就沒有資訊化的必要了。這樣的想法沒有錯，但電腦不具人工智慧，對它輸入及提取資料都需經由控制，而控制不能無比複雜與模糊，所得的結果當然僅能囿於有限條件之下。然而對學術研究及專題論著目錄的編製來說，所需求的絕不僅是特定詞彙下的資料，它可能使用其他詞彙或者相關但不具特定字詞。新階段的資料庫製作者應該考量到這些需求。目前已有不少資料庫在檢索上加進了目錄點選及分類瀏覽的項目，以前述古典文獻學研究所編製的「館藏臺灣近六十年文史哲



論文集篇目資料庫（1945-2005）」為例，為了提高查全率，關鍵詞內容的建置儘可能求多，不足之處就以分類瀏覽來補足。儘管編製過程時間倉促，分類還有許多問題，但大項的區分仍值得信賴。此外，未來也打算加進所收論文集一覽表，再藉助系統的靈活性，恣意切換全數論文集、單部論文集及論文集的各篇篇目，讓紙本工具書曾發展出來的功能亦表現於資料庫中，而資料庫再利用它的優點活化傳統的著錄方式。

### （三）資料庫製作或重編例、製作過程之說明

目前的資料庫製作，思考重心幾乎都在檢索功能，以實用為主。而資料是數位形式的、虛擬的，前端使用者看不到全部的內容，也不需要看到內容，當然也不一定知道是誰主導編製，在什麼規畫下，有哪些細部的資料去取原則與安排。使用者多半搜尋完就離開，對所搜得的資料具有何種背景與限制卻不甚關心。紙本時代的編輯，一定有編序、凡例乃至附錄等內容，說明編輯收錄等原則，讓讀者明瞭資料的情況與定位。但資料庫時代，這些內容卻被忽視了。大部分資料庫只有無比簡略的說明，而且還多為初發布時的訊息，後來更新的情況與未來計畫常都未見。紙本工具書因可見到內容，還可以藉收錄情況進行判斷，資料庫則毫無掌握辦法。對專題目錄的編製，只能說交叉使用、多比對幾個資料庫，但難以確切查出它們彼此空缺的部分。固然讀者不喜網路上連篇累牘的說明文字瀏覽，但製作者可以檔案加掛的方式，提供研究者及特定使用者掌握相關訊息。它的內容應該包含主題說明、收錄時間範圍、地域範圍、資料型態、資料來源、收錄標準、去取原則、更新速度、發展方向、製作起迄、資料內容編製者、系統製作者、系統設

計等等，應自此發展出一種資料庫型態工具書的完整說明體例。

### （四）設計目錄資料庫專門系統，便利輸入與整合查詢

紙本目錄時代即有各目錄著錄情況、項目差異的問題，造成使用者必須各別適應。工具資料庫的製作，則應有能力改善這種不足並更上層樓。前一個階段的工具性資料庫製作，由於各有不同的檔案格式與系統形式，所以呈現為各自獨立的情況，資料庫之間無法整合查詢，即使同一個單位製作的前後資料庫也經常如此。例如國家圖書館館藏圖書目錄系統、期刊論文系統與臺灣近六十年文史哲論文集篇目系統，分別由國圖不同單位製作，由於輸入過程中彼此對符號型態與題名認定的規範不同，產生目前進行自動整合查詢時的障礙。文史哲論文集資料庫所收錄的內容全為國圖典藏，但整合查詢時卻多半無法聯繫上館藏圖書資料。再者，這些資料庫製作常是先建為 excel 檔，再轉入系統，轉檔時碰上 unicode 字碼或檔案欄位局部錯移等問題，便形成前端訊息的錯誤。又資訊設計者若不了解資料內容的特殊性，往往僅以一般的檢索功能來展現結果，對於彰顯特定材料內容各部位的關係便有不足，這些都嚴重影響資料庫的品質。下一個階段的資料庫製作，應發展出專門的建置系統，完善、縝密設計後端資料線上輸入系統，控制欄位字串的輸入，讓輸入者在當下就能修正錯誤，減少校對時的瑣細勘驗。而完善的欄位，可保證訊息的完整與詳盡，多元、靈活的提取與資料切換，則可滿足不同的使用需求，整合查詢的問題也得以在相同的系統下獲得解決的辦法。



資料庫解決了部分舊問題，但也形成了新的問題。在未來的中國文史哲工具資源編製中，資料庫肯定將進行著主流的發展趨勢。我們今天致力於改善資料庫的種種不足，思考的方向或許可以轉回紙本形式，一則紙本工具書仍然有它存在的價值，對於學術資料的結構性呈現與感知，資料庫無論如何都無法取代。二則紙本資料的編輯與使用思維不見得是一種守舊，它具有穩定性與

潛在可能性，這些目前求諸於資料庫都還未能滿足。然而未來，不論是資料的查詢或是各種工具書的編製乃至資料內容的重新組構，都將適應於新生代的使用習慣，以資料庫發展為主，那麼，值此資訊潮流往下一個高峰邁步的同時，我們有必要檢視這一路探索過來的腳步，而給予更通透、清明的目標指引。

〔專科工具書編輯研討會 工具書編輯之未來展望座談會〕



▲主持人林慶彰教授（左三）、劉春銀主任（右三），發表人張錦郎教授（左一）、須山蔚教授（左二）、劉寧慧教授（右二）、余崇生教授（右一）（編輯組提供）