



「中國古籍叢書目錄資料庫」 的建置與效益

劉寧慧*

一、前言

中國古籍發展了兩千多年，留存到今天的主要是宋代雕版印刷後的成果，即使如此，也有一千年的歷史，而清代距離又最近，具體保存下來的刻本、抄本與稿本等，佔了傳世古籍中的最大量。這些廣泛的典籍作品，主要存藏在中國大陸，也分散在世界各地圖書館或私人收藏，實際的存在總數，沒有辦法統計。典籍豐富，從古典時期開始就展開有目錄的著錄，像官方藏書目錄、私人藏書目錄、史書藝文志等等，一直到民國以來至今，各典藏單位編輯的館藏目錄、聯合目錄、總目錄及學者專家編纂的專科目錄等，記載著歷代著作與流傳的古書，就是這些記錄的文獻也堪稱汗牛充棟了。但是想要通盤總結及完全收錄歷來古籍，卻是從古至今莫不有的理想與難以實現的夢想。古典時期受限典籍與著錄的私有與分散，尚無條件進行大規模總結；時至近現代，古籍多匯流公立典藏單位，相對於過去是集中，且多編輯出版有館藏古籍目錄，大眾認識的機會較過去便捷，但今天仍有新的課題要面對，那就是分散在全世界各漢籍典藏單位的中國古書要如何進行記錄的蒐羅、著錄的彙整與便捷的檢尋？將古籍的整理再提昇至新的境界，促進古典研究更上層樓。

近二十年，中國大陸《中國地方志總目提要》、《中國古籍善本書目》、《中國家譜總目》、《中國中醫古籍總目》，乃至《中國古籍總目》，都陸續發表問世了。這些成果引領、鼓舞了各古典學問領域的研究發展，本身也締造了

* 國立臺北大學古典文獻與民俗藝術研究所副教授

古籍整理新的高峰。然而古籍整理是一項沒有終點的工作，不僅因為它的領域眾多，有待總結的主題還所在多有，古籍持續發現與發表的內容，亦等待不斷地擴充與完善舊編。這些工作的理想性要求愈來愈高，課題的艱辛程度不下過往。所幸近二十年開發了電腦資訊技術，用在古籍整理、著錄編目、全文檢索這些工具本質的業務，使得古籍目錄的發展飛躍成長，迭創新猷。當然，我們又面臨了新的考驗與挑戰，必須再深入檢討古籍的情況，靈活運用時代新技術，以期發揮它更具深度的價值。而古籍叢書在這個有待致力的課題結構中，是一個既關鍵且切要的項目。說明如下。

二、古籍叢書的文獻難題

古籍叢書從宋代開始發展，是一個受到雕版印刷技術鼓舞起來的後起編纂文獻。今天了解的情況是，南宋末年出現了彙編雜纂性的典型叢書體裁，元代繼承不豐，明代後期才大量發展，清代則達到了古典時期的高峰。民國以來迄今，古籍叢書的編印風氣未歇，是一種仍不斷發展中的文獻體裁。南宋迄今九百年，民國初年以前的古籍叢書粗估六千種，每種叢書收錄數十、數百種子目不等，去除叢書間重複的子目，不相重的作品至少十萬種以上，是傳世中國古籍約二十萬種的半數。或許這其中多數的典籍還有單刻本，叢書本不見得有多可貴，但十萬種中不乏僅存叢書本的作品，這對於古籍文獻來說是重要成分，之於古典研究而言是寶貴材料。可是古籍叢書的著錄與介紹，成績還沒有達到理想。雖然從明代開始，官私藏書目錄有所反映叢書的流通，清嘉慶四年（1799）顧脩編有第一部叢書專門目錄，迄今叢書專目已發展至 80 餘種，綜合式古籍目錄則著錄更繁多，但叢書目錄還是問題重重，檢查叢書還是麻煩不便，這些原因要歸之於古籍叢書文獻本身的問題。分析其難題，有如下情況：

1. 古籍叢書文獻龐雜，分布廣泛。古籍叢書六千種的約數，是經由目前已整理的主要目錄成果所大致推定的，事實上，要蒐錄它的總體文獻，有相當難度。叢書部頭大，一部之中又是數十百千種的作品，一個古籍典藏單位能收藏的有限，叢書分散在一地的各圖書館，全球各地公私漢籍古籍典藏單位，除非大多數的單位都整理有精詳的目錄，叢書著錄亦完善無缺，否則要通盤調查與彙整，比單行本要複雜上許多。
2. 叢書編纂過程狀況百出。單行本作品即單一部圖書，書籍內容在完成及刊



刻後不易大有更動，不同版本間可能存在字句等些許變化，但少見內容的大幅差異。叢書卻不然，它由眾多作品彙集而成，初刻是一套組合，再刻就可能有所變動，增加或減少其中作品，到了重刻、重編，內容的變化更大，而叢書名稱可能並未改變。這些前後子目更易、形成過程錯綜複雜的情況，並不見得有具體資料詳述或難從叢書本身看得出道理，造成今天在整理叢書的時候一團迷霧與現象紛雜不堪。

3. 古籍典藏單位著錄叢書未盡規範。典藏單位叢書著錄一定要包含子目，而且是實藏子目，不能據目錄子目。前已言及，叢書內容在編纂及流傳的過程中經常發生變化，子目情況有所增減新異，各典藏單位題為同一部叢書間的子目都可能有差別，不同館間同一部名稱的叢書更可能有所參差，在著錄時必須詳依實藏子目內容完善記錄，如此在採錄、彙整為總目時才可能發現問題，進而釐清可能情況。然而叢書狀況實在過於複雜，不是每一個典藏單位都能特別為之考訂，有些甚至連基本的子目都未著錄，僅憑叢書名、編者、版本訊息，是無法確實明白叢書實質的，在文獻整理時可說幾近無效用。
4. 叢書認定標準不盡相同，分類不一。叢書是多種作品彙集的體裁，有另外的總名，可是這樣的描述也適用像類書、別集、總集、雜纂等彙編體裁及合刻等版刻形式，叢書非常容易與它們相混淆。或許其他的體裁都還可以再嚴格限縮範圍，但叢書就顯得籠統、無絕對獨特性，是以屬於叢書的及不屬於其他體裁的灰色地帶文獻，都歸攏到叢書這裡來，因為它的包含性邊緣最模糊。但如此一來，叢書認定就有見仁見智的部分了，各種目錄中的範圍與標準都不太一致。除此之外，叢書文獻本身的分類架構也不盡相同，各目錄對叢書歸屬自有一套判準及選擇，這些都對叢書文獻進行統整與析理時產生一定的為難。
5. 叢書零種識別有限。叢書在流傳過程中極容易分散或甚至遭到割裂，原因是叢書部頭較大，流通過程中經藏家輾轉遞藏，每一個藏家存放的方式不一。今天從明清私人藏書目錄中常可發現，藏家是打散歸藏與著錄的，然而後來的藏家可能就只獲得其中一部分子目，在前後無序跋表現原書的情況下，往往形成叢書零種，數量多寡不一，子目亦失原次序。而近現代圖書館收藏時，也沒有根據或明白它們是何種作品，但在著錄時必須有一個稱呼，就形成像「某某數種」這樣的後世自擬名稱，與原書完全無關。這使

得今天要統整叢書文獻時，必須能有效匯聚類同、子目或有關連之作品，方便進行比較與比對，盡最大可能發現其中的關連，還原零種叢書的真貌。

6. 叢書目錄編輯未精，檢尋未便。叢書文獻雖龐雜不易析理，但從清代中期迄今，仍不乏目錄整理者對之做出努力。可是各目錄編輯條件不一，呈現方式有異，索引的使用方法也嫌複雜，整體來說，還沒有一部叢書目錄能夠達到收錄完整、著錄精詳、分類統一、查詢簡便及功能完整的期望。

叢書文獻遲未能完善析理，也代表了學術界對這份龐大古籍遺產的陌生與未重視，叢書發展史與文獻內容的研究，當然也還難以進步。在過去把古籍整理的重點優先擺在善本，現在則可以細膩地討論與分析其他，叢書便是一個代表品項。

三、古籍叢書目錄資料庫的解決方案

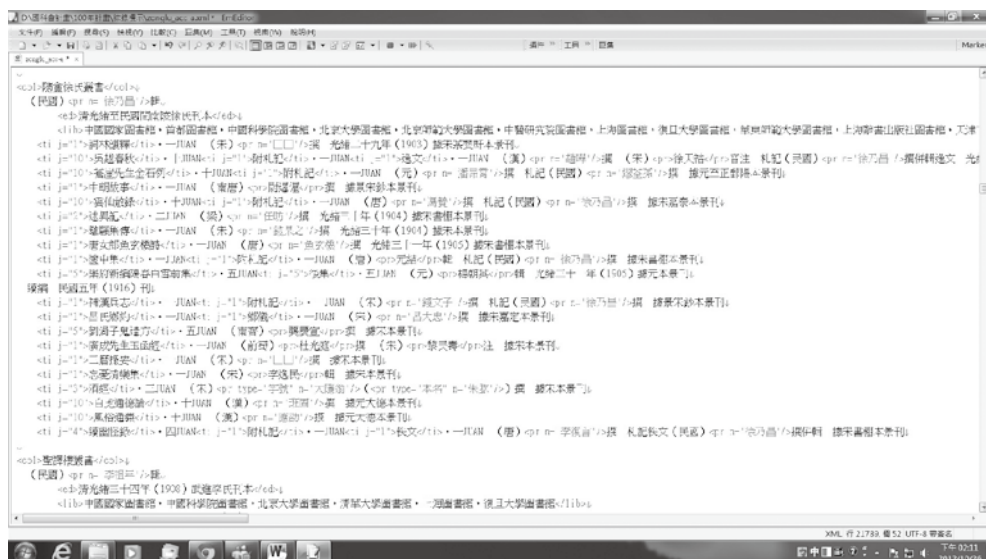
電腦資訊的本質功能與強項便是工具性資料的分析、整理與檢索，這些在如今為數不少的館藏古籍目錄中已見卓越成果，也成了今人查詢相關資料倚賴的新工具與利器。古籍叢書在這裡的發揮卻還幾等於零。第一，各館藏古籍目錄資料庫雖亦著錄有叢書，但能有子目列示的不多，具分類瀏覽叢書的更加有限，何況僅單館所藏。第二，目前僅日本京都大學「日本藏中文古籍數據庫」收錄有較多的叢書並幾均列有子目，臺灣與大陸都還沒有建置出類似規模與形式的大型聯合目錄。第三，就是 1959 與 1999 年分別出版的古籍叢書核心目錄《中國叢書綜錄》與《中國叢書廣錄》都迄無資料庫檢索系統，二目的檢查賴紙本四角號碼，於叢書書名、子目書名與著者人名間的查詢來回「總目」、「子目」與「索引」數冊，十分複雜不便，非初接觸者能夠理解。筆者所建置「中國古籍叢書目錄檢索系統」即期望能改善這些問題並朝資料庫式目錄全盤取代紙本目錄的方向發展，甚至比紙本目錄更加優越，解決叢書文獻的整理與目錄利用問題。下文說明此資料庫的特色。

1. 資料庫逐步彙整古籍叢書目錄。今天可見的古籍叢書專門目錄約 60 種，而增補《中國叢書綜錄》的《中國叢書廣錄》已採錄過其中泰半，二目匯合可達古籍叢書近六千種的核心內容。上二目匯入資料庫以後，逐步以二目未採錄之目錄進行擴充，再取二目之後新編成之《中國古籍總目·叢書部》展開比對，未來則陸續以各館藏古籍目錄增補。由於資料庫不懼內容龐大，



無紙本的篇幅問題，輸入擴充的工作可無限延伸，又資料庫可先檢索已輸入資料，時間快速，方式便利，毋須以眾多紙本反覆查尋，亦不擔心各叢書目錄歸類不一，難以檢查。

2. 對叢書各著錄項目進行標示。輸入資料庫之所有書目資料須進行標示，給予叢書名、編者名、版本、典藏地、子目書名、子目作者個別的代稱，未來進行系統檢索時，相關的訊息能夠同時顯示，關連式的資料亦得層列展開，而不僅止於全文檢索的搜尋。此有利於叢書書目中重疊資料的獲取，於是同書名不同書及不同書名而同書者，訊息可毫無遺漏；不同叢書而子目略同者可一併勾稽，叢書零種的考訂增加不少成功機會，卻免除了在多部目錄中茫無頭緒查尋的苦惱。
3. 以檢索系統方式勾稽資料，以資料排比形式見出異同與關連。檢索系統採關鍵字詞搜尋，一致與模糊依序顯現，提供使用者選擇——點閱瀏覽及相互比較，察見各書目資料間異同或可能之相關性，對於叢書目錄中著錄單位自擬書名的叢書及分散割裂的叢書，可得到紙本目錄無法達成的自動匯聚與簡便比對功能，不再受紙本目錄歸類不一、著錄分散及查詢困難之圍限。



圖一 叢書目錄標示系統示例

四、古籍叢書目錄資料庫的功能設計與加值

(一) 缺字處理

古籍資料庫製作上相當困難的一部分是大量缺字。叢書雖僅約六千部，但其中涵藏的子目書，不去重複的話在數十萬種之譜，這幾等於傳世古籍的大半，製作古籍叢書目錄資料庫，與製作全數古籍目錄，在缺字的問題面臨上差不多。本系統採取龍淵文創開發設計動態組字領先技術，全盤解決缺字問題，按古籍目錄原著字樣造字呈現，並且不須下載特別字體程式，只要使用 Google Chrome 瀏覽器，即可完整顯示文字。

(二) 主要檢索功能

1. 檢查叢書名。輸入叢書名關鍵詞，可獲取包含此關鍵詞之所有內容，含書名著錄及書目中相關說明。點擊叢書名，右欄列示此叢書完整記錄，含編者、版本、典藏地、子目書名、子目作者及其他。
2. 檢查編、著者名。不論是叢書編著者或是子目書編著者，本系統一併處理，一次檢尋。輸入人名，此作者所編之叢書及所著的單種作品皆全數列示，系統亦同時呈現該子目書的歸屬叢書，欲觀看歸屬叢書的完整資料，



圖二 叢書目錄檢索系統首頁示例

<http://nccu.predragon.org/ntpubooks/index.html>



圖三 叢書目錄檢索系統叢書名查詢示例



圖四 叢書目錄檢索系統人名查詢示例

點擊叢書名可展開顯示。

3. 檢查子目書名。輸入任意書名關鍵詞，左欄列示資料庫中全數相關訊息，在呈現上亦表現所歸屬叢書名及子目書作者等基本情況，點擊叢書名可見



圖五 叢書目錄檢索系統子目書名查詢示例

叢書完整書目資料，點擊作者名可變換成以人名為搜索條件，靈活切換成人物編著資料。亦即在總目、子目與人名中任意來回切換，方便使用者多層面之觀察需要。

(三) 主要瀏覽功能

以上檢索功能之設計，取代了各紙本目錄多冊編輯及查尋法之不便，但紙本目錄仍有檢索系統不能達到的功用，比如總目的分類呈現與子目類編。總目分類除了作為檢查，還可以提供觀察某一類叢書的全數書目，效用仍十分重要，對於叢書文獻研究、發展考查等有關鍵影響。子目類編則重新整理叢書收錄子目，不按各叢書歸屬著錄，而是依四部分類重做組織，對於需求叢書具體子目資料的使用者來說，可離開叢書形式，輕鬆獲取其中各類別古籍資料。本系統期望全面取代紙本叢書目錄，亦將叢書分類瀏覽及子目書分類瀏覽作為後續製作重點之一。

(四) 加值規劃

古籍叢書目錄檢索結果，只是書目之記錄，對於此作品更進一步的認識與相關訊息的解釋，都還有無限發展與想像空間。本系統後續加值規劃為：一、人物傳記資料的自動附屬。目前維基百科已發展為相當豐富及詳盡的網路資料庫，維基百科提供協議文本，在註明引用的情況下可連接個別資料



庫，本系統盼相關技術及授權問題解決下，可連接維基百科成果於叢書編著者、子目書作者項下，方便資料庫使用者即時閱覽相關人物傳記，深入理解問題。二、歷代帝系表之自動附屬。同樣採合法連接維基百科之方法，呈現各叢書書目版本中提及之歷代帝王年號相關資料。三、《四庫全書總目提要》之附屬與點閱。古籍目錄雖一時間無法繫連各著錄之完整原書，但精簡陳述該作品的《四庫全書總目提要》應是相當重要的參考資料。本系統期望建置《四庫全書總目提要》內容，未來附於各叢書書目及子目書書目中，便利使用者即時查尋點閱清人介紹該作之文字，而對於四庫全書或四庫館臣所收或所見，與本系統叢書目錄資料同書名不同版本者，亦可起比較之功，古籍目錄檢索之效益勢必擴增。四、突顯臺灣公藏古籍叢書典藏情況。目前臺灣典藏古籍叢書主要分布於國家圖書館、中央研究院傅斯年圖書館、故宮博物院、中央圖書館臺灣分館、臺灣大學圖書館、臺灣師範大學圖書館及東海大學圖書館等單位，然各館獨立製作古籍目錄資料庫或館藏目錄，欲查詢臺灣藏叢書須分別開啓各資料庫，殊為不便，且叢書書目資料有的典藏單位著錄不盡完整。本系統希望全面收錄臺灣藏古籍叢書資料，既可與多數大陸資料作比對，明瞭其中差異，亦能增補大陸資料的未足。尤其典藏地之標明，期能便利臺灣學術研究使用，能於單次檢索中獲取多館之資料。

五、結語

在資料庫技術發展之前，紙本目錄要求全求備，要持續無限制擴充補完，要編製各種索引滿足不同檢查功能，是非常不容易的事。紙本需考慮篇幅，著錄需著眼精簡，編輯工作亦有定期與結束之日，所以過去的紙本目錄進展速度有限。資料庫技術不僅解決了諸多人工的辛勞及種種條件之限制，更重要的是，它令原本難度相當大的收錄內容比對比較的問題得到突破，使資料庫式目錄發展一日千里。古籍叢書總目亦將在電腦技術的支持下，逐步積累、擴增並持續成長，達到過去無法完成的夢想。不過，今天仍有新的課題與困難，亦即紙本叢書目錄要作為資料庫的基礎內容，存在版權問題，在未獲授權之前，不能直接使用具保護資料。本系統使用前此目錄成果僅作為資料參考，製作目標是發展成古籍叢書總目資料庫，盼於加值及再現叢書資料之後，叢書目錄資料庫能成為嶄新的內容與使用經驗，並以此成果帶引古籍叢書的討論，使古籍叢書學科研究臻至成熟與完善境界。