

參加「擊古鑠今--數位時代下的文化傳承與知識發現論壇」心得報告

流通組 謝鶯興

會議日期：民國 104 年 5 月 27 日

主辦單位：國家圖書館、聯合線上公司

會議地點：國家圖書館 1F 簡報室

本館近些年來，一直在試著將館藏各類文獻掃描，希望在逐步數位化中提供更豐富的內容，進行無遠弗屆的服務。但受限於技術層面的欠缺，目前在「虛擬校史館」中仍未能具有全文檢索的功能，僅能藉由輔助的建置，稍得以搜尋資料庫中的部份資料。

接到館長室轉寄國家圖書館辦理的「擊古鑠今--數位時代下的文化傳承與知識發現論壇」訊息，得知此論壇為因應資料數位化及網路技術日益演進的時代，分別在文學、史料及音藝資源層面，探究中國古籍文獻、古典藝術音樂、民初近代報刊之數位化資源保存與加值的應用，以及兩岸現代數位資源的發展趨勢，共同面向數位時代的文化傳承與知識發現，提供他們各自的經驗以及研究成果的分享。希望能藉由這次的研習，汲取他人的經驗或技術，以為館藏文獻數位化的藉資。

壹、議程

本次的議程，依發表的主題分為四個單元(節)，分別從古籍文獻、古典藝術、近代報刊的使用及現代書刊資源等內容。茲轉錄於下。

日期/時間	主題	主持人/主講人
08:00~08:45	報到	
09:00~09:10	致詞	國家圖書館曾淑賢館長 聯合線上股份有限公司李彥甫總經理
第一節 古籍文獻巡禮		
09:10~09:50	坐擁書城：漢字古籍的現代命運	中國陝西師大學袁林教授
09:50~10:30	國家圖書館古籍文獻數位資源與加值服務	國家圖書館張圍東博士
10:30~10:35	Q & A 時間	
10:35~10:50	茶敘	
第二節 古典藝術大觀園		
10:50~11:20	中國藝術數字資源的保存和利用	雅昌文化集團出版事業部 副總經理 無線事業部總經理

		理 劉志國先生
11:20～11:50	數位時代的古典音樂之旅	金革唱片 林柏杰總監
11:50～12:00	Q & A 時間	
12:00～13:30	中午餐敘/用餐地點 B1 讀者餐廳	
第三節 重現近代報刊		
13:30～14:10	不純情羅曼史：從《臺灣民報》系列看臺灣人的婚戀愛欲	國立臺灣圖書館 蔡蕙頻編輯
14:10～14:50	臺灣研究的百寶箱--談近代報刊資料庫的運用	國立中興大學台灣文學與跨國文化研究所 廖振富教授
14:50～15:00	Q&A 時間	
15:00～15:30	茶點／Break Time	
第四節 現代書刊資源		
15:30～16:00	從《人大複印報刊資料》看：大資料時代學術期刊功能變革	中國人民大學複印報刊資料中心編輯、人大數媒科技(北京)有限公司產品總監 張璐先生
16:00～16:30	與時俱進--兩岸數位閱讀資源發展趨勢	聯合線上股份有限公司 讀書吧 陳芝宇經理
16:30～16:40	Q&A 時間	
16:40～16:50	抽獎時間	
16:50	滿載而歸	

貳、內容摘要

從議程的規劃與現場參與所知，全程不僅依主題的不同來區分場次，發表者也依其身份立場：主持規劃、參與工作、部門主管、介紹推廣及研究使用等等的差異，進行不同層次的內容發表。

一、古籍文獻巡禮

第一節是「古籍文獻巡禮」，分別由中國陝西師範大學袁林教授與台北國家圖書館張圍東博士主講。

袁教授曾先後主持(或完成)幾個漢籍數據庫，因此從其實務談論對於古籍數位化的問題。在「坐擁書城--漢字古籍的現代命運」題目中，他從書本時代的困難、漢字古籍新定義、總體設計與已有進展、技術要求與總體目標等四個方向談論。

袁教授認為，中國古代書籍製作與利用有六難：得書難、藏書難、傳書

難、編書難、用書難、改書難，在整個紙本書籍時代並未根本改觀。但是因古籍的數位化技術與發展，帶來了一些根本性的變化。

提出世人應先對於漢字古籍(不包含少數民族所使用的文字書寫)定義需要重新確定，並且涉及到對於文獻種類、時間斷限、國別區分、計算單位等方面的界定。

在需要重新確定漢字古籍製作與利用的整個體系中，並不從圖書管理的角度著眼，而是從圖書的利用思考，希望能為使用者提供最全面、便捷、強大的服務。因此認為開發出來的新體系，至少應包含四個部分：1.目錄數據庫，提供所有文獻目錄信息；2.圖版數據庫，提供文獻的原始圖版；3.全文數據庫，對文獻提供深度檢索和加工利用，由此生成新的信息資源；4.高級服務平臺，提供工具，實現標點、校勘、攷證等初級研究工作不同程度的自動化，並在一定程度上實現文史論文初稿撰寫自動化，大數據為這些功能的實現提供了條件。

張圍東博士在台北國家圖書館特藏組服務，就地利之便，主講「國家圖書館古籍文獻數位資源與加值服務」。

張博士先說明國家圖書館簡介、館藏源流與特色，接著說明該館的數位典藏與資料庫，在數位合作與國際推廣上的努力與成果，最後說明他們在數位加值與出版以及永續經營的理念和堅持。

如：自 2001 年起參與數位典藏國家型科技計畫，進行典藏古籍數位化作業，並建置多種資料庫。在國際推廣文化方面，近幾年來積極主導國際合作古籍數位化古籍聯合目錄之建置計畫。自 1990 年代後期起，國圖為掌控散佚海外重要中文善本古籍文獻，逐年調查海外中文善本古籍文獻存藏及數位典藏現況，建「中文古籍書目資料庫」(<http://rbook2.nc1.edu.tw/Search/Index/2>)，採互惠合作，與更多善本古籍典藏單位建立實質合作關係。

於 2005 年起與美國國會圖書館、美國華盛頓大學圖書館、美國柏克萊加州大學圖書館及加拿大多倫多大學圖書館等重要圖書館進行古籍數位合作，古籍數位成果上傳國圖古籍與特藏文獻資源系統(<http://rbook2.ncl.edu.tw/>)，提供使用者檢索與利用等。

在第一節兩場之間，出現不在議程的小插曲，即由華裔人板橋先生介紹「雕龍數據庫」。號稱廣納中日古籍，收錄詳實完備，提供全文關連檢索，似乎和「四庫全書電子版」的「關聯詞」相似，具有繁簡字、異體字、古今字及形近等漢字的關聯檢索，但和袁教授所說「關聯詞」將同人不同稱號的關聯不同。

二、古典藝術大觀園

第二場次由雅昌文化靳菊紅副總監和金格唱片林伯杰總監兩位擔綱。兩位都是藝術業界者。前者講「讓藝術走進每個人的生活--中國藝術教育與藝術資料庫的未來」，後者談「數位時代的古典音樂之旅」。

靳副總監提出：藝術教育是國民素質教育不可或缺的重要內容，也是引領社會風尚--弘揚中華民族精神、培養藝術人才的基礎工程。應該從「精英教育」自「大眾教育」轉型。長期處於教資體系邊緣化的大眾藝術素養的提升，則有賴於更多知識儲備和更高欣賞能力的引導。

藝術資源的數位化，可以降低藝術書籍難以持有的自由性，並具有四種特性：1.資料性，涵蓋各個藝術門類資源；2.指導性，技法、作品演變、藝術特性等一整套藝術欣賞方式的傳授；3.工具性，系統地梳理藝術史脈絡以供查詢；4.實踐性，提供藝術市場資料資源，走向拍賣、收藏等藝術產業的下游環節的未來發展。

林總監則以以深入淺出的專業解說，從明星的誕生、唱片世代、影像世代談起，音樂載體的誕生與革新改變人們對音樂的接受度，配合精采多元的影片觀賞，讓我們進入古典音樂興的大觀園，一窺數位時代的古典音樂發展。

三、重現近代報刊

第三單元的主題為「重現近代報刊」，由臺灣圖書館蔡蕙頻博士與中興大學廖振富教授，從使用與研究者的角度討論近代報刊數位化的使用。蔡蕙頻博士以「不純情羅曼史：從《臺灣民報》系列看臺灣人的婚戀愛欲」為講題，以其使用臺灣圖書館典藏的早期報紙，特別談論日治時期，近代法律引進臺灣後，過去習用民間習慣法的臺灣人，是如何來適應這種新觀念，針對「臺灣人的婚戀愛欲」的主題，從《臺灣民報》經由數位化建置資料庫檢索這些詞句進行研究。當然還是先說明臺灣近代報業的發展，接著對照《臺灣青年》和《臺灣民報》編輯群的差異，才拉到日治時期從《臺灣民報》刊載訊息討論當時臺灣人對婚姻的觀念，結論是呈現臺灣文化的向上，重視女性追求平等的個人主義。

廖教授以「臺灣研究的百寶箱--談近代報刊資料庫的運用」為題，呈現他近些年來大量利用這些資料庫所提供的檢索點與圖像檔進行研究，得到：日治時期監獄文學、安重根刺殺伊藤博文事件、樸社與瀛社對照觀察、台中公園文學地景、台灣教育會雜誌的日台漢文交流、中村櫻溪的北台灣山水遊記等成果。

講演中鎖定幾個議題及關鍵字，舉：(1)吳鳳、廖添丁、蔡惠如，(2)剪辮、斷髮，(3)櫟社、瀛社，(4)台中公園，(5)林獻堂等為例，說明如何運用近代報刊資料庫，作為研究的輔助材料。認為這些議題，包括臺灣人物傳說、日治時期詩社活動、社會革新、政治運動、城市地景、臺灣人物與近現代史。都可以從《臺灣日日新報》得到許多報導。頗值得我們效法與學習的經驗。

四、現代書刊資源

最後一場是「現代書刊資源」為主題，由人大數媒張璐總監主講「從《人大複印報刊資料》看：大數據時代學術期刊功能變革」，聯合線上陳芝宇經理主講「與時俱進--兩岸數位閱讀資源發展趨勢」。兩位都是從負責經營資料庫的立場說明各自產品的特色。

張總監分別說明：大數據的定義、特色及人們對大數據理念的轉變，是要全部資料、要及時有效和揭示相關的關係。這些正在引發為學術服務而生的學術期刊的功能變革。接著介紹人大書報資料中心，論文篩選過程，論文質量評估指標。主要指標從學術創新程度、論證完備程度、社會價值及難易程度，佐以輔助指標的課題立項與發表載體來判斷。證明學術評價的作用，學術期刊數位化的轉型，透過互聯網，可以進行深度的挖掘與分析數據關聯，最後將使得學術期刊的功能變革，從產品價值導向轉為用戶價值導向。

陳經理著重兩岸數位閱讀資源的發展，先以幾個熱門的人物如柴靜、吳奇隆等人為例，試著從大陸雜誌資料庫看到的訊息，拉到「華文雜誌知識庫」，可以了解中國，貼近當地文化脈動的功效。接著介紹早期大陸的數字閱讀，當前的數字閱讀，百度閱讀的免費策略，到網路小說改變習慣等等，總結大陸在網路世界的靈活與多變，促進了華文市場的多元性。

參、心得與感想

眾所皆知，文獻的數位化，並非僅是文獻儲存與保存的另一種方式，它主要藉由數位化的成果，在無遠弗屆的網路世界能方便取用瀏覽，不需花費巨大的時間與金錢，大老遠地千里奔波。數位化後的全文檢索，則是提供瀏覽文獻之外的資料蒐集的便利性，並給有意研究某一主題者，特別是初學者，可以進行更深入研究的資助。誠如本次論壇主講者之一的袁林教授所說的：「提供工具，實現標點、校勘、攷證等初級研究工作不同程度的自動化，並在一定程度上實現文史論文初稿撰寫自動化。」亦即協助研究者從蒐集資料到資料的編排，可以完成研究論文的初稿。

或許這種說法，在學者的眼中會認為不可完全依賴的，但初學者，可以

藉由資料庫全文檢索的功能，不管是檢索相關議題研究的論文，或是針對某一主題的文本蒐輯，內容比對，甚至是資料的真贗，將會有莫大的助益。這也是圖書館工作人員可以盡力進行推廣的服務項目，或是訓練讀者如何來使用這些功能之處。

然而，資料庫要提供全文檢索，就漢文古籍而言，是相當困難的，首先要面對古人的一人多種稱呼，一物不同命名，一事不同名稱，同名不同人，或是一人不同譯名等問題。

圖書館界一直都面對這種困擾，因此有「主題詞」、「關鍵詞」、「權威檔」等規劃，冀望能協助使用者取得最大的資料。可惜效果似乎不彰。我想，這和計劃製作這種「關聯詞」的單位預算，參與製作者的學科背景、文化修養等，應有很大的關係。

以本館目前試著進行「善本古籍數位化」，請廠商到館掃描，再帶回後製的工程。然而廠商的對於這類的文化修養不足，必需由館內先行編製目錄提供他們參考，但又面臨無法順利對照目錄與內容的關連性。加上館內尚缺乏這種數位化全文檢索的後製技術，除非是將掃描檔轉為 PDF 檔，雖然可提供檢索功能，但若沒有「權威檔」(或「關聯詞」)的輔助，仍無法進行全面性(如一人多稱謂問題)的取得相關資料，相信這也是資料庫廠商所面對的事。袁林教授提到：「史學論文的基礎是與論點相關聯的史料，而關聯史料之間又會有相關聯的詞彙，在海量全文數據庫中檢索這些詞彙，按一定順序排列檢索得到的段落或文句，實際上就形成了論文的初稿。」雖是一種很進步的檢索功能，但是仍與理想的「權威檔」(或「關聯詞」)編製有一大段的差距，有賴各學科的專家們合作編訂，才能達到這種境界。



國家圖書館曾淑賢館長致詞

廖振富教授與蔡蕙頻博士 Q & A