

創造、管理及使用數位館藏：臺大圖書館經驗談
**Creating, Managing and Using Digital Collections: The National
Taiwan University Library Experience**

項潔 Jieh Hsiang

國立臺灣大學資訊工程學系教授兼圖書館館長

University Librarian,
Professor of Computer Science and Information Engineering,
National Taiwan University
hsiang@csie.ntu.edu.tw

洪筱盈 Hsiao-ying Hung

國立臺灣大學資訊網路與多媒體研究所博士生

Doctoral Student, Graduate Institute of Networking and Multimedia,
National Taiwan University
hyhung@turing.csie.ntu.edu.tw

【摘要 Abstract】

大學圖書館蒐集數位資源或進行館藏數位化工作，已經是普遍的趨勢。但是有愈來愈多的數位館藏之後，圖書館該如何有效管理、進而提供一個便利的使用方式，成為必須解決的問題。本文從數位館藏的創造、管理及提供使用的角度，逐一分享臺大圖書館的理念、實作經驗與看法，並提及因應數位館藏而產生圖書館管理層面的可能改變。

The ever increasing volume of digital objects brings dramatic changes to modern libraries. In this paper we classify a library's digital collections into four categories according to their relationship with physical resources. We also point out issues related to the collection, preservation, and managements of these digital collections, and how the management of a library should adjust accordingly.

【關鍵詞 Keywords】

數位館藏、館藏發展、圖書館管理

Digital Collections, Collection Development, Library Management

大學圖書館蒐集數位資源或進行館藏數位化工作，已經是普遍的趨勢。但是圖書館有愈來愈多的數位館藏之後，該如何有效管理、進而提供一個便利的使用方式，成為必須解決的問題。本文將從數位館藏的創造、管理及提供使用的角度，逐一分享臺大圖書館的理念、實作經驗與看法，其討論的數位館藏的範圍以圖書館館藏數位化及免費數位資源為主，不包括付費電子資料庫與電子期刊。

一、數位館藏的四種類型

臺大圖書館在過去數年間，進行了大量數位化工作，從實作經驗中，可將數位資源依據資料的特徵差異，分為四種數位館藏類型（註1）：

1. 實體館藏的數位化；
2. 實體館藏中新創的數位館藏；
3. 未擁有實體館藏的數位館藏；
4. 原生數位型態（born digital）的數位館藏。

不同性質的數位館藏需要有不同的處理與服務方式，以下分別對這些不同的數位館藏類型，闡述其特性，並以臺大經驗為例做說明。

（一）實體館藏的數位化

最常見的一種數位館藏就是把實體的館藏以數位方式一對一重製、保存並提供使用，也就是圖書館同時擁有資料的實體館藏與數位館藏，且兩者在資訊組織的分類編目上可以是相同的一筆資料，只需在書目紀錄上同時標明存放在圖書館

的實體書架位置及數位檔案位置。這種將實體館藏數位化的數位館藏又可分為兩類，一類是「數位典藏 (Digital Archives)」，另一類則是館藏圖書數位化。

數位典藏是將有價值的文物，透過數位化技術，將文物轉化成文字、影像、聲音或視訊等數位物件。圖書館進行數位典藏多以特藏資料為主，因為特藏資料往往年代久遠而脆弱易損，禁不起多次的翻閱搬動，最有保存的急迫性，所以用數位方式典藏保存與提供使用，不但能減少因不當使用或災害造成完全喪失的風險，也可以做為往後若有修復必要時的參考依據。以臺大圖書館為例，已經數位化大量的珍貴典藏，包括淡新檔案、伊能嘉矩手稿、岸裡大社文書、古契書、臺灣古拓碑、歷代寶案等館藏：淡新檔案是清代臺灣北部的行政與司法檔案，為現存清代臺灣省、府、州、縣廳屬檔案中最具規模且最完整的史料 (註2)；伊能嘉矩手稿是以臺灣平埔族為對象的田野調查資料 (註3、註4)；岸裡大社文書則是有關臺灣中部開發的文書契字 (註5)，這些都是研究早期臺灣史的重要一手資料。這些臺灣早期發展的珍貴史料，除了建立詳細詮釋資料及數位典藏影像外，並進行全文數位化與標註校對工作，提供全文檢索與分類，提高使用的便利性(註6)。

將館藏書籍數位化是圖書館第一個大規模的數位化工作，美國與中國的百萬冊圖書數位化 (註7) 計畫就是此工作的一個範例。隨著光學文字辨識 (Optical Character Recognition, OCR) 技術的進步及 Google 圖書搜尋的快速發展，使用者愈來愈習慣以全文檢索來查找資料。哈佛大學、加州大學、紐約公共圖書館等十

二個圖書館(註8)與 Google 圖書搜尋合作,多倫多大學、康乃爾大學等多所大學及公共圖書館則與微軟的 Live Search Books 合作(註9),可見到在不久的未來,圖書館的自動化系統在使用層面亦必將產生革命性的變化。這類型的實體館藏數位化強調使用上的目的,可以打破時空因素的限制,除了讓使用者可以透過網路搜尋引擎來搜尋圖書館館藏外,大量的圖書不再侷限於只能用書目資料查找,能夠以全文檢索的方式被使用,這個突破將把圖書館資料的使用觀念與方式提升到一個新的境界。

(二) 實體館藏中新創的數位館藏

實體館藏最大的限制在於一件藏品只能有一個單一的書架位置,而數位化則能打破這個限制。除此之外,圖書館甚至可以依據特定主題,以數位的方式從現有的實體館藏中聚集創造出新的數位藏品集合(collection),賦予館藏新的意義。

這種數位館藏的特徵在於數位的物件有實質館藏的存在,但可能並無資訊組織的分類編目上有實質的對映。此處以一個實際的例子做說明,臺大圖書館在2002年時將館內收藏的日治時期舊藏書做了一次整理,從這些舊出版品中找出與臺灣相關的老照片四萬餘張,並將這些照片數位化後集成「臺灣老照片資料庫」(註10)這個資料庫非常珍貴,一方面因為1945年以前的臺灣照片數量原本就稀少(中央社亦有大量同一時期的照片,但大部分是中國大陸的照片,與臺灣相關的甚少),此資料庫收錄四萬多張照片已經是收藏臺灣老照片的集結中資料量

最大的；另一方面則是因為這些照片來源是絕版書籍中的插圖與攝影圖集，書籍內文多有對照片源由做說明解釋，故能夠據以建立完整的詮釋資料，相較於一般的照片資料庫因為來源不明而無法登錄詮釋資料而難以使用的問題，此資料庫提供了研究者更精確且豐富的說明資訊。

這類型的數位館藏是根據特定主題，由專家學者及圖書館館員從圖書館分散的眾多實體館藏中整理集結而「創造」出來的，能夠賦予既有的舊館藏新生命，從老照片資料庫的例子可以發現，這些照片原本分散在不同時代、不同主題的書籍文獻中，使用者即使在書籍中翻閱到幾張照片，可能也無法對這些照片產生什麼特別的聯想，但是一旦將這些照片集結在一起，藉由方便的搜尋、比對與分析，說不定會產生新的發現或不同的解讀意義。

除了臺灣老照片資料庫是一個良好的使用範例外，臺大圖書館還集結了包括漢文日日新報、明清臺灣行政檔案全文資料庫、臺灣宗教民俗圖錄（註11）等多種新創的主題性數位化館藏。再以明清臺灣行政檔案為例，圖書館與歷史系教授合作，從館藏各種書籍與檔案文件中，蒐集且數位化超過 3500 萬字以上在明朝、清朝時期有關臺灣的第一手官方檔案文書，包括給皇帝的奏折、起居注、宮中檔、軍機檔等，這樣的規模已經涵蓋了臺灣在明清時期 80% 以上的官方文件，是研究臺灣社經發展非常重要的第一手資料（註12）。

(三) 未擁有實體館藏的數位館藏

徵集有文化價值的特有館藏是每個大型圖書館努力的目標之一，但其過程往往不是那麼容易。我們的經驗發現資料擁有者不見得願意贈送或販賣這些實體藏品給圖書館，但他們願意出借藏品讓圖書館數位化，甚至圖書館能以圖資專業協助他們建立完整的藏品清單，或對實體資料進行修復與分類整理。圖書館將資料數位化後，實體資料歸還資料擁有者，且雙方能夠各取得一份數位化檔案，藉由這樣互惠的雙贏模式，圖書館能夠增加許多具有價值的數位化館藏，提供研究者更豐富的研究素材，收藏者也得以繼續保有其實體收藏，並透過使用者的研究增加其價值。

此類型數位館藏的特徵在於圖書館沒有擁有實體的館藏資料，但在資訊組織的層次上則與實體資料的分類編目方法相同，差別在於館藏地登錄為數位檔案。這樣的數位化館藏在臺大圖書館內愈來愈多，包括臺北市文獻委員會古契書、竹塹北門鄭利源號古契書（註13）、日治時期法院檔案等都是用這樣雙方互惠的數位化方式取得的。圖書館不一定絕對要擁有那些實體館藏，只要設法讓資料進行數位化，且透過網路來讓使用者搜尋與使用，而不是僅被收藏存封在某個儲藏室中，這對研究者而言有莫大的助益，且也能夠達成圖書館保存文獻以支援研究與教學的功能性責任。

(四) 原生數位型態 (born digital) 的數位館藏

隨著資訊科技與電腦網路的發達，愈來愈多資料原生僅有數位型態 (born

digital)，例如網站、網路文獻、電子書等，這類數位資料在資訊組織的層面上，與前述三種類型數位館藏的意義及實務操作相當不同，沒有與實體館藏的對映關係，甚至於用什麼單位來計量數位館藏量都沒有一致的看法。根據調查，網頁的平均壽命是 44 天至兩年，網頁建立的四年後大約只剩下四分之一的網頁仍可存取(註14、註15)。由於原生數位資料具有增加快速、變異性大、時效短等特性，造成許多圖書館不將此納入館藏發展範圍，然而對未來的研究者而言，這是研究當代文化極為重要的材料。

目前已有許多國家進行原生數位資料的保存工作，尤其是網路資源的保存著力最多，包括澳洲國家圖書館的 PADI (Preserving Access to Digital Information, <http://www.nla.gov.au/padi/>)、美國 Internet Archive (<http://www.archive.org/>)、英國網路保存聯盟 (UK Web Archiving Consortium, <http://info.webarchive.org.uk/>)、以及臺大圖書館之網頁典藏庫(<http://webarchive.lib.ntu.edu.tw/>) 保存臺灣大學相關的網站及臺灣政經社會重大事件網站，這些計畫都是為保存數位網路資源以供未來永續使用的計畫。

原生數位資料不只是大量的網路資源，以大學圖書館來說，大學內的學術產出也大多都是以數位型態存在，包括期刊與會議論文、研究報告、研究記錄、演講素材等各類資料，這些研究成果分散在各處，不但無法長久保存，更無法展示出學校整體的研究發展脈絡。近年來許多學校積極發展的機構典藏 (Institutional Repository) 就是扮演著保存學校研究成果的數位儲藏中心，臺大機構典藏

(<http://ntur.lib.ntu.edu.tw/>) 自 2006 年 12 月成立，至今已蒐羅臺灣大學三萬多篇過去的研究成果，這些研究資料都是圖書館數位館藏相當重要的一部份(註16)。

二、圖書館面對數位館藏的營運變革

面對愈來愈多類且大量的數位館藏，圖書館需要思考如何調整組織架構、服務流程與發展策略，以承載龐大的數位資料處理量，不但要將數位資源納為館藏的一部份，還要更進一步提供便利的數位使用環境。茲就不同層面分述如下：

(一) 落實實體與數位館藏並重的館藏發展政策

圖書館館藏不僅是實體書籍與期刊而已，數位典藏、原生數位資料等廣大的數位資源也應該都是館藏發展的一部份，圖書館應重新思考什麼是館藏，讓數位館藏也有具體明確的發展目標與策略，據以落實實體與數位館藏並重的館藏發展政策。在檢討館藏發展政策的同時，也應調整過去偏重實體館藏的館藏計量方式，增加合理的數位館藏計量指標，讓圖書館投入在各種數位館藏的成果能充分反映在館藏量與使用量的統計分析中。

臺大圖書館的館藏發展政策希望在臺灣研究史料收藏能夠達到世界領先的水準，因此近年來積極由實體與數位管道並進，廣泛納入不同來源與性質的臺灣史料為館藏；本文前述四種數位館藏之類型與內容舉例，就是近年來以數位方式蒐羅臺灣研究史料的經驗歸納而得。

(二) 調整圖書館之技術服務流程與任務編組

數位資源浩瀚無邊，為能順利且確保各種納入館藏的數位資源都是有學術價值的，需要有一套數位館藏的採訪流程與資料篩選機制，並且以適當的編目方式呈現於圖書館館藏目錄中。有鑑於數位館藏與實體館藏特性差異頗大，從採訪、編目、閱覽等服務流程更是隨資料性質而截然不同，因此需要指派特定人員或新增電子資源部門專職負責數位館藏的整體發展，訂定各種辦法與作業流程，並且增加資訊技術人才以處理及管理大量的數位資料、網站資料庫以及儲存設備，進而著手針對數位館藏開發新的使用與應用模式，如此才能使圖書館對數位館藏的整體營運方針趨於穩定且永續發展。調整圖書館組織架構或服務流程，能將數位館藏納入日常的館務處理程序中，避免讓圖書館人員以短時期性的專案計畫或特例形式來看待數位館藏，長此以往，不但能漸漸拓展圖書館的數位服務範圍，也能讓館員認知到蒐羅數位館藏並提供使用，本來就是圖書館業務中很自然的一部份。

以臺大圖書館為例，採訪、期刊、編目與閱覽等部門的館員任務應數位館藏的需求做了相當的調整，並為了配合大量數位館藏的業務，資訊技術人員從編制的 5 名擴增到目前約有 17 個專職人力負責數位館藏之儲存管理、系統維護與各類資訊系統開發工作。擴增圖書館的資訊技術人力，能使圖書館內的數位館藏能夠得到最適切的維護，且能站在圖書館第一線瞭解使用者的真正需求，以開發出更有助於研究與教學的數位館藏資訊服務。

(三) 開發各種數位館藏的使用服務

圖書館對內將實體館藏數位化，對外則篩選高品質的數位資料納為數位館藏，已經是一個普遍的趨勢。然而當圖書館累積大量的數位館藏之後，如何管理、如何提供便利的使用方式、進而如何作為教學或研究上的應用，成為新的課題。

使用者對數位館藏與實體館藏所能提供的服務的期待往往截然不同。以實體館藏為例，圖書館採訪、編目與上架之後，使用者查詢書目記錄，然後順利到書架上取得並借閱，就算完成一次良好的服務程序；但若將同樣的流程邏輯套用到提供數位館藏服務則不適宜——如果仍然只能讓讀者用題名、作者等書目資料欄位檢索，則未利用到數位資源所提供的豐沛資訊，但從另一方面來說，對所有資訊均沒有權重而一視同仁的全文檢索又會造成資訊過載，查回過多資料反讓讀者無所適從。所以數位資料不同於實體資料，不是圖書館採訪編目納入成為館藏後工作就算結束，數位資料的關鍵在於納入館藏後該如何提供使用的機制，唯有營造出一個良好的數位使用環境，提供更多有助於教學與研究數位輔助工具，才能充分發揮數位資料便於聚集與查找的特性，讓數位館藏更具實用性。以下提出幾種可利用數位館藏提供的加值服務：

1. 與外界網路服務結合，發展更彈性與即時的服務

若能利用現有的網路資源提供整合服務，可以開發出更有彈性與即時的服務方式，這類服務包括搜尋引擎、網路書店、以及 Google 與 Yahoo! 的圖書館館藏圖書數位化等，都能讓圖書館服務更加貼近讀者需求。舉例來說，臺大圖書館與

Yahoo!搜尋引擎合作，篩選高素質的免費學術網路資源，除了建立獨立的臺大學術資源網 (<http://sg.lib.ntu.edu.tw/>) 外，亦可在訪客流量龐大的 Yahoo!首頁直接進行學術搜尋 (<http://tw.search.yahoo.com/academia>)；臺大圖書館亦與圖書比價網站合作，在 Webpac 上顯示圖書封面影像，增加讀者查詢館藏圖書的額外資訊，若欲購買，亦可連結至比價網站，察看書店販售的圖書價格。

2. 與館藏書目資料結合，提供數位資源搜尋分類加值服務

圖書館的數位館藏與網路上眾多的數位資料最大的不同點，在於圖書館館藏內容具有詳細的分類編目資訊。網路搜尋引擎受限於紛雜無組織的網路資料內容，僅能提供簡單的關鍵字檢索，然後只能憑靠使用者自己的判斷來篩選數千萬查回的網頁；圖書館若能善用資料量龐大的書目資料庫或數位館藏的詮釋資料 (metadata)，在檢索後幫助使用者分類與統計查回的資料類別，提供更進一步檢索的關聯線索，這將更能顯示出圖書館在數位資訊組織上的優勢，臺大數位典藏資源中心 (<http://darc.ntu.edu.tw/>) 就是利用此概念，整合圖書館與臺大多個系所的數位典藏成果，提供豐富的瀏覽與查詢後分類資訊，讓這些數位典藏內容更易被查找使用與解讀。

3. 提供更多的網路個人化服務

圖書館個人化服務也是近年來很受重視的一個趨勢。如何將 Web 2.0 的觀念及技術加入服務是臺大圖書館讀者服務發展的重要方向，舉例來說，臺大圖書館

首頁(<http://www.lib.ntu.edu.tw/>)及臺大新聞剪報系統(<http://news.lib.ntu.edu.tw/>)

均提供 RSS 資訊服務，讀者可選擇訂閱自己所關心的各類高教新聞、公告、新進館藏選介與新進電子資源等頻道，讓圖書館的各種服務與館藏資源都可以周知給有需要的讀者，提高圖書館資源與數位資訊服務的利用率。

4. 提供數位館藏的教學與研究輔助工具

圖書館集合了大量不同來源、類型、領域等豐富多樣的數位研究素材，為發揮數位資料便於聚集、比較、統計、檢索與查閱的特性，使更多的研究者與學生樂意使用這些數位館藏（註17），圖書館應嘗試利用資訊技術，發展各種便於教學與研究的輔助工具。舉例來說，臺大圖書館數位學習網（<http://elearning.lib.ntu.edu.tw/>）提供近 30 個導覽與主題單元，每個月平均有逾 33 萬個點擊次數，不但大幅減少館員進行利用教學指導的工作量，亦提高讀者使用滿意度；另外，臺大圖書館已經收集了大量的數位化臺灣研究史料，為了鼓勵臺灣史學的研究與教學都能夠使用這些數位館藏，除了發展一個能夠全文檢索的臺灣史料之數位圖書館外，更持續開發提供線上的歷史年表、官職表、沿革地圖等研究輔助工具，擷取歷史文件的地名與人名等詞彙進行分析統計，使史料的呈現、查找與各類統計輔助更貼近史學研究人員的特殊需求，這將有助於提高學術研究從量變進而產生質變的可能性，造成新的學術研究典範。

三、結語

著網路與數位環境日趨成熟，使用者查找資訊的行為逐漸改變成優先在網路環境下搜尋。順應這樣的趨勢，圖書館的服務必須隨之轉型，保存、管理與提供數位館藏服務給那些不進圖書館的使用者，是圖書館實體與數位館藏並重的營運思維中重要的一環。

大學圖書館需要有更積極的作為，為了創造與管理各類數位館藏，進而主動透過網路提供高素質的數位服務，必須有一連串從圖書館組織架構、館藏發展政策、技術服務與讀者服務等各層面的整體改造，唯有圖書館館務架構與人力調配均能充分發揮，產生加乘的功效，才能在網路時代成為資訊與知識集散整合的知識中心。

致謝

本文得以完成，感謝教育部及國科會對臺大圖書館進行館藏數位化與開發數位服務的經費支持，感謝臺大圖書館的所有同仁及參與臺大機構典藏、網頁典藏庫、學術資源網、數位典藏國家型科技計畫臺大機構計畫（NSC 95-2422-H-002-001，NSC 95-2422-H-002-007）之研究人員的努力。

註釋

註 1: 項潔、洪筱盈，「大學圖書館數位保存與館藏發展策略：以臺大圖書館為例」，大學圖書館 10 卷 2 期（2006 年 9 月），<

<http://www.lib.ntu.edu.tw/General/publication/univj/uj10-2/sc19-01.pdf> >。

註 2: 國立臺灣大學圖書館, 「淡新檔案」, <

http://www.lib.ntu.edu.tw/CG/resources/Taiwan/taiwan_ds1.htm > (2007 年 2 月 10 日)。

註 3: 吳密察, 「台灣大學藏『伊能文庫』」, 大學圖書館 1 卷 3 期(1997 年 7 月),

頁 4-23, < http://www.lib.ntu.edu.tw/General/publication/univj/uj1-3/uj3_2.html > (2007 年 2 月 10 日)。

註 4: 國立臺灣大學圖書館, 「伊能嘉矩與臺灣資料」, <

<http://www.lib.ntu.edu.tw/CG/resources/Taiwan/Ino/ino.htm> > (2007 年 2 月 10 日)。

註 5: 國立臺灣大學圖書館, 「岸裡大社文書簡介」, <

http://www.lib.ntu.edu.tw/CG/resources/Taiwan/taiwan_al_a.htm > (2007 年 2 月 10 日)。

註 6: 國立臺灣大學圖書館, 「臺灣大學數位典藏資源中心」, <

<http://darc.ntu.edu.tw/> > (2007 年 2 月 10 日)。

註 7: 高等學校中英文圖書數字化國際合作計畫, < <http://www.cadal.cn/> > (2006

年 12 月 2 日)。

註 8: Google Book Search: Library Partners, <

<http://books.google.com/googlebooks/partners.html> > (8 February 2007).

註 9: Live Search's WebLog, "Live Search Books Beta Release,"

<<http://blogs.msdn.com/livesearch/archive/2006/12/05/live-search-books-beta-r>

[elease.aspx](#)> (9 December 2006).

註 10: 國立臺灣大學圖書館, 「臺灣舊照片」, <

http://www.lib.ntu.edu.tw/CG/resources/Taiwan/taiwan_op.htm > (2007 年 2 月

10 日)。

註 11: 國立臺灣大學圖書館, 「臺灣宗教民俗圖錄資訊網」, <

<http://ci6.lib.ntu.edu.tw:8080/religion/> > (2007 年 2 月 10 日)。

註 12: 邱婉容, 「明清臺灣行政檔案全文建檔及標注計畫」, 中華民國九十五年圖

書館年鑑 (臺北市: 國家圖書館, 民 95 年 10 月), 頁 128-129。

註 13: 鄭華生口述、鄭炯輝整理, 新竹鄭利源號典藏古文書。南投市: 臺灣文獻

館, 民 94 年。

註 14: Kenny, A.R., McGovern, N.Y, Botticelli P., Entlich R., Lagoze C., and Payette S., "Preservation risk management for web resources: Virtual remote control in Cornell's project prism," *D-Lib Magazine* 8:1, January 2002, <

<http://www.dlib.org/dlib/january02/kenney/01kenney.html> > (2 November 2006).

註 15: 洪淑芬, 「網路資源備份保存技術與應用探析」, 國家圖書館館刊 95 年第

2 期 (2006 年 12 月), 頁 75-121。

註 16: 項潔、洪筱盈, 「臺灣機構典藏發展芻議」, 教育資料與圖書館學 43 卷 2

期 (2005 年 12 月), 頁 173-189, <

<http://research.dils.tku.edu.tw/Joemls/43/43-2/173-190.pdf> > (2007 年 2 月 10

日)。

註 17: 洪淑芬、羅雅如, 「典藏數位化建構的歷史研究網路: 臺灣大學典藏古碑

拓本與《淡新檔案》的參照研究」, 大學圖書館 10 卷 2 期 (2006 年 9 月),

< <http://www.lib.ntu.edu.tw/General/publication/univj/uj10-2/sc19-05.pdf> >。