

# 化虛擬短暫為實質永恆

## ——我國網站典藏系統的建置與現況

■ 蔡佩玲

### 一、前言

處現今數位資訊與網路時代之際，各界對於網路資訊的運用及倚重，其數量及影響性與日俱增。許多國際性統計數據，均呈現出世界各國網際網路使用人數逐年增加的趨勢。以臺灣為例，迄2008年底為止，網路使用人數達1,514萬人，佔總人口數之66.1%，其成長率相較於2000年高達141.9%。<sup>[1]</sup>網路人口的迅速成長，與臺灣近幾年來對於網路資訊環境的改善及網路資源內容的建置充實，有密切關係。如國家資訊基礎建設，除積極架構全國網路軟硬體環境，同時亦投注文化資產的數位化。一般民眾運用網際網路便捷容易，加以網路的資源豐富多元，自然促使網路使用人數持續成長。這相對也刺激許多政府機關、私人企業、單位或團體，甚至個人，莫不積極掌握並運用網際網路這項新式虛擬平臺，作為資訊發布、傳播與儲存的媒介。網際網路上各式資訊以驚人的速度迅速成長，目前已成為重要的資訊來源管道之一。

數位資訊與網路時代的來臨，對一向以讀者資訊服務為重要職掌的圖書館而言，具有相當的衝擊與挑戰。讀者可輕易透過網路主動獲取各式資訊，或透過搜尋引擎及各式入口網站，查檢或瀏覽多元的資訊。故向來自許以擔任資訊提供者與資訊領航員角色的圖書館，在數位時代下必然需由傳統圖書

館，轉型成為數位圖書館。

隨著資訊科技的不斷發展，圖書館由傳統以紙本為館藏重心轉而成為紙本與電子媒體並行的圖書館，更朝著虛擬圖書館邁進。電腦網路的日益普及，使得圖書館可瀏覽、檢索、取得位於館外遠端的資訊，所謂館藏不再侷限於館內所有的資料，加上資訊的類型也由紙本、微捲、線上書目資料庫、光碟資料庫擴展至各式網路資源，網路科技使得圖書館對館藏及館藏發展的定義不同於以往。圖書館需要蒐藏的文獻資源必須滿足讀者需求，除既有紙質媒體外，尚且包括各種新興的網路資源及電子媒體。面對數量逐年增加的電子資源，圖書館也多配合修訂電子資源的館藏發展政策予以因應，以其作為徵集、規畫數位館藏發展的依據。故近年來數位出版的發展成為熱門的議題，且數位出版的成果豐富多元，已有凌駕傳統出版的趨勢，故成為圖書館徵集館藏時的重要標的，無論是電子書、電子期刊、電子報、電子資料庫等，均已成為現代化圖書館館藏的重要項目。

此外，網際網路上存在著數量驚人的網頁，刊載各式各樣的資訊；如前述許多政府或民間單位、個人，也紛紛透過網站建置，作為資訊的彙集與傳播媒介。故各網站及其組成的網頁內容，具有訊息性、知識性、文化性意義，同時更具有累積資訊的歷史性意義。<sup>[2]</sup>但因網路發表的管道迅速便捷，故許



多資訊與文件內容直接以數位檔案形式存在於網路上，不必然一定會以實體媒材形式發行出版。但網站雖成長快速，但消失也快，今日所看到的網站內容可能隨時因伺服器關閉、網站移除、網站名稱變更等因素無法再次被利用。研究認為一般網頁的生命週期有44天，有44%於1998年建置的網站於1999年已不復見<sup>[3]</sup>；也有認為網頁壽命為100天的不同說法。<sup>[4]</sup>無論如何這均呈現網站存在時間短暫易逝的特性。故許多學者或專業機構發佈具文化、科學意義的重要文獻，或具時事議題性質的網站內容，可能因時間更迭而消失，因此「網站典藏」(Web Archiving)的意義，即是為留存這部分可能消失無蹤的文獻。所謂「網站典藏」係指蒐集部分網際網路的網站內容，並確保將蒐集所得之網站內容予以檔案典藏，以提供未來研究者、歷史學家及公眾再次查檢運用。因網站數量及內容龐大，故網站典藏多採用程式設定以快照(Snapshot)的擷取方式進行。<sup>[5]</sup>目前世界最大的網站典藏系統「Wayback Machine」是由非營利組織Internet Archive所建置。在其「Archive It」服務項下標榜的口號為「Archiving the internet for future generations: collect it, manage it, search it ARCHIVE IT!」(為未來世代珍藏網路資源：蒐集、管理、檢索與典藏)，已清楚揭示網站典藏的意義與內涵。

## 二、其他國家建置網站典藏系統概況

臺灣地區於近幾年才開始網站典藏的建置作業，而國外許多國家已有多年的發展經驗。除前述Internet Archive的Wayback Machine屬非營利組織負責外，其他國家大多由各國的國家圖書館、檔案館及政府資訊機關負責網站典藏系統的建立與執行，如澳洲的

「Preserving and Accessing Networked Documentary Resources of Australia, PANDORA」、美國的「Library of Congress Web Archives, MINERVA」、加拿大的「Government of Canada Web Archive, GCWA」、歐洲的「European Archive」、英國的「UK Web archiving consortium, UKWAC」、丹麥的「Netarchive.dk」、葡萄牙的「Tombo, Portuguese web archive」、日本的「Web ARchiving Project, WARP」、韓國的「Online Archiving & Searching Internet Sources, OASIS」、新加坡的「Web Archive Singapore, WAS」、中國大陸的「中國web信息博物館」(Web Infomall)等。臺灣地區目前有兩項網站典藏系統，分別是國立臺灣大學圖書館建置的「臺灣大學網站典藏庫」(NTU Web Archiving System)及國家圖書館建置的「臺灣網站典藏系統」(Web Archive Taiwan)。

以下簡介部分國際發展網站典藏系統之發展內涵及成果現況，並分項臚列如下。

(一) Wayback Machine (<http://www.archive.org/index.php>)

屬於Internet Library下計畫之一，整體計畫除網頁之外，包括文字、聲音檔、影像檔及軟體等。Internet Archive為一非營利組織於1996年成立於美國舊金山。Wayback Machine從1996年開始經由Alexa Internet及其他單位贈予典藏的網頁開始，自1999年正式開始，目前已累積超過達85兆頁的網頁資料，該網站提供網址輸入查詢典藏的網站內容，尚未提供關鍵字檢索功能。<sup>[6]</sup>

(二) Library of Congress Web Archives (Mapping the Internet Electronic Resources Virtual Archive, MINERVA) (<http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html>)

The Library of Congress Web Archives (LCWA) 是美國國會圖書館數位典藏計劃中的一個分項計畫，自2000年開始進行，以主題性方式的深入蒐集美國網站資料，完整蒐集主題相關的網站，並進行分類。至2009年5月中，瀏覽功能共提供3,008項主題，2,023項人名及3,878項網站，檢索部分可提供單一欄位如人名、網站名稱、主題、摘要、年代及語文的檢索。單一蒐藏網站除了呈現相關基本描述欄位 (Bibliographic Information) 之外，並對各網站予以概述 (Collection Overview)。該網站現與Internet Archive合作，由該單位代為擷取網站內容，技術面依據功能採取不同軟體，如網站選擇與確認 (Selection and Permission) 方面採用 Web Capture Tool (WCT)，網站徵集 (Acquisition) 方面採用Heritrix Web Crawler，在檢索服務方面因與相關單位合作，採用 WERA (Web Archive Access)、Wayback Access Tools 及 NutchWAX Searchengine。

(三) Government of Canada Web Archive, GCWA (<http://www.collectionscanada.gc.ca/index-e.html>)

2004年4月22日，加拿大圖書館暨檔案法 (The Library and Archives of Canada Act) 正式立法。而從2005年12月開始，為了保存的目的，加拿大圖書館暨檔案館 (Library and Archives Canada，簡稱LAC) 以每半年為期，開始搜集具代表性的加拿大網站，並將所搜集的資料存放在一個名為加拿大政府網站典藏庫 (The Government of Canada Web Archive，簡稱GCWA) 中。而在加拿大政府網站典藏庫裡，使用者除了可以透過關鍵字、政府部門名稱、網址列來檢索外，也可以透過以PDF等檔案存放格式的方式來檢

索。 [7]

(四) UK Web archiving consortium, UKWAC (<http://www.webarchive.org.uk/ukwa/>)

UKWAC (UK Web Archiving Consortium) 於2004年6月成立至今，已搜集約4千個網站，由British Library、The National Archives、National Library of Wales、National Library of Scotland (至2007年底)、JISC、Wellcome Trust等六個機構合作，由2005年開始有選擇性的收錄英國相關的網站，每個機構可以訂定自己的典藏政策，各有其收藏的主題。像Wellcome Library專著於醫學的網站，British Library專注於文化歷史或重大政治議題，National Library of Wales則蒐集能夠反應現代Wales生活的相關網站。UKWAC使用澳洲國家圖書館PANDORA計畫的開發工具PANDAS，因此其系統架構與澳洲PANDORA相似。此外，UKWAC在收錄網站前會先經過該網站的擁有者同意才收錄。至2009年3月底該網站已搜集4,878個網站15,173個子項，累積資料空間約達3.2 Terabytes。 [8]

(五) Preserving and Accessing Networked Documentary Resources of Australia, PANDORA (<http://pandora.nla.gov.au/>)

該計畫由澳洲國家圖書館與澳洲其他九個圖書館與文化典藏機構共同合作，自1996年開始建置PANDORA系統，為確保澳洲的網路出版品和其他網際網路上的重要資源的長期保存，故選擇性的擷取保存具長期研究及保存價值之網站資源。在檢索服務網頁上，依網站性質分為15大類，各類別下再次分小類，大類包括藝術、商業、國防、教育、環境、歷史、人文、澳洲本土、政治、科學、藝術等，採集之網站均經原單位授權公開。澳洲國家圖書館建置PANDORA網站



典藏的時間較長，故相關制度之建立較為完善，如針對網站擷取選擇原則、保存原則、詮釋資料記錄等，均訂有規範。2005年5月PANDORA為測試更大範圍的澳洲網站擷取及其他擷取程式的網站擷取功能及效益，與Internet Archive簽訂協議共同合作。

(六) Web ARchiving Project, WARP (<http://warp.ndl.go.jp/>)

日本國會圖書館(NDL)自2002年開始進行之實驗型計畫，並於2006年完成執行。該計畫中選取擬收錄之網站，透過授權書機制，取得該網站單位同意後才開始進行擷取存藏作業，設定日後開放範圍。蒐集網站對象包括全國性政府機關及機構、地方政府、合併前或合併後之城市、鄉鎮、公眾感興趣之單位組織、大學、重大議題等；網路出版品部分則以較具規劃長期出版者並清楚著錄卷期及出版日期者為對象。該網站典藏系統提供簡易查詢、電子出版品查詢及瀏覽查詢三部分。瀏覽功能分為收錄網站機關性質及網站主題類別兩種，主題分為10大類，以下系分小類，類似一般圖書之十進分類法形式。

(七) 中國Web信息博物館 (<http://www.info-mall.cn/>)

中國Web信息博物館(Web InfoMall)屬國家數位專案支援進行，由北京大學網路實驗室開發的中國網頁歷史資訊存儲與展示系統。依據該網頁文字說明，目前系統自2001年以來已累積30億以中文為主的網頁，並以平均每月4千5百萬網頁的速度擴大規模，但由於圖像檔的大小通常較大，故目前Web InfoMall並沒有保存圖檔，相關圖檔的顯示均以連線現有網頁來呈現。該系統網站目前僅提供完整網址輸入查詢。

### 三、我國建置網站典藏系統概況

目前國內已著手建置網站典藏系統的機構為臺灣大學圖書館及國家圖書館。兩者均採取選擇性網站典藏，所使用的擷取程式均為HTTRACK，兩者發展經過與現況如下：

(一) 臺灣大學網站典藏庫 (NTUWAS, <http://webarchive.lib.ntu.edu.tw/>)

國立臺灣大學圖書館自2006年7月開始規畫，初期屬該館機構典藏計畫項下，至2007年12月建置完成。系統設計包括前端讀者介面及後端管理介面。前端主要提供網站提供展示、分類、查詢、推薦四大主功能。展示部分包括典藏縮影、時光迴廊、最新典藏、熱門點選四項；分類部分包括典藏類目、主題探索、典藏時代三項；查詢部分提供網站檢索、全文檢索及進階搜尋三項；另提供讀者主動推介網站功能。依據該網站內容說明，該網站之收集對象係以臺大相關網站(完整收藏)、政府機構網站(完整收藏)、教育學術網站、藝文相關網站(各教學資源及史料檔案等)、族群相關網站、非營利社團網站。此外，為典藏時代性網站資料，設定重大事件和重要人物兩類為典藏重點。該網站典藏系統至2009年5月底止已累積蒐錄網站數量達4,259個網站。

(二) 國家圖書館—臺灣網站典藏系統

國家圖書館近十年來對館藏期刊、善本古籍及臺灣研究等文獻已進行數位化及資訊服務。但隨著資訊科技迅速發展，數位出版日趨蓬勃。網路科技使圖書館對館藏及館藏發展的定義不同於以往，亦需兼顧數位形式，以滿足讀者對數位資訊需求並兼顧全方位國家文獻徵集與典藏之核心任務。

國家圖書館於96年1月於館藏發展諮詢會議中提出建置「國家圖書館網站典藏先導

系統」(以下簡稱先導系統)的議題,並開始進行相關前置諮詢規劃。首先邀請國立臺灣師範大學陳昭珍教授及中央研究院計算中心陳亞寧組長進行兩場以網站典藏系統為主題的專題演講;96年4月至7月間,另邀請學者專家們如陳雪華教授、王梅玲教授、宋雪芳教授、陳光華教授、歐陽崇榮教授等諮詢委員,提供系統建置的建議及指導;另並參訪當時已著手進行網站典藏規劃之臺灣大學圖書館,承蒙其不吝分享建置系統的經驗。國家圖書館遂於96年9月期間開始規劃先導系統規格。

該先導系統功能係對網站定期擷取內容,並根據擷取時間建立個別的典藏版本,存放於伺服器中,除建構完備之模擬、封裝、更新、轉置等永久典藏機制外,並提供讀者查詢運用。先導計畫完成3百餘網站之擷取與典藏,選擇中央及地方政府機關、教育及學術機構、圖書館、國家圖書館及重大社會議題網站為對象,主要藉此測試、建立系統前端讀者服務及後端管理運作機制。先導系統成果於97年5月間正式開放供讀者運用。97年經諮詢委員建議更名為「臺灣網站典藏系統」(Web Archive Taiwan,簡稱WAT),另強化系統前端服務介面及後端管理功能,並新增收錄2千個網站。

該系統目前於前端服務介面主要以瀏覽及檢索兩項功能為主。瀏覽部分有主題瀏覽、機構瀏覽、地區瀏覽及重大議題瀏覽等四部份;展示部分呈現最新及熱門網站供點選;典藏檢索部分提供網址輸入、網站名稱及網頁全文檢索功能;進階查詢部分,提供更多欄位之檢索與組合查詢功能。讀者或機關單位如有建議本系統收錄的網站也可線上直接推薦,透過後臺管理端之測試、審核、

檢驗機制,可決定是否新增;首頁並將系統實際典藏網站數、檔案數統計及時呈現。

該系統所收錄網站均發函授權,徵詢各網站單位典藏及授權開放範圍意願,依據授權內容逐一進行開放範圍設定。兩年之作業經驗,深感授權作業之艱鉅與辛勞,雖然擬典藏之網站內容均為已開放於網際網路無償供各界點閱之網頁資訊,但或許國內對網站典藏觀念較為陌生,且著作權之歸屬確實非常複雜,故需對相關單位多加溝通說明。本系統於後臺管理端也建置授權管理機制,除設定開放授權範圍外,對於網站內部分網頁應授權單位要求擬不開放者,亦能逐一設定隱藏;各份回覆之授權書均予全彩掃描存檔,並建置各網站聯絡人連繫資訊。依〈圖書館法〉規定及近期經濟部智慧財產局公文內容,98年系統對於網站典藏授權主要在於徵詢各網站同意開放的範圍屬館內區域網路(Intranet)或網際網路(Internet),至98年5月底止寄發之近1,700份授權書回收率已逾58%以上,相較於文獻中記載其他國家低於20%之回收率高出甚多;已回覆之授權書中同意開放至網際網路者達80%以上。

98年度該系統仍將持續發展,除新增收錄網站外,主要在改善並強化系統作業功能,包括調整前端使用介面、增加擷取速度效益、訂定收錄網站準則、撰述標準作業流程(SOP)、持續典藏網站內容異地備份等,期使系統除充分落實國家重要網站網頁資源典藏外,在讀者檢索服務亦能更臻完善,提供研究者及各界參考運用,將網路上虛擬短暫之文獻資源,透過網站典藏系統的建置,轉化成為實質永恆。

#### 四、問題與結論



在國家圖書館建置網站典藏系統過程中，面臨許多挑戰；而審視國立臺灣大學圖書館對於該館建置網站典藏系統的篇章，兩者所歷經的困難與經驗非常類似，約歸納如下幾項：

#### (一) 網站典藏擷取之困難性與完整性

使用擷取程式時，係以網站網址設定，由程式依排程頻率自動進行。但面對鑲嵌在FLASH程式撰寫之網頁連結，則無法擷取，造成下載不完整；另對於資料庫及具權限控管之網頁資源亦無法在擷取之列，此為全世界網站典藏系統均需面對的問題。另屬於Javascript程式碼動態產生的連結，透過擷取程式無法正確抓取連結資訊，造成網頁資料遺漏。因此對於各個網站的典藏內容效益，很難有全面的典藏保障，檢閱現有之各項網站典藏系統，連結失效或內容不完整的情形很多。因此如何改善擷取程式之效益，為目前積極討論的議題。

除了網頁限制外，另一同一網站下可能同時具備其他版本（如英文版、兒童版、民眾版、學術人員版等），故進行擷取設定需由館員逐一檢視後，新增設定其他版本的擷取網址，否則會造成典藏擷取之不完整性。

#### (二) 典藏空間需求龐大

網站典藏所需資訊硬體之儲存空間龐大，且因擴增迅速，因此需相對投入巨額的經費購置設備，對置放硬體及備份資料之空間等也需納入考量。尤其屬於大型性綜合網站，網頁中包含文字、網頁、圖片、聲音、動畫及多媒體檔等多元檔案類型，不只檔案所佔空間大，每次擷取新版本時所需耗費時間也長。部分網站因均屬同一主機關下之支網，易造成擷取內容重複，致典藏空間浪費，如我國駐外單位網站則有類似情形發

生；另如政府公報網站其全文pdf檔與文字檔並存；或部分學校網站之電子相簿項下收藏有數以千計的照片檔，擷取後均需極大的儲存空間。目前國家圖書館的網站典藏系統，已能做到版本比對功能，亦即網站新版本於擷取完成後，系統會自行比對前一版本的差異性，僅保留有變更的網頁，重複者自動刪除，以節省儲存空間。但因各網站設定為完整擷取，故空間成長速度仍然驚人，如欲節省空間，尚需人工介入檢視，重新設定恰當的擷取層次。

#### (三) 網站典藏編目問題

世界各國對圖書館的圖書編目格式較有共同規範，便於資料交換與交流。但對於網站網頁所需採取之詮釋資料格式（Metadata Format）則尚未有定論。較具規模者為澳洲PANDORA依據《典藏電子資源編目手冊》（PANDORA Archive Electronic Cataloguing Manual）完成目錄資料，並透過國家書目資料庫提供檢索，同時也能在各參與合作館的館藏目錄查詢。<sup>[9]</sup>各國現有之網站典藏前端服務功能繁簡不一，很難全面檢驗後端對各網站建置的詮釋資料格式。國家圖書館正著手規畫由網站典藏系統轉出詮釋資料，匯入圖書館館藏目錄系統的可行性。

#### (四) 授權作業之必要性檢視

世界其他國家進行網頁典藏計畫時，授權問題是一項無法忽略的重要課題。目前舊金山非營利組織Internet Archive的網站典藏系統，採直接擷取未經授權；但依文獻記載澳洲、英國、日本則逐一經原網站授權。授權徵詢是一項艱鉅且長期性的作業，因網站典藏具不同日期版本擷取的特性，故進行授權連繫作業時，網站性質相異，該網站負責單位的哪一層級具有決定授權與否及程度的權

利？國外文獻均稱取得「Webmaster」同意，但以政府機關網站為例，網站維護者未必具決定權，故需由哪一層級長官決定，亦是國家圖書館進行授權連繫時，經常被問及，卻難有標準答案的議題。另網頁內容多元，如內容另包括其他個人或單位之資料，是否還需徵詢原提供者的授權？此外，各網站授權是一次授權即可，還是因網站內容新增更易，每一版本都需經過授權？如果各版本均需授權，可想見更是一項極浩大的工程。故網站典藏之授權作業，如何在資訊典藏、資訊傳播與著作權尊重三者間取得平衡與共識，仍是計畫執行時的一項重大挑戰。

#### （五）圖資界與資訊界的合作

網站典藏屬數位典藏之一環，對圖書館而言是一項新的發展領域，依性質需結合圖書資訊界與資訊科技界的專業知識，攜手合作才能圓滿達成目標。以國家圖書館為例，圖資界對於讀者資訊服務面，在介面設計方面能提供較多的數位服務經驗，但是在網站典藏所需之基本核心擷取功能、系統架構、系統運作、資料儲存及備份機制方面，則需大幅倚重資訊專業的同仁協助。故網站典藏系統之成敗與否，需資訊專業與圖資界兩方充分搭配合作，與時俱進，才能發揮最佳效益。而世界許多網站典藏專案計畫之進行，也多由資訊專業單位開始。如中國之中國信息博物館（Web InfoMall），由北京大學網路實驗室開發建設；國立臺灣大學圖書館網站典藏系統之規劃，主要也是由圖書館的資訊單位開始進行。

#### （六）建立跨機構單位之網站典藏合作機制

因網站資源豐富，網站及網站內容更易消長頻繁，目前各國建置網站典藏系統之目的，莫不以保存各國重要網路資源為首要任

務。如欲達此目的，則需兼顧橫向典藏網站對象之全面性及縱向網站版本之及時性與完整性，故擬應納入蒐集典藏之網站內容數量將非常可觀。因此，網站典藏計畫之推動，除了前述圖資界及資訊界之共同合作外，更需全國跨機構單位之共同攜手合作。目前國際間如澳洲及英國等國家型的網站典藏計畫，均屬跨機構單位間之合作機制，各機構單位依其職掌或區域特質，分工擔任不同主題對象之網站典藏，眾志成城，攜手合作，將所有資源充分運用，擷取成果整合分享，透過共建共享，如此當能發揮最大的效益，此點頗值得國內進行網站典藏計畫時借鏡參考。

#### 註釋

- [1] Internet World Stats: Usage & Population Statistics, Internet Usage in Asia, access at Apr.19, 2009, from <http://www.internet-worldstats.com/stats3.htm#asia>.
- [2] 陳光華，〈臺灣大學網站典藏系統之研發〉，《大學圖書館》11卷2期（民96.9），頁2。
- [3] Library of Congress(2003), Web Capture & Archiving, Retrieved May 17, 2009, from <http://www.loc.gov/acq/devpol/webarchive.html>.
- [4] Bibliotheca Alexandrina, International School of Information Science (ISIS), Web Archive, Retrieved May 17, 2009, from [http://www.bibalex.org/isis/frontend/archive/archive\\_web.aspx](http://www.bibalex.org/isis/frontend/archive/archive_web.aspx).
- [5] 參考Wikipedia, Web Archiving, Retrieved May 10, 2009, from [http://en.wikipedia.org/wiki/Web\\_archiving](http://en.wikipedia.org/wiki/Web_archiving).
- [6] 查詢Wayback Machine網頁資訊，上網日期：2009.05.17。網址：<http://www.archive.org/web/web.php>。
- [7] Government of Canada Web Archive, Posted November 21, 2007, from <http://agiig.wordpress.com/2007/11/21/government-of-canada-web-archive/>.



- [8] 有關UKWAC計畫可參閱“UKWAC: Building the UK's First Public Web Archive”一文, Bailey & Thompson, 2006, from <http://www.dlib.org/dlib/january06/thompson/01thompson.html>。
- [9] PANDORA, Australia's Web Archive, PANDORA Overview, from <http://pandora.nla.gov.au/overview.html>。

#### 參考書目

- 楊道玲。〈Web資源採集與保存研究〉(武漢大學檔案學科碩士論文。2005.5)。
- 王志庚。〈Web Information Preservation at National Library of China〉,「中歐數字資源長期戰略保存研討會」簡報資料。2004.7.13-17。
- 王麗蕉。〈網路資源長期保存:以多元層次描述模式建構之探討〉,《教育資料與圖書館學》44:4 (Summer 2007),頁455-471。
- 林巧敏。〈各國國家圖書館網站資源典藏計畫之探討〉,「由圖書館學到圖書資訊學—跨越兩個世紀」研討會會議論文。頁30。
- 林巧敏。〈澳洲國家圖書館數位館藏發展概述〉,《國家圖書館館刊》,96年第1期(民96.6)。頁169-197。
- 林巧敏、陳雪華。〈國家圖書館電子館藏發展策略之研究〉,《國家圖書館館刊》97年第1期(民97.6)。頁25-61。
- 陳力、郝守真、王志庚。〈網絡信息資源的採集與保存—國家圖書館的WICP和ODBN項目介紹〉,《國家圖書館學刊》2004年第1期。頁2-6。
- 陳光華。〈網站典藏的現況與發展〉。引自<http://jalis.njau.edu.cn/wp-content/uploads/2009/04/20090331webarchiving.pdf>。上網日期:98.6.21。
- 陳光華。〈臺灣大學網站典藏系統之研發〉,《大學圖書館》11卷2期(民96.9)。頁1-18。
- 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(民95.9)。頁2-18。